



Statistik Austria Qualitätsrichtlinien

Version 2.1
Stand 14.03.2017



Inhaltsverzeichnis

VORWORT	3
1. Qualität in der Amtlichen Statistik	5
1.1 Definition	5
1.2 Rahmenbedingungen	8
2. Einsatz fundierter Methoden	10
3. Qualitätsstandards im statistischen Produktionsprozess	13
3.1 Statistischer Produktionsprozess	13
3.2 Grundlagenarbeiten	15
3.2.1 Planung statistischer Projekte.....	15
3.2.2 Konzepte, Definitionen und Klassifikationen	18
3.2.3 Statistische Register	21
3.3 Datenerhebung und -beschaffung	24
3.3.1 Arten der Datenbeschaffung	24
3.3.2 Verwendung von Verwaltungsdaten	29
3.3.3 Fragebogendesign und –test	33
3.3.4 Maßnahmen zur Steigerung der Rücklaufquote	38
3.4 Datenaufbereitung (Datenbearbeitung)	41
3.4.1 Datenerfassung und Signierung.....	41
3.4.2 Plausibilitätsprüfung	43
3.4.3 Imputation.....	45
3.4.4 Verknüpfung verschiedener Datenquellen	48
3.5 Auswertung und Analyse	51
3.5.1 Arbeits- und Analysetabellen als Grundlage von Publikationen	51
3.5.2 Schätzung aus Stichprobenerhebungen	53
3.5.3 Berechnung von Indikatoren – statistische Maßzahlen	57
3.5.4 Indizes	59
3.5.5 Gesamtrechnungen.....	62
3.5.6 Modellbasierte Schätzungen und Prognosen	65
3.5.7 Zeitreihen und deren Bereinigung.....	67
3.5.8 Flash Estimates.....	70
3.5.9 Sonderauswertungen	73
3.6 Veröffentlichung statistischer Ergebnisse	75
3.6.1. Darstellung als Tabelle, Text und in visualisierter Form	75
3.6.2. Open Data	79
3.6.3 Statistische Datenbank STATcube	81
3.6.4 Datenübermittlung an Eurostat	83
3.6.5 Internationale Anfragen.....	85
3.6.6 Revisionen.....	87
3.6.7 Geheimhaltung	89
3.6.8 Dokumentation und Qualitätsberichterstattung.....	92
4. Maßnahmen zur Qualitätssicherung	96
Anhang 1: Europäische Statistiken: Verhaltenskodex	99
Anhang 2: Fundamentale Prinzipien der Amtlichen Statistik der Vereinten Nationen	107
Anhang 3: Task Force zur Überarbeitung der Qualitätsrichtlinien	108
Glossar	109
Bibliographie	119

VORWORT

In ihrem [Leitbild](#) versteht sich die **Bundesanstalt Statistik Österreich** (im Folgenden kurz „Statistik Austria“ oder „Bundesanstalt“ genannt) als führender Informationsdienstleister Österreichs, der sich zur Erstellung hochwertiger Statistiken und Analysen verpflichtet. Um dem im vollen Umfang nachkommen zu können, ist zu gewährleisten, dass die zu erstellenden Produkte nach wissenschaftlich anerkannten Methoden und Standards erarbeitet werden. Das vorliegende Dokument ist als Erweiterung und Ergänzung des Leitbildes zu sehen. Für die Produzentinnen und Produzenten Amtlicher Statistiken in Österreich sind die Richtlinien eine wesentliche Orientierung in dem Sinn, dass die einzelnen Produktionsschritte unter Beachtung der in diesem Dokument festgeschriebenen Standards und Regeln erfolgen. Den Benutzerinnen und Benutzern statistischer Ergebnisse soll kommuniziert werden, dass Produkte von Statistik Austria unter Anwendung modernster statistischer Methoden erstellt werden und Standards garantiert sind, die eine hinreichende Qualität des statistischen Datenmaterials sicherstellen. Derartige Standards können einerseits durch internationale Vorgaben und Rechtsakte gegeben sein – so wie etwa in den Volkswirtschaftlichen Gesamtrechnungen. Andererseits gibt es auch implizit gegebene Standards, die in der Verwendung bestmöglicher Methoden („Current Best Practices“) liegen oder durch den Anspruch im internationalen Vergleich ein entsprechendes Niveau zu erreichen, gegeben sind. Arbeiten und Vorgehensweisen, die eine hinreichende Ergebnisqualität garantieren sowie ein hohes Qualitätsbewusstsein haben, haben in der Amtlichen Statistik in Österreich eine lange Tradition.

Das Bekenntnis zur Qualität ist auch Bestandteil für das statistische System wichtiger rechtlicher Normen. National betrifft dies das Bundesstatistikgesetz 2000 in der geltenden Fassung (BStatG 2000). Die Arbeit hinsichtlich Qualität auf dem Gebiet der Amtlichen Statistik ist stark durch die internationale Ebene geprägt. So ist auch die für die Bundesanstalt gültige mehrdimensionale Definition der Produktqualität jene, die durch Eurostat entwickelt wurde. Des Weiteren ist der Verhaltenskodex („European Code of Practice“, siehe Anhang 1) in diesem Zusammenhang zu erwähnen, der im Prinzip 4 (Verpflichtung zur Qualität) ein eindeutiges Bekenntnis zu Qualität einfordert. Dies umfasst auch die im Qualitätssicherungsrahmen des Europäischen Statistischen Systems (ESS) festgelegte Notwendigkeit der Schaffung und Veröffentlichung von Qualitätsrichtlinien.

Entscheidend für die Validität statistischer Ergebnisse ist es, dass die betreffenden Arbeitsschritte methodisch fundiert zumindest nach „Good Practices“ erfolgen. Nach Maßgabe ist die Realisierung von „Best Practices“ anzustreben. Das vorliegende Dokument, das für Mitarbeiterinnen und Mitarbeiter von Statistik Austria eine wichtige Referenz darstellt, soll Datennutzerinnen und -nutzern vermitteln, dass die gelebte Praxis bei der Produktion der Ergebnisse auf größtmögliche Qualität unter Berücksichtigung gegebener Randbedingungen abzielt. Darunter fallen:

- Verfügbarkeit ausreichender Ressourcen, was sowohl die personellen Verfügbarkeiten als auch die Sachmittel betrifft.
- Größtmögliche Schonung der Respondenten
- Gegebene budgetäre Bedeckung

Im Zentrum der Erstellung statistischer Produkte liegt der statistische Produktionsprozess, der im Detail hoch komplex sein kann, im Wesentlichen jedoch zumeist aus drei Teilen besteht: Datenbeschaffung, Datenbearbeitung und Datenpräsentation.

In diesem Zusammenhang muss erwähnt werden, dass die Rahmenbedingungen für die Amtliche Statistik einem laufenden Wandel unterworfen sind. Die Einsatzmöglichkeiten neuer Technologien und die Änderung rechtlicher Gegebenheiten, sowie die Verknappung der Ressourcen bei gleichzeitig steigender Nachfrage nach statistischen Produkten bringen es

mit sich, dass Innovationen im statistischen Produktionsprozess eine absolute Notwendigkeit darstellen. Die damit verbundene Erweiterung der Typen statistischer Produkte - vor allem was die Vielfalt der Produktionsmöglichkeiten darstellt - bildet eine entscheidende Herausforderung für nationale statistische Institute.

Kapitel 1 geht auf rechtliche Rahmenbedingungen ein und definiert Qualität in einer für statistische Produkte geeignete Weise. Im Zuge dessen wird auch der Rahmen („Quality Framework“) erläutert, der auch für Statistik Austria stark durch den europäischen Qualitätsrahmen geprägt ist. Der Nachweis des Einsatzes fundierter Methoden ist, wie bereits erwähnt, ein auch im internationalen Vergleich wichtiger Bewertungsschritt, der das Vertrauen in die veröffentlichten Ergebnisse stärken soll. In Kapitel 2 finden sich eine Reihe möglicher Bewertungskriterien, an Hand derer die Feststellung, wann Methoden als geeignet betrachtet werden können, vorgenommen werden kann.

Die eigentlichen Richtlinien für die Produktion von Statistiken werden in Kapitel 3 entwickelt. Es werden jene Arbeitsschritte betrachtet, die für die Produktion Amtlicher Statistiken von Wichtigkeit sind. Wie bereits erwähnt, bilden sich die entwickelten Richtlinien entlang des statistischen Produktionsprozesses ab. Die Leserinnen und Leser werden bemerken, dass die formulierten Richtlinien von unterschiedlicher Imperativität getragen sind. Mit dem bei einigen Richtlinien verwendeten Zeitwort „muss“ wird zum Ausdruck gebracht, dass ein Unterschreiten der in der Richtlinie geforderten Standards nicht möglich ist. Zum Unterschied dazu beschränken sich einige Richtlinien darauf, Vorgangsweisen einzufordern („etwas soll geschehen“), sie aber nicht als unbedingt verpflichtend anzusehen. Dies hat den Hintergrund, dass bei manchen Arbeitsschritten, Vorgänge bzw. Methoden zwar wünschenswert, jedoch durch praktische Gegebenheiten nicht immer vollständig realisierbar sind. Diese formal schwächere Stringenz ändert nichts an der Tatsache, dass verantwortliche Mitarbeiterinnen und Mitarbeiter von Statistik Austria mannigfaltige Anstrengungen unternehmen, die Einhaltung auch derartiger Richtlinien zu gewährleisten.

Um die Einhaltung der ambitionierten Ziele und hohen Maßstäbe aus Kapitel 3 erreichen zu können, müssen die Arbeitsschritte und relevanten Prozesse durch geeignete Maßnahmen zur Qualitätssicherung und Überprüfung begleitet werden. In diesem Sinne sind die Verfahren und Methoden, mit denen die Einhaltung der zuvor beschriebenen Richtlinien und Standards erreicht bzw. überprüft werden, Gegenstand des Kapitels 4.

Die erste Version der Qualitätsrichtlinien wurde im Jahr 2010 erarbeitet. Um der fortschreitenden Modernisierung und den Entwicklungen im statistischen Produktionsprozess Rechnung zu tragen, wurden die Richtlinien im Jahre 2015 einer Überarbeitung unterzogen und 2016 neu aufgelegt. Dies geschah im Rahmen einer hausinternen Task-Force (siehe Anhang 3), in der sämtliche Fachbereich des Hauses durch erfahrene Mitarbeiterinnen und Mitarbeiter vertreten waren. Das Resultat dieser Überarbeitung ist die nunmehr vorliegende Version 2.1.

1. Qualität in der Amtlichen Statistik

1.1 Definition

Hinterfragt man den Begriff der ‚Qualität in der Statistik‘, muss bewusst werden, dass es sich dabei um **keine eindimensionale Begrifflichkeit** handeln kann. Vielmehr gilt es, eine Vielzahl von Aspekten und damit verbundene Komponenten darunter zu subsumieren. Das gegenständliche Dokument setzt sich mit der Qualität statistischer Produkte auseinander. Die gleichermaßen bedeutsame Qualität in Relation zum statistischen Produktionsprozess wird im Rahmen des bei Statistik Austria institutionalisierten Internen Kontrollsystems (IKS) behandelt und in regelmäßigen Zeitabständen einer Evaluierung unterzogen. In der Folge bezieht sich der Begriff „Qualität“ immer auf Produktqualität.

Die ISO Norm 8402 aus 1994 definiert:

„**Qualität** ist die Gesamtheit der **Eigenschaften und Charakteristiken** eines Produkts oder einer Dienstleistung, die darauf basieren, **festgestellte oder implizierte Anforderungen** zu erfüllen.“

Gemäß dieser Definition lassen sich für die Praxis der statistischen Produktion zwei Fragen ableiten:

- Was sind die wesentlichen **Eigenschaften und Charakteristiken** eines statistischen Produktes?
- Welche **Anforderungen** können festgestellt werden und welche sind implizit gegeben?

Erste Überlegungen, wie Qualität im Kontext der Amtlichen Statistik zu definieren ist, wurden seitens Eurostat bereits Mitte der 90er Jahre angestellt. Mit der 1998 ins Leben gerufenen „**Arbeitsgruppe Qualität**“ wurde die Definition der „Qualität“ in der Amtlichen Statistik erstmals den Mitgliedstaaten kommuniziert. Im Rahmen der im Jahr 2000 etablierten Leadership Group Quality (LEG – Quality) kam es schließlich zu der seit dem Jahr 2003 etablierten gültigen Version der Definition. Demnach bezieht sich Qualität in der Statistik auf folgende sechs Dimensionen (auch Kriterien genannt):

- Relevanz
- Genauigkeit
- Aktualität und Rechtzeitigkeit
- Zugänglichkeit und Klarheit
- Vergleichbarkeit
- Kohärenz

Dabei bedeuten die einzelnen Dimensionen:

Relevanz

Relevanz reflektiert, in wie weit statistische Produkte mit potentiellen Wünschen von Benutzerinnen und Benutzern übereinstimmen. Wird durch das Produkt die Gesamtheit der Wünsche abgedeckt?

Genauigkeit

Genauigkeit im engen statistischen Sinn beschreibt die Nähe von Berechnungen oder Schätzungen zum exakten oder wahren Wert.

Aktualität und Rechtzeitigkeit

Aktualität spiegelt den Zeitunterschied zwischen dem Auftreten eines statistischen Phänomens und dessen Beschreibung wider (Referenzperiode vs. Veröffentlichung).

Rechtzeitigkeit bezieht sich auf die zeitliche Distanz zwischen der geplanten Fertigstellung des Produktes (Bezug nehmend auf Zeitpunkte die vorher durch rechtliche Grundlagen wie z. B. Vereinbarung, Vertrag, Gesetz, aber auch durch Projektpläne etc. festgelegt wurden) und dessen Veröffentlichung.

Zugänglichkeit und Klarheit

Zugänglichkeit bezieht sich auf die Möglichkeiten, wie Nutzerinnen und Nutzer auf Ergebnisse bzw. Daten zugreifen können: Veröffentlichungsmedien, -plattformen und -formate, Zugang zu Mikrodaten etc.

Klarheit beschreibt die Informationsumgebung des Produktes. Sind die Daten ausreichend durch Metadaten dokumentiert? Ist die tabellarische, textuelle und grafische Darstellung ausreichend und verständlich? Gibt es Informationen über die Qualität der Daten?

Vergleichbarkeit

Vergleichbarkeit zielt - Bezug nehmend auf **ein statistisches Produkt** - auf die Beschreibung etwaiger Auswirkungen von Unterschieden in den verwendeten Konzepten, Messinstrumenten und Verfahren hinsichtlich zeitlicher und/oder räumlicher Komponenten bzw. nach sonstigen Vergleichskriterien ab.

Kohärenz

Kohärenz zwischen Statistiken beurteilt die Vergleichbarkeit **verschiedener statistischer Produkte**, die prinzipiell zu unterschiedlichen Zwecken generiert wurden, aber Schnittmengen bezüglich gemeinsamer Gegenstände aufweisen.

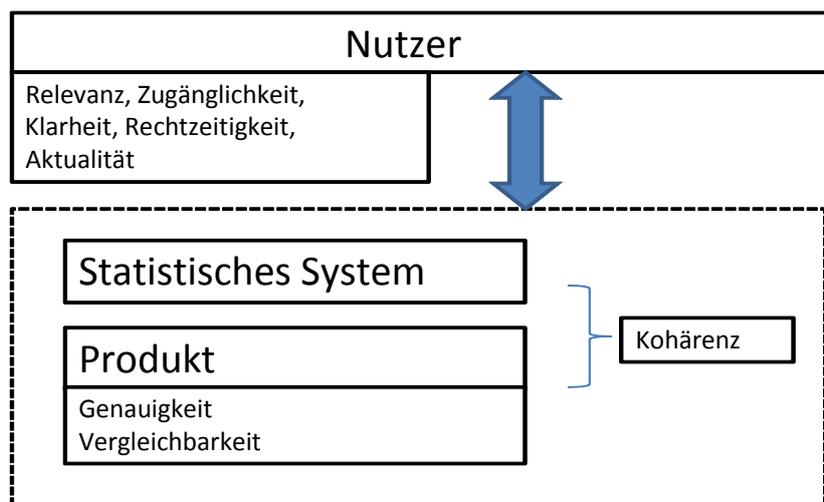


Abbildung 1: Die Qualitätskriterien im Verhältnis zu Nutzerinnen und Nutzern sowie in der Verankerung in ein statistisches System

Abbildung 1 zeigt, dass Amtliche Statistiken nicht stand alone - quasi um ihrer selbst willen - produziert werden, sondern einerseits eingebettet in ein statistisches System existieren und

andererseits massiv durch **Nutzerbedürfnisse** bestimmt werden. Die Qualitätskriterien sind auch in dem Dreigestirn Produkt-Nutzer-Statistisches System interpretierbar. Während Genauigkeit und Vergleichbarkeit klassische produktinhärente Qualitätsdimensionen sind, beurteilen Kriterien wie Relevanz, Zugänglichkeit und Klarheit sowie Rechtzeitigkeit und Aktualität genau das Verhältnis dieser 3 Ebenen. Die **systemische Integration** ist insofern erwähnenswert, da im Falle der Amtlichen Statistik kaum ein Produkt denkbar ist, das keine inhaltliche Relation zu einer Familie ähnlicher Statistiken aufweist. Kohärenz ist eine Komponente, die diesen Verbindungen dahingehend Rechnung trägt, den Datennutzerinnen und -nutzern zu vermitteln, in wie weit Ergebnisse verschiedener Produkte zu gemeinsamen Gegenständen **kombiniert nutzbar** sind.

Wie bereits erwähnt, umfasst Qualität in der Statistik mehrere Gesichtspunkte. Diese **Multidimensionalität** spiegelt sich auch in der Definition wider. Betrachtet man die Dimensionen, so wird deutlich, dass jede einzeln für sich einen wichtigen singulären Einflussfaktor auf die Gesamtqualität darstellt. Geht man jedoch einen Schritt weiter und betrachtet die Kriterien im wechselseitigen Zusammenspiel, erkennt man, dass für das Erreichen gewisser Zielvorgaben eine Konkurrenzsituation gegeben ist. Diese „**Trade off**“ Situation wird bei der gemeinsamen Betrachtung von Genauigkeit versus Aktualität am deutlichsten. Dies ist aber nur ein plakatives Beispiel. Man kann sogar so weit gehen, zu sagen, dass bei beliebiger Auswahl zweier oder mehrerer Dimensionen es zu konkurrierenden Zielsetzungen kommen wird. Diese Situation macht die Erreichung einer optimalen Ergebnisqualität für die Amtliche Statistik nicht einfacher.

Ziel ist es jedenfalls, Qualität hinsichtlich sämtlicher der zuvor angesprochenen Kriterien zu optimieren. In diesem Kontext spielt naturgemäß der Aspekt der **Qualitätsmessung** eine wichtige Rolle. Betrachtet man auch hier die Multidimensionalität, ist es leicht einsehbar, dass Qualität in der Statistik nicht durch eine einzelne singuläre quantitative Zahl darstellbar ist. Es ist vielmehr die gedankliche Synthese aller evaluierten Komponenten, die sich zu einem Gesamtbild verdichten sollen. Eine Gesamtbeurteilung eines statistischen Produktes hängt auch davon ab, wie die einzelnen Dimensionen **gewichtet** werden. Dies ist nicht nur vom jeweiligen Produkt abhängig, sondern kann von Anwendung zu Anwendung differieren und ist im Extremfall sogar von Nutzer zu Nutzer unterschiedlichen subjektiven Einschätzungen unterworfen. Neben der Problematik **unterschiedlicher Gewichtung** gilt die Annahme, dass der Qualitätsbegriff durch die sechs beschriebenen Dimensionen zu einem hohen Maß abgedeckt ist. Jedoch ist nicht auszuschließen, dass ein gewisser, **nicht erkannter Rest** bleibt, der auch zur Qualität beiträgt, jedoch nicht in den sechs Dimensionen enthalten ist. *Abbildung 2* verdeutlicht anhand zweier Beispiele das zuvor Gesagte. Dabei entspricht die Breite der Rechtecke der jeweiligen Gewichtung der Qualitätsdimensionen.



Abbildung 2: Die Gewichtung für die einzelnen Qualitätsdimensionen kann unterschiedlich sein.

Die Qualitätsdefinition ist ein wichtiger konzeptiver Beitrag zur **Produktqualität** im Rahmen der Arbeit der Amtlichen Statistik. Produktqualität nimmt im Rahmen der Qualitätsarbeit von Statistik Austria eine zentrale Position ein. Auch die in diesem Dokument vorliegenden Qualitätsrichtlinien sind als Beitrag zur Produktqualität zu verstehen. Es soll nicht unerwähnt bleiben, dass auch die anderen Säulen der Qualitätsarbeit entscheidenden Einfluss auf die Arbeit und Unternehmenskultur von Statistik Austria haben. Diese Eckpfeiler sind durch Reproduzierbarkeit, Effizienz, Kundenorientierung, Respondentenentlastung und Mitarbeiterorientierung gegeben.

Reproduzierbarkeit umfasst die Automatisierung bzw. Dokumentation sämtlicher Prozessschritte vom Rohdatensatz bis zur veröffentlichten Statistik. So soll sichergestellt werden, dass die Ergebnisse unabhängig von der bearbeitenden Person sind bzw. dass auch andere Mitarbeiterinnen und Mitarbeiter in einer Wiederholung dieselben Ergebnisse produzieren würden.

Im Bereich der **Effizienz** können die Implementierung von Tools und Methoden zur Projektplanung genannt werden, die für die Erreichung vor allem betriebswirtschaftlicher Ziele unerlässlich sind. Permanente Weiterentwicklung der Informationsumgebung und des Datenzugangs sowie die Einführung eines Veröffentlichungskalenders sind **kundenorientierte Maßnahmen**, wozu auch die Nutzerinnen- und Nutzerbefragung zu zählen ist, die im 2-jährigen Abstand durchgeführt wird.

Die **Entlastung der Respondentinnen und Respondenten** ist in Zeiten, in denen Bürgerinnen und Bürger sowie Unternehmen hohen administrativen Belastungen ausgesetzt sind, ein besonders wichtiges Anliegen für Statistik Austria. In diesem Sinne erfolgen permanent Maßnahmen zur Optimierung der Datenbeschaffung, wobei in erster Linie bereits vorhandene Daten genutzt, sowie neue Datenquellen erschlossen werden müssen, die nicht mit Direktbefragungen verbunden sind. Um die durch die Amtliche Statistik verursachte Belastung der Unternehmen bzw. Betriebe zu messen, ist seit 2001 das Belastungsbarometer ein fester Bestandteil bei allen verpflichtenden wirtschaftsstatistischen Erhebungen.

Letztendlich sind die **Mitarbeiterinnen und Mitarbeiter von Statistik Austria** das wichtigste Kapital, mit dem hochwertige statistische Produkte erstellt werden. Daher wurde ein umfangreiches Aus- und Weiterbildungsprogramm implementiert, das auch laufend überarbeitet wird, wodurch den Mitarbeiterinnen und Mitarbeitern die Möglichkeit geboten wird, ihre Qualifikation und ihr Know-how auf hohem Niveau zu halten bzw. zu verbessern. Dies ist insbesondere auch vor dem Hintergrund eines sich wandelnden Anforderungsprofils von StatistikexpertInnen von Bedeutung, das für die Erschließung neuer Datenquellen (Verwaltungsdaten, Big Data etc.) erforderlich ist. Des Weiteren wird in der alle zwei Jahre durchgeführten Mitarbeiterinnen- und Mitarbeiterbefragung die Zufriedenheit der Arbeitnehmerinnen und Arbeitnehmer hinsichtlich relevanter Determinanten erfragt.

1.2 Rahmenbedingungen

Arbeiten, die die Qualität der Amtlichen Statistik betreffen, sind unter dem Blickwinkel bestimmter geltender Rahmenbedingungen zu sehen, die teilweise rechtlich bindend sind. Auf nationaler Ebene ist vor allem das BStatG 2000 zu erwähnen.

Im § 24 werden dort besondere Grundsätze bei der Aufgabenwahrnehmung angesprochen, die da lauten:

1. Objektivität und Unparteilichkeit bei der Erstellung der Statistiken;
2. Anwendung statistischer Methoden und Verfahren nach international anerkannten wissenschaftlichen Grundsätzen und Standards und deren Offenlegung;
3. Laufende Überprüfung der Statistiken auf Qualitätsverbesserungen;

4. Sicherstellung einer möglichst hohen Aktualität der Statistiken;
5. Minimierung der Belastung und ausreichende Information der Betroffenen und Auskunftspflichtigen;
6. Wahrung der Grundsätze der Veröffentlichung;
7. Vertraulichkeit von personenbezogenen Daten.

Die explizite Aufnahme dieser Grundsätze in die nationale statistische Rechtsgrundlage unterstreicht die Wichtigkeit aller Qualitätsaspekte, deren Beachtung in der Amtlichen Statistik in Österreich auch schon davor eine lange Tradition hat.

Rahmenvorgaben für Qualität in der Amtlichen Statistik sind naturgemäß auch im internationalen Kontext vorhanden. Dass dies auf sehr hoher Ebene im Sinne umfassender internationaler Abdeckung geschieht, ist an den **Fundamentalen Prinzipien** für Amtliche Statistik der Vereinten Nationen (UN) zu sehen (siehe Anhang 2). Diese insgesamt 10 Grundsätze umfassen wichtige Aspekte wie Einhaltung wissenschaftlicher Standards, Reaktion auf missbräuchliche Verwendung, Geheimhaltung vertraulicher Daten und Transparenz von Maßnahmen zur Erstellung von Statistiken. Qualität ist explizit im Grundsatz 5 angesprochen, der die Nationalen Statistischen Institutionen (NSI's) verpflichtet, möglichst hochwertige Datenquellen heranzuziehen.

Wie bereits erwähnt, ist die Konsistenz im Europäischen Statistischen System (ESS) auch für Statistik Austria eine entscheidende Vorbedingung. Nicht zuletzt deshalb sind die Rahmenbedingungen auf europäischer Ebene von großer Wichtigkeit. Das legislative Rückgrat für die Erstellung von Statistiken innerhalb des ESS ist die Statistik **Verordnung (EG) Nr. 223/2009** des Europäischen Parlaments und des Rates vom 11. März 2009 über europäische Statistiken – kurz Statistik-Verordnung genannt.. Durch diese Verordnung werden wichtige Aspekte der statistischen Governance angesprochen. So ist in Artikel 12 die zuvor gegebene Definition der für die Statistik **gültigen Qualitätskriterien**¹ verankert, der die Mitgliedstaaten verpflichtet, den Einsatz von „**Best Practices**“ einzufordern, um Qualitätsstandards gerecht zu werden, die entlang der Kriterien zu definieren sind. Die Verordnung wurde 2013 bis 2015 nach einem eingehenden Diskussionsprozess einer Revision unterzogen, bei der auch die Stärkung der Rolle des Verhaltenskodex als Instrument der Sicherstellung des Vertrauens in amtliche Statistiken ein wichtiges Anliegen war. Die Verordnung 2015/759 zur Änderung der Verordnung 223/2009- auch was die Artikel 11 und 12 anlangt wurde am 29.04.2015 veröffentlicht.

In Artikel 11 der Statistik-Verordnung ist der **Verhaltenskodex („Code of Practice“)** verankert. Dieser 15 Prinzipien umfassende, in der Erstversion durch eine Empfehlung der Kommission vom 25. Mai 2005 veröffentlichte Governance-Träger wurde in den Jahren 2009 bis 2011 einer Revision unterzogen und in der Neuversion nach Annahme durch den Ausschuss für das Europäische Statistische System (AESS) im Jänner 2012 durch Eurostat veröffentlicht. Die Prinzipien dienen der Einhaltung der Unabhängigkeit, Integrität und Professionalität der nationalen und gemeinschaftlichen Stellen innerhalb des ESS (siehe Anhang 1). Qualität spielt innerhalb des Verhaltenskodex eine zentrale Rolle. Dies manifestiert sich sowohl in den Prinzipien 11 bis 15, die die Qualitätsdimensionen widerspiegeln, als auch im Prinzip 4 „**Verpflichtung zur Qualität**“. Statistik Austria bekennt sich voll inhaltlich zu den Prinzipien des Verhaltenskodex und wurde bisher auch zwei sogenannten Peer Reviews unterzogen. Der erste im Jahr 2006 behandelte die Indikatoren der Prinzipien 1 bis 6 und 15 wohingegen bei der zweiten Überprüfung 2014 alle Prinzipien inkludiert waren. Mit der Vorlage dieses Dokuments wird dem Indikator 4.1 Rechnung getragen, der fordert:

„Die Qualitätspolitik ist festgelegt und öffentlich zugänglich. Für das Qualitätsmanagement sind Organisationsstruktur und -instrumente vorhanden.“

¹ Die Verordnung definiert im Artikel 12 die Qualitätsdimensionen. Die Textierung zu den einzelnen Dimensionen ist in der Verordnung etwas anders gewählt als in diesem Dokument.

Das vorliegende, der Öffentlichkeit zugängliche Dokument ist somit auch als ein Ausdruck des Bekenntnisses von Statistik Austria zum Verhaltenskodex und im Speziellen zum Prinzip 4 zu sehen.

2. Einsatz fundierter Methoden

Im Leitbild von Statistik Austria sind Grundsätze, wie Produktion hochwertiger Statistiken und Analysen, möglichst hohe Produktqualität und optimale Effizienz bei der Erstellung von Statistiken verankert. Dies mündet letztendlich in den Anspruch, der führende Informationsdienstleister Österreichs zu sein, und bringt die Verpflichtung mit sich, die Art und Weise der Produktion von Statistiken auf Basis des neuesten Standes der wissenschaftlichen Entwicklung permanent zu hinterfragen. In diesem Sinne ist die Erweiterung der in den jeweiligen Fachbereichen historisch gewachsenen Kernkompetenzen ein vorrangig zu unterstützendes Ziel für die Mitarbeiterinnen und Mitarbeiter von Statistik Austria. Dies wird auch durch die unter dem Aspekt der Mitarbeiterorientierung ausgesprochenen Bekenntnisse zur beruflichen Weiterbildung und damit Weiterentwicklung zu innovativem Denken, Eigenverantwortung, team- und projektorientiertem Arbeiten sowie dem gelebten kooperativen Führungsstil explizit zum Ausdruck gebracht..

Die angesprochenen Kernkompetenzen betreffen sämtliche Schritte des **statistischen Produktionsprozesses** (auf ihn wird im Kapitel 3.1 näher eingegangen), der im Wesentlichen aus Planung, Datenbeschaffung, Datenverarbeitung und Datenpräsentation besteht. Vorgangsweisen, mit denen einzelne Arbeitsschritte innerhalb des Produktionsprozesses in Angriff genommen und durchgeführt werden, können sehr allgemein unter dem Begriff (**statistische**) **Methoden** subsummiert werden. Das Adjektiv „statistische“ wird hier deshalb in Klammer gesetzt, da nicht alle eingesetzten Verfahren und Methoden, die im Zuge der Erstellung Amtlicher Statistiken zum Einsatz kommen, wissenschaftlich dem Fachgebiet der Statistik zuzuordnen sind. Neben statistischen- und betriebswirtschaftlichen Methoden kommen auch Methoden des Projektmanagements und der Informationsverarbeitung zum Einsatz. Die Wissenschaftlichkeit der eingesetzten Methoden hängt stark vom jeweiligen Arbeitsschritt und dem involvierten Produkt ab. Im Kontext dieses Dokuments betrifft der Begriff „Methoden“ sämtliche Verfahren, die im Zuge der Produktion Amtlicher Statistiken zum Einsatz kommen, um die während der Planung erstellten **Konzepte** bestmöglich umzusetzen.

Wenn man bedenkt, dass die Anforderungen an die Amtliche Statistik, auch was die Erschließung von nicht erhebungsbasierten Datenquellen betrifft, sich in den letzten Jahren stark erweitert haben, ist Konzepttreue nicht immer a priori als selbstverständlich anzusehen. Die für die Erstellung der Statistik zu definierenden und festzulegenden Konzepte stellen die Basis für die Auswahl geeigneter Methoden dar. Nicht zuletzt deshalb ist der **Konzeptentwicklung** bei der Planung statistischer Projekte besondere Bedeutung beizumessen.

Nachdem die verwendeten Methoden bei der Erstellung Amtlicher Statistiken eine entscheidende Rolle spielen, finden sie naturgemäß auch im Verhaltenskodex Eingang. Im Grundsatz 7, der den Gebrauch solider Methoden einfordert, heißt es:

„Qualitativ hochwertige Statistiken müssen auf einer soliden Methodik basieren. Dies erfordert geeignete Instrumente und Verfahren sowie ein entsprechendes Know-how“,

Grundsatz 8 bezieht sich hingegen auf die Adäquatheit der verwendeten statistischen Verfahren:

„Qualitativ hochwertige Statistiken müssen auf der Anwendung geeigneter statistischer Verfahren - von der Erhebung bis zur Validierung der Daten - basieren“

Auch hier wird der bereits erwähnten Tatsache Rechnung getragen, dass sich Methoden über den gesamten statistischen Produktionsprozess erstrecken. Überdies wird in beiden Grundsätzen darauf Wert gelegt, dass die eingesetzten Instrumente und Verfahren **geeignet** (also fundiert) sind.

Grundsätzlich können Methoden, die für eine spezifische Problemstellung einsetzbar sind, wie folgt klassifiziert werden:

Verfügbare Methoden: Methoden, mit denen geplante Produktionsschritte grundsätzlich umgesetzt werden können.

Geeignete, gute Methoden („Good Practices“): Methoden, die sich bei der Umsetzung als geeignet erwiesen haben.

Momentan beste Methode(n) („Current Best Practice(s“): Methoden, die nach gegenwärtigem Status die beste Möglichkeit zur Umsetzung darstellen.

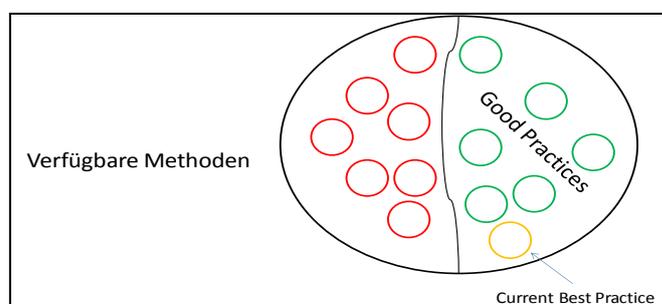


Abbildung 3: Good Practices sind eine Teilmenge der verfügbaren Methoden, eine Methode ist als Current Best Practice ausgezeichnet

Abbildung 3 zeigt, dass die geeigneten Methoden („Good Practices“) eine Teilmenge sämtlicher verfügbarer Methoden darstellen. Unter den geeigneten Methoden gilt eine als gegenwärtig beste („Current Best Practice“). Denkbar ist auch, dass es nicht bloß eine beste Methode gibt, sondern einige wenige quasi gleichwertig nebeneinander existieren.

Wann etwas als „Good“ bzw. „Best Practice“ gilt, ist jedoch nicht eindeutig. Man kann aber ohne Anspruch auf Vollständigkeit zu erheben für den Bereich der Amtlichen Statistik zumindest einige implizit gegebene Anhaltspunkte nennen, wann eine Methode gegenwärtig als fundiert und geeignet gilt („Good Practice“):

Langjährige Erprobtheit: Viele Verfahren, die in NSI's zum Einsatz kommen, sind dort seit langer Zeit in Verwendung. Dadurch ist sichergestellt, dass sich im Laufe der Zeit auf entsprechendes Know-how und Erfahrung aufbauen lässt. Aufgrund der laufenden Bewertung der durch die Methode gelieferten Ergebnisse und die damit verbundenen Verbesserungen ist anzunehmen, dass eine hinreichende Ergebnisqualität erzielt wird.

Internationale Verbreitung: Methoden und Verfahren, die in einer Vielzahl von Nationalen Statistischen Instituten (NSI's) zur Anwendung kommen, haben den Vorteil, dass ein Wissensaustausch ermöglicht wird, der eine optimierte Weiterentwicklung gestattet.

Ausreichende Abtestung: Vor allem neu entwickelte Methoden müssen, bevor sie in den Echtbetrieb gehen, ausreichend getestet werden. Nur wenn anhand hinreichender Dokumen-

tation der Testergebnisse sichergestellt ist, dass eine Methode verlässliche Ergebnisse liefert, darf sie auch eingesetzt werden.

Approbation: Die Methode wurde unter anerkannten nationalen und internationalen Experten diskutiert. Neue- bzw. weiterentwickelte Methoden werden in internationalen Gremien bzw. auf Forschungsveranstaltungen vorgestellt. Der dort stattfindende Diskussionsprozess soll dazu führen, eventuell vorhandene Schwächen bzw. Verbesserungspotentiale aufzuzeigen. Eine andere Form der Approbation ist, wenn die Methode Eingang in ein international anerkanntes Handbuch oder Methodeninventar findet.

Nachweisliche Verbesserung: Die Methode bewirkt eine nachweisliche Verbesserung einer oder mehrerer Qualitätskomponenten. Dabei muss beachtet werden, dass dies unter vernünftigem Einsatz von Ressourcen geschieht. In diesem Sinne sollten auch Methoden Beachtung finden, die eine Effizienzsteigerung bei gleichbleibender Qualität bewirken.

Ob alle oder nur einige der oben angeführten Aspekte erfüllt sein müssen, um von einer geeigneten fundierten Methode zu sprechen, ist jedenfalls nicht eindeutig. Wenn mehrere Methoden zur Wahl stehen, hängt die Auswahl letztendlich von der jeweiligen Problemstellung ab. Dabei spielen Überlegungen hinsichtlich der zu erzielenden Ergebnisqualität ebenso eine Rolle wie faktische Gegebenheiten hinsichtlich verfügbarer zeitlicher bzw. personeller aber auch finanzieller Ressourcen.

Bei vielen Schritten des statistischen Produktionsprozesses ist in diesen Qualitätsrichtlinien verankert, dass bestimmte Organisationseinheiten der Statistik Austria verpflichtend zu konsultieren sind, z.B. der Bereich Methodik zum Design einer Stichprobenerhebung. Dies soll sicherstellen, dass das in der Statistik Austria vorhandene Know-How möglichst optimal genutzt wird und die am besten geeigneten Methoden zum Einsatz kommen.

Statistik Austria stellt den Anspruch, dass – wenn möglich – best practices, zumindest jedoch geeignete Methoden („Good Practices“) zur Anwendung kommen. Daher ist es notwendig, die Forschungsaktivitäten in den relevanten Fachgebieten laufend zu beobachten und auch aktiv mitzugestalten. Dies betrifft einerseits Kontakte zu anderen NSI's und zu relevanten internationalen Organisationen, als auch die Präsenz bei entsprechenden Fachveranstaltungen und wissenschaftlichen Konferenzen. Auch die proaktive Zusammenarbeit mit der Wissenschaft bietet in diesem Zusammenhang eine wichtige Möglichkeit zur Weiter- bzw. Neuentwicklung leistungsfähiger Methoden. Von dieser Symbiose kann einerseits die Wissenschaft profitieren, die unter Wahrung der Vorgaben des BStatG 2000 und des Datenschutzgesetzes auf den reichhaltigen Datenschatz der Amtlichen Statistik zugreifen kann. Andererseits ergibt sich für die Amtliche Statistik der Vorteil, die Forschungskapazität der akademischen Welt für sich nutzbar zu machen.

Neben der Beobachtung des Forschungsmarktes ist die aktive Schulung und regelmäßige berufliche Fortbildung der Mitarbeiterinnen und Mitarbeiter eine weitere Notwendigkeit, um den Einsatz von fundierten Methoden zu gewährleisten. Nicht nur die bereits erwähnten Kernkompetenzen müssen dabei vermittelt werden, sondern auch, welche Methoden und Verfahren von der Bundesanstalt als geeignet (also als „Best Practices“ bzw. „Good Practices“) angesehen werden. Dabei werden den Mitarbeiterinnen und Mitarbeitern im Rahmen des Schulungsprogramms nicht nur nationale Weiterbildungsmöglichkeiten geboten, sondern auch die Teilnahme an internationalen Schulungsveranstaltungen (z. B. im Rahmen des European Statistical Training Program – ESTP) ermöglicht. Des Weiteren ist die gezielte Förderung von Mitarbeitermobilität zwischen Bereichen und Direktionen, aber auch in facheinschlägigen Institutionen außerhalb von Statistik Austria ein Bestandteil des modernen Personalentwicklungskonzepts von Statistik Austria. Auch die Kooperation mit Universitäten und anderen wissenschaftlichen Einrichtungen kann zu einer Steigerung der Kompetenzen der Mitarbeiterinnen und Mitarbeiter führen.

3. Qualitätsstandards im statistischen Produktionsprozess

3.1 Statistischer Produktionsprozess

Statistische Produkte im Rahmen der Amtlichen Statistik decken eine Vielfalt von gesellschafts- und wirtschaftspolitisch relevanten Themen ab. Als statistisches Produkt verstehen sich sämtliche Daten und Studien, die deskriptive sowie analytische Information über die wirtschaftlichen, demographischen, sozialen, ökologischen und kulturellen Gegebenheiten der Gesellschaft als Entscheidungsgrundlage für soziale und/oder ökonomische Fragestellungen, wissenschaftliche Forschung oder internationale Vergleiche beinhalten.

Die Methoden der Erstellung, die verwendeten Konzepte und Definitionen sowie die Darstellung der Ergebnisse können abhängig von der Themenstellung verschieden sein. Wie bei jeder Produktionsform gibt es aber gemeinsame fundamentale Produktionsschritte, die den meisten Amtlichen Statistiken zugrunde liegen. In *Abbildung 4* sind schematisch die wesentlichen Teile des statistischen Produktionsprozesses dargestellt.

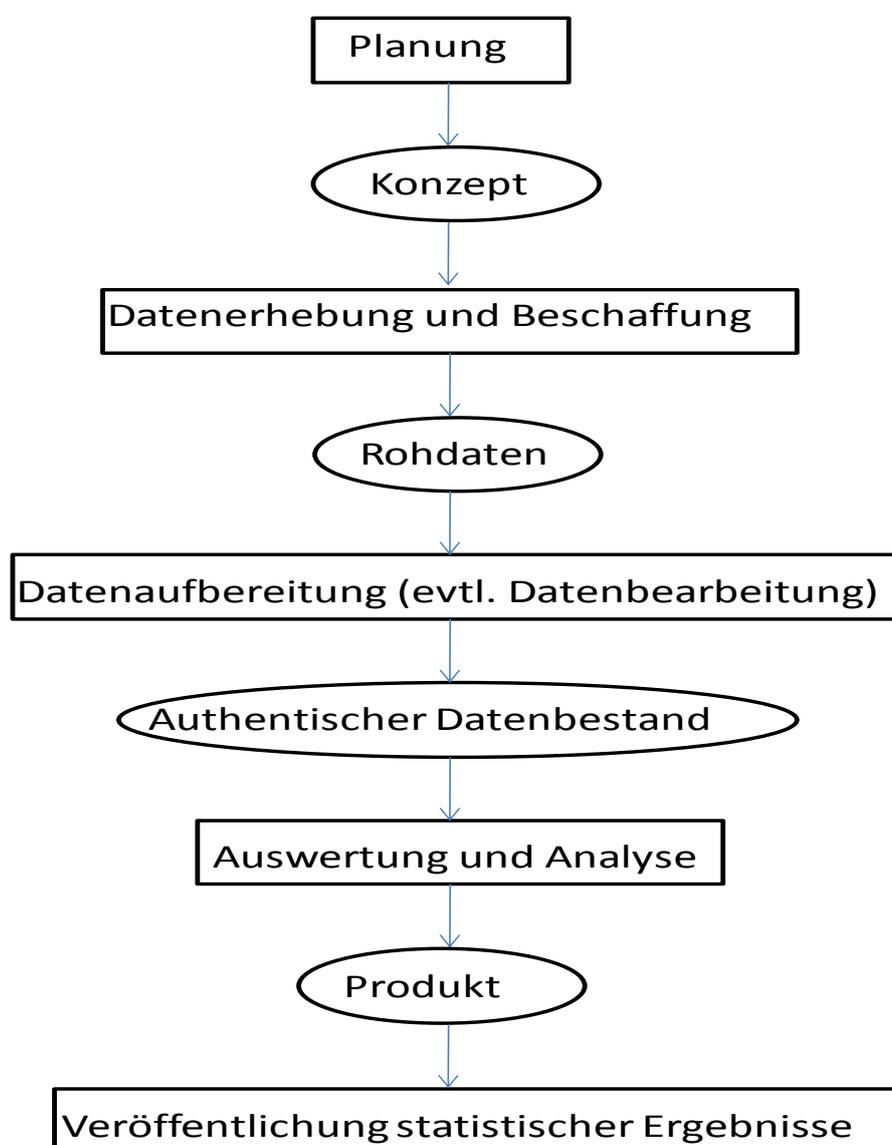


Abbildung 4: Der statistische Produktionsprozess

Am Beginn der Produktionskette steht die **Planung**, die gedankliche Vorwegnahme zukünftigen Handelns zur Vorbereitung des strukturierten Ablaufs des statistischen Produktionsprozesses. Sorgsame, ergebnis- bzw. zielorientierte Planung ist die Grundvoraussetzung für einen ressourcenschonenden sowie kosteneffizienten Projektablauf und sichert den Erhalt qualitativ hochwertiger Ergebnisse.

Die **Beschaffung** der zur Projektrealisierung benötigten Daten ist jener Schritt, der die Basis zur weiteren Verarbeitung legt. Mit dem Ziel der Sparsamkeit sowie der Entlastung der Respondentinnen und Respondenten müssen gemäß § 6 BStatG 2000 – sofern möglich – bereits vorhandene Daten vor primärstatistischen Erhebungen genutzt werden. Des Weiteren ist Stichprobenerhebungen der Vorrang gegenüber Vollerhebungen zu geben. Bei primärstatistischen Erhebungen ist die Datenbeschaffung meist zeit- und ressourcenintensiver als bei sekundärstatistischen Erhebungen. Am Ende dieses Teilprozesses, der bei primärstatistischen Erhebungen auch Feldphase genannt wird, stehen die **Rohdaten**.

Je nach Art der Datenbeschaffung kann der Arbeitsschritt der **Datenbe- bzw. Datenverarbeitung** unterschiedlich aufwändig gestaltet sein. Über diesen Transformationsprozess wird aus den Rohdaten der finale Datenkörper generiert. Man spricht in diesem Zusammenhang auch vom **Authentischen Datenbestand**.

Ist der Authentische Datenbestand einmal erstellt, bietet sich aufgrund der Größe des Merkmalskranzes, der in der Amtlichen Statistik zumeist gegeben ist, ein breites Feld von **Auswertungs- und Analysemöglichkeiten**. Die Ergebnisermittlung erfolgt zumeist nicht mittels rein deskriptiver Statistik, viel mehr kommen Methoden der Inferenzstatistik zum Einsatz, wobei sich der Bogen von freier Hochrechnung über die Anwendung statistischer Modelle bis hin zu multivariaten Analysen spannt.

Am Ende des statistischen Produktionsprozesses steht die **Veröffentlichung statistischer Ergebnisse**, welche nicht nur ein explizites Ziel von Statistik Austria, sondern auch eine rechtliche Verpflichtung, die im § 19 des BStatG 2000 verankert ist, darstellt.

In weiterer Folge werden in diesem Kapitel die einzelnen Teile der statistischen Produktionsprozesse näher erläutert, die sich zum Einen in der Organisationsstruktur der Bundesanstalt und zum Anderen in diversen bei Statistik Austria implementierten Systemen widerspiegelt. So erfolgt etwa die budgetäre Planung anhand dieses Produktionsprozesses, der wiederum in weitere Teilprozesse untergliedert ist.

3.2 Grundlagenarbeiten

3.2.1 Planung statistischer Projekte

Gründliche **Planung** steht am Beginn jedes Produktionsprozesses. Sie dient der Vorbereitung und Strukturierung des Ablaufs der einzelnen Teilprozesse und ist notwendig, um qualitativ hochwertige Endprodukte zu erhalten. Diese allgemein gültige Aussage trifft auch für den statistischen Produktionsprozess zu. Auch hier muss sich die Planung nach **Zielvorgaben** richten, die einerseits durch gesetzliche Vorgaben und andererseits implizit durch den Anspruch bestimmte Qualitätsstandards zu erfüllen, abgesteckt werden. Das Ziel der Planung ist die **Operationalisierung des Projektgegenstands, der intendierten statistischen Beschreibung eines bestimmten Phänomens**. Die planungsrelevanten Aktivitäten lassen sich in eine Grob- und unmittelbar daran anschließende Detailplanung unterteilen:

- Die **Grobplanung** dient dazu eine Entscheidungsgrundlage zu schaffen, um eine fundierte Aussage über die Durchführbarkeit des Projektes treffen zu können. Zu diesem Zweck muss das Projekt hinsichtlich Projektgegenstand und Projektziel charakterisiert werden. Basierend auf Erfahrungen wird für vordefinierte Projektphasen ein Projektgerüst erstellt. Besonderer Beachtung bedarf dabei die Frage nach den für das Projekt notwendigen Ressourcen, die auch verfügbar sein müssen.
- Ist die Entscheidung für das statistische Projekt gefallen, wird im Zuge der **Detailplanung** eine weiter gehende Differenzierung der in der Grobplanung definierten Projektphasen vorgenommen. Sämtliche für die Durchführung des Projektes wesentlichen Aspekte, wie z.B. Zeitpunkt und Dauer benötigter personeller und technischer Ressourcen, aber auch die Abklärung methodischer Einzelfragen, werden behandelt und die Prozessschritte, wie sie im Kapitel 3.1 angeführt sind, detailliert durchgeplant. Ziel der Detailplanung ist eine am „Reißbrett“ durchgeführte, möglichst detailgetreue Abbildung des statistischen Produktionsprozesses resultierend in einem genauen Projektplan.

Der Planungsprozess bei statistischen Projekten umfasst - wie in *Abbildung 5* gezeigt - mehrere Aspekte.

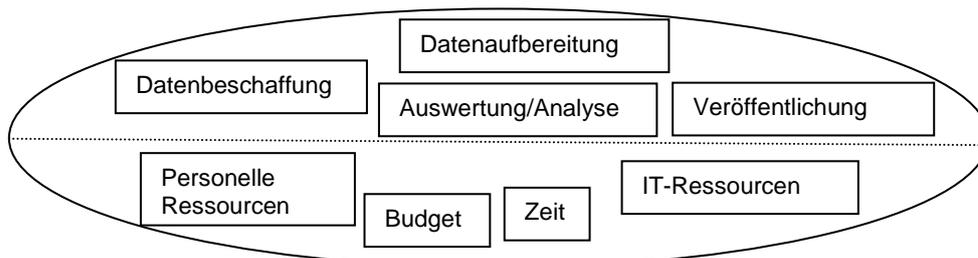


Abbildung 5: Planungsrelevante Aspekte

Bei den planungsrelevanten Aspekten muss der Prozess der **Datenbeschaffung** spezielle Beachtung finden. Die Möglichkeit, dass benötigte Daten bereits in anderen Fachbereichen innerhalb von Statistik Austria zur Verfügung stehen könnten, sollte bereits in der Planungsphase abgeklärt werden. Ist dies nicht der Fall, so muss die Datenbeschaffung aus externen Quellen geplant werden. Insbesondere sollte überlegt werden, ob durch die Erschließung und eventuelle Nutzung neuer Datenquellen Qualitätsverbesserungen bzw. Effizienzgewinne erzielt werden können.

Im § 6 BStatG 2000 wird eindeutig festgelegt, dass zunächst alle bereits vorhandenen Daten (sowohl innerhalb der Bundesanstalt als auch in anderen Institutionen) unbedingt zu nutzen sind, sofern dies möglich ist, bevor Direktbefragungen durchgeführt werden.

Richtlinien:

Grobplanung

- Für alle projektrelevanten Planungsaktivitäten ist eine Projektleiterin bzw. ein Projektleiter zu benennen, die bzw. der als Ansprechperson fungiert.
- Für jedes statistische Projekt muss im Rahmen einer Grobplanung ein Grobkonzept erstellt werden, welches Projektgegenstand und Projektziel mit seinen wesentlichsten Eigenschaften spezifiziert.
- Für jedes statistische Projekt müssen im Rahmen einer Grobplanung alle in Frage kommenden Datenquellen aufgelistet, auf deren Eignung überprüft und deren Nutzung bzw. Nicht-Nutzung begründet werden.
- Die Verfügbarkeit sämtlicher Sachmittel, die zur Durchführung des Projekts notwendig erscheinen, ist im Zuge der Grobplanung zu überprüfen. Dies betrifft insbesondere alle innerhalb der Bundesanstalt aus anderen Organisationseinheiten stammenden Sachleistungen, wie z. B. IT-Ressourcen, aber auch extern bezogene Sachmittel.
- Es ist zu überprüfen, ob eine termingerechte Verfügbarkeit der notwendigen Personalressourcen innerhalb der Bundesanstalt gegeben ist.
- Anhand der Ergebnisse der Grobplanung muss überprüft werden, ob das angestrebte Projekt grundsätzlich durchführbar ist.

Detailplanung

- Der Datenbeschaffung muss besonderes Augenmerk geschenkt werden. Dabei soll nicht-befragungsbasierten Datenquellen, sowie elektronischen Übermittlungsformen - wenn immer möglich - der Vorzug gegeben werden.
- Die Erschließung jener Datenquellen, die im Rahmen der Grobplanung als geeignet erkannt wurden, ist bei der Detailplanung genau zu beschreiben. Externe Dateneigner, sowie im Falle sekundärstatistischer Datennutzung die Datenhalter, müssen in den Planungsprozess einbezogen werden.
- Bereits während der Planung soll festgelegt werden, welche Methoden im Aufarbeitungsprozess Anwendung finden werden. Dabei ist sicherzustellen, dass die geplanten Methoden „Best Practices“, zumindest aber den „Good Practices“ (siehe Kapitel 2) entsprechen.
- In den Planungsprozess sind neben dem befassten Fachbereich all jene organisatorischen Einheiten einzubeziehen, die für die Erstellung des Endproduktes von Relevanz sind.
- Jedes Projekt bei Statistik Austria ist vorab mittels einer Vorkalkulation budgetär zu kalkulieren. Dies muss stets in Kooperation mit dem Bereich Controlling erfolgen. Um zu realistischen Planungswerten der zu leistenden Aufwände zu gelangen, sollen Erfahrungswerte aus ähnlichen Projekten – idealerweise aus Vorgänger- bzw. vergleichbaren Projekten - herangezogen werden.
- Es ist ein Projektplan zu erstellen, der den Verlauf des statistischen Produktionsprozesses anhand der einzelnen Projektphasen bzw. bei der Durchführung größerer statistischer Projekte anhand von Meilensteinen (Zwischenzielen) abbildet.
- Um Nachvollziehbarkeit zu gewährleisten, müssen sämtliche Aktivitäten im Zuge der Planung eines statistischen Projektes genau dokumentiert werden.

3.2.2 Konzepte, Definitionen und Klassifikationen

Das Informationsangebot von Statistik Austria beinhaltet eine breite Palette gesellschaftspolitisch **relevanter Themen**. Die **Konzepte** der Amtlichen Statistik bilden die Grundlage für die numerische Beschreibung allgemeiner Fragestellungen, die sich auf soziale, wirtschaftliche oder andere gesellschaftspolitisch relevanten Phänomene beziehen. Klassische Beispiele für Fragestellungen von Statistik Austria sind z.B. Beschäftigung, Bevölkerungsentwicklung, Wirtschaftsentwicklung oder die Berechnung des Verbraucherpreisindex (Inflation).

Der statistischen Untersuchung dieser Themengebiete liegen Konzepte zugrunde, die von Statistik Austria als geeignet identifiziert wurden bzw. wenn möglich international harmonisiert sind. Ein Beispiel dafür ist das Konzept der Volkswirtschaftlichen Gesamtrechnungen.

Definitionen sind operationale Beschreibungen innerhalb der Konzepte, die zur konkreten Erstellung von Statistiken notwendig sind (z.B. Merkmalsdefinitionen). Während ein Konzept die allgemeine Grundlage für eine Vielzahl statistischer Projekte bildet, legt eine Definition eindeutig fest, wie relevante Bausteine (z.B. welche Merkmale werden erhoben) zur Realisierung eines statistischen Projekts operationalisiert werden.

Für die Beurteilung der Qualität einer Statistik sind in diesem Zusammenhang zwei Aspekte zu berücksichtigen:

- **Validität:** Die Validität gibt Auskunft darüber inwieweit gelingt es, die aufgeworfene Fragestellung mit den gewählten Konzepten und Definitionen abzubilden und zu beantworten.
- **Kohärenz:** Bestimmte Themen werden innerhalb der Statistik Austria auf vielfältige Art und Weise betrachtet. In unterschiedlichen Konzepten können Merkmale (Variablen) zwar dem Wortlaut nach gleich sein, aber mit unterschiedlichen Definitionen umgesetzt werden. Ein typisches Beispiel ist der Erwerbsstatus bei der Arbeitsmarktstatistik, dem unterschiedliche Definitionen zugrunde liegen können.

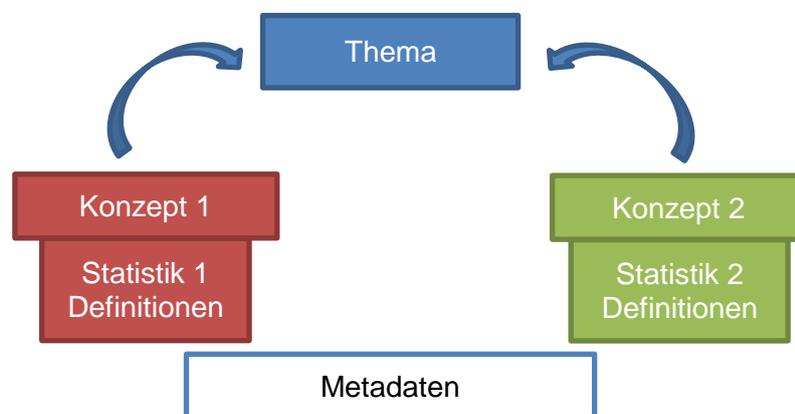


Abbildung 6: Verschiedene Statistiken beschreiben mit eigenen Definitionen basierend auf unterschiedlichen Konzepten ein spezifisches Thema

Um die verwendeten Konzepte und Definitionen für die Nutzerinnen und Nutzer statistischer Produkte transparent zu machen, müssen ergänzend ausreichend **Metadaten** zur Verfügung gestellt werden. In diesem Zusammenhang sei auf das Kapitel 3.6.7 („Dokumentation und Qualitätsberichterstattung“) verwiesen.

Eine Reihe von Merkmalen ist bereits durch Klassifikationen systematisch - und in den meisten Fällen international standardisiert - beschrieben und damit einheitlich operationalisiert. Klassifikationen sind damit eines der grundlegenden Werkzeuge für die Erstellung von Amtlichen Statistiken. Sie gewährleisten dabei die vollständige Erfassung der Grundgesamtheit und verfügen über sich ausschließende und damit eindeutige Kategorien.

Bei Statistik Austria kommen eine Reihe von nationalen und internationalen Klassifikationen, Systematiken sowie Verzeichnissen zur Anwendung. Diese werden in regelmäßigen Abständen sowohl auf internationaler als auch auf nationaler Ebene in koordinierter Weise revidiert und den sich ändernden gesellschaftlichen und wirtschaftlichen Gegebenheiten angepasst.

Um den Datennutzerinnen und -nutzern sämtliche relevanten Informationen über solche Klassifikationen und Verzeichnisse zu geben, stellt Statistik Austria die [Klassifikationsdatenbank](#), die eine große Anzahl nationaler und internationaler Klassifikationen und Systematiken beinhaltet, unentgeltlich im Internet zur Verfügung.

Vor allem wirtschaftsstatistisch wichtige Klassifikationen wie die NACE oder die CPA müssen in regelmäßigen Abständen revidiert werden, um das aktuelle Wirtschaftsgeschehen valide abzubilden. Derartige Revisionen können zu Brüchen in Zeitreihen führen (siehe Kapitel 3.5.7). Daher ist es notwendig, die Aggregate neu entstehender Klassifikationsausprägungen für Perioden, die vor der Revision liegen, zu schätzen (**Backcasting**).

Richtlinien:

- Definitionen müssen Nutzerinnen und Nutzern verfügbar gemacht werden und es soll überprüft werden, ob sie für die Nutzerinnen und Nutzer auch verständlich sind (Feedback, Nutzerbefragung).
- Sämtliche Konzepte und Definitionen sollen sich an internationalen Standards orientieren. Abweichungen müssen begründbar sein. Das Augenmerk soll dabei auf zeitliche und räumliche Vergleichbarkeit sowie größtmögliche Kohärenz zu anderen Produkten gelegt werden.
- Alle für ein statistisches Produkt relevanten und aktuellen Definitionen, Konzepte und Klassifikationen müssen in der Standard-Dokumentation im Kapitel 2.1 „Statistische Konzepte, Methodik“ erläutert werden.
- In der Standarddokumentation ist insbesondere auf allfällige durch Definitionen verursachte Unschärfe, Zeitreihenbrüche und Kohärenzprobleme einzugehen.
- Wenn immer möglich, soll auf die bei Statistik Austria verfügbaren Standard-Klassifikationen zurückgegriffen werden.
- Alle für Statistik Austria relevanten Klassifikationen sowie deren Revisionen und Modifikationen müssen unverzüglich in die Klassifikationsdatenbank eingepflegt werden.
- Sämtliche klassifikationstechnischen Fragen sind mit dem Bereich „Klassifikationen“ in der Abteilung „Register, Klassifikationen und Geoinformation“ zu koordinieren.
- Respondentinnen und Respondenten sind auf Anfrage über ihre klassifikatorische Zuordnung zu informieren. Im Speziellen ist Statistik Austria durch das Bundesstatistikgesetz dazu verpflichtet, alle österreichischen, wirtschaftlich aktiven Unternehmen hinsichtlich ihrer aktuellen ÖNACE-Zuordnung schriftlich und kostenlos zu informieren (Klassifikationsmitteilung) Diese beinhaltet die Haupt- und Nebentätigkeiten für rechtliche Einheiten/Unternehmen und ihrer Standorte.

3.2.3 Statistische Register

Die Führung **Statistischer Register** ist eine wichtige Grundlagenarbeit für die Amtliche Statistik. Der Zweck ist die Abbildung einer bestimmten Grundgesamtheit (beispielsweise Personen, Unternehmen, Gebäude) durch die Sammlung und Wartung für die Registerseinheiten relevanter Daten. Im Gegensatz zu einer Erhebung ist die Führung eines Registers ein permanenter Prozess, der darauf abzielt, die betrachtete Grundgesamtheit möglichst vollständig und zeitaktuell abzubilden. Statistische Register sind von Administrativen Registern dadurch abzugrenzen, dass sie nicht für Verwaltungszwecke, sondern zu statistischen Zwecken geführt werden und intern durch Statistik Austria betrieben und gewartet werden. Einerseits sind Statistische Register ein wichtiges **Planungsinstrument**, andererseits werden sie aber auch zur **Produktion von Statistiken** verwendet.

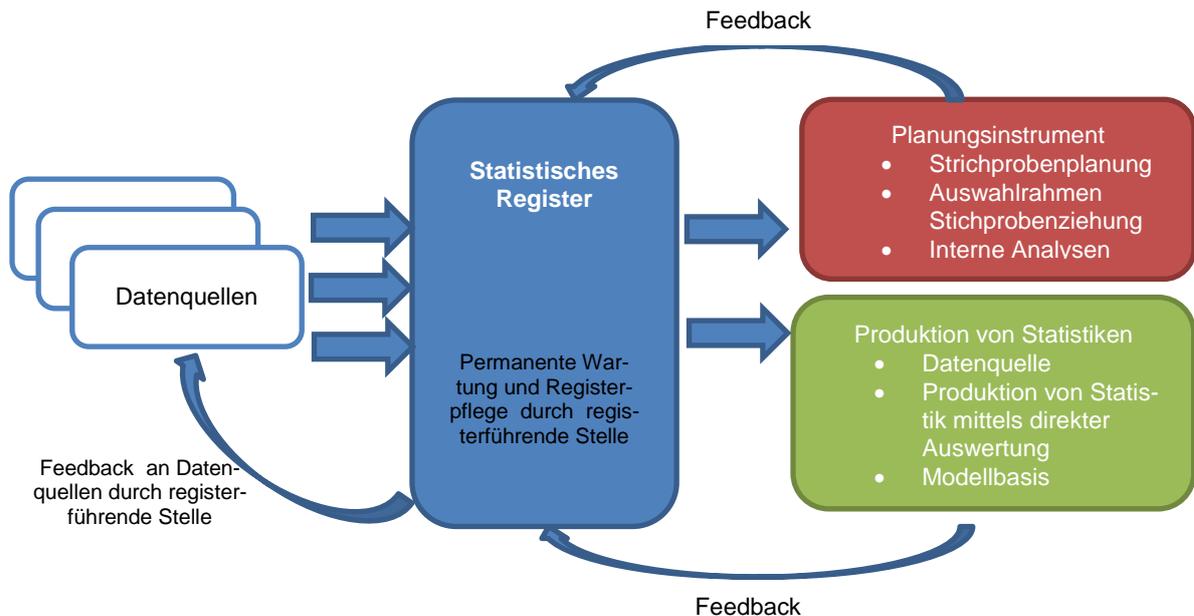


Abbildung 7: Statistische Register und deren Nutzung

Die laufende **Wartung von Statistischen Registern** erfolgt zumeist unter Einbeziehung externer administrativer Quellen. Dabei kommt dem Transformationsprozess, der die ursprünglich zu administrativen Zwecken erhobenen Daten in statistisch nutzbare Einheiten überführt, entscheidende Bedeutung zu. Berücksichtigt werden muss, dass ein Register nicht nur aus Einheiten auf einer Ebene, sondern auch aus Einheiten auf mehreren Ebenen bestehen kann. (z.B. Haushalte-Familien-Personen oder Unternehmen-Standorte). Die Wartung eines Statistischen Registers umfasst die Neuaufnahme und Löschung von Registerseinheiten, sowie Merkmalsänderungen bei bestehenden Registerseinheiten.

Als **Planungsinstrument** bilden hausintern verfügbare Register eine zur Ziehung von Stichproben oftmals unverzichtbare Basis. Dies betrifft einerseits die Planung der Stichprobe und andererseits den direkten Ziehungsprozess, wo Register als Auswahlrahmen herangezogen werden. Durch die Reichhaltigkeit des verfügbaren Datenmaterials dienen Register auch häufig dazu, interne Analysen durchzuführen, die als Input bei der Planung statistischer Projekte Verwendung finden.

Die zweite Nutzungsmöglichkeit ist jene, Statistische Register direkt als Datenquellen zur **Produktion von Statistiken** heranzuziehen. So kann man durch die Verfügbarkeit aktueller Registerdaten diese anstelle von erhobenen Daten in die Datensammlung mit einbeziehen, wodurch ein Beitrag zur Respondentenentlastung erzielt wird. Aber auch die direkte Auswertung von Registerbeständen ohne Durchführung einer Erhebung und die damit verbundene Veröffentlichung von Ergebnissen ist denkbar. Da - vor allem im wirtschaftsstatistischen Be-

reich - statistische Erhebungen gewisse Teile der Grundgesamtheit (z. B. kleine Unternehmen) ausklammern, kommt Statistischen Registern als Modellbasis zur Schätzung nicht erhobener Merkmale große Bedeutung zu (siehe Kapitel 3.5.6).

Richtlinien

- Vor der Nutzung von Statistischen Registerdaten ist zu überprüfen, ob die Konzepte und Definitionen dem intendierten statistischen Zweck in ausreichendem Maße entsprechen. Dabei soll jedes benötigte Merkmal betrachtet werden.
- Die interne Nutzung von Registerdaten bedarf der Einbeziehung der registerführenden Fachbereiche. Die Organisationseinheit, die die Registerdaten nutzt, muss dem registerführenden Fachbereich ein Feedback über die Qualität der genutzten Registereinheiten geben.
- Statistische Register gelten als statistische Produkte. Die Definition der Registereinheiten und die Konzepte der Führung und Wartung des Registers müssen dokumentiert und nachvollziehbar sein. Zu jedem von Statistik Austria geführten Register ist eine Standard-Dokumentation zu erstellen.
- Sämtliche gesetzlich angeordneten und verfügbaren administrativen Datenquellen müssen zur Registerwartung herangezogen werden. Die Wartung eines Registers soll so erfolgen, dass höchstmögliche Vollzähligkeit erzielt wird und die zeitliche Verzögerung der Abbildung von Änderungen möglichst gering ist.
- Die Nachvollziehbarkeit sämtlicher Schritte, die zu einer Aufnahme, Löschung bzw. Aktualisierung einer Registereinheit geführt haben, muss gegeben sein.
- Jede Registereinheit muss mit einem eindeutigen Identifikationsmerkmal (Kennzahl) versehen sein, das die Identifizierbarkeit für die interne Verwendung gewährleistet, aber in keinem sonstigen sachlichen Zusammenhang zur Registereinheit steht. Ein einmal vergebenes Identifikationsmerkmal soll nicht mehr geändert werden.
- Die Speicherung und Aktualisierung der Beziehungen der Registereinheiten zu anderen bei Statistik Austria bestehenden Registern bzw. zu wichtigen externen Datenkörpern über ein eindeutiges Identifikationsmerkmal ist neben der laufenden Aktualisierung der fachspezifischen Registervariablen Gegenstand der Registerwartung.
- Bei Registerdaten handelt es sich um Einzeldatensätze. Die Einhaltung der Grundsätze des Datenschutzes und der Geheimhaltung ist in diesem Zusammenhang zu gewährleisten.
- Um die Entwicklung einzelner Registereinheiten bzw. der abzudeckenden Grundgesamtheit (oder Teile davon) beobachten zu können, sollen zu jedem Register historische Registerstände rekonstruierbar sein.
- Statistische Register können als Auswahlrahmen für statistische Erhebungen dienen. Damit die Belastung potentieller Respondentinnen und Respondenten beobachtbar ist, soll innerhalb eines Registers Evidenz darüber geführt werden, bei welchen Erhebungen welche Registereinheit mit einbezogen war.

3.3 Datenerhebung und -beschaffung

3.3.1 Arten der Datenbeschaffung

Daten bilden die Arbeitsgrundlage zur Erstellung statistischer Produkte. Die **Verarbeitung von Daten** („Data Processing“) geht von **Rohdaten** aus und endet mit der Erstellung eines **Authentischen Datenbestandes**¹. Um zu den Rohdaten zu gelangen, muss bereits im Planungsprozess (siehe Kapitel 3.2.1) die Art der Datenbeschaffung festgelegt werden. Es gibt mehrere Möglichkeiten, wie der Prozess der Datenbeschaffung gestaltet werden kann:

- I. Verwendung von bereits bei Statistik Austria **verfügbaren Daten**,
- II. Verwendung von extern verfügbaren Daten
 - o **Verwaltungsdaten**,
 - o **Neue Datenquellen (z.B Big Data)**,
- III. Beschaffung von Daten mittels **primärstatistischer Erhebungen**.

Die Zielsetzungen des BStatG 2000 sind im Sinne der Respondentenentlastung dahingehend, dass das traditionelle Instrument der primärstatistischen Erhebung (Befragung der Auskunftspflichtigen) nur dann anzuwenden ist, wenn adäquate Daten nicht anderweitig beschafft werden können. Insofern ist die oben getroffene Auflistung auch als Prioritätenreihung zu verstehen. Dadurch hat die Nutzung von Verwaltungsdaten (auch in aggregierter Form) zunehmend an Bedeutung gewonnen. In diesem Zusammenhang soll auf die Kapitel 3.3.2 „Verwendung von Verwaltungsdaten“ und 3.4.4 „Verknüpfung verschiedener Datenquellen“, aber auch auf 3.5.2 „Schätzung aus Stichprobenerhebungen“ verwiesen werden.

Als moderner Informationsdienstleister ist Statistik Austria gefordert, aktuelle Entwicklungen betreffend Datenquellen zu beobachten und folglich zu prüfen, inwieweit diese statistisch nutzbar gemacht werden können. Dazu gehört die Verwendung „neuer“ Datenquellen, u. a. auch sogenannter Big Data, stets unter gleichzeitiger Berücksichtigung der Grundsätze Qualität, Respondentenentlastung, Datenschutz und Wirtschaftlichkeit.

Bei **primärstatistischen** Erhebungen können folgende Methoden zur Anwendung kommen:

- a) Persönliche Befragung (PB) vor Ort durch Interviewer – ohne Computer
- b) Persönliche Befragung (PB) vor Ort durch Interviewer computerunterstützt „CAPI“
- c) Persönliche Befragung (PB) telefonisch durch Interviewer computerunterstützt „CATI“
- d) Merkmalerhebung in gedruckter Form („PAPI“)
- e) Merkmalerhebung in elektronischer Form („Webfragebögen“, „CAWI“)
- f) Merkmalerhebung ohne Persönliche Befragung (PB) (z.B. Preiserhebung)

Während die persönliche Befragung eher im sozial- und bevölkerungsstatistischen Bereich (Personen/Haushalte) Verwendung findet, werden die anderen Erhebungsformen eher bei wirtschaftsstatistischen Erhebungen (Unternehmen/Betriebe) eingesetzt.

Eine Erhebung mittels Interviews wird zumeist entweder computerunterstützt vor Ort („CAPI“) oder telefonisch („CATI“) durchgeführt. Dabei steht die Schulung der Interviewenden im Vordergrund.

¹ Definition "Authentischer Datenbestand": Finaler Datenbestand, von dem Publikations- bzw. Ergebnistabellen erstellt werden. Die im authentischen Datenbestand enthaltenen (Mikro)Daten werden geplanter Weise in Ihren Merkmalswerten nicht mehr verändert oder um weitere Merkmale angereichert. Nachgelagerte Verfahren, für die der authentische Datenbestand als Input dient, dienen der Aggregation der (Mikro)daten bzw. der Analyse von resultierenden statistischen Ergebnissen.

Im Falle einer Erhebung mittels Fragebögen liegt das Augenmerk auf der Qualität der mitgereichten Instruktionen (z.B. substantielle Erläuterungen, präzise Identifizierungshilfen zu Fragen der Klassifikation oder die unmissverständliche Beschreibung elektronischer Meldewege).

Es ist das erklärte Ziel von Statistik Austria den Einsatz elektronischer Meldeschienen zu forcieren und wenn immer möglich eine vollständige Abdeckung der Erhebungsmasse ohne den Einsatz nicht-elektronischer Meldemedien zu erreichen.

Bei einzelnen statistischen Projekten können im Zuge der Datenbeschaffung auch mehrere Methoden zum Einsatz kommen (ein häufiger Fall ist die Unterstützung primärstatistischer Erhebungen durch die Nutzung von externen Datenquellen).

Die nachfolgende Tabelle fasst für die zuvor beschriebenen Verfahren die wichtigsten Vor- bzw. Nachteile zusammen.

Vorteile	Nachteile
I. Verwendung von bereits bei Statistik Austria verfügbaren Daten	
<ul style="list-style-type: none"> + Kein Zusatzaufwand für den eigentlichen Beschaffungsprozess + Keine weitere Belastung von Auskunftgebenden + Know-How über diese Daten bereits vorhanden 	<ul style="list-style-type: none"> - Daten wurden für einen anderen Zweck erhoben und genutzt, daher entsteht für den intendierten Zweck ein Adaptierungsaufwand.
II. Verwendung von extern verfügbaren Daten	
<ul style="list-style-type: none"> + Zumeist keine bzw. deutlich geringere Belastung von Respondenten als im Falle einer Primärerhebung + Zumeist hohe Abdeckung erzielbar + Zusätzlicher Erkenntniswert + Zumeist höhere Aktualität der verfügbaren Daten 	<ul style="list-style-type: none"> - Daten zumeist nur nach Bearbeitung statistisch nutzbar - Abhängigkeit von Veränderungen innerhalb der externen Datenquelle - Bei den Datenhaltern fallen Kosten und Arbeitsaufwände an - Bei Statistik Austria können evtl. Kosten aus einer intensiven Datenaufbereitung oder aufgrund höherer Speicherplatznutzung anfallen.
III. Beschaffung von Daten mittels primärstatistischer Erhebungen	
a – Persönliche Befragung vor Ort durch Interviewer – ohne Computer	
<ul style="list-style-type: none"> + Fehlinterpretationen kann unmittelbar gut entgegengewirkt werden 	<ul style="list-style-type: none"> - Eventuelle Verzerrungen durch Interviewer-Effekte - Wegzeiten für Interviewer fallen an - teure Form der Erhebung
b – Persönliche Befragung vor Ort durch Interviewer computerunterstützt “CAPI”	
<ul style="list-style-type: none"> + Technische Möglichkeiten können ausgeschöpft werden (Verlinkung, automatische Korrekturen, Plausibilitätschecks etc.) + Fehlinterpretationen kann unmittelbar gut entgegengewirkt werden 	<ul style="list-style-type: none"> - Eventuelle Verzerrungen durch Interviewer-Effekte - Wegzeiten für Interviewer fallen an - Technische Ausstattung und Support für Interviewende sind notwendig

+ Drucksortenbedarf erheblich verringert

- teuerste Form der Erhebung

c – Persönliche Befragung telefonisch durch Interviewer computerunterstützt „CATI“

+ Technische Möglichkeiten können ausgeschöpft werden (Verlinkung, automatische Korrekturen, Plausibilitätschecks, etc.)
+ Fehlinterpretationen kann entgegengewirkt werden
+ Drucksortenbedarf erheblich verringert
+ Keine Interviewerwegzeiten
+ Kontrollierte Interviewersituation

- Problematik der Verfügbarkeit von Telefonnummern
- hoher organisatorischer Aufwand

d – Merkmalerhebung in gedruckter Form („PAPI“)

+ Unabhängigkeit von Computernetzwerken und deren Verfügbarkeit.
+ Kosten gut kalkulierbar
+ Ist selten von Unabwägbarkeiten des Alltags betroffen (plötzliche Erkrankung, kurzfristiger Termin etc.).

- Nachgelagerte Datenerfassung ist notwendig
- Hotline ist notwendig

e – Merkmalerhebung in elektronischer Form („Webfragebögen“, „CAWI“)

+ Möglichkeit des Eingriffs in laufende Erhebungen (z.B. Update der FAQ).
+ Nachgelagerte Datenerfassung ist nicht notwendig.
+ Technische Möglichkeiten können ausgeschöpft werden (Verlinkung, automatische Korrekturen, Plausibilitätschecks etc.)

- Technischer Support für Externe wird nötig (vergessene Zugangscodes, Kompatibilitätsprobleme etc.)
- Hotline ist notwendig
- Computernetzwerke notwendig

f – Merkmalerhebung ohne persönliche Befragung

+ Hohe Treffsicherheit weil kein zusätzlicher Interpretationsspielraum für Respondenten gegeben ist.

- Abhängig vom freien Zugang zu den jeweiligen Merkmalen
- Nachgelagerte Datenerfassung ist evtl. notwendig

Richtlinien:

- Die Art der Datensammlung muss bereits im Zuge der Planung festgelegt werden. Dabei soll ein ausgewogenes Verhältnis zwischen Gehalt bzw. Nutzbarkeit der gewonnenen Information, vorhandenen budgetären Mitteln und Belastung der Respondentinnen und Respondenten angestrebt werden.
- Bei der Wahl der Art der Erhebung sind jene Datenquellen zu bevorzugen, die nicht mit einer direkten Befragung von Respondentinnen und Respondenten verbunden sind. Neben der Verwendung von Verwaltungsdaten sind insbesondere auch neue Datenquellen, die einen Erkenntnisbeitrag zur Erstellung der entsprechenden Statistik leisten können, zu identifizieren.
- Im Falle persönlicher oder telefonischer Interviews müssen die Interviewenden - unterstützt durch einen Interviewleitfaden - vorab geschult und ggf. nachgeschult werden. Ein Pre-Test des Fragekatalogs mit einer ausreichenden Anzahl simulierter Interviews (inkl. mit Personen außerhalb des Fachbereichs), sollte zweckmäßigerweise durchgeführt werden. Jedenfalls sollen persönliche und telefonische Interviews, wenn möglich, computerunterstützt durchgeführt werden.
- Bei Befragungen, die von Respondentinnen und Respondenten selbst auszufüllen sind, wird diesen, sofern möglich, eine nutzerfreundliche elektronische Form der Antwortmöglichkeit zur Verfügung gestellt. Diese elektronische Antwortmöglichkeit soll neben dem PC auch über gängige mobile devices (smart phones, tablets etc.) und softwarekompatible Apps gegeben sein.
- Der Zeitpunkt der Kontaktaufnahme zu Respondentinnen und Respondenten soll im Verhältnis zum Zeitpunkt der Meldungen sorgsam bzw. möglichst zeitnah zum Erhebungszeitpunkt gewählt werden. Zur Kontaktaufnahme benötigte Informationen müssen vorab gesammelt, überprüft und bei Bedarf aktualisiert werden.
- Im Zuge einer Erhebung soll der Rücklaufquote besondere Beachtung zukommen. Liegt sie hinter den Erwartungen zurück, ist dieser Entwicklung mit geeigneten Maßnahmen frühestmöglich entgegenzusteuern (siehe Kapitel 3.3.4 „Urgenzen“).
- Bei jeder Erhebung sollten Respondentinnen und Respondenten mittels Ankündigungs- oder Begleitschreiben über den Zweck und die Intention der Erhebung informiert werden. Jedenfalls ist ihnen auch eine konkrete Anlaufstelle zu nennen, an die sie sich im Falle auftretender Probleme wenden können. Des Weiteren ist ggf. darauf hinzuweisen, welche Sanktionen bei einer Verweigerung drohen.
- Bei den im Zuge einer primärstatistischen Erhebung gesammelten oder sekundärstatistisch genutzten Daten handelt es sich zumeist um Einzeldaten. Daher ist dem Datenschutz und der Geheimhaltung besonderes Augenmerk zu widmen.
- Wann immer möglich, sollen den Respondentinnen und Respondenten elektronische und insbesondere nutzerfreundliche Meldemöglichkeiten zur Verfügung gestellt und deren Handhabung erläutert werden. Die Vorteile einer Nutzung elektronischer Meldeschienen müssen erläutert werden. Ebenso muss der technische Support durch einen Helpdesk gewährleistet sein.

- Die Datenbeschaffung ist genau zu dokumentieren. Dabei ist zu beachten, dass die Übermittlung von Register-, Verwaltungs- oder Statistikdaten grundsätzlich auf elektronischem Weg erfolgen soll (inkl. zugehöriger Metadaten). Weiters soll die Erfassung relevanter Paradata (z.B. Interviewdauer, Feedback der Interviewenden) zur Sicherung eines hohen qualitativen Standards aktueller wie auch künftiger Beschaffungsmaßnahmen dienen.

3.3.2 Verwendung von Verwaltungsdaten

Verwaltungsdaten sind (gemäß Begriffsbestimmung laut BStatG 2000) Daten, die bei Stellen in Wahrnehmung von bundes- oder landesgesetzlich übertragenen Aufgaben oder in Vollziehung unmittelbar anwendbarer gemeinschaftsrechtlicher Vorschriften angefallen sind. Es sind also Daten, die ursprünglich zu nicht statistischen Zwecken gesammelt wurden, aber als **sekundärstatistische Quelle** genutzt werden können. Zumeist dient der administrative Zweck dazu, gewisse Gegebenheiten oder Ereignisse, die sich auf Einzeldaten beziehen, festzustellen und festzuhalten. Im Unterschied dazu ist die Identifizierbarkeit von Personen für statistische Zwecke nicht notwendig und im Sinne des Datenschutzes unerwünscht. Daher wurde im § 15 des BStatG 2000 die Anonymisierung von personenbezogenen Daten angeordnet sobald sie nicht mehr aus den in § 5 Abs. (2) BStatG 2000 genannten Gründen oder für eine weitere angeordnete statistische Erhebung erforderlich sind; dies ermöglicht eine effiziente Statistikerstellung unter Wahrung eines umfassenden Datenschutzes.

Die Inhaber von Verwaltungsdaten sind in Entsprechung des BStatG 2000 neben der Verpflichtung, diese Daten der Bundesanstalt zu übermitteln, auch dazu angehalten, im Zuge der Datenübermittlung auf Verlangen von Statistik Austria auch die Merkmalsdefinitionen, auf welche Art die Daten angefallen sind, sowie welche Berechnungsmethoden allenfalls verwendet wurden, bekannt zu geben. Derartige Metadaten bilden die Grundlage für die Analyse und Beurteilung, in wie weit diese Verwaltungsdaten für Zwecke der Amtlichen Statistik geeignet sind, respektive welche interpretativen Restriktionen bei deren Verwendung zu berücksichtigen sind.

Die Nutzung von administrativen Daten für die Zwecke der Amtlichen Statistik bringt eine Reihe von **Vorteilen** mit sich:

- Sofern die Befragung vollständig durch Verwaltungsdaten substituiert werden kann, können die Kosten reduziert werden.
- Die Respondentinnen und Respondenten werden erheblich entlastet.
- Die Übermittlung der Verwaltungsdaten im Falle der Verfügbarkeit ist unentgeltlich und erfolgt auf elektronischem Weg.
- Die fortschreitende technologische Entwicklung erlaubt auch die leichte Handhabung sehr großer Datenmengen.
- Statistiken, die auf Verwaltungsdaten basieren, welche laufend gewartet werden, können meist in kürzeren Intervallen produziert werden und ermöglichen es, Trends rascher zu erkennen.
- Die Möglichkeit zur Verknüpfung mehrerer Datenquellen mittels geeigneter Schlüssel (siehe Kapitel 3.4.4) bietet eine Vielzahl von Analysemöglichkeiten und qualitätssichernden Maßnahmen.

Nicht zuletzt wegen dieser Vorteile ist Statistik Austria gemäß BStatG 2000 verpflichtet - wann immer möglich - Verwaltungsdaten für die Produktion von Statistiken zu verwenden. Bedingt durch all diese Gegebenheiten ist die Nutzung administrativer Datensätze unverzichtbar für eine Vielzahl von Produkten der Amtlichen Statistik geworden.

Die **Anwendungsmöglichkeiten** für den Gebrauch von Verwaltungsdaten sind vielfältig und beinhalten:

- Direkte Nutzung als Datenquelle
- Nutzung als unterstützende Datenquelle bei Erhebungen zur Ergänzung spezifischer Merkmale und für Plausibilitätsprüfungen
- Nutzung als Auswahlrahmen
- Indirekte Nutzung als Input zur gebundenen Hochrechnung, Schätzung und Modellbildung
- Vergleichsmöglichkeit für Erhebungsergebnisse

Um administrative Daten effektiv zur Statistikproduktion verwenden zu können, ist es notwendig auch **potentielle Schwächen** der Datenquellen zu kennen. Einer der wichtigsten Unterschiede zur klassischen Erhebung ist, dass Statistik Austria keine Kontrolle über die Erhebung und Produktion der Daten hat. Die dadurch auftretenden Probleme können sein:

- Unterschiedliche Konzepte, Definitionen und Klassifikationen
- Abdeckungsprobleme (z.B. Über- und Untererfassung)
- Bei Merkmalen, die für den administrativen Zweck nicht von Relevanz sind, sind die Daten oft ungeprüft bzw. es gibt nur lange oder unregelmäßige Wartungszyklen.
- Nicht rechtzeitig verfügbare Daten
- Rechtliche Änderungen können Auswirkungen auf die statistischen Ergebnisse haben

Unterschiedliche Konzepte, Definitionen und Klassifikationen können dazu führen, dass erheblicher Aufwand erforderlich ist, um die Merkmale aus Administrativdaten für statistische Zwecke nutzbar zu machen. Es können dadurch Einschränkungen in der statistischen Nutzbarkeit der administrativen Datenquelle entstehen oder deren Nutzung gänzlich ausschließen. Um derartigen Herausforderungen begegnen zu können, ist die Kommunikation mit den externen Dateninhabern von entscheidender Bedeutung. Die genaue Dokumentation von Divergenzen jeglicher Art durch Statistik Austria ist in diesem Zusammenhang unabdingbar, um Fehler zu vermeiden und Verbesserungen erzielen zu können.

Da die Nutzung administrativer Datenquellen in den letzten Jahren stark zugenommen hat, wurde 2006 bei Statistik Austria die Position der **Verwaltungsdaten-Koordination** eingerichtet. Deren Ziel ist die Sicherstellung, Verbesserung und den Ausbau der derzeitigen Nutzung von Verwaltungsdaten zu gewährleisten, wie dies im BStatG 2000 vorgesehen ist. Eine der zentralen Aufgaben der Verwaltungsdatenkoordination ist die Wartung der **Datenbank zur Verwaltungsdatenkoordination (VDK-Datenbank)**, die einen Überblick über jene statistische Projekte gibt, die Verwaltungsdaten und andere, bei Statistik Austria bereits existierende Daten verwenden. Die wichtigste Funktion der Datenbank ist die Übersicht der in den jeweiligen Projekten verwendeten Datenquellen. Hier sind der Name der Quelle und der Dateninhaber angeführt. Für jede dieser Quellen sind die internen und externen Ansprechpersonen projektbezogen ersichtlich. Bezüglich der externen Ansprechpersonen können auch Informationen über die Organisationen, die sie vertreten, abgerufen werden. Die Möglichkeit einer Suche nach Projektnamen, Projektleitungen und Datenquellen soll die Arbeit mit der Datenbank erleichtern.

Bei vielen statistischen Projekten (z.B. Registerzählung) findet man mit der Nutzung einer einzigen administrativen Datenquelle nicht das Auslangen. Dort sind auch die Richtlinien zur Verknüpfung mehrerer Datenquellen zu beachten (siehe Kapitel 3.4.4). Verwaltungsdaten spielen eine maßgebliche Rolle bei der Ergänzung primärstatistischer Datenkörper zur Darstellung der Grundgesamtheit, einerseits in ihrer Funktion als Teil der sekundärstatistischen Daten, andererseits beispielsweise als Eckdaten für die an sie gebundene Modellrechnung (siehe Kapitel 3.5.6) oder die gebundene Hochrechnung (siehe Kapitel 3.5.2).

Richtlinien:

- Plant man Verwaltungsdaten zu verwenden, muss zunächst abgeklärt werden, ob für Statistik Austria jene Daten schon zur Verfügung stehen (VDK-Datenbank). Ist dies der Fall soll mit der im Haus verantwortlichen Person Kontakt aufgenommen werden.
- Sind bei Statistik Austria keine Datenquellen vorhanden, müssen potentielle Verwaltungsdateninhaber bereits im Rahmen der Planung in den Analyseprozess hinsichtlich der Datenbeschaffung einbezogen werden.
- Vor der Nutzung einer administrativen Datenquelle ist zu überprüfen, ob Konzepte und Definitionen dem intendierten statistischen Zweck im ausreichenden Maße entsprechen. Dabei soll jedes benötigte Merkmal des administrativen Datenkörpers betrachtet werden.
- Der Prozess, der dazu dient, administrative Daten statistisch nutzbar zu machen, soll geplant und getestet werden. Alle Maßnahmen, die die Transformation des Datenmaterials betreffen, müssen angemessen dokumentiert werden.
- Ergänzt man administrative Daten um zusätzliche Merkmale, soll auf vorhandenes Datenmaterial zurückgegriffen werden. Eine Aufschätzung mittels Massenimputation (in etwa mehr als 20 % fehlende Werte) soll nur dann vorgenommen werden, wenn keine anderen adäquaten Möglichkeiten gegeben sind.
- Kontakte zu externen Dateninhaber müssen geknüpft und laufend gepflegt werden. Gegebenenfalls sollen Vereinbarungen getroffen werden, die eine regelmäßige und zeitgerechte Lieferung der Daten garantieren.
- Sollten Daten seitens der Dateninhaber revidiert werden, muss die Übermittlung der korrigierten Datenbestände gewährleistet sein. Zusätzlich sollte im Falle von periodischen Datenlieferungen ein Prüfprozess zum Erkennen jeglicher Änderung in der Datenstruktur vorhanden sein (z.B. Metadaten, Codierung etc.).
- Kontakte zu Dateninhaber sollen dazu genutzt werden, regelmäßiges Feedback über die Qualität der gelieferten Daten zu geben.
- Bei Administrativdaten handelt es sich zumeist um Einzeldatensätze, die sich auf natürliche oder juristische Personen beziehen. Somit ist durch Anonymisierung bei der Datenverarbeitung der Datenschutz einzuhalten.
- Die Nutzung von Verwaltungsdaten, die bisher noch nicht für Statistik Austria verfügbar waren, muss der Verwaltungsdaten-Koordination zur Kenntnis gebracht werden. Es muss angestrebt werden, die VDK-Datenbank auf einem möglichst aktuellen Stand zu halten.
- Für besonders wichtige externe Datenquellen (z.B. Daten des Hauptverbandes der österreichischen Sozialversicherungsträger bzw. Daten der Finanzbehörde) sollen die Daten zentral zur Verfügung stehen. Die Abläufe der Datenlieferung sowie der Satzaufbau sollen standardisiert sein.

- Die Nutzung externer administrativer Datenquellen für ein statistisches Produkt muss gegenüber Datennutzerinnen und -nutzern transparent gehandhabt werden. Sämtliche Implikationen, die sich aus der Nutzung von Verwaltungsquellen ergeben, sind in der Standard-Dokumentation zu beschreiben. Dies betrifft insbesondere Fragen der Abdeckung, Rechtzeitigkeit und Aktualität sowie Verzerrungen aufgrund konzeptueller Unterschiede.
- Die Übermittlung administrativer Daten an Statistik Austria sollte elektronisch über eine gesicherte Schnittstelle erfolgen.

3.3.3 Fragebogendesign und -test

Fragebögen spielen eine zentrale Rolle im Datensammlungsprozess der Statistik Austria. Sie stellen – wie in Kapitel 3.3.1 beschrieben – eine der wichtigsten Datenquellen für Statistiken dar. Die Erstellung eines Fragebogens erfordert hochspezialisiertes Fachwissen und einschlägige Erfahrung, zumal die Qualität der Ergebnisse einer Erhebung in hohem Ausmaß von der Qualität des Fragebogens abhängt. Während bei anderen Schritten im Datenaufbereitungsprozess- wie etwa bei Modellberechnungen - Wiederholungen und Revisionen möglich sind, können Befragungen nicht wiederholt und Fragebogenfehler nicht oder nur schwer korrigiert werden.

Ziel eines Fragebogens ist es, vordefinierte **Konzepte** und **Merkmale** zu messen. Im europäischen statistischen System werden dabei mehrheitlich output-harmonisierte Merkmale vorgegeben. Darunter wird verstanden, dass Konzepte und Merkmale dasselbe messen sollen, wobei das Messinstrument (in diesem Fall der Fragebogen) jedoch nicht vorgegeben ist. Da die Erhebungen von Statistik Austria oftmals einen sehr umfangreichen Merkmalskatalog abbilden, können Fragenbögen entsprechend komplex aufgebaut sein. Ein zu messendes Konzept kann dabei durch ein oder mehrere Erhebungsmerkmale, also Fragen oder Variablen, abgebildet werden. Die Erhebungsmerkmale sind zudem vom Vorhandensein bestehender Verwaltungsdaten abhängig.

Im Rahmen der Unternehmensstatistiken ist es darüber hinaus notwendig zu evaluieren, inwieweit die zu erhebenden Merkmale (z.B. Umsatzerlöse, Personalaufwand) überhaupt in geeigneter Form bei den Unternehmen vorliegen; bzw. ob Merkmalsdefinitionen entsprechend angepasst werden müssen.

Verwendet werden dabei ausschließlich standardisierte Fragebögen, die in der Regel aus vorgegebenen Fragen bzw. Variablendefinitionen und Antwortkategorien bestehen. Durch diese standardisierte Vorgehensweise soll sichergestellt werden, dass die Antworten vergleichbar und Unterschiede nicht auf die Messmethode zurückzuführen sind. Ein Fragebogen ist auch ein Kommunikationsinstrument. Es wird dabei angestrebt, die Interview- bzw. Erhebungssituation standardisiert zu halten, um unerwünschte Einflüsse auf die Ergebnisse zu minimieren.

Fragebogendesign und -tests sind damit ein Teilstück eines Prozesses, der von der Formulierung der Konzepte und Merkmale bis hin zur Evaluation des gesamten Erhebungsprozesses führt. Dieser Prozess beinhaltet im Idealfall Schleifen, wie etwa jene zwischen Fragebogentests und der Revision des Fragebogens. Die Beobachtung und Evaluation des Erhebungsprozesses sollte letztendlich wieder Eingang in das Fragebogendesign und möglicherweise sogar in die Konzepte und Merkmale selbst finden.

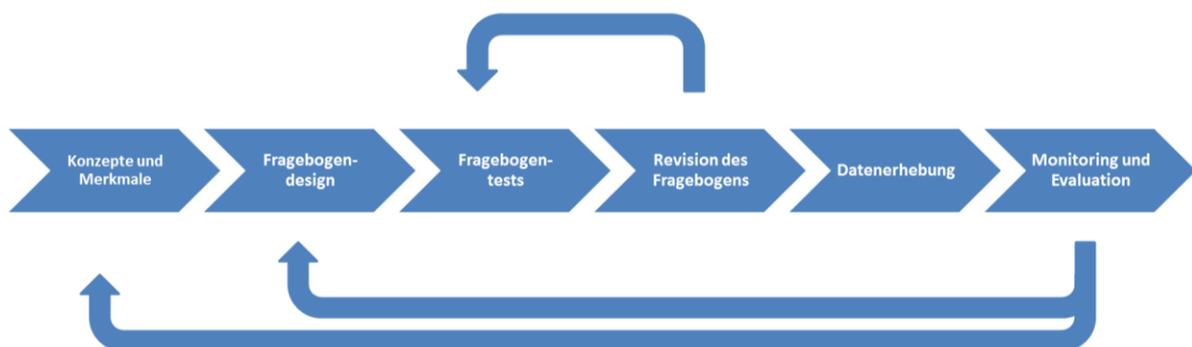


Abbildung 1 Prozess von Fragebogendesign und -test (in Anlehnung an Brancato et al., 2006)

Fragebogendesign:

Unter Fragebogendesign ist die gesamte Ausgestaltung eines Fragebogens zu verstehen. Dies beinhaltet die Formulierung der Fragen, die Erstellung von Erläuterungen und sonstigen Hilfetexten, die konkrete Darstellung der einzelnen Fragen, die Gruppierung zu Merkmals-

blöcken bzw. des Gesamtlayouts sowie die Funktionalitäten innerhalb des Fragebogens (z.B. Filterführung, Plausibilitätsprüfungen).

Ein gut designter Fragebogen erfüllt die folgenden Gütekriterien einer Messung:

- **Objektivität:** Die Ergebnisse einer Messung sind unabhängig von den Akteuren und Institutionen, welche die Erhebung durchführen oder den Auftrag dafür geben.
- **Reliabilität:** Die Werte sind zuverlässig, d.h. unabhängig von Zeit, Ort und Person wird das Gleiche gemessen.
- **Validität:** Das Messinstrument misst tatsächlich das, was es messen soll.
- **Vergleichbarkeit** über nationale oder kulturelle Kontexte hinweg.
- Ökonomie, Nützlichkeit, Zumutbarkeit und Akzeptanz.

Bei elektronischen Fragebögen, die von den Respondenten bzw. Respondentinnen voll- oder teilautomatisch befüllt werden können (z.B. eQuest), ist zudem auf den Faktor **Kontinuität der Fragebogenstruktur** zu achten.

Damit diese Gütekriterien tatsächlich erreicht werden können, müssen potentielle Fehlerquellen durch das Design verhindert werden. Fehler bzgl. der **Validität**, **Messfehler** als solches sowie **Non-Response-Fehler** gelten dabei als die drei größten Fehlerquellen.

Test des Fragebogendesigns:

Nur durch die systematische Evaluation des Fragebogendesigns mittels eines Mixes an Testmethoden kann festgestellt werden, inwiefern die oben genannten Gütekriterien einer Messung tatsächlich erfüllt werden.

Beim Testen des Fragebogens kommt dem Gütekriterium Validität eine besondere Bedeutung zu, da dieses ausschließlich durch ein Überprüfen der gestellten Fragen inhaltlich beurteilt werden kann. Hierbei gilt es zu evaluieren, ob (1) alle wichtigen Aspekte des theoretischen Konzepts erfasst werden, (2) die Fragen bzw. die Merkmalsdefinitionen gleich verstanden werden und (3) die Fragen im Sinne der intendierten Art und Weise interpretiert werden. Eine Vernachlässigung von Fragebogentests insbesondere in diesem Bereich führt zu maßgeblichen Fehlern in der Analyse der mit dem Fragebogen gewonnenen Daten.

Die Verwendung verschiedener Befragungsmodi bei einer Erhebung (Mixed-Mode Design) und die immer stärker werdende Technologisierung der Modi bedeuten, dass auch das Fragebogentesten komplexer wird: Neben dem Testen der Fragen auf Validität (s. oben) muss auch der jeweilige Befragungsmodus berücksichtigt werden. Es gilt zu überprüfen, ob der verwendete Modus die Antwort auf eine Frage beeinflusst, ob die Fragebogenstruktur innerhalb eines Modus logisch aufgebaut ist, ob die An- bzw. Abwesenheit eines Interviewers bzw. einer Interviewerin (Selbst- bzw. Fremdauskunft) Einfluss auf die Antwort hat sowie ob die technischen Aspekte der verwendeten Hard- und Software den Ansprüchen des Fragebogendesigns genügen.

Aus diesem Grund, und weil sich die Gütekriterien eines Erhebungsinstruments auf verschiedenen inhaltlichen Ebenen bewegen, besteht ein guter Fragebogentest immer aus einem Mix verschiedener Testmethoden (Triangulation).

Dies ist insofern besonders wichtig, da die Testmethoden als komplementär zueinander zu verstehen sind. Beispielsweise identifizieren kognitive Pretests potentielle Probleme und schlagen Lösungen vor, während Experimente die Wirkung dieser Lösungen überprüfen (vgl. Presser et al. 2004). Jeder Fragebogen der Statistik Austria durchläuft daher - im Idealfall - systematisch einen - der jeweiligen Problemstellungen angepassten - Mix aus Testmethoden.

Für Befragungen, die zum Zweck der Erstellung von **Sozial- und Bevölkerungsstatistiken** durchgeführt werden, ist das korrekte Verständnis der Fragen bei persönlichen Interviews von großer Wichtigkeit. Je nach Komplexität der Fragestellung sind detaillierte Testmethoden vonnöten, welche die Qualität des Fragebogens bzw. einzelner Fragen in die Tiefe gehend beleuchten (z.B. Kognitives Testen).

Im Rahmen der **Datenerhebung von Unternehmensstatistiken** kommen ausschließlich standardisierte Webfragebogen (bzw. eingeschränkt Papierfragebögen) zum Einsatz. Der Schwerpunkt der Testmethoden wird daher in der Evaluation durch Experten (z.B. Buchhaltungspersonal von Unternehmen, Steuerberater) bzw. in der technischen Überprüfung der

Qualitätsrichtlinien der Statistik Austria

Funktionsweise der Erhebungsinstrumente liegen. Die Testmethoden entsprechen dabei den für die einzelnen Fachbereiche gültigen wissenschaftlichen Kriterien.

Richtlinien:

Fragebogendesign

- Der Fragebogen soll klar strukturiert und mit einem einheitlichen und passenden Layout versehen werden. Wenn nötig sollen Fragen durch Hilfetexte und Erläuterungen ergänzt werden, sodass diese für RespondentInnen wie Erhebungspersonen gleichermaßen gut verständlich sind. Routings und Filterführungen sorgen für einen leichten und effizienten Fluss durch den Fragebogen. Die Reihenfolge der Fragen muss mit Bedacht gewählt werden.
- Die Gestaltung des Fragebogens soll zum Ausdruck bringen, dass es sich um eine Erhebung von Statistik Austria für die Zwecke der Erstellung amtlicher Statistik handelt (u.a. durch Logo, Beschreibung der Statistik).
- Der Fragebogen soll durch Instruktionen begleitet sein, die das Antworten bzw. Ausfüllen so weit wie möglich erleichtern. Es soll der Respondentin bzw. dem Respondenten klar vermittelt werden, welche Frage als nächstes in Abhängigkeit welcher Antwort zu beantworten ist. Ebenso soll durch die Instruktionen vermittelt werden, welche Ausfüllungen (bei quantitativen Fragen Wertebereiche) zulässig sind.
- Ein Fragebogen soll sich an den jeweiligen Zielgruppen orientieren und den Respondenten bzw. die Respondentin leicht durch die gesamte Befragung leiten. Ein Fragebogen verwendet der Zielgruppe angepasste Begrifflichkeiten und ist, wenn möglich, in der Muttersprache des Respondenten bzw. der Respondentin vorhanden.
- Werden Erhebungspersonen im Datensammlungsprozess eingesetzt, ist eine detaillierte Projektvorstellung anzubieten. Etwaige Rückfragen, die für die Erfüllung des Werkvertrages relevant sind, sind zu beantworten. Um die Datenqualität bewerten zu können empfiehlt sich ein Debriefing der Erhebungspersonen.
- Die Interview- bzw. Erhebungssituation soll so standardisiert sein, dass unerwünschte Einflüsse auf die Ergebnisse (z.B. durch Verhalten des Interviewers/der Interviewerin) minimiert werden.
- Bei der Gestaltung des Fragebogens sind wissenschaftliche Gütekriterien wie Objektivität, Reliabilität, Validität, Vergleichbarkeit sowie Ökonomie, Nützlichkeit, Zumutbarkeit und Akzeptanz einzuhalten.
- Es werden ausschließlich standardisierte Fragebögen, die in der Regel aus exakt vorgegebenen Fragen und Antwortkategorien bestehen, verwendet. Dadurch soll sichergestellt werden, dass die Antworten vergleichbar und Unterschiede nicht auf die Messmethode zurückzuführen sind.
- Fragebogendesign und –test sind als Teilstück eines Prozesses zu verstehen, der Schleifen beinhaltet. Die Ergebnisse der Tests, Erkenntnisse aus der Beobachtung und Evaluation des Erhebungsprozesses sollen Eingang in das Fragebogendesign finden.
- Die Entwicklung vom eigentlich zu erhebenden Konzept zur Spezifikation dessen, was konkret gemessen werden soll, soll begleitend dokumentiert werden.

- Es soll angestrebt werden, konsistente Fragen zu bilden, die einerseits dem statistischen Konzept und andererseits dem Alltagsverständnis der Respondentinnen und Respondenten entsprechen. Jede Frage soll neutral und balanciert formuliert werden und passende Antwortmöglichkeiten anbieten.
- Bei der Übersetzung der Konzepte und Zielvariablen aus dem Englischen in die deutsche Sprache dürfen die nationalen Gegebenheiten nicht außer Acht gelassen werden.
- Im Fragebogendesign ist auf das Informationssystem (z.B. entsprechende IT-Systeme, Unterlagen, Gedächtnis) der Respondentinnen und Respondenten Rücksicht zu nehmen, z.B. mittels Hinweise auf bestimmte Unterlagen.

Testen von Fragebögen

- Jeder Fragebogen ist daraufhin zu testen, inwiefern durch dessen Verwendung die Gütekriterien einer Messung – Objektivität, Reliabilität, Validität, Ökonomie, Nützlichkeit, Zumutbarkeit und Akzeptanz – erfüllt werden. Zudem soll durch Testen das Ausmaß an potentiellen Befragungsmode-Effekten evaluiert werden.
- Das Testen von Fragebögen erfolgt systematisch und unter Verwendung eines – der jeweiligen Problemstellung angepassten und effizienten – Mix aus Testmethoden (Triangulation).
- Die Ergebnisse des Fragebogentests werden dokumentiert und zugänglich gemacht.
- Erkenntnisse aus Fragebogentests führen zu einer kontinuierlichen Weiterentwicklung des Fragebogens.

3.3.4 Maßnahmen zur Steigerung der Rücklaufquote

Statistik Austria ergreift eine ganze Reihe von Maßnahmen, um die Rücklaufquote bei Erhebungen zu steigern: Bereits vor der Erhebung erhalten die Betroffenen Vorabinformationsschreiben, welche sie über Inhalt, Freiwilligkeit der Teilnahme bzw. Auskunftspflicht und Zeitpunkt der Erhebung informiert. Bei freiwilligen Erhebungen werden teilweise Anreize sowohl in Form von finanziellen als auch in Form von nicht-finanziellen Aufwandsentschädigungen vergeben. Ein anderer Anreiz, RespondentInnen zur Teilnahme zu bewegen, ist die Übersendung von zur Verfügung gestellten Auswertungen (z.B. eigene Daten im Vergleich zum Durchschnitt, etc.) nach Beendigung der Erhebungsergebnisse.

Ein besonders wichtiges Instrument zur Vermeidung bzw. Minimierung von Antwortausfällen und somit zur Steigerung der Rücklaufquote sind Erinnerungen an den (gesetzlichen) Einsendetermin bzw. Urgenzen. Diese Maßnahmen sind ein wesentlicher Faktor zur Entsprechung der Grundsätze der Rechtzeitigkeit und Genauigkeit und daher eine der maßgeblichen qualitätssichernden Maßnahmen.

Generell ist davon auszugehen, dass der Erfolg der Datenerhebung stark von dem Umstand abhängig ist, ob die Erhebung **gesetzlich verpflichtend** oder auf **freiwilliger Basis** durchgeführt wird. Die Rücklaufquote von Datensammlungen auf Basis freiwilliger Erhebungen wird erheblich geringer anzunehmen sein als jene gesetzlich verpflichtender Erhebungen. Im Falle von Erhebungen auf freiwilliger Basis wird man mit einem schriftlichen bzw. telefonischen Ersuchen die potentiell Betroffenen neuerlich um ihre Mitwirkung bitten, um eine ausreichende Responserate und damit die Qualität der Ergebnisse sicherzustellen. Für derartige Erinnerungsschreiben gewinnen automatisierte Prozesse unter Verwendung der elektronischen Schiene (z. B. via Email oder zu einem festgelegten Zeitpunkt ausgelöste Warnhinweise über das Fehlen der Meldung im Rahmen der elektronischen Meldemedien) mehr und mehr an Bedeutung.

Formale Urgenzverfahren werden im Fall angeordneter Statistiken und Erhebungen grundsätzlich dann eingeleitet, wenn bis zum Zeitpunkt des durch die Anordnungsverordnung vorgegebenen Termins durch die von der Auskunftspflicht Betroffenen keine Meldung an Statistik Austria übermittelt wurde. Im Falle freiwilliger Erhebungen muss auf formale Urgenzverfahren verzichtet werden.

Statistik Austria ist bemüht, wenn immer möglich, die Einleitung formaler Urgenzverfahren zu vermeiden. Das bedeutet, dass in einem ersten Schritt versucht wird, die meldepflichtigen Betroffenen vor Fälligkeit mittels E-Mail an den Abgabetermin der noch ausstehenden Meldungen zu erinnern.

Das formale Urgenzverfahren besteht im Fall angeordneter Statistiken aus einem mehr-(zumindest zwei-)stufigen, in der Regel schriftlichen Prozess der Kontaktaufnahme mit den Respondentinnen und Respondenten, nämlich:

- einer ersten Mahnung in Form einer Erinnerung, welche vorwiegend mittels eMail durchgeführt wird
- einer zweiten Mahnung mittels Schreiben, welches nochmals auf die gesetzlichen Konsequenzen einer Nichtentsprechung der Meldeverpflichtung hinweist (verbunden mit einer nachweislichen Übernahme des Schriftstückes durch einen Bevollmächtigten der Meldeeinheit).

Darüber hinaus können weitere Urgenzen bezüglich jener Meldefälle erfolgen, für die zwar um eine Erstreckung der Meldefrist angesucht, diese jedoch nicht eingehalten wurde (Urgenz(en) nach Frist).

Im Falle der Erinnerungsschreiben gewinnen automatisierte Prozesse unter Verwendung der elektronischen informellen Schiene mehr und mehr an Bedeutung.

Die Termine für die einzelnen Verfahren sind im Einklang mit der Erhebungsperiodizität, dem Umfang des Aufarbeitungsprozesses sowie den Aufarbeitungs- und Veröffentlichungsterminen bereits in der Planungsphase festzulegen sowie als Kostenfaktor im Rahmen der budgetären Bedeckung zu kalkulieren.

Grundsätzlich ist zu berücksichtigen, dass Statistik Austria nach bestehender Rechtslage nicht verpflichtet ist, Urgenzverfahren durchzuführen. Die Einleitung eines Verwaltungsstrafverfahrens bei der in erster Instanz zuständigen Bezirksverwaltungsbehörde wegen Verletzung der Mitwirkungspflicht oder wissentlich unvollständiger oder nicht dem besten Wissen entsprechenden Angaben wird als ultima ratio angesehen. Damit soll ein möglichst umfassendes Verständnis der Respondentinnen und Respondenten für die Notwendigkeiten angeordneter Erhebungen geweckt und erhalten werden. Ein durchgeführtes Verwaltungsstrafverfahren entbindet nicht von der bestehenden Auskunftspflicht.

Richtlinien:

- Maßnahmen zur Steigerung der Rücklaufquote sind als Instrumente zur Vermeidung bzw. Minimierung der Antwortausfälle und damit zur Sicherung einer ausreichenden Responserate unerlässlich.
- Art, Umfang, Zahl und Intensität der Urgenzen sind bei der Planung und insbesondere der Sicherstellung der budgetären Bedeckung einer statistischen Erhebung zu berücksichtigen. Dabei ist unter Berücksichtigung der Grundsätze der Minimierung der Belastung der Respondentinnen und Respondenten sowie der Kostenminimierung vorzugehen.
- Bevor formale Urgenzverfahren zur Anwendung kommen, soll versucht werden, durch neuerliche Kontaktaufnahme mit der Respondentin bzw. dem Respondenten (e-Mail-Erinnerung an den Abgabetermin etc.) die Erhebungseinheit zur Lieferung der gewünschten Informationen zu bewegen.
- Bei Verwendung von Webfragebögen durch die Respondentinnen und Respondenten im Falle einer vorhandenen Respondentendatenbank (und damit einer Email-Adresse) sollte insbesondere im Fall von angeordneten Erhebungen ein elektronisch automatisiertes 1. Mahnverfahren forciert werden.
- Eine 2. Mahnung mittels Schreiben, welches nochmals auf die gesetzlichen Konsequenzen einer Nichtentsprechung der Meldeverpflichtung hinweist, soll aus Gründen der Beweiskraft grundsätzlich schriftlich mittels Rsb-Schreiben erfolgen.
- Die Wahl der Art und des Umfangs sowie der Häufigkeit hängt vor allem von der Responserate, die man erreichen will, aber auch von den projektspezifischen Fristen hinsichtlich der Veröffentlichungs- und Übermittlungspflichten ab. Dabei müssen die zur Verfügung stehenden Ressourcen des mit der Erhebung beauftragten Organs der Statistik beachtet werden.

3.4 Datenaufbereitung (Datenbearbeitung)

Der Prozess der Datenaufbereitung stellt das Herzstück der statistischen Produktion dar. Es sind jene Schritte zu vollziehen, die aus den Rohdaten den authentischen Datenbestand erzeugen. Im Zentrum der Arbeiten stehen Korrekturen und Ergänzungen der Daten mit der Zielsetzung eventuell auftretende systematische Verzerrungen („Bias“), die die Qualität der von Statistik Austria den Nutzerinnen und Nutzern zur Verfügung gestellten Ergebnissen beeinflussen könnten, zu beseitigen. Sämtliche Arbeitsschritte müssen jedenfalls im Sinne der Nachvollziehbarkeit genau dokumentiert werden.

3.4.1 Datenerfassung und Signierung

Um Daten sinnvoll weiterverarbeiten zu können, müssen sie in elektronischer Form zur Verfügung stehen. **Datenerfassung** bezieht sich auf den Vorgang, der verfügbare Informationen in ein elektronisches Format umwandelt. Im Falle einer elektronischen bzw. computerunterstützten Erhebung bzw. bei der Übermittlung administrativer Datenbestände ist dies durch die Art der Datenbeschaffung bereits implizit gegeben; dadurch werden Erfassungsfehler deutlich reduziert. Im Falle postalischer Erhebungen muss die Datenerfassung manuell erfolgen. Eventuell können Fragebögen auch mittels Scanning-Techniken eingelesen werden. In den seltenen Fällen von offenen Antworten können Text Mining Algorithmen verwendet werden, um die Informationen, welche für Statistik Austria von Interesse sind, aus den textuellen Antworten zu erkennen.

Unter **Signierung** (Coding) wird das Zuweisen alphanumerischer Schlüssel zu einer textuellen Antwort verstanden. Zumeist erfolgt der Signiervorgang automatisch oder zumindest automationsunterstützt. Für spezielle Anwendungen ist teilweise auch eine manuelle Signierung erforderlich. Um Antworten erfolgreich automatisch signieren zu können, ist das Vorhandensein hinreichend entwickelter Signierschlüssel eine unabdingbare Voraussetzung. Darüber hinaus müssen jene Methoden, die die Zuordnung der Schlüsselausprägungen zu den textuellen Antworten vornehmen, fundiert sein. Algorithmen zur automatischen Signierung innerhalb von Statistik Austria bedienen sich verschiedener Methoden des Textabgleichs. Der Signiervorgang erfolgt oftmals in zwei Stufen. Zunächst werden so viele Datensätze wie möglich automatisch signiert. In einem zweiten Schritt werden die übrigen Datensätze entweder automationsunterstützt oder manuell signiert.

Die Signierung unter Verwendung von komplexen Klassifikationen bedarf einer besonderen Qualifikation der Mitarbeiterinnen und Mitarbeiter, damit eine hinreichende Ergebnisqualität erzielt wird. Eine gezielte Schulung ist notwendig, um eine korrekte Identifizierung und Zuordnung in Entsprechung der klassifikatorischen Regeln sicherzustellen und die notwendige Vergleichbarkeit zu gewährleisten.

Richtlinien:

Datenerfassung

- Für jedes Merkmal muss im Vorhinein der Wertebereich bzw. der Wertevorrat festgelegt werden. Ebenso muss entschieden werden, welche Ausprägungen bzw. welche Werte für nicht numerisch behandelbare Situationen (z. B. fehlende Werte, „trifft nicht zu“, etc.) herangezogen werden. Derartige Werte/Ausprägungen müssen deutlich erkennbar sein.
- Im Falle manueller Datenerfassung soll das Erfassungspersonal auf kritische Eingabefelder vorab aufmerksam gemacht werden. Eingaben außerhalb von Wertebereichen sollen nicht zugelassen werden. Durch doppelte Erfassung der Datensätze und anschließendem Vergleich der Eingaben sollen Fehleingaben auf ein Minimum reduziert werden.
- Bei offenen Antworten soll der Einsatz von Text Mining Algorithmen angedacht und nach erfolgreichem Testen angewandt werden.

Signierung

- Bei der Planung der Erhebung soll auf offene Fragen möglichst verzichtet werden, damit Signierung nur dort notwendig ist, wo sie unbedingt erforderlich ist.
- Signierung soll, wenn immer möglich, automatisch durchgeführt werden. Die Rate der automatischen Zuordnungen sollte möglichst hoch sein. Die Parameter zur automatischen Signierung sollen so gewählt werden, dass einerseits eine möglichst hohe Zahl an Datensätzen signiert, aber andererseits die Zahl der Fehlsignierungen minimiert wird.
- All jene Datensätze, die nicht automatisch signiert werden können, sollen möglichst automationsunterstützt durch Expertinnen und Experten signiert werden. Durch ausreichende Schulung sollen diese in die Lage versetzt werden, die verbleibenden Datensätze korrekt zu signieren.
- Im Falle der Verwendung komplexer Klassifikationen sollen jene Mitarbeiterinnen und Mitarbeiter, die den Signiervorgang durchführen, gezielt geschult werden, um eine korrekte Identifizierung und Zuordnung in Entsprechung der klassifikatorischen Regeln sicherzustellen.
- Nach Finalisierung der Signierung sollte eine Substichprobe aller Datensätze betrachtet werden, um abzuschätzen, wie hoch der Anteil der Fehlsignierungen ist.

3.4.2 Plausibilitätsprüfung

Um statistische Daten sinnvoll zu Ergebnissen verarbeiten zu können, muss sichergestellt werden, dass die verwendeten Datensätze **fehlerfrei**, logisch **konsistent** und inhaltlich **plausibel** sind. Die Durchführung der Mikroplaus findet zumeist IT-technisch automatisiert statt. Manchmal kann es jedoch aufgrund der Sachlage notwendig sein, zumindest einen Teil der Datensätze (vorliegende Fragebögen) manuell zu überprüfen. Bei einer manuellen Überprüfung erfolgt die Korrektur sehr oft dadurch, dass eine Sachbearbeiterin bzw. ein Sachbearbeiter im Zuge telefonischer Kontaktaufnahme mit der Respondentin bzw. dem Respondenten zu einer Lösung gelangt. Die Abgrenzung der Plausibilitätsprüfung („Data Editing“) ist je nach verwendeter Definition unterschiedlich. Während laut [UN-Definition](#) auch der Korrekturprozess inbegriffen ist, versteht [Eurostat](#) (in Einklang mit der Definition von Statistics Canada) unter Plausibilitätsprüfung („**Statistical Editing**“) lediglich das Aufspüren von Fehlern. Die Korrektur wird dann dem Prozess der Imputation zugerechnet. In der Praxis sind beide Vorgangsweisen vorstellbar. Gleichzeitige Fehlerkorrektur im Zuge der Fehlererkennung und andererseits das Löschen fehlerhafter Werte und Auffüllen durch anschließende Imputation. Wegen der inhaltlichen Nähe werden Plausibilitätsprüfung und Imputation gemeinsam unter dem Begriff **Datenbereinigung** („**Data Cleaning**“) subsumiert.

Der Vorgang der Plausibilitätsprüfung umfasst zwei Teilbereiche. Bei der **Mikroplaus** („**Micro Editing**“) werden Einzeldatensätze auf inhaltliche Konsistenz überprüft. Zu diesem Zweck werden **Plausibilitätsregeln** („**Edit Rules**“) aufgestellt, die auf jeden einzelnen Mikrodatensatz angewandt werden. Ein besonderer Aspekt bei der Mikroplaus ist die **Erkennung von Ausreißern** („**Outlier Detection**“), extremen Werten bei quantitativen Variablen, die die Ergebnisse stark beeinflussen können. Im Falle computerunterstützter Datensammlung (z. B. CAPI) können Plausibilitätsprüfungen und Korrekturen bereits während der Dateneingabe durchgeführt werden („**Editing near the source**“). Vor allem im Falle manueller Überprüfung kann sich der Prozess der Mikroplaus als sehr aufwendig erweisen. In den letzten Jahren wurden daher Strategien entwickelt, die eine Steigerung der Effizienz ermöglichen. So kann zum Beispiel bei inhomogenen Massen die Mikroplaus nur auf solche Einheiten beschränkt werden, deren Größe über einer zuvor definierten Grenze liegt und daher angenommen werden muss, dass deren Einfluss auf die Gesamtaggregate relevant ist („**Selective Editing**“).

Im Zuge der **Makroplaus** („**Macro Editing**“) werden nach der eigentlichen Datenbereinigung Aggregate und Verteilungen hinsichtlich ihrer inhaltlichen Glaubhaftigkeit überprüft (Vergleich zu Vorperioden, Vergleich mit anderen Quellen etc.). Dazu kann es bereits im Zuge des Planungsvorgangs erforderlich sein, geeignete Kennzahlen bzw. Indikatoren und Quoten als Messgrößen zu entwickeln, zu testen und festzulegen. Nicht glaubwürdige Ergebnisse müssen näher analysiert werden und nach Abklären der Ursache sind Mikrodaten gegebenenfalls zu korrigieren.

Der Bereich der Plausibilitätsprüfung ist auch Gegenstand intensiver Forschungstätigkeit, die von Kooperation universitärer Einrichtungen mit der Amtlichen Statistik geprägt ist. Hier ist vor allem die [UNECE Work Sessions on Data Editing](#) zu nennen, wo eine Vielzahl von theoretischen Beiträgen sowie praktischer Anwendungen verfügbar ist.

Obzwar Plausibilitätsprüfungen einen wichtigen Beitrag zur Qualitätssteigerung statistischer Produkte liefern, kann übertriebenes iteratives Korrigieren zu kontraproduktiven Effekten und Verzerrungen kommen.

Richtlinien:

- Bevor es zur Veröffentlichung statistischer Ergebnisse kommen kann, müssen sowohl auf Mikro- als auch auf Makroebene Plausibilitätsprüfungen durchgeführt werden.
- Im Falle elektronischer Datensammlung sollen geeignete Plausibilitätsregeln nach Maßgabe der Respondententlastung bereits in der Erfassungssaplikation integriert sein, um fehlerhafte Eingaben nach Möglichkeit zu vermeiden.
- Bei der Formulierung der Plausibilitätsregeln auf Mikroebene sollen folgende fehlerhafte Situationen erkannt werden:
 - Werte außerhalb des definierten Wertebereichs
 - nicht konsistente Wertekombinationen
 - unplausible Wertekombinationen
 - unvollständige Datensätze
 - unplausible Werte aufgrund von Vergleichen mit Vorperiodenergebnissen
- Für die Makroplaus sollen möglichst zuverlässige Vergleichsreferenzen herangezogen werden.
- Je nach Anwendung müssen in geeigneter Form Überlegungen hinsichtlich Aufwand/Qualität Eingang in das Design der Plaus finden.
- Die im Rahmen der Plausibilitätsprüfung einzusetzenden Methoden sollen grundsätzlich im Vorfeld getestet werden, um den Einfluss auf die resultierenden Ergebnisse abschätzen zu können.
- Es soll vermieden werden, durch übertriebenes iteratives Korrigieren der Daten die Ergebnisse einem implizit gegebenen Modell anzunähern, weil dadurch Verzerrungen in den resultierenden Schätzern entstehen können („Over-Editing“).

3.4.3 Imputation

Imputation ist jener Vorgang, der **fehlende** oder **ungültige Werte** (die nicht durch Datenkorrektur lt. Kapitel „Plausibilitätsprüfung und Datenkorrektur“ hergeleitet werden können, bzw. **inkonsistente Werte**, die während der Plausibilitätsprüfung entdeckt wurden, ersetzt. Dies geschieht durch das Ändern bzw. Einfügen von Merkmalsausprägungen auf einer als fehlerhaft und/oder unvollständig identifizierten Beobachtung mit dem Ziel, eine **vollständige, plausible und intern kohärente** Beobachtung zu erhalten. Geht man von der Praxis aus, ist mit Imputation die Ergänzung fehlender Werte („**Missing Values**“) bei Erhebungsdaten (statistischen Produkten) auf **Mikrodatenniveau** gemeint. Imputation ist nicht mit der Zuschätzung nicht erhobener Teilpopulationen gleichzusetzen. Oftmals wird der Vorgang der Imputation auch als zweiter Teil der **Datenbereinigung** gesehen und steht daher in engem inhaltlichen Zusammenhang zur Plausibilitätsprüfung (siehe Kapitel 3.4.2).

Es gibt verschiedenste Ursachen für fehlende Werte, bei klassischen Befragungen ist der häufigste Fall die **Antwortverweigerung (Non Response)**. Andere Ursachen können zum Beispiel Übermittlungsfehler technischer Natur, fehlende Verknüpfungen zu Verwaltungsdaten, Interviewerfehler oder gelöschte Werte im Zuge der Plausibilitätsprüfung sein.

Imputation soll vor allem dem Zweck dienen, die durch Antwortverweigerung entstandene **Verzerrung („Non Response Bias“)** auszugleichen. Dabei wird meist nur das Fehlen einzelner Merkmale auf einer Beobachtung („Item-Non Response“) durch Imputation korrigiert. Es ist auch möglich, ganze Beobachtungen, die nicht vorhanden sind („**Unit-Non Response**“), durch Imputation zu ergänzen. Dies sollte jedoch vermieden werden, meist wird das Problem des Ausfalls ganzer Einheiten durch Gewichtung (siehe Kapitel 3.5.2) ausgeglichen („**Non Response Weighting**“). Allerdings kann es auch das Erhebungsdesign erfordern, einzelne fehlende Beobachtungen durch Imputation zu ergänzen (z.B. einzelne, fehlende Personeninterviews im Zuge von Haushaltserhebungen). Ein Charakteristikum der Imputation besteht darin, dass der für den Einzelfall eingesetzte Wert mit **Unsicherheit** verknüpft ist, also mit einer gewissen Wahrscheinlichkeit nicht dem wahren Wert entsprechen wird. Eine mögliche Methode diese Unsicherheit in die Fehlerrechnung miteinzubeziehen ist **Multiple Imputation** (Rubin, 2004), wobei imputierte Werte als Realisation eines Zufallsprozess definiert sind. Dementsprechend wird das Einsetzen eines Wertes, das durch gesicherte Verknüpfung mit einer externen sehr zuverlässigen Verwaltungsquelle erfolgt, nicht als Imputation bezeichnet.

Es ist wichtig zu beachten, dass die Methoden der Imputation nur dann sinnvoll anwendbar sind, wenn die Quantität der fehlenden Werte **klein im Verhältnis zur Gesamtzahl** der Beobachtungen bleibt. In manchen Situationen kann es jedoch notwendig sein, große Teile oder sogar alle Werte zu einem spezifischen Merkmal des Datenkörpers zu ergänzen. Diese auch als **Massenimputation** bezeichnete Vorgangsweise sollte allerdings nur gewählt werden, wenn es keine anderen sinnvollen Optionen gibt.

Der Prozess der Imputation durchläuft mehrere Stufen:

1. Analyse der Struktur der fehlenden Werte
2. Imputation
3. Analyse der imputierten Werte

Am Beginn steht eine Analyse der fehlenden (und unplausiblen) Werte, um die Datenprobleme so gut wie möglich zu verstehen. Diese ersten Analysen sollten auch Visualisierungen beinhalten, da grafische Analysen oft zu einem guten Verständnis der Situation führen, hierfür kann z.B. das R-Paket VIM, welches im Bereich Methodik entwickelt wurde, verwendet werden. Des Weiteren ist es notwendig die Struktur der fehlenden Werte in Bezug auf Merkmale, die für die jeweilige Erhebung von Interesse sind, genauer zu verstehen. Ein gängiges Konzept, das hier zur Hilfe genommen werden kann, ist die Einteilung von Little & Rubin (1987) in 3 verschiedenen Szenarien:

- „missing completely at random“ (MCAR) – Das Fehlen eines Werts ist rein zufällig.
- „missing at random“ (MAR) geschehen. - Das Fehlen eines Werts hängt nur von den beobachteten Merkmalen ab.
- „missing not at random“ (MNAR) – “Worst Case”: Das Fehlen hängt von dem Wert selber ab.

Von dieser Analyse und von der Verteilung der Daten hängt im Wesentlichen die Wahl der Methode ab, die im Folgenden angewandt wird.

Prinzipiell lassen sich die Methoden nach Art der Ermittlung des imputierten Wertes klassifizieren:

- Methoden, die auf Merkmalen der jeweiligen Beobachtung fußen,
- Methoden, die auf Spenderdaten zurückgreifen,
- Methoden, die Verteilungsparameter des betrachteten Merkmals verwenden,
- Zeitreihenmethoden.

Für die Wahl der Methode und die Festsetzung damit verbundener Parameter ist auch entscheidend, welche Hypothese man über die durch Non Response vermutete Verzerrung in den Ergebnissen annimmt. Oft ist der Ablauf der Imputation auch ein iteratives Verfahren, in dem versucht wird, den Imputationsprozess zu verbessern, vorallem mit Hilfe der Analyse der imputierten Werte und der Ergebnisse. Ist die Qualität der Imputation schlecht, so wird nach möglichen Fehlerquellen in der Durchführung und der IT-Umsetzung gesucht oder gegebenenfalls eine andere Imputationsmethode gewählt.

Richtlinien:

- Falls nicht zwingende Gründe dagegen sprechen, sollen bei jedem Datenkörper fehlende und ggf. als unplausible definierte Werte ersetzt werden. Dabei ist zu beachten, dass im Zuge der Plausibilitätsprüfung definierte Regeln nicht verletzt werden.
- Imputation soll nur für den Fall des Fehlens einzelner Merkmalswerte („Item-Non Response“) durchgeführt werden. Es muss für jedes Projekt definiert werden, ab wann eine Beobachtung als gänzlich fehlend eingestuft wird.
- Die Kompensation von Totalausfällen von Beobachtungen („Unit-Non Response“) sollte wenn möglich durch Gewichtung (siehe Kapitel 3.5.2) erfolgen.
- Massenimputationen einzelner Merkmale (in etwa >20%) sollen nur dann durchgeführt werden, wenn keine anderen Optionen (wie z.B. Gewichtung) machbar erscheinen.
- Die Wahl der Imputationsmethode muss mit Bedacht auf die jeweilige Problemstellung erfolgen und ist für jedes Merkmal einzeln zu definieren
- Das Ergebnis eines Imputationslaufes ist zu analysieren und es ist zu beurteilen, ob die durch Imputation bewirkten Effekte sich im Rahmen des Erwünschten bewegen. Insbesondere sind Verteilungen und errechnete Aggregate vor und nach einer Imputation zu vergleichen. Um Nachvollziehbarkeit zu gewährleisten, müssen all jene Werte, die imputiert wurden, markiert werden (z.B. mit Flag-Variablen).
- Bei periodischen Erhebungen sollen die gewählten Imputationsmethoden regelmäßig evaluiert werden. Besondere Beachtung sollen dabei Änderungen in der Struktur der fehlenden Werte finden.
- Die Durchführung von Imputationen erfordert entsprechendes Know-how. Bei jedem neuen Projekt bezüglich Imputation muss der Bereich Methodik einbezogen werden. Die Wahl der Methoden, die Ergebnisbeurteilung und die Entscheidung über eine etwaige Adaptierung der Methode sollen in Absprache zwischen Fachbereich und Methodik geschehen.
- Um dem Anspruch des Einsatzes bestmöglicher Methoden („Current Best Practices“) gerecht zu werden, soll die internationale Forschungstätigkeit auf dem Gebiet der Imputation laufend beobachtet werden. Dies betrifft sowohl fachspezifische anwendungsorientierte Aktivitäten in anderen NSI's als auch allgemeine theoretische Betrachtungen, die zur Entwicklung generalisierter Lösungen verwendbar sind.

3.4.4 Verknüpfung verschiedener Datenquellen

Der statistische Produktionsprozess war in den letzten Jahren einem Wandel unterworfen. In vielen Fällen bietet die Heranziehung einer einzelnen Datenquelle zu wenig Information, um den analytischen Zwecken eines statistischen Projektes gerecht zu werden. Bei Erhebungen ist es aufgrund der Entlastung der Respondentinnen und Respondenten sinnvoll, Merkmale aus administrativ verfügbaren Quellen zuzuführen. Die effiziente und primäre Nutzung bestehender Verwaltungs- und Registerdaten für Zwecke einer statistischen Datenproduktion stellt auch eine der zentralen Forderungen des BStatG 2000 dar. Allerdings müssen oftmals mehrere externe Datenkörper in Betracht gezogen werden, um einen für ein bestimmtes statistisches Produkt geeigneten **Authentischen Datenbestand** erzeugen zu können. Nicht zuletzt aus diesen Gründen ist für viele Projekte die **Verknüpfung verschiedener Datenquellen** ein wichtiger Produktionsschritt im Zuge des Aufarbeitungsprozesses.

Die Rahmenbedingungen des Verknüpfungsprozesses können dabei unterschiedlich sein, folgende Verknüpfungsarten können unterschieden werden:

- Exact Record Linkage
- Record Linkage
- Statistical Matching

Exact Record Linkage

Im einfachsten Fall werden zwei Datenkörper, die beide mit denselben Identifikationsmerkmalen („Schlüsselvariable“) versehen sind, miteinander in Beziehung gesetzt und die Datensätze werden paarweise zusammengeführt, wobei wichtig ist, dass die zusammengeführten Einheiten einander inhaltlich entsprechen. Wie *Abbildung 9* zeigt, können nicht immer sämtliche Datensätze eine Entsprechung finden.

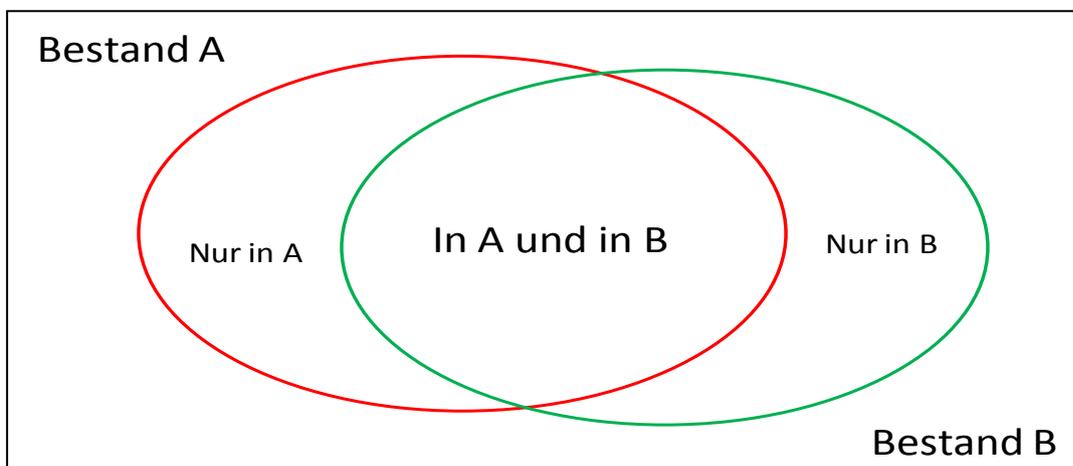


Abbildung 9: Zusammenführung zweier Datenkörper. Es ist sowohl ein „Inner Join“ (in A und in B), als auch ein „Outer Join“ (in A oder in B) möglich.

Bei der Zusammenführung von Einzeldaten muss besonders auf die Datenschutzbestimmungen geachtet werden. Aus diesem Grund werden die Identitätsdaten bei natürlichen Personen durch das bereichsspezifische Personenkennzeichen Amtliche Statistik (bPK-AS) und bei Daten von Unternehmen durch nicht-umkehrbare Unternehmenskennzahlen ersetzt, sobald sie nicht mehr aus den in § 5 Abs. (2) BStatG 2000 genannten Gründen oder für eine weitere angeordnete statistische Erhebung erforderlich sind. Dadurch ist eine vollständige Anonymisierung gewährleistet (§ 15, Abs. (1) BStatG 2000).

Wenn Beziehungen zwischen Schlüsselvariablen nicht eindeutig sind, d.h. es existieren zu einem Datensatz mehrere Entsprechungen in einem anderen Datenkörper (z. B. zwischen Einheiten des Unternehmensregisters und den Dienstgeberkonten des Hauptverbands), sind vor der Verknüpfung umfassende Analysen zu erstellen. Bei solchen 1:n- oder auch bei m:n-Beziehungen muss untersucht werden, wie die Aggregation der mehrfach gefundenen Datensätze erfolgen kann.

Record Linkage

Etwas komplexer ist die Datenverknüpfung, wenn die Datensätze nicht über dieselben Identifikationsmerkmale verfügen und vor Zusammenführung der Datensätze die Beziehung zwischen den unterschiedlichen Schlüsselvariablen implementiert werden muss. In manchen Datenquellen existiert keine Schlüsselvariable, über die eine Zusammenführung erfolgen kann. In diesem Fall hat man die Möglichkeit, über textuelle Abgleiche einzelner Identifikationsmerkmale (z.B. die Adresse), mittels ähnlicher Methoden wie bei automatischer Signierung (siehe Kapitel 3.4.1), die Daten sinnvoll zu verknüpfen. Daraus ergibt sich die Notwendigkeit, Beziehungen zwischen wichtigen Schlüsselvariablen geeignet zu speichern und laufend zu warten.

Statistical Matching

Ist auch eine Verknüpfung durch Record Linkage nicht möglich, kann - wo dies inhaltlich geeignet erscheint - versucht werden, jedem Datensatz ein möglichst ähnliches Duplikat zuzuordnen. Grundvoraussetzung für diese, auch „Statistical Matching“ genannte Technik ist, dass die zu verknüpfenden Datenkörper möglichst viele gemeinsame Merkmale aufweisen. Aus diesen kann dann eine Distanzfunktion gebildet werden, auf deren Basis die Zusammenführung durchgeführt wird.

Qualitätsbewertung der Datenquellen (Fehlerrechnung)

Die Qualität eines aus Verwaltungs- und Registerdaten erstellten statistischen Produkts kann mithilfe folgender Schritte beurteilt werden:

- Bewertung der spendenden Register insbesondere hinsichtlich
 - Kompatibilität der Definitionen
 - Administrativer Zweck
 - Datenpflege
 - Befüllungsgrad der Schlüsselvariable und der benötigten Merkmale
 - Übereinstimmung mit einem geeigneten Referenzdatensatz
- Bewertung des zusammengeführten Datensatzes, beispielsweise im Falle von einem Merkmal in mehreren Quellen
- Bewertung der imputierten Daten mithilfe geeigneter Bewertungskennzahlen (z.B. Klassifikationsrate)

Wann immer ein statistisches Produkt aus Verknüpfungen von verschiedener Datenquellen erstellt wird, sollte eine Qualitätsbewertung mittels des skizzierten Ablaufs in Betracht gezogen werden.

Richtlinien:

- Bevor die Verknüpfung von Daten aus unterschiedlichen Datenquellen durchgeführt wird, sind genaue Informationen über die verwendeten Konzepte und Definitionen der involvierten Datenkörper einzuholen und zu dokumentieren. Dies betrifft sowohl die Massen und Einheiten als auch die enthaltenen Merkmale.
- Bei der Zusammenführung von Datensätzen auf Individualebene müssen stets sämtliche, für die statistische Geheimhaltung relevanten Rechtsvorschriften beachtet werden.
- Wenn möglich, soll die Zusammenführung unterschiedlicher Datenquellen mittels einer gemeinsamen, dem Datenschutz entsprechenden, Schlüsselvariable durchgeführt werden.
- Um einen möglichst hohen Zusammenführungsgrad zu gewährleisten ist eine mehrstufige Zuordnung anzuwenden. Je nach Fragestellung können unterschiedliche Strategien sinnvoll sein. Gelingt es nicht, alle Einheiten über ein Identifikationsmerkmal zusammenzuführen, kann zunächst versucht werden, mittels Textabgleichen zu eindeutigen Zuordnungen zu gelangen. In zweiter Linie kann versucht werden, mittels „Statistical Matching“ möglichst ähnliche Einheiten zu finden. Bleibt ein Rest von nicht verknüpfbaren Einheiten, können geeignete Schätzverfahren angewandt werden.
- Sind Zuordnungen nicht eindeutig und existieren daher zu einem Ausgangsdatsatz mehrere Zuordnungen im Zieldatenkörper, müssen klar nachvollziehbare Regeln definiert werden, nach welchen der endgültige Zieldatensatz zu selektieren ist.
- Sind die Einheiten zwischen den zu verknüpfenden Datenbeständen auf unterschiedlicher Ebene, so muss klar sein, wie dies im Zieldatensatz berücksichtigt wird. Ist eine Disaggregation (z.B. Unternehmensdaten hinunterzuberechnen auf die Arbeitsstättenebene) notwendig, muss dokumentiert werden, mit welchen Modellen dies erfolgt.
- Statistical Matching ist nur dann eine Option, wenn die zu verknüpfenden Datenkörper eine hinreichend große Menge gemeinsamer Merkmale aufweisen. In diesem Fall muss eine geeignete Methode angewandt werden.
- Sämtliche Aspekte, die mit der Verknüpfung von Datenquellen verbunden sind, müssen in der Standard-Dokumentation beschrieben werden. Dies betrifft sowohl die einbezogenen Datenquellen wie auch die verwendeten Methoden und quantitativen Informationen über den Verknüpfungsprozess (z. B. Verknüpfungsraten).

3.5 Auswertung und Analyse

3.5.1 Arbeits- und Analysetabellen als Grundlage von Publikationen

Arbeits- und Analysetabellen begleiten den kompletten statistischen Produktionsprozess und somit auch den Schritt zur Publikation. Der Einsatz dieser Tabellen ist vielschichtig, wobei deren Inhalte und Ziele maßgeblich von der Charakteristik des jeweiligen statistischen Projekts abhängen.

Bevor über die Art und den Umfang einer Publikation (siehe Kapitel 3.6) endgültig entschieden wird, ist durch die Produktion geeigneter, gut definierter und strukturierter Arbeits- bzw. Analysetabellen eine Grundlage zur soliden Beurteilung des vorliegenden Datenmaterials zu schaffen. Diese werden zwar ebenso wie spätere Publikationstabellen aus dem Authentischen Datenbestand generiert, bleiben aber unveröffentlicht.

Arbeitstabellen dienen vor allem der finalen Kontrolle von Ergebnissen vor deren Veröffentlichung. Unter Vorwegnahme des geplanten Veröffentlichungs- bzw. Lieferprogramms stellen sie das Zahlenmaterial mittels detaillierter Merkmalskombinationen und Aggregationen möglichst vollständig dar.

Analysetabellen dienen hingegen vorrangig der Abklärung weiterführender Fragestellungen, beispielsweise zur:

- Geheimhaltung
- Genauigkeit von Ergebnissen
- Sinnhaftigkeit von Gliederungen
- Eignung von Ergebnissen für die Einlagerung in diversen Datenbanken
- Erstellung von abgeleiteten Datenbeständen
- Machbarkeit von Sonderauswertungen

Richtlinien:

- Vor einer Veröffentlichung sollen begleitende Arbeits- bzw. Analysetabellen zumindest das Publikationsprogramm vollständig simulieren und damit alle zur Veröffentlichung geplanten Zellen vorab erzeugen.
- Beim Entwurf von Arbeits- bzw. Analysetabellen soll zwecks leichter Betrachtung und Interpretierbarkeit zweidimensionalen Tabellen der Vorzug gegeben werden.
- Die Tabellen sollen auf Authentischen Datenbeständen beruhen.
- Mit der Erstellung der Tabellen einhergehende Probleme sollen im Team bzw. unter Einbeziehung der Projektleitung diskutiert werden.
- Im Zusammenhang mit den Tabellen stehende Aktivitäten sollen zumindest so dokumentiert bzw. archiviert werden, dass daraus Prozesse der Entscheidungsfindung gut ersichtlich sind.

3.5.2 Schätzung aus Stichprobenerhebungen

Von einer **Stichprobe** spricht man, wenn nicht Daten der **Grundgesamtheit einer Population** sondern lediglich Teile der Population in einer Erhebung betrachtet werden. Diese Teilmenge der Population wird aus einem Auswahlrahmen gezogen, welcher idealerweise mit der Grundgesamtheit ident ist. Abweichungen zwischen Auswahlrahmen und Grundgesamtheit können sich jedoch aufgrund von kurzfristigen Veränderungen der Zusammensetzung der interessierenden Grundgesamtheit sowie nicht erfassten Elementen ergeben.

Zumeist ist die Größe der Stichprobe wesentlich kleiner als die der Grundgesamtheit, womit für die ausführende Organisation eine wesentliche **Arbeits- und Kostenersparnis** gegenüber einer Vollerhebung gegeben ist. Demgegenüber steht der Nachteil des Genauigkeitsverlustes, der durch das Nichtbefragen von Einheiten entsteht. Da jedoch die oben genannten Vorteile deutlich überwiegen und die Ungenauigkeiten größtenteils kontrollierbar sind, sind Stichprobenverfahren in der **Amtlichen Statistik weit verbreitet und unverzichtbar**. Nicht zuletzt wegen der gegenüber Vollerhebungen deutlich geringeren Belastung der Respondentinnen und Respondenten sind Stichproben gemäß BStatG 2000 Vollerhebungen vorzuziehen, wenn immer dies möglich ist.

Es gibt zwei Arten von Stichproben, die in der Amtlichen Statistik Verwendung finden, nämlich **Zufallsstichproben** und **Nicht-Zufallsstichproben**.

Zufallsstichproben

Die wichtigste Eigenschaft einer Zufallsstichprobe ist, dass man jeder Einheit der Grundgesamtheit, welche auch im Auswahlrahmen erfasst ist, eine von Null verschiedene Auswahlwahrscheinlichkeit zuordnen kann. Von einer reinen Zufallsstichprobe spricht man, wenn alle Einheiten dieselbe Wahrscheinlichkeit haben in die Stichprobe gezogen zu werden und die Ziehung in der ganzen Grundgesamtheit gleichzeitig stattfindet. Reine Zufallsstichproben finden in der Offiziellen Statistik eher selten Verwendung. Im Gegensatz zu reinen Stichproben ist die Auswahlwahrscheinlichkeit bei geschichteten Zufallsstichproben nur innerhalb einer Gruppe von Respondenten (einer „Schicht“) mit denselben Eigenschaften (wie z.B, Geschlecht, Bundesland), gleich groß. Bei Statistik Austria sind nahezu ausschließlich geschichtete (stratifizierte) Zufallsstichproben in Verwendung. Manchmal kommen auch mehrstufige Zufallsstichproben zum Einsatz. Eine Zufallsstichprobe ist durch ihr Design charakterisiert, welches direkt die Wahrscheinlichkeiten in die Stichprobe zu gelangen bestimmt. Die Ergebnisse aus Zufallsstichproben sollen Rückschlüsse auf die zu untersuchenden Fragestellungen über die Grundgesamtheiten ermöglichen. Da dies mit Ungenauigkeiten behaftet ist, spricht man auch von **statistischen Schätzungen**. Ziel der Schätzungen ist es, ein möglichst unverzerrtes Bild der Realität zu geben.

Bei Zufallsstichproben bildet die **Hochrechnung** den entscheidenden Aufarbeitungsschritt zur Vorbereitung der Schätzung. Dabei wird jeder Stichprobeneinheit ein **Hochrechnungsgewicht** zugewiesen, weshalb man auch von einer **Gewichtung** der Daten spricht („weighting“). Der Gewichtungsprozess kann mehrere Stufen beinhalten, die 3 häufigsten Stufen sind:

1. Designgewichtung
2. Durch Antwortausfall bedingte Gewichtung („Non Response Weighting“)
3. Gebundene Hochrechnung / Kalibrierung

Designgewichte sind durch die Stichprobenplanung vorgegeben und entsprechen für jede Stichprobeneinheit der reziproken Auswahlwahrscheinlichkeit. Sie können bereits nach der Stichprobenziehung berechnet werden.

Vor allem bei freiwilligen Erhebungen kann es zu **Ausfällen durch Non-Response** kommen. Um Verzerrungen („Non-Response Bias“) entgegenzuwirken, muss der Rücklauf analysiert werden und Erkenntnisse aus der Feldphase bei der Konstruktion der **Non-Response-Faktoren** Eingang finden.

Abschließend kann man ausgehend von den zuvor ermittelten Gewichten (Designgewichte \times Non Response Faktoren = **Basisgewichte**) einen weiteren Schritt durchführen, um die Hochrechnungsergebnisse an **externe Quellen** anzupassen („**gebundene Hochrechnung**“). Dies kann mittels verschiedener Methoden wie z. B. Differenzschätzung, Regressionschätzung oder Kalibrierung geschehen.

Bei jeder Zufallsstichprobe sind die durch Schätzer ermittelten Ergebnisse mit einer Ungenauigkeit, nämlich dem **Stichprobenfehler**, behaftet. Für jede Stichprobenerhebung von Statistik Austria werden für möglichst viele der wichtigsten Variablen und Auswertungsdomains Stichprobenfehler errechnet und publiziert (**Fehlerrechnung**).

Zur laufenden Verbesserung der Stichprobenerhebung empfiehlt es sich auch mittels geeigneter Analysen (z.B. Nonresponse-Analysen, Simulation verschiedener Gewichtungsverfahren) zu evaluieren, ob es systematische Verzerrungen ("Bias") in den auf der Stichprobe basierenden Schätzern gibt. Sollte dies der Fall sein muss die Gewichtung dahingehend optimiert werden, um diesen Verzerrungen entgegenzutreten.

Nicht-Zufallsstichproben

Nicht-Zufallsstichproben lassen zwar keinen unmittelbaren Rückschluss auf die gesamte Grundgesamtheit zu, bilden jedoch häufig gemeinsam mit verfügbaren sekundärstatistischen Datenquellen (Register- und Verwaltungsdaten wie auch Ergebnisse anderer bereits vorliegender Statistiken) in Kombination mit einer Modellrechnung die maßgebliche Grundlage für die Ergebnisdarstellung für die statistische Grundgesamtheit. Oftmals kommen dabei insbesondere im Bereich der Wirtschaftsstatistiken sogenannte **Konzentrationsstichproben** zur Anwendung, wobei nur Einheiten über einem bestimmten Schwellenwert unter Berücksichtigung eines in der Regel standardisierten Repräsentanzkriteriums zur Sicherung eines qualitativ als ausreichend angesehenen Deckungsgrads ausgewählt werden. Man spricht in diesem Zusammenhang auch von einer „**Vollerhebung mit Abschneidegrenzen**“ („**Cut-off-Census**“). Dem Vorteil der Entlastung „kleinerer“ Respondenten (zumeist wirtschaftsstatistische Einheiten) steht zwar der grundsätzliche Nachteil gegenüber, dass ein gewisser Teil der Population nicht erfasst wird bzw. keine konventionelle Hochrechnung und damit auch keine Fehlerabschätzung möglich ist, allerdings kommt es im Allgemeinen zu keiner Verschlechterung der Ergebnisqualität.

Wie sich gezeigt hat, ist die Kombination einer primärstatistischen Nicht-Zufallsstichprobe und eines **modellbasierten Ansatzes zur Datenergänzung (MDE)**, der einerseits den Informationsvorrat der Primärerhebung voll ausschöpft und andererseits zusätzlich Informationen aus vorhandenen Verwaltungsquellen effizient nutzt, eine Alternative zur konventionellen Hochrechnung. Diese Kombination ermöglicht vor allem auf regionalen und detaillierten Darstellungsebenen eine wesentlich genauere Ergebnisdarstellung der Grundgesamtheit als die konventionelle Hochrechnung selbst. Die Struktur der Grundgesamtheit kann durch die MDE als Kombination aus Primär-, Sekundärdaten und Modellrechnung ohne grundsätzlichen Informationsverlust dargestellt werden. Da die Datenerstellung auf Mikroebene erfolgt, steht für jede statistische Einheit der Grundgesamtheit somit ein vollständiger Datensatz für die Ergebnisdarstellung zur Verfügung.

Richtlinien:

- Wenn immer möglich, soll einer Stichprobenerhebung gegenüber einer Vollerhebung der Vorzug gegeben werden. Wenn immer die Projektzielsetzung dies zulässt, soll Zufallsstichproben gegenüber Nicht-Zufallsstichproben der Vorzug gegeben werden.
- Ergebnisse aus Stichprobenerhebungen sind immer als Schätzungen zu verstehen, die mit Ungenauigkeiten behaftet sind. Die Schätzmethoden und Verfahren der Stichprobentheorie sind nur im Falle von Zufallsstichproben anwendbar.
- Um Ergebnisse aus Stichprobenerhebungen sinnvoll schätzen und vorgegebene Genauigkeitsvorgaben erfüllen zu können, ist eine hinreichende Datenmenge erforderlich. Aus diesem Grund muss bereits bei der Planung der Stichprobe der erwartete Rücklauf berücksichtigt werden.
- Schätzungen aus Stichprobenerhebungen basieren auf einer Hochrechnung (Gewichtung). Dabei muss zumindest das Stichprobendesign und der Antwortausfall berücksichtigt werden (freie Hochrechnung). Es sollen die Ergebnisse in einem finalen Schritt an eine externe Quelle angepasst werden (gebundene Hochrechnung).
- Im Falle einer gebundenen Hochrechnung muss bei der Auswahl der externen Quellen auf die Verlässlichkeit der daraus stammenden Werte geachtet werden. Nach einer finalen Anpassung müssen die Resultate kontrolliert werden. Jedenfalls ist darauf zu achten und durch geeignete Methoden (z.B. Trimming, Unter-/Obergrenzen für Anpassungsgewichte) sicherzustellen, dass die Variabilität der Gewichte möglichst gering gehalten wird.
- Für jede Stichprobenerhebung müssen zumindest für Schlüsselvariablen Stichprobenfehler ausgewiesen werden. Die Methode, mit denen die Fehlerrechnung durchgeführt wurde, muss dokumentiert werden und das Stichprobendesign und die Gewichtung berücksichtigen.
- Um Verständlichkeit und Klarheit zu gewährleisten, muss verdeutlicht werden, in welcher Form der Stichprobenfehler dargestellt wird (relativer Standardfehler, absoluter Standardfehler, Konfidenzintervall etc.). Bei der Darstellung ist jedenfalls auch das entsprechende Signifikanzniveau (in der Regel 95%) anzugeben. Bei periodisch wiederkehrenden Erhebungen soll die Darstellung über die Zeit konstant gehalten werden.
- Die Nicht-Zufallsstichprobe (Cut-off-Census) stellt in Verfolgung des Grundsatzes der Entlastung von Klein- und Kleinstrespondenten in Kombination mit den Konzepten der modellbasierten Datenergänzung (MDE) die maßgebliche Alternative zur konventionellen Hochrechnung für Zwecke der Ergebnisdarstellung über Grundgesamtheiten dar. Die Anwendung eines Cut-off-Census setzt jedoch die Definition neuer Qualitätskriterien (wie z.B. eines Repräsentanz-/Deckungsgrades) voraus, da eine konventionelle Fehlerrechnung nicht möglich ist.
- Für sämtliche Aspekte hinsichtlich Stichprobendesigns und -planung, Hoch- und Fehlerrechnung bei Stichprobenerhebungen muss der Bereich Methodik eingebunden werden.

- Es muss gegenüber den Benutzerinnen und Benutzern vermittelt werden, bis zu welcher Gliederungstiefe der Stichprobenfehler einen vertretbaren Wert nicht übersteigt. Dies gilt vor allem für regionale Gliederungen. Sollten Werte publiziert werden, die einen ein gewisses Maß übersteigenden Stichprobenfehler aufweisen, müssen diese entsprechend gekennzeichnet werden. (Als mögliche Einteilung der Zuverlässigkeit von Stichprobenergebnissen kann folgende Einteilung für den relativen Stichprobenfehler verwendet werden: $CV \leq 16.5\%$; $16.5\% < CV \leq 33.3\%$; $CV > 33.3\%$).
- Der Rücklauf einer Stichprobenerhebung muss dokumentiert werden. Folgende Größen sind dabei vor allem von Interesse: die ursprüngliche Brutto-Stichprobengröße, der erzielte Netto-Stichprobenumfang und die Antwortraten für wichtige Merkmale.
- Werden bei Nichtzufallsstichproben („Konzentrationsstichproben“) nicht erhobene Populationsteile zugeschätzt, muss die Methode, nach der dies geschieht, dokumentiert sein. Jedenfalls müssen Datennutzerinnen und -nutzer genau darüber informiert werden, auf welche Grundgesamtheit sich die Ergebnisse einer derartigen Erhebung beziehen.
- Sämtliche Stichprobenaspekte, insbesondere der Stichprobenplan, die Methode der Hoch- und Fehlerrechnung sowie alle relevanten quantitativen Maße müssen in der Standard-Dokumentation beschrieben werden.

3.5.3 Berechnung von Indikatoren – statistische Maßzahlen

Statistische Maßzahlen

Die statistische Auswertung von Datenbeständen erfolgt im einfachsten Fall mittels gebräuchlicher Funktionen, die auf die Daten angewandt werden. Man spricht auch von einfachen **statistischen Maßzahlen**. Dabei muss unterschieden werden, ob es sich bei der Auswertung um quantitative (**intervallskalierte** und **rational skalierte**) oder qualitative (**nominale** und **ordinale**) Merkmale handelt. Während für quantitative Größen die Berechnung von Total- und (arithmetischen) Mittelwerten im Vordergrund steht, sind bei qualitativen Variablen absolute und relative Häufigkeiten (Anteilswerte) von Interesse. Statistische Maßzahlen verfolgen den Zweck, die Verteilung eines Merkmales innerhalb einer **statistischen Masse** zu charakterisieren. Nachdem sich der Mittelwert bei bestimmten Datensituationen als ein ungeeignetes Lagemaß erweisen kann, wird für manche Anwendungen auf den **Median** (mittlerer Wert) zurückgegriffen.

Werden statistische Maßzahlen für eine tiefere Gliederungsebene (z.B. Bundesländer, Altersgruppen) angegeben, so ist besonders auf die Geheimhaltung und Genauigkeit zu achten.

Die Auswertung von Datenmaterial geht naturgemäß über die Betrachtung eines Merkmales hinaus. Oftmals ist es notwendig, zwei Merkmale miteinander in Beziehung zu setzen. Dabei werden häufig Quotienten aus Merkmalssummen bzw. im qualitativen Fall aus Häufigkeiten gebildet. Dabei ist es sowohl möglich, zwei gleichartige bzw. gleichrangige Massen miteinander zu vergleichen (z.B. aktueller Kfz-Bestand im Verhältnis zum Bestand am 01.01.2010, Umsatz im laufenden Monat im Verhältnis zum Umsatz im Vormonat), als auch wesensmäßig verschiedene Massen miteinander in Beziehung zu setzen (z.B. Lebendgeborene im Verhältnis zur Bevölkerung, Bevölkerungsdichte, Ertrag pro Hektar).

Indikatoren

Oftmals werden sozio-ökonomische Fragestellungen mithilfe von Indikatoren gemessen. Dabei ist der Indikator die Operationalisierung eines Konzepts, durch welches auch theoretische Begriffe messbar gemacht werden, welche selbst nicht direkt ermittelbar sind. Durch geeignete Operationalisierungen ermöglichen es Indikatoren komplexe sozio-ökonomische Entwicklungen über die Zeit und über Länder hinweg zu beobachten. Indikatoren spielen gesellschaftspolitisch eine hochrelevante Rolle. Statistische Maßzahlen können dann als sozio-ökonomische Indikatoren aufgefasst werden, wenn sie „eine quantitative Abbildung gesellschaftlich bzw. gesellschaftspolitisch relevanter Sachverhalte darstellen“ (vgl. Endrueit G., Trommsdorff G. 2002).

Im ökonomischen Kontext ist ein volkswirtschaftlicher Indikator (auch Konjunkturindikator oder makroökonomische Kennzahl genannt) eine Messgröße, die Aussagen über die konjunkturelle Entwicklung oder die wirtschaftliche Situation im Allgemeinen von Volkswirtschaften erlaubt. Wichtige Indikatoren der Bevölkerungsstatistik sind beispielsweise die Fertilitätsrate oder die Lebenserwartung. Soziale Indikatoren sollen v.a. sozialpolitisch relevante Konzepte messen, welche durch ökonomische und demografische Indikatoren alleine nicht fassbar sind.

In manchen Bereichen der Amtlichen Statistik werden unterschiedliche Maßzahlen gebündelt und als (zusammengesetzte) Indikatoren bezeichnet. Beispiele dafür sind:

- Strukturindikatoren
- Konjunkturindikatoren

- Indikatoren zu Armut und sozialer Eingliederung

Die Bereitstellung solcher Gruppen von Indikatoren, deren Berechnungsalgorithmus oft komplex ist (z.B. Europa 2020 Indikatoren), und die auch zu internationalen Vergleichen herangezogen werden, ist in den letzten Jahren zu einem wichtigen Aufgabenfeld der Amtlichen Statistik geworden.

Richtlinien:

- Die Anwendung einer statistischen Maßzahl muss der Art des Merkmals angemessen sein. Mittelwerte und Totalwerte können nur bei quantitativen Merkmalen angewandt werden. Für qualitative Merkmale sind Häufigkeiten und Anteilswerte die in der Amtlichen Statistik gängigsten Maßzahlen.
- Die Gliederungstiefe, mit der statistische Maßzahlen ausgewiesen werden, soll so gewählt werden, dass eine ausreichende Genauigkeit gegeben ist, d.h. beispielsweise dass der Stichprobenfehler ausreichend klein bleibt. Des Weiteren ist darauf zu achten, dass eine Identifizierbarkeit von Einzeldatensätzen nicht möglich ist (siehe Kapitel 3.6.6).
- Werden Größen zweier Massen miteinander in Beziehung gesetzt, muss darauf geachtet werden, dass die Definitionen und Konzepte der betroffenen Merkmale stimmig sind.
- Sollten sich Berechnungsmethoden eines periodisch publizierten Indikators ändern, so müssen Nutzerinnen und Nutzer darüber in Kenntnis gesetzt werden. Wenn immer möglich, sollen entsprechende Rückrechnungen durchgeführt werden.
- Bei der Berechnung komplexer Indikatoren soll der verwendete Algorithmus genau dokumentiert werden. Sollte es internationale Vorgaben zur Berechnung eines Indikators geben, sind diese einzuhalten. Es soll darauf geachtet werden, dass die Berechnungsmethode über die Zeit stabil bleibt.
- Bei Indikatoren aus Stichprobenerhebungen ist immer auch ein Stichprobenfehler anzugeben. Doch auch für Indikatoren, die nicht aus einer Stichprobenerhebung resultieren bzw. deren Berechnungsmethoden eine einfache Ermittlung des Stichprobenfehlers nicht möglich machen, sollen Genauigkeitsabschätzungen für die errechneten Werte angegeben werden.
- Werden Gruppen von Indikatoren periodisch berechnet und publiziert, sollen sie auf der Webseite von Statistik Austria zugänglich sein. Die Hintergrundinformation über den intendierten Sinn und Zweck der Indikatoren, die dahinterstehenden Konzepte und die Methoden der Berechnung soll Nutzerinnen und Nutzern zur Verfügung gestellt werden.
- Sämtliche statistischen Maßzahlen und Indikatoren eines statistischen Produktes müssen in den entsprechenden Standard-Dokumentationen beschrieben werden.

3.5.4 Indizes

Wie bereits im Vorkapitel erwähnt, stellt die möglichst frühzeitige Bereitstellung von Indikatoren eine der Hauptaufgaben der Amtlichen Statistik dar. Vor allem im wirtschaftsstatistischen Bereich dienen **Indizes**, die meist in sehr kurzen Zeitabständen (monatlich, quartalsweise) erstellt und veröffentlicht werden müssen, als wichtige Indikatoren zur Beobachtung der konjunkturellen Entwicklung. Ein Index ist dabei ein statistischer Messwert, der die wirtschaftlichen Veränderungen (von einer Periode zur anderen) anzeigt. Je nachdem ob dabei die Veränderung von Mengen (z.B. das produzierte, umgesetzte oder exportierte Gütervolumen) oder von Preisen gemessen werden soll, unterscheidet man zwischen **Mengen- bzw. Volumenindizes** und **Preisindizes**.

Ein Index stellt eine Beziehung zwischen gleichartigen Größen zu unterschiedlichen Zeitpunkten her. Die interessierende Größe wird für eine Berichtsperiode erhoben und zu den entsprechenden Größen in einer **Basisperiode** in Bezug gesetzt. Im Allgemeinen interessiert dabei meist die Entwicklung mehrerer, sachlich zusammengehörender Größen (z.B. Warenkorb beim Verbraucherpreisindex). Die Aggregation der einzelnen Größen erfolgt dann zumeist nicht rein summarisch, sondern basiert auf einem **Gewichtungsschema**. Gewichtungen können methodisch auf zwei Arten vorgenommen werden. Zum einen kann man die Gewichtung für die Basisperiode erfassen und für die einzelnen Berichtsperioden konstant lassen (**Laspeyres-Index**), zum anderen verfolgt das Konzept des **Paasche-Index** die Strategie, die Gewichtung für jede Berichtsperiode gesondert vorzunehmen und somit den Wert des Warenkorbes aus aktuellen Verbrauchszahlen zu ermitteln.

Das Laspeyres-Konzept hat den Nachteil, dass durch die größer werdende zeitliche Distanz zur Basisperiode sich die Gewichtungsannahmen von der Realität entfernen. Deshalb ist es notwendig, die Gewichtungen mit einer gewissen Periodizität zu überprüfen und zu adaptieren. Adaptionen setzen dabei auf eine neue Basisperiode auf, sodass ein zeitlicher Vergleich zu vor diesem Zeitpunkt gelegenen Perioden problematisch ist. Es ist daher erforderlich, den neuen Index am Niveau des alten beginnen zu lassen. Durch diesen **Verkettung** genannten Vorgang wird die zuvor vorhandene Sprungstelle aus der Zeitreihe entfernt und die beiden Indexreihen stetig miteinander verbunden.

Wenn man den Warenkorb in kurzen Abständen (etwa jährlich) anpasst und jeden neuen Index mit dem Auslaufenden verkettet, spricht man von einem **Kettenindex**. In diesem Fall knüpft man allerdings in der Regel als Basismonat an den Monat Dezember und nicht an den Jahresdurchschnitt des Vorjahrs an.

Einer der wichtigsten volkswirtschaftlichen Indizes ist der **Preisindex**. Preisindizes gelten als adäquates Instrument der Inflationsmessung, die die Preisrelationen der Berichtsperiode zur Vor- bzw. Basisperiode darstellen. Im Falle des räumlichen Vergleiches stellen sie die Relation der Preisniveaus einer Region zu einer Vergleichsregion dar. Eine spezielle Rolle spielen die Preisindizes als Deflatoren, wenn monetäre Größen (wie etwa die Komponenten des Bruttoinlandsproduktes) zu konstanten Preisen (real) dargestellt werden sollen. Darüber hinaus werden sie für makroökonomische Analysen verwendet und bilden die Grundlage für viele wirtschafts- und geldpolitische Entscheidungen. In der längerfristigen Betrachtung (Jahresabstand, längere Zeitreihen) geben sie Auskunft über die Stabilität des Geldwertes, weshalb sie auch zur Wertsicherung in Verträgen herangezogen werden (Wertsicherungsklauseln).

Bei der Erhebung von Preisen ist eine Reihe von Problemen zu berücksichtigen. So kann es bei im Warenkorb befindlichen Produkten zu einem **Qualitätswechsel** kommen. Nur wenige Waren werden über längere Perioden in der gleichen Ausführung angeboten. In vielen Fällen

ändert sich eine Reihe von Parametern, die preisbestimmend sind. Um die reine Preisveränderung zu erfassen, muss die Qualitätsänderung aus der Preisänderung heraus gerechnet werden (sogenannte **Qualitätsadaptierung**). Ist die Qualitätsänderung quantifizierbar, liegen die Dinge relativ einfach: Man korrigiert den Preis entsprechend, als hätte die Qualitätsänderung nicht stattgefunden. In den meisten Fällen ist eine Qualitätsadaptierung aber ein schwieriges Unterfangen, weil z. B. oft nicht einmal klar ist, ob eine Verbesserung oder eine Verschlechterung für den Konsumnutzen vorliegt. Auf europäischer Ebene ist man daher dazu übergegangen, auf der Basis von Forschungsprojekten, verbindliche Methoden nach Produktgruppen festzulegen.

Ein weiteres Problem liegt in der Tatsache begründet, dass die tatsächlich bezahlten Preise (die sogenannten Transaktionspreise) nicht immer eruiert sind. Neben den Listenpreisen müssen daher auch Informationen über Rabatte, Skonti und Ermäßigungen (z.B. Seniorenkarten) erhoben werden. Ein Großteil der Aktionspreise geht üblicherweise in den Index ein, nicht aber ungesetzliche (z. B. Ablösen für Wohnungen) und unechte Preise (z. B. Schulbücher). Alle diese Dinge spielen keine Rolle, wenn die Konditionen im Zeitablauf die gleichen bleiben, werden aber problematisch, wenn sie sich ändern. Mitunter gibt es überhaupt keinen erhebenden (Netto)Preis, weil z.B. die Konditionen je nach Kunden verschieden sind oder Produkte auf Grund saisonaler Umstände gar nicht verkauft werden.

Richtlinien:

- Das Konzept und die Methoden eines Index müssen transparent sein. Die Grundgesamtheit der zu messenden Ausgaben bzw. Mengen muss definiert sein. Etwaige Abdeckungsprobleme müssen festgehalten werden.
- Abweichungen von anerkannten Methoden und Definitionen/zu Standardmethoden – sowohl national als auch international – müssen begründet und transparent festgehalten werden.
- Bei Methodenänderungen sollte geprüft werden, ob es internationale Vorgaben gibt, die einzuhalten sind und ob die geplanten Änderungen mit diesen übereinstimmen.
- Die Basisperiode des Warenkorbs sowie der Gewichtung muss transparent sein. Weiters muss die Basis der Indexwerte der Aggregate explizit festgehalten werden. Dabei kann es sich um unterschiedliche Perioden handeln.
- Über Änderung der Basisperioden soll rechtzeitig und so vollständig wie möglich informiert werden.
- Das gewählte Berechnungskonzept (Laspeyres, Paasche oder Kettenindex) und der Aggregationsprozess müssen dokumentiert sein. Die Wahl des Konzeptes hängt von der Zielsetzung des Index, sowie der zeitgerechten Verfügbarkeit der Daten und den europäischen Vorgaben ab.
- Die Quellen der Gewichtung müssen, soweit möglich, transparent sein (unter Einhaltung von Geheimhaltungsrichtlinien). Diese sollen möglichst verlässlich sein.
- Die Gewichtung und der Warenkorb sollen regelmäßig auf Aktualität überprüft werden und allfällige Änderungen möglichst bald implementiert werden.
- Bei Preisindizes ist dem Problem der Qualitätsänderungen einzelner Waren/Produkte und Dienstleistungen Rechnung zu tragen.
- Bei häufigen Änderungen der Methoden ist die Dokumentation in regelmäßigen Intervallen zu aktualisieren.
- Indizes gelten als statistische Produkte. Daher muss zu jedem relevanten Index, der von Statistik Austria veröffentlicht wird, eine Standard-Dokumentation erstellt werden. Ist der betrachtete Index Teil einer Gruppe von Indikatoren, die im Zuge eines anderen Projekts erstellt werden, soll er in der dortigen Dokumentation beschrieben werden

3.5.5 Gesamtrechnungen

Gesamtrechnungen verknüpfen und ergänzen statistische Basisdaten in systematischer Weise. Sie dienen vor allem Analysezwecken, werden aber zunehmend auch für administrative Vorgänge (z.B. Festlegung und Überwachung der Wirtschaftspolitik) verwendet.

Gesamtrechnungen lassen sich im Wesentlichen in zwei Gruppen einteilen:

- a) Volkswirtschaftliche Gesamtrechnungen (VGR)
- b) Satellitenkontensysteme.

Das umfassendste Gesamtrechnungssystem ist jenes der **Volkswirtschaftlichen Gesamtrechnungen** (VGR). Die **VGR** haben die statistische Darstellung der Gesamtwirtschaft in einem Kontenschema zum Gegenstand, basierend auf der Vorstellung eines Wirtschaftskreislaufs. Ihre zentrale Rechtsgrundlage bildet das Europäische System volkswirtschaftlicher Gesamtrechnungen (ESVG 2010). Im Unterschied zum international vereinheitlichten System of National Accounts (SNA 2008) der Vereinten Nationen hat das ESVG für Österreich als EU-Verordnung rechtsverbindlichen Charakter. Damit wird ein kohärentes, konsistentes und integriertes System, das auf europaweit akkordierten Konzepten, Definitionen und Klassifikationen beruht, gewährleistet.

Das ESVG 2010 enthält zwei Hauptdarstellungsformen: Zum einen **Sektorkonten**, welche für die einzelnen institutionellen Sektoren (z.B. Kapitalgesellschaften, Staat, Private Haushalte) eine Beschreibung der verschiedenen Phasen des Wirtschaftskreislaufs liefern, d.h. der Produktion, der Einkommensentstehung, -verteilung, -umverteilung und -verwendung von nichtfinanziellem und finanziellem Vermögen. Zum anderen die Tabellen nach **Wirtschaftsbereichen** und das **Input-Output-System**, welche tiefer gegliederte Beschreibungen des Produktionsprozesses liefern, wobei die Darstellung von Produktionsstruktur, Einkommen und Erwerbstätigkeit sowie die Abbildung von Güterströmen erfolgen.

Für die beiden Hauptdarstellungsformen gibt es jeweils weiterführende räumliche und zeitliche Differenzierungsmöglichkeiten. So werden die Tabellen nach Wirtschaftsbereichen sowohl für das gesamte österreichische Wirtschaftsgebiet als auch für regionale Untergliederungen ermittelt. Teilbereiche der Sektorkonten und der Tabellen nach Wirtschaftsbereichen werden auch quartalsweise publiziert.

Darüber hinaus werden für einige Variablen (z.B. Bruttowertschöpfung, Konsumausgaben, Investitionen, Exporte) Preis- und Volumenänderungen erfasst. Der Umfang und Detailgrad der Differenzierungen ist ebenso wie die Liste der preis- und volumenmäßig zu erfassenden Variablen durch das verbindliche Datenlieferprogramm des ESVG 2010 genau festgelegt.

Mit regional untergliederten Tabellen nach Wirtschaftsbereichen befassen sich die **Regionalen Gesamtrechnungen**. Zur Erstellung räumlich differenzierter Tabellen bestehen methodisch prinzipiell zwei Möglichkeiten. Bei der Bottom-up-Methode werden Informationen über Einheiten einer Region regionalhierarchisch „von unten nach oben“ addiert, bis der regionale Wert der Variable festgestellt ist. Im Gegensatz dazu ordnet die Top-Down-Methode nationale Variablen mittels geeigneter verfügbarer Indikatoren den einzelnen Regionen zu. Sie kommt nur dann zum Einsatz, wenn mangelnde Datenverfügbarkeit bzw. -qualität einen Bottom-up-Ansatz nicht erlauben.

Für Gebiete, deren Informationsbedarf im Rahmen der VGR nur unzureichend abgedeckt ist, werden sogenannte **Satellitenkonten** erstellt. Auch diese Satellitenkonten beruhen auf europa- und weltweit akkordierten Konzepten, Definitionen und Klassifikationen. Sie sind mit den zentralen Konten der VGR konsistent verknüpft und heben bestimmte Bereiche des sozialen oder wirtschaftlichen Lebens heraus. Beispiele für nationale und z.T. regionale Satelliten-

tenkontensysteme, die bei Statistik Austria implementiert sind, umfassen die Umwelt-Gesamtrechnungen, das System der Gesundheitsausgaben, die Tourismus-Satellitenkonten und die Land- bzw. Forstwirtschaftliche Gesamtrechnung.

Die Erstellung von Gesamtrechnungen involviert eine Vielzahl verschiedener Datenquellen. Diese umfasst einerseits von Statistik Austria durchgeführte statistische Projekte und Erhebungen und andererseits externe Verwaltungsquellen. Darüber hinaus ist es bei manchen Problemstellungen notwendig, Einzelinformationen (z. B. aus Geschäftsberichten) heranzuziehen. Oftmals ist es nicht möglich, eine Transaktion oder Kontenposition mittels einer Datenquelle befriedigend abzudecken und es muss auf eine Kombination verschiedener Datenquellen – gelegentlich von unterschiedlicher Typologie – zurückgegriffen werden. Die zur Verfügung stehenden Quellen sind oft unterschiedlich hinsichtlich der Begrifflichkeiten und Definitionen sowie des zugrunde liegenden Einheitenkonzepts. Sie werden entweder so transformiert, dass eine unmittelbare Verwendung für die Gesamtrechnungen erfolgen kann, oder dienen als Input für Modelle, mit denen bestimmte Teilaggregate geschätzt werden.

Prinzipiell werden bei der Erstellung von Gesamtrechnungen exakte Methoden gegenüber Methoden, die Schätzungen beinhalten, bevorzugt. Die Wahl der Verfahren und ihre Genauigkeit hängen von der Qualität und der Vergleichbarkeit der Daten ab.

Bei Gesamtrechnungen werden neben üblichen laufenden Revisionen, die sich auf kleinere Korrekturen einzelner Jahre/Quartale beziehen, auch umfassendere, sogenannte „große“ Revisionen durchgeführt. Letztere bedeuten die grundlegende Überarbeitung der gesamten Rechnung bzw. sehr langer Zeitreihen. Gründe für umfassende Revisionen können sein:

- Einführung neuer Konzepte, Definitionen, Klassifikationen
- Einbau neuer, bislang nicht verwendeter statistischer Berechnungsgrundlagen
- Anwendung neuer Berechnungsmethoden
- Modernisierung der Darstellung und Einführung neuer Begriffe
- Erhöhung der internationalen Vergleichbarkeit

Richtlinien:

- Gesamtrechnungen sowie Satellitensysteme sollen immer gemäß einem klar definierten Konzept erstellt werden. Existieren international verbindliche Richtlinien, müssen die dort definierten Konzepte, Definitionen und methodischen Vorgaben eingehalten werden.
- Die Publikation und Meldung von Ergebnissen aus Gesamtrechnungen sind sehr oft an stringente Termine gebunden. Der besonderen Relevanz – vor allem der Indikatoren der Volkswirtschaftlichen Gesamtrechnungen - Rechnung tragend, muss der Rechtzeitigkeit der Fertigstellung eine hohe Priorität eingeräumt werden.
- Die Wahl der Methoden zur Ermittlung von Aggregaten zur Gesamtrechnung soll immer gemäß „Best Practices“ sein. Daher sollte in internationalen Kontakten der Wissensaustausch laufend gepflegt werden.
- Die zur Erstellung von Gesamtrechnungen verwendeten Verfahren sollen transparent und nachvollziehbar beschrieben und der Öffentlichkeit zugänglich gemacht werden.
- Da die Verfügbarkeit von Inputdaten für die Qualität von Gesamtrechnungen von entscheidender Bedeutung ist, soll laufend beobachtet werden, ob neue Datenquellen für eine spezifische Problemstellung zur Verfügung stehen. Ebenso soll proaktiv auf interne und externe Datenhalter eingewirkt werden, um die Nutzbarkeit bestehender Datenquellen zu erhöhen.
- Prinzipiell soll bei der Erstellung von Gesamtrechnungen von einer möglichst tiefen Gliederung ausgegangen und zu Gesamtwerten aggregiert werden. Diese Methode („bottom up“) ist gegenüber dem Aufteilen von Totalwerten auf einzelne Gliederungspositionen („top down“) zu bevorzugen. Die Auswahl der anzuwendenden Verfahren ist entsprechend der Qualität und der Verfügbarkeit der Daten durchzuführen.
- Manchmal ist es nötig, gewisse Teilaggregate einer Gesamtrechnung modellhaft zu schätzen. In diesem Fall gelten sämtliche Richtlinien und Qualitätsstandards zur modellbasierten Schätzung (siehe Kapitel 3.5.6).
- Ist der Bedarf einer großen Revision gegeben, sollen Rückrechnungen unter der Voraussetzung der Aufrechterhaltung der Datenqualität in dem Maße erfolgen, dass es Datenutzerinnen und -nutzern möglich ist, auf hinreichend lange Zeitreihen zurückzugreifen.
- Gesamtrechnungen gelten als statistische Produkte. Daher muss zu jeder von Statistik Austria veröffentlichten Gesamtrechnung eine Standard-Dokumentation erstellt werden. Ein Schwerpunkt der Dokumentation soll dabei bei der Beurteilung und Beschreibung der verwendeten Datenquellen und der genauen Verfahrensbeschreibung zur Errechnung der einzelnen Aggregate liegen. Gegebenenfalls soll in speziellen Methodenhandbüchern sehr detailliert auf komplexe Berechnungsmethoden eingegangen werden.

3.5.6 Modellbasierte Schätzungen und Prognosen

Die Publikation von Ergebnissen in der Amtlichen Statistik basiert zumeist auf der direkten Auswertung erhobener (oder auf andere Art beschaffter) und im Stichprobenfall hochgerechneter Daten.

Man spricht von einer **modellhaften (modellbasierten) Schätzung**, wenn für Teilpopulationen keine Einzeldatensätze zur direkten Auswertung vorhanden sind oder die Problemstellung eine Ergebnisschätzung aus den vorhandenen Einzeldaten durch Anwendung komplexer Methoden notwendig macht. Insbesondere im Bereich der Wirtschaftsstatistik werden modellbasierte Schätzungen mehr und mehr zum signifikanten Bestandteil der bestehenden Konzepte zur Ergebnisdarstellung über Grundgesamtheiten (z.B. MDE – modellbasierter Ansatz zur Datenergänzung). Auch die Anwendung von Hochrechnungsmethoden bei Stichprobenerhebungen lässt sich als modellhaftes Vorgehen interpretieren (vgl. auch Kapitel 3.5.2). Typisch für eine modellbasierte Anwendung ist, dass Daten in ein Modell eingehen („Input“) und diese dann mittels einer klar definierten Methode und beruhend auf bestimmten Einstellungen („Parameter“) ein Resultat liefern („Output“). Modelle können sehr einfach sein und sich auf die Durchführung einiger weniger Rechenschritte beziehen. Mit steigender Komplexität von Problemstellungen kann es aber auch sein, dass Modelle sich nicht trivialer mathematisch-statistischer Methoden bedienen.

In diesem Dokument wird, neben den bereits oben erwähnten Anwendungen, in einigen Kapiteln auf weitere spezielle und für die Amtliche Statistik wichtige Modelle eingegangen. Darunter sind die saisonale Bereinigung von Zeitreihen (Kapitel 3.5.7) und Flash Estimates (Kapitel 3.5.8) zu verstehen. In diesem Kapitel wird die generelle Anwendung von Modellen betrachtet. In einigen Fällen dienen Modelle auch dazu, aus bestehenden Aggregaten Werte zu schätzen, die in der Zukunft liegen. Solche **Prognosen** (z. B. für die Entwicklung der Bevölkerung) bilden eine wichtige Entscheidungsgrundlage für Datennutzerinnen und -nutzer.

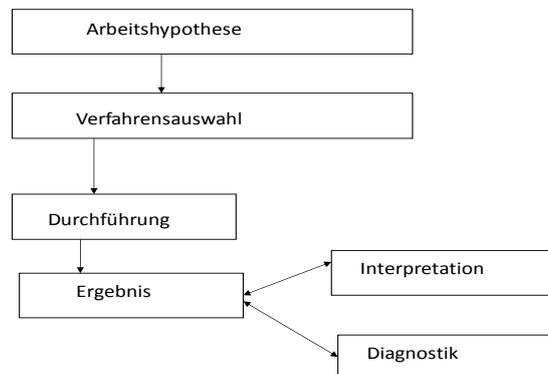


Abbildung 12 : Ablauf bei modellhafter Schätzung

Abbildung 12 zeigt schematisch die Arbeitsschritte bei modellhafter Schätzung. Ausgehend von einer Arbeitshypothese und der jeweiligen Problemstellung gilt es, ein geeignetes Modell zu finden. Häufig bedient man sich bei der Modellwahl sogenannter **multivariater Verfahren**. Dabei werden multivariate statistische Verteilungen untersucht. Man betrachtet hier nicht eine Variable isoliert (univariat verteilt), sondern das Zusammenwirken mehrerer Variablen zugleich.

Bevor man jedenfalls die Modellrechnung durchführen kann, ist es notwendig die für das Verfahren notwendigen **Methoden** und **Parameter** zu wählen. Nach der Durchführung ist die **Beurteilung und Interpretation** des Ergebnisses von großer Bedeutung. Hierfür gibt es für die meisten Verfahren geeignete Statistiken – sogenannte **Modelldiagnostiken** –, die eine

Beurteilung der Resultate ermöglichen. Bei der Wahl der richtigen Methode spielt das Skalenniveau der Zielvariable(n) eine zentrale Bedeutung und darf nicht ignoriert werden.

Die gängigen Statistikpakete beinhalten eine Vielzahl von Routinen und Prozeduren zur Berechnung und Anwendung von Modellen. Zu beachten ist dabei, dass auch bei Anwendung nicht adäquater Methoden Ergebnisse ermittelt werden. Daher ist das Verständnis der Diagnosestatistiken und der Modellvoraussetzungen eine notwendige Voraussetzung bei der Anwendung von Modellen.

Unter dem Terminus **Machine Learning** – als Oberbegriff für die künstliche Generierung von Wissen aus Erfahrung - werden Methoden zusammengefasst, die zur automatischen Klassifizierung oder Vorhersagen verwendet werden können. Diese erhalten vor allem im Zusammenhang mit Big Data eine größere Bedeutung in der amtlichen Statistik. Da solche Methoden überwiegend als heuristisch ohne analytische Herleitung einzustufen sind, müssen die Ergebnisse mit besonderer Sorgfalt betrachtet werden, da es keine klar zu fassenden Modellvoraussetzungen bzw. Modelldiagnostiken gibt.

Richtlinien:

- Bei der Anwendung von Modellen muss auf die jeweilige Problemstellung eingegangen werden. Das Problem muss genau formuliert sein und sollte mit dem Bereich Methodik hinsichtlich der optimalen Modellwahl diskutiert werden.
- Es soll prinzipiell nur dann modellhaft vorgegangen werden, wenn die Ergebnisse mittels direkter Datenauswertung nicht mit zumindest derselben Qualität erzielbar sind.
- Die Auswahl eines bestimmten Verfahrens soll nur in Absprache mit Experten geschehen, die bezüglich des angedachten Verfahrens methodische Kenntnisse und praktische Erfahrungen haben.
- Die Ergebnisse einer modellhaften Schätzung müssen genau beurteilt werden. Nur wenn sich die Werte sämtlicher Diagnoseparameter in einem vertretbaren Rahmen befinden, können die Ergebnisse akzeptiert werden.
- Neben der Beurteilung der Modellgüte anhand der Diagnostik müssen die Ergebnisse auch inhaltlich auf Korrektheit und Plausibilität geprüft werden.

3.5.7 Zeitreihen und deren Bereinigung

Unter einer **Zeitreihe** wird die periodische, sequentielle Beobachtung einer Variablen über die Zeit verstanden. Aufgrund der Tatsache, dass in der Amtlichen Statistik periodische Erhebungen eine lange Tradition haben und die Beobachtung der langfristigen Entwicklung von Indikatoren von hoher Relevanz ist, stellt die analytische Aufarbeitung des über die Zeit zur Verfügung stehenden Datenmaterials eine wichtige Aufgabe dar. Die Periodizität, mit der Zeitreihen zur Verfügung gestellt werden, ist unterschiedlich. Am häufigsten sind monatliche, quartalsweise und jährliche Daten. Zeitreihen in der Amtlichen Statistik können sehr lange zurückreichen. Oftmals unterliegen Erhebungen oder Berechnungen zur Ermittlung bestimmter Maßzahlen Veränderungen. Diese können dann zu **Brüchen in der Zeitreihe** führen. Solche Zeitreihenbrüche stören die Vergleichbarkeit über die Zeit und Datennutzerinnen und -nutzer müssen mittels adäquater Metadaten darüber informiert werden, wie sich die konzeptuellen Veränderungen auswirken bzw. bis wann vergleichbare Zeitreihen zurückreichen. Bei manchen Indikatoren (vor allem Indizes) werden Techniken angewandt, die die Erstellung einer durchgängigen Zeitreihe ermöglichen (Verkettungen, siehe auch Kapitel 3.5.4).

In der Zeitreihenanalyse geht man von der Grundvorstellung aus, dass sich eine Ursprungsreihe in mehrere Komponenten zerlegen lässt. Im Wesentlichen erfolgt die Einteilung dabei in

- **Trend- und Konjunkturkomponente:** mittel- und langfristige Entwicklungsrichtung einer Zeitreihe,
- **Saisonkomponente:** periodische Schwingungen mit einer Periode von einem Jahr, die dem Zyklus der Jahreszeiten entspricht,
- **irreguläre Komponente:** all jene Einflüsse, die nicht den anderen beiden Komponenten zugeordnet werden können, etwa außergewöhnliche wirtschaftliche oder klimatische Entwicklungen (Erhalt von Großaufträgen, Streiks, ungewöhnlich stark ausgeprägte jahreszeitliche Entwicklungen aufgrund extremer Witterungsverhältnisse)

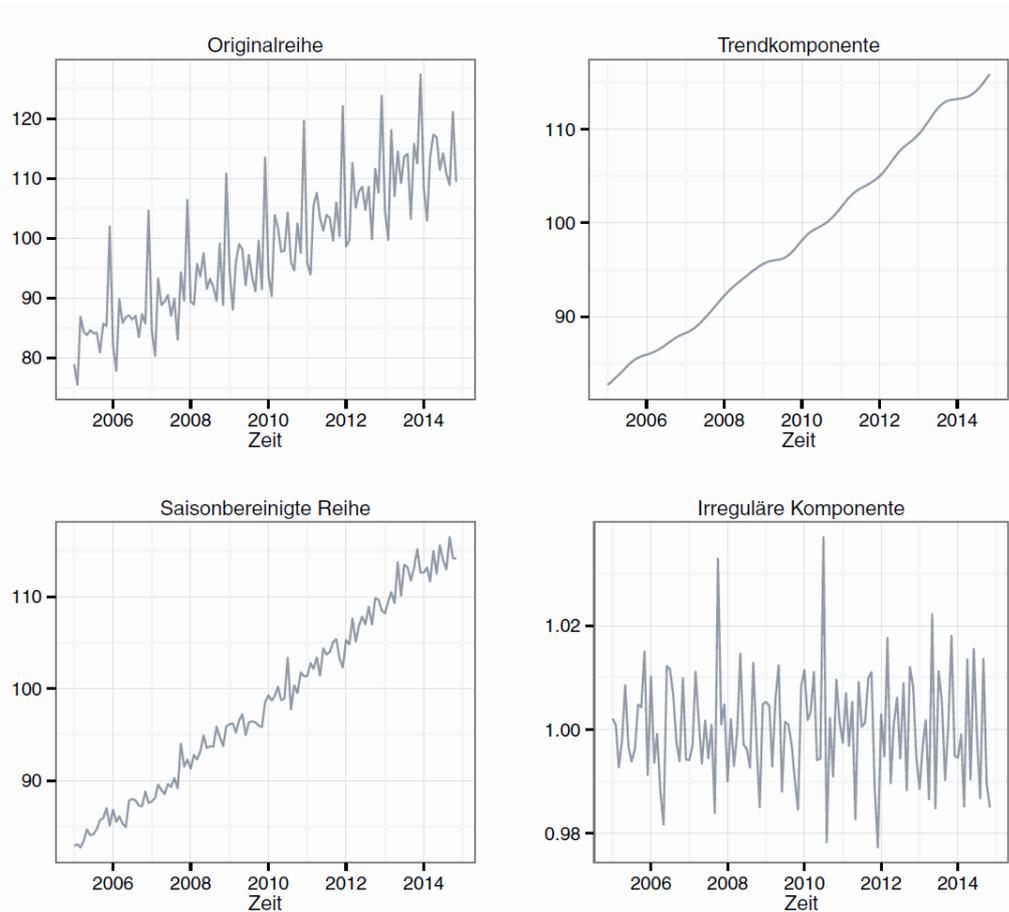


Abbildung 13: Zeitreihe mit Komponenten

Diese Komponenten sind nicht direkt beobachtbar und müssen geschätzt werden. Der Fokus liegt dabei insbesondere in der langfristigen Entwicklungsrichtung der Zeitreihe, wohingegen periodische Schwingungen, die jedes Jahr in etwa derselben Intensität auftreten, herausgefiltert werden. Die Eliminierung von saisonalen Einflüssen bezeichnet man als **Saisonbereinigung**.

Grundsätzlich unterscheidet man hinsichtlich der Art der Beobachtungen einer Zeitreihe zwischen sogenannten Stromgrößen (zeitraumbezogene Daten) und Bestandsgrößen. Erstere sind häufig von der Kalenderkonstellation des Beobachtungszeitraums abhängig, so steigt etwa die Produktion mit der Zahl der Arbeitstage. Bestandsgrößen werden mitunter von der Lage des Beobachtungsstichtages beeinflusst. Unregelmäßigkeiten des Kalenders können durch eine arbeitstägige Bereinigung ausgeglichen werden, welche einer Saisonbereinigung vorgelagert ist. Neben arbeitstägigen Einflüssen sind gegebenenfalls auch andere Kalendereffekte (z.B. Schaltjahreseffekt, Ostereffekt) zu berücksichtigen.

Für die Bereinigung selbst (arbeitstägige- und saisonale Bereinigung) gibt es eine relativ große Auswahl an Verfahren wobei als die beiden wichtigsten X12-Arima und Tramo/Seats zu nennen wären.

Bereinigungen sollen nur dann durchgeführt werden, wenn keine fachlich begründeten Einwände bestehen. Saisonale Bereinigung dient zunächst Analysezwecken. Ob die Publikation einer bereinigten Reihe sinnvoll ist, hängt von der jeweiligen Sachmaterie ab. So kann die Bereinigung von Preisindizes zum Wegglätten von saisonalen Schwankungen sowie Kalendereffekten kommen, deren Abbildung durchaus erwünscht ist.

Richtlinien:

- Bei periodisch häufig wiederkehrenden Projekten entsteht langfristig eine Zeitreihe. Um zeitliche Vergleichbarkeit möglichst gut zu gewährleisten, sollen Konzepte, Definitionen und Methoden - wenn immer möglich - über die Zeit konsistent sein.
- Lassen sich Zeitreihenbrüche nicht vermeiden, müssen Datennutzerinnen und -nutzer über die Ursachen und Auswirkungen informiert werden.
- Es soll für jede Zeitreihe, die unterjährige Ergebnisse liefert, untersucht werden, ob ein arbeitstägiger und/oder saisonaler Einfluss gegeben ist.
- Bevor die Saisonbereinigung durchgeführt wird, müssen sämtliche Kalendereffekte bereinigt werden.
- Zeitreihen, bei denen ein saisonaler Einfluss festgestellt wird, sollen wenn keine fachlichen Gründe dagegen sprechen bereinigt werden. Prinzipiell sollen sowohl die originalen als auch die bereinigten Werte publiziert werden. Datennutzerinnen und -nutzer müssen auf die Tatsache aufmerksam gemacht werden, dass die bereinigten Werte bei Veröffentlichung des nächsten Wertes revidiert werden.
- Um Saisonbereinigung durchführen zu können, muss eine ausreichende Länge der Zeitreihe (etwa 5 Jahre) gegeben sein.
- Die Ergebnisse der Zeitreihenbereinigungen müssen kontrolliert und die Modellannahmen und Parameter periodisch überprüft werden.
- Vorperiodenvergleiche bei unterjährigen Daten sollen auf saisonbereinigten Werten basieren. Vorjahresvergleiche sind mit arbeitstägig bereinigten Daten anzustellen.
- Die Bereinigung von Zeitreihen soll mit X12-Arima oder mit Tramo/Seats durchgeführt werden.
- Bei allen Aspekten der Zeitreihenbereinigung ist der Bereich Methodik einzubinden.

3.5.8 Flash Estimates

Der Bedarf an **möglichst schneller** statistischer Information, die sich auf einen bestimmten Berichtszeitraum bezieht, ist in den letzten Jahren massiv gestiegen. Unter diesem Gesichtspunkt ist die Intention zu sehen, sofort nach Ende einer Berichtsperiode eine erste Schätzung der Ergebnisse vorzunehmen, dies auch im Bewusstsein der Tatsache, dass die Datenbasis, auf die man sich stützt, nicht vollständig vorhanden ist. Eine derartige Schätzung wird **Flash Estimate** genannt. Gemäß Eurostat (siehe etwa Kapitel 16 im „Handbook on quarterly national accounts“ (Eurostat 1999)) ist ein Flash Estimate definiert als ein Schätzwert, der so früh wie möglich nach Ende der Referenzperiode produziert und veröffentlicht wird, basierend auf einer, im Vergleich zu herkömmlichen Schätzungen, reduzierten Informationsmenge. Wie aus dieser Definition zu ersehen ist, entstammt die Idee aus der Wirtschaftsstatistik, lässt sich aber im Prinzip auch auf andere Bereiche der Amtlichen Statistik übertragen. Zumeist sind Flash Estimates bei unterjährigen Statistiken mit monatlicher bzw. quartalsweiser Periodizität sinnvoll.

Flash Estimates beziehen sich immer auf eine gegenwartsnahe, jedoch bereits vergangene Berichtsperiode und grenzen sich dadurch zu Prognosen („Forecasts“) ab, deren Zielsetzung in der Beschreibung zukünftiger Perioden liegt.

Das Problem bei Flash Estimates ist, dass durch die **Unvollständigkeit der Datenbasis** – bei manchen Problemstellungen ist möglicherweise erst ein sehr geringer Anteil der Daten verfügbar – die Schätzung mit Unsicherheiten behaftet ist. Im Prinzip lassen sich Flash Estimates als erste vorläufige Ergebnisse interpretieren. Sie sind somit jedenfalls Gegenstand einer Revision, sobald die Datenbasis vollständig oder zumindest vollständiger zur Verfügung steht. Um der zuvor angesprochenen Ungenauigkeit entgegenzuwirken, bedient man sich bei der Berechnung von Flash Estimates modellhafter Schätzungen.

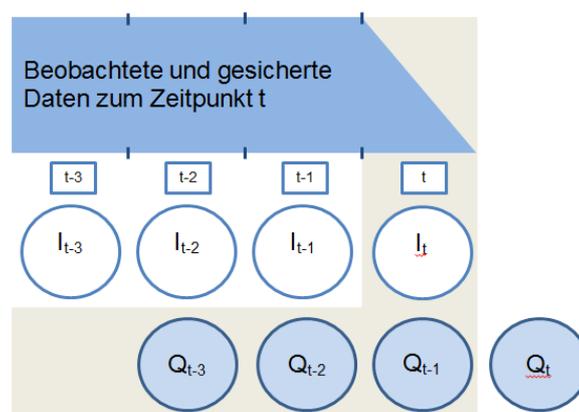


Abbildung 14: Der Flash Estimate für die aktuelle Periode t beruht auf der unvollständigen Datenbasis und den Erfahrungen aus Vorperioden

Wie *Abbildung 14* veranschaulicht, dient als Input für ein Modell zur Schnellschätzung einerseits die reduzierte Datenbasis für die aktuelle Berichtsperiode t . Andererseits gewinnt man aus Vorperioden laufend Erkenntnisse über die Abweichungen der endgültigen Ergebnisse zu den Flash Estimates, was den zweiten wesentlichen Input für die Berechnungsmodelle darstellt. In der Abbildung bedeuten die Q_j jeweils die endgültigen Datenwerte aus k Vorperioden. Die I_j stehen für die korrespondierenden Werte der Flash Estimates. In der Praxis werden oftmals bereits die endgültigen Werte zur letzten Vorperiode existieren. Jedoch kann es vor allem bei monatlichen Zeitreihen vorkommen, dass der endgültige Vorperiodenwert zum Zeitpunkt der Schnellschätzung noch nicht vorliegt. Theoretisch ist es denkbar, das Modell

noch mit anderen Informationen anzureichern (z. B. externe Daten, die mit der Materie in Zusammenhang stehen).

Aus *Abbildung 14* und dem Gesagten erkennt man, dass die Erfahrungen aus den Vorperioden von Bedeutung sind. Somit ist ein permanenter Lernprozess gegeben und das Modell wird sich verbessern lassen, je mehr Vorperioden zur Verfügung stehen. Daraus ergibt sich der Bedarf, die modellhaften Zusammenhänge regelmäßig zu überdenken.

Da Modellbildung bei Flash Estimates eine entscheidende Rolle spielt, besteht ein unmittelbarer Zusammenhang zu Kapitel 3.5.6 („Modellbasierte Schätzungen und Prognosen“). Alle Richtlinien zur Modellbildung, die dort Anwendung finden, haben auch für die Berechnung von Flash Estimates ihre Gültigkeit. Ebenso sind die Richtlinien des Kapitels 3.5.7 („Zeitreihen und deren Bereinigung“) von Relevanz, da Flash Estimates – wie bereits gesagt – bei periodisch wiederkehrenden, zumeist unterjährigen Produkten anfallen und somit Zeitreihen bilden.

Richtlinien:

- Flash Estimates sollen nur für jene Maßzahlen oder Indikatoren eines statistischen Produktes berechnet werden, deren hohe Relevanz eine unverzügliche Veröffentlichung rechtfertigt.
- In die Modellbildung sollten Werte aus den Vorperioden und Schätzungen aufgrund der reduzierten Datenbasis der aktuellen Periode als auch aus den Vorperioden eingehen. Dabei sollten möglichst viele Vorperioden berücksichtigt werden - allerdings nur insoweit dies eventuelle Zeitreihenbrüche zulassen.
- Das Modell zur Berechnung der Flash Estimates soll in periodischen Abständen überprüft und gegebenenfalls modifiziert werden.
- Bei der Veröffentlichung von Flash Estimates sollen Nutzerinnen und Nutzer über die Verfügbarkeit und die verwendete Methodik (inklusive zugrundeliegender Indikatoren) informiert werden. Es ist darauf hinzuweisen, dass die Ergebnisse vorläufig sind und jedenfalls revidiert werden.
- Es sollen laufend Qualitätsindikatoren berechnet werden, die eine Genauigkeitsabschätzung der Flash Estimates ermöglichen. Jedenfalls muss die absolute und relative mittlere Abweichung der Schätzungen von den endgültigen Werten über die Zeit beobachtet werden. Datennutzerinnen und -nutzern sollen diese Informationen zugänglich gemacht werden, damit das Risiko der Verwendung der Flash Estimates einschätzbar wird.
- Sämtliche Aspekte über Flash Estimates für ein bestimmtes statistisches Produkt, welche Indikatoren betroffen sind, die verwendeten Methoden und deren Qualität müssen in der Standard-Dokumentation beschrieben werden.

3.5.9 Sonderauswertungen

Grundsätzlich richtet sich die Publikation statistischer Ergebnisse (siehe Kapitel 3.6) nach einem bestimmten, im Vorfeld definierten Programm. Darüber hinaus äußern Datennutzerinnen und -nutzer auch Anforderungen und Fragestellungen, die Auswertungen notwendig machen, welche nicht im Standardprogramm enthalten sind.

Solche **Sonderauswertungen** können unter gewissen Voraussetzungen von Statistik Austria durchgeführt und einem Auftraggeber in der Regel entgeltlich geliefert werden. In manchen Fällen beschränkt sich die Anfrage nicht auf eine zusätzliche einfache deskriptive Auswertung eines Datenbestandes, sondern die beauftragte Arbeit hat einen analytischen Charakter, der die Anwendung spezieller Verfahren erfordert. Des Weiteren ist denkbar, dass für bestimmte Fälle die Auswertung mehrerer Datenbestände notwendig ist. Somit können Arbeiten zur Erfüllung spezieller Anforderungen auch Analysen und Verknüpfungen beinhalten. Für Methoden, die bei Sonderauswertungen angewandt werden, haben alle in diesem Dokument andernorts formulierten Richtlinien volle Gültigkeit. Jedenfalls gilt für all diese Arbeiten, dass keine gesonderten Erhebungen oder sonstige Datenbeschaffungen durchgeführt werden, sondern auf bereits existierende Authentische Datenbestände zurückgegriffen wird.

Spezielle Anfragen sind primär dahingehend zu beurteilen, ob die Durchführung einer Sonderauswertung im Einklang mit den bei Statistik Austria geltenden Standards und Vorgaben machbar ist. So gilt es Fragen, wie Einhaltung von Geheimhaltungsregeln, Verfügbarkeit von Ressourcen und Sinnhaftigkeit der herangetragenen analytischen Fragestellung im Vorfeld abzuklären.

Manche Sonderauswertungen können durchaus wiederkehrenden Charakter haben und im Laufe der Zeit standardmäßig in den Planungs- und Erstellungsprozess der betroffenen statistischen Projekte eingebunden werden, prinzipiell aber werden Sonderprojekte erst auf Anfrage geplant und bei festgestellter Machbarkeit durchgeführt.

Da Sonderauswertungen naturgemäß auch nicht Bestandteil der budgetären Gesamtplanung für das laufende Produktionsjahr sind, muss im Vorfeld die Frage der Finanzierung mit dem Auftraggeber abgeklärt werden. Dabei ist festzuhalten, dass Statistik Austria bei der Erstellung von Statistiken an gewisse (auch in diesem Dokument formulierte) Qualitätsstandards gebunden ist. Sämtliche Qualitätsrichtlinien zu eingesetzten Methoden, die an anderen Stellen in diesem Dokument beschrieben werden, gelten auch für alle Sonderauswertungen sowie gesonderte Analysen und Verknüpfungen. Eine Unterschreitung dieser Standards zugunsten einer geringeren finanziellen Vergütung steht nicht im Einklang mit den Grundsätzen von Statistik Austria.

Richtlinien:

- Sonderauswertungen können nur auf Anfrage eines potentiellen Auftraggebers durchgeführt werden. Es ist dafür Sorge zu tragen, dass Anfragen zu kompetenten und entscheidungsbefugten Mitarbeiterinnen und Mitarbeitern weitergeleitet werden.
- Sonderauswertungen sollen immer anhand bereits existierender standardmäßig zu erstellender Authentischer Datenbestände durchgeführt werden.
- Im Vorfeld der Durchführung gesonderter Arbeiten muss mit dem Auftraggeber der zeitliche, inhaltliche und finanzielle Rahmen des Projektes abgeklärt werden. Zu diesem Zweck soll dem Auftraggeber ein durch Statistik Austria erstellter Kostenvoranschlag und ein Zeitplan zur Verfügung gestellt werden.
- Durch Mitarbeiterinnen und Mitarbeiter von Statistik Austria muss festgestellt werden, ob die intendierte Arbeit unter Einhaltung der gültigen Qualitätsstandards durchführbar ist. Dies betrifft insbesondere:
 - Fragen der Geheimhaltung
 - Inhaltliche Sinnhaftigkeit
 - Verfügbarkeit des notwendigen Know-hows
- Erst nach Auftragserteilung kann mit der Durchführung des Sonderprojektes begonnen werden. Die Durchführung erfolgt durch Mitarbeiterinnen und Mitarbeiter von Statistik Austria. Keinesfalls werden Einzeldatensätze übermittelt.
- Für Methoden und Verfahren, die bei Sonderauswertungen zum Einsatz kommen, gelten die gleichen Richtlinien und Qualitätsstandards wie für alle anderen Arbeiten von Statistik Austria. Diese geltenden Standards dürfen auch bei Sonderauswertungen nicht unterschritten werden.
- Für gesonderte Projekte sind nur dann Standard-Dokumentationen zu erstellen, wenn dies mit dem Auftraggeber vereinbart wurde, andernfalls ist die Art und Weise der Dokumentation Bestandteil der Vereinbarung mit dem Auftraggeber. Jedenfalls sollte seitens Statistik Austria ab einer gewissen Projektgröße proaktiv die Erstellung einer möglichst umfangreichen Dokumentation (wenn möglich im Format der Standard-Dokumentation) betrieben werden. Die interne Dokumentation sämtlicher Arbeitsschritte ist jedenfalls schon aus Gründen der Nachvollziehbarkeit unumgänglich.

3.6 Veröffentlichung statistischer Ergebnisse

3.6.1. Darstellung als Tabelle, Text und in visualisierter Form

Die Ergebnisse des statistischen Produktionsprozesses sind dem Auftraggeber bzw. der Auftraggeberin und der Öffentlichkeit zeitgleich zugänglich zu machen, wie im Bundesstatistikgesetz vorgesehen. Bei der Wahl der geeigneten **Publikationsform** (z. B. Druck, Datenträger, Internet, etc.) und des **Publikationstyps** (z. B. Pressemitteilung, Standardpublikation, Tabellenband, wissenschaftlicher Artikel, Twitter-Nachricht, etc.) ist es sinnvoll, vorab die Zielgruppe und den Zweck der Publikation in Betracht zu ziehen. Allenfalls vertraglich oder hausintern festgelegte Vorgaben sind dabei selbstverständlich zu beachten.

Aktuell ist eine Printpublikation statistischer Ergebnisse ohne gleichzeitige Bereitstellung eines druckfähigen Dokuments (z. B. im Portable Document Format, kurz PDF) auf der Website kaum mehr denkbar, während umgekehrt der Trend zur reinen Internetpublikation (bis hin zur Smartphone-App) anhält. Dafür sind nicht nur Kostengründe verantwortlich, sondern auch der technische Fortschritt, der sich in rasch ändernden Bedürfnissen der Nutzerinnen und Nutzer spiegelt. Diesen Rechnung zu tragen, ist auch für die amtliche Statistik unabdingbar.

Unabhängig von der Publikationsform stehen unterschiedliche Darstellungsmöglichkeiten für die Datenpräsentation zur Verfügung.

Tabellen

Sie gelten als klassische Darstellungsform statistischer Ergebnisse. Für ihre formal richtige Erstellung gilt die ÖNORM A 6195 (1989 09 01).

Von statistischen Tabellen wird größte Einfachheit bei gleichzeitig eindeutiger und übersichtlicher Darstellung von Informationen gefordert. Für den korrekten Aufbau solcher Tabellen müssen die einzelnen Elemente zueinander in logische und graphische Beziehung gebracht werden (z. B. Zeile, Spalte, Feld; Lineatur, Nummerierung; Textgestaltung und Typographie; Schreibung, Anordnung und Reihung von Zahlen; Gliederung, Layout, etc.)¹.

Die ÖNORM A 6195 stammt aus Vor-Internet-Zeiten und erläutert die korrekte Darstellung von Tabellen für Print-Publikationen. Die gleichen Regeln gelten selbstverständlich auch für die zum Zweck des Drucks erzeugten PDF-Dateien, die ins Web gestellt werden. Das Tabellenangebot auf den Webseiten von Statistik Austria soll nach Möglichkeit in drei Formaten erfolgen: HTML zur raschen Anzeige ohne Download, PDF als Druckversion und Excel zur weiteren Verarbeitung. Zur korrekten Auszeichnung von Excel-Tabellen für das Web dient der [STF-Konverter](#). Für jene Sonderform von Tabellen, die im Internet nur in den ersten Ebenen des Statistikzweiges präsentiert werden, steht eine geeignete Word-Vorlage zur Verfügung.

Neben den vorgefertigten Webtabellen auf www.statistik.at finden Nutzerinnen und Nutzer in der Datenbank STATcube die Möglichkeit, ihren individuellen Anforderungen entsprechende, hochdimensionale Tabellen interaktiv zu erstellen (siehe Kapitel 3.6.2) bzw. sich auf <http://data.statistik.gv.at/web/> für ausgewählte Themen des maschinenlesbaren Open Data-Formats zu bedienen (siehe Kapitel 3.6.2.1).

¹ https://shop.austrian-standards.at/action/de/public/details/10857/OENORM_A_6195_1989_09_01

Text

Texte spielen bei der Publikation statistischer Ergebnisse eine unterstützende Rolle. Prinzipiell sollte beschrieben werden, was zum Verständnis der tabellarisch präsentierten Ergebnisse notwendig ist. Das inkludiert auch eine kurze Beschreibung der verwendeten Konzepte und Methoden. Ausführliche Methodenbeschreibungen und Metadaten sind in den Standard-Dokumentationen (siehe auch Kapitel 3.6.7) verfügbar.

Es versteht sich von selbst, dass im Zusammenhang mit statistischer Ergebnispräsentation ganz allgemein ein präziser, sachlicher Stil zur Anwendung kommen sollte. Je nach Zielgruppe und Publikationstyp wird das Verhältnis zwischen Text, Tabellen und Grafiken jedoch unterschiedlich sein. So gelten etwa für Pressemitteilungen eigene Regeln, die von der Stabsstelle Medien- und Informationspolitik im „Handbuch Medienarbeit“ bzw. unter „Hilfestellungen“ im [Intranet](#) ausführlich beschrieben werden. Zu den Anforderungen hinsichtlich der Erstellung von Standardpublikationen, Artikeln für die „Statistischen Nachrichten“, Updates von „Österreich.Zahlen.Daten.Fakten“ oder Texten für andere, standardisierte Publikationen geben die Mitarbeiterinnen der „[Informationsverbreitung](#)“ gerne Auskunft.

Visualisierung

Nicht zuletzt aufgrund der Verbreitung mittels Internet und in sozialen Medien sind statistische Daten heute einem immer größeren Nutzerkreis zugänglich. Sie „lesen“ und auch „verstehen“ zu können, zählt zu den Kompetenzen, die von Statistikämtern weltweit unter dem Schlagwort „statistical literacy“ gefördert werden. Während sich die Aussage einer komplexen Tabelle nicht unbedingt auf den ersten Blick erschließt, kann eine bildhafte Darstellung leichter erfasst werden. Auch der Übungseffekt kommt nicht zu kurz: Je mehr man selbst durch interaktive Tools zum Ausprobieren animiert wird, desto besser versteht man die präsentierten Aussagen der Statistik. Der Visualisierung von statistischen Daten kommt daher in jüngster Zeit zunehmend Bedeutung zu.

Neben **Diagrammen**, und **Kartogrammen** wurden von Statistik Austria in jüngster Vergangenheit immer mehr **interaktive Anwendungen** zur Verfügung gestellt. Dazu zählen etwa [i.MAPs](#), die als Kartendarstellung oder auch in Form einer Bevölkerungspyramide angeboten werden. Daneben gibt es aber auch Anwendungen wie den „[Persönlichen Inflationsrechner](#)“, der neben der tabellarischen Darstellung der Ausgabenstruktur auch Linien- und Balkendiagramme sowie Treemaps enthält.

Die modernste Form kartographischer Auswertung bieten die sogenannten **Geographischen Informationssysteme (GIS)**. Sie ermöglichen es, statistische Daten nicht nur auf Basis vorgegebener räumlicher Einheiten (wie der Verwaltungsgliederung oder der Zählsprenkel) darzustellen, sondern für beliebig definierte räumliche Gebilde wie

- Gitternetzmaschen beliebiger Weite;
- Einzugsbereiche von Schulen und Spitälern;
- Gefahrenzonen wie Lawenstriche oder Hochwassergebiete;
- Zonen erhöhter Schadstoff- oder Lärmemission.

Statistik Austria war in den letzten Jahren bestrebt, das Angebot an kartographischen Darstellungen laufend zu erhöhen. Ein Beispiel ist die „[Verkehrsunfallskarte](#)“, aber auch der seit vielen Jahren bestehende [Wirtschafts- und Regionalatlas Österreichs](#). Mit dem OnlineAtlas können viele wirtschaftsstatistische Merkmale anhand von mehr als 5.300 Europa- bzw. Bundesländerkarten interaktiv abgefragt und anschaulich visualisiert werden.

Richtlinien:

- Vor Veröffentlichung statistischer Ergebnisse, insbesondere in Pressemitteilungen, ist das Vier-Augen-Prinzip einzuhalten. Dies bedeutet, dass die Korrektheit der zu veröffentlichenden Daten (und der Abgleich zwischen Text und Tabellenwerten) von einer anderen Person als dem/der jeweiligen Autor/Autorin geprüft und bestätigt werden muss.

Tabellen

- Alle Tabellen sollen zumindest folgende Komponenten beinhalten:
 - Tabellenüberschrift
 - Kopfzeile
 - Vorspalte
 - Zahlenbereich
 - Fußzeilenbereich
- Die Tabellenüberschrift soll den im Zahlenbereich präsentierten Gegenstand und die in der Vorspalte getroffene Gliederung sowie den Berichtszeitraum (oder Stichtag), auf den sich die Daten beziehen, beinhalten. Die zeitliche, sachliche und räumliche Abgrenzung der der Tabelle zugrunde liegenden Masse soll aus der Überschrift hervorgehen.
- Aus der Kopfzeile soll klar hervorgehen, auf welche Teilmasse sich die Zahlen der jeweiligen Spalten beziehen. Des Weiteren soll die Einheit angegeben sein, in der die Zahlen präsentiert werden.
- Die Gliederung in der Vorspalte soll klar strukturiert sein. Es soll hervorgehen, in welcher Beziehung zueinander sich einzelne Gliederungspositionen befinden. Gegebenenfalls soll dies durch das Ausweisen von Summenzeilen verdeutlicht werden.
- Die in den Zellen präsentierten Zahlen sollen gut lesbar sein. Bei der Wahl der Einheit ist darauf zu achten, dass nicht zu viele Ziffern zur Darstellung notwendig sind. Gegebenenfalls soll mit (mathematisch korrekter) Rundung gearbeitet werden. Überdies gelten folgende Konventionen für die Präsentation in Tabellen:
 - Ist ein Feld unbesetzt, schreibt man einen Querstrich (-)
 - Ist der erhobene Wert kleiner als die Hälfte der vorgesehenen kleinsten Einheit, schreibt man eine Null (0 oder 0.0)
 - Liegen die betreffenden Informationen noch nicht vor, sind aber zu erwarten, schreibt man drei Punkte (...)
 - Werden Zahlen überhaupt nicht ermittelt oder verstößt ihre Veröffentlichung gegen die Bestimmungen der statistischen Geheimhaltung, so wird diese Zelle mit dem Buchstaben „G“ markiert und der eigentliche Wert unterdrückt.
- Im Fußzeilenbereich soll in der Quellenangabe immer „Statistik Austria“ angegeben sein. Gegebenenfalls sollen dort auch Anmerkungen in Fußnotenform zu einzelnen Zellwerten oder Gliederungspositionen innerhalb der Tabelle zu finden sein.

Text

- Die Texte als Ergänzungen der tabellarischen Publikationen sollen alle für das Verständnis der präsentierten Zahlen notwendigen Aspekte kurz und prägnant zusammenfassen. Keinesfalls sollen präsentierte Ergebnisse mit einer Aufzählung im Text wiederholend dargeboten werden. Sinnvoll ist allerdings, auf die wichtigsten Hauptergebnisse bzw. auffälligsten Änderungen einzugehen.

- Für alle statistischen Publikationen sollen folgende Grundsätze gelten:
 - Der Stil soll präzise, eindeutig, und sachlich sein.
 - Die Wortwahl soll zielgruppengerecht sein und ohne Fachjargon auskommen,
 - Unvermeidbare Fachausdrücke sind zu definieren.
 - Die Verwendung von Zitaten hat in einheitlicher Form zu erfolgen.
 - Für häufig verwendete Abkürzungen ist ein entsprechendes Verzeichnis zu erstellen.
 - Vor jeder Freigabe einer Publikation muss der Text einer intensiven redaktionellen und inhaltlichen Prüfung unterzogen werden. Die Freigabe einer Publikation obliegt dem fachlich zuständigen Direktor/Abteilungsleiter.

Visualisierung

Die Entwicklung und Implementierung eines Inhouse-Grafik-Standards, der modernen und methodisch bedingten Anforderungen gerecht wird, ist Teil der Strategie 2020. Sobald ein Regelwerk vorliegt, werden ausschließlich die dortigen Vorgaben gelten. Inzwischen gilt, was bisher geübte Praxis war:

- Bei einer grafischen Präsentation soll der ausgewählte Grafiktyp (Balkendiagramm, Liniendiagramm, Kreisdiagramm etc.) der Problemstellung angepasst sein.
- Grafiken müssen valide sein, so dürfen keine perspektivischen Verzerrungen auftreten, Größen sollen flächentreu präsentiert werden und Achsen dürfen nur unvollständige dargestellt werden, wenn daraus kein täuschender Eindruck entstehen kann.
- Grafiken oder Kartogramme müssen ausreichend beschriftet sein (Achsen, Datenreihen, wichtige Datenpunkte und die Legende). Die Überschrift muss „sprechend“ sein und den intendierten Inhalt der Grafik bzw. des Kartogramms beschreiben.
- Interaktiven Darstellungen ist, wenn möglich, gegenüber statischen der Vorzug zu geben.
- Wann immer möglich, sollen geobasierte Darstellungen zum Einsatz kommen.
- Bei der Erstellung kartografischer Darstellungen soll der „Bereich Kartographie und Geographische Informationssysteme“ einbezogen werden.

3.6.2. Open Data

Open Data (Offene Daten) sind sämtliche Datenbestände, die im Interesse der Allgemeinheit der Gesellschaft ohne jedwede Einschränkung zur freien Nutzung, zur Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden.

Für die öffentliche Verwaltung spielen in diesem Zusammenhang „Open Government Data“ eine wichtige Rolle. Darunter versteht man jene nicht-personenbezogenen und nicht-infrastrukturkritischen Datenbestände, die im Interesse der Allgemeinheit ohne jedwede Einschränkung zur freien Nutzung, zur Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden sollen

Statistik Austria als Teil der öffentlichen Verwaltung bietet seit 2012 auf dem Portal data.statistik.gv.at eine Reihe von Datensätzen aus dem Datenschatz der amtlichen Statistiken in Österreich an. Die angebotenen Daten entsprechen den [Open Data-Prinzipien](#), erfüllen also folgende Kriterien:

1. Vollständigkeit
2. Primärquelle
3. Zeitnahe Zurverfügungstellung
4. Leichter Zugang
5. Maschinenlesbarkeit
6. Diskriminierungsfreiheit
7. Verwendung offener Standards
8. Lizenzierung
9. Dokumentation (Dauerhaftigkeit)
10. Keine Nutzungskosten

Insbesondere der kostenfreie Zugang sowie die Maschinenlesbarkeit ermöglichen es hier einem entsprechenden Kreis von Nutzerinnen und Nutzern zeitnah Daten für die Entwicklung spezifischer Anwendungen (z.B. „Apps“ oder andere automatische Weiterverarbeitung) zu verwenden.

Zu erwähnen ist auch, dass das Open Data Portal der Bundesanstalt vollständig mit dem Portal des Bundes data.gv.at synchronisiert ist, was bedeutet, dass sämtliche Open Data-Datensätze von Statistik Austria auch über dieses Portal erreicht werden können.

Statistik Austria hat als führender Informationsdienstleister den Anspruch, alle gesellschafts- und wirtschaftspolitisch relevanten Themenfelder abzudecken – das gilt auch für den Bereich Open Data. Um keiner Fehleinschätzung hinsichtlich des Bedarfs der Zielgruppe zu unterliegen, sollte den Rückmeldungen potentieller Nutzerinnen und Nutzer dieses Angebotes besonderes Augenmerk geschenkt werden. Daher ist die Beobachtung von Meldungen der Community auf entsprechenden Plattformen (hier vor allem Twitter) ein wichtiges Instrument der Bedarfserhebung.

Die hausinterne Koordination bezüglich Open Data liegt in Händen der Stabsstelle Qualitätsmanagement und Methodik in enger Kooperation mit der Stabsstelle Medien- und Informationspolitik.

Richtlinien:

- Bevor ein Datensatz auf dem Portal statistik.data.at zur Verfügung gestellt werden kann, muss ein entsprechendes, durch die IT Abteilung zur Verfügung gestelltes Metadatenfile befüllt werden. Dieses definiert einerseits klar, welche Daten davon betroffen sind, und enthält andererseits auch jene Metainformationen, die Nutzerinnen und Nutzer für die Verwendung des Datensatzes benötigen.
- Die Veröffentlichung eines neuen Open Data Datensatzes bedarf der Genehmigung der zuständigen Direktions- bzw. Abteilungsleitung. Des Weiteren sind die stabstellen Qualitätsmanagement und Methodik sowie Medien- und Informationspolitik zu informieren.
- Es ist darauf zu achten, dass nur solche Daten als Open Data angeboten werden, die nicht an anderer Stelle, etwa in der Datenbank STATcube, ausschließlich kostenpflichtig verfügbar sind. Dies ist auch insbesondere hinsichtlich der angestrebten Gliederungstiefe für den geplanten Datensatz zu beachten.
- Es muss gewährleistet sein, dass sämtliche Open Data Datensätze regelmäßig gewartet werden, was insbesondere die Aktualisierung auf den neuesten Datenbestand betrifft. Die Aktualisierung erfolgt durch den Fachbereich in Kooperation mit der IT-Abteilung
- Bevor Daten auf das Open Data Portal gestellt werden muss überprüft werden, ob sämtliche relevanten Geheimhaltungsrichtlinien eingehalten werden
- Die Bestrebungen sollten dahingehen, dass Statistik Austria eine möglichst hohe Abdeckung mit seinem Open Data Angebot erreicht.

3.6.3 Statistische Datenbank STATcube

STATcube ist die zentrale, öffentlich zugängliche Datenbank von Statistik Austria und damit das Herzstück des Datenangebots des Hauses für alle Nutzerinnen und Nutzer. Ziel ist, dass die Inhalte von STATcube das gesamte Angebotsspektrum von Statistik Austria, insbesondere von periodisch wiederkehrenden Projekten, abdecken.

Die Publikation statistischer Ergebnisse in vordefinierter Form (gedruckt oder auch als Webtabellen) ist statisch, bietet nur ein limitiertes Angebot und schränkt Auswertungsmöglichkeiten ein. Die wachsende Nachfrage nach interaktiv erstellbarer, statistischer Information erfordert, dass Benutzerinnen und Benutzern die Möglichkeit offen steht, eigene Auswertungen aus vorhandenen Daten selbsttätig zusammenzustellen.

Statistik Austria betreibt mit **STATcube** eine **Statistische Datenbank**, in der fachstatistische Datenbestände eingelagert sind. STATcube, das Nachfolgeprodukt der seit den 1970er Jahren bestehenden Datenbank ISIS, ist seit 2011 im Einsatz. Den Datennutzerinnen und Datennutzern stehen mit STATcube Auswertungsmöglichkeiten für eine große Datenfülle zur Verfügung. Die Nutzung von Schnittstellen gestattet den Export in gängige Dateiformate, was eine Weiterverarbeitung mit anderen, weitverbreiteten Software-Produkten ermöglicht.

Statistische Daten in STATcube sind immer eindeutig einem bestimmten statistischen Gegenstand (Personen, Unternehmen, etc.) zugeordnet, für den dann dessen Anzahl oder ein auf ihn bezogenes Merkmal (z.B. Umsätze, Beschäftigte) aggregiert wird. Die Anreicherung der Datenbestände um sogenannte Gliederungsmerkmale ermöglicht es Benutzerinnen und Benutzern einerseits die betrachtete Masse zeitlich, räumlich und inhaltlich einzugrenzen und andererseits mehrdimensionale Tabellen zu definieren. Aufgrund der Möglichkeit der mehrdimensionalen Aufgliederung bezeichnet man die in STATcube vorhandenen Datenbestände auch als Datenwürfel.

Die Möglichkeit der Erzeugung hochdimensionaler Tabellen mit einer großen Anzahl von Datenzellen bringt jedoch auch problematische Aspekte mit sich. So können aufgrund der Erzeugung von Zellen mit geringer Besetzung Geheimhaltungsprobleme auftreten (siehe Kapitel 3.6.6). Ein weiteres Problem stellen Daten dar, die aus Stichprobenerhebungen stammen und mit Unsicherheiten behaftet sind. Aufgrund der starken Untergliederungsmöglichkeiten einer Datenbank können Zellen entstehen, deren Werte mit so großen Unschärfen versehen sind, dass ihre Verwertbarkeit problematisch erscheint (siehe Kapitel 3.5.2).

Die Wartung einer Datenbank erfordert auch eine Benutzerverwaltung, die einen Überblick darüber ermöglicht, wer auf welche Datenbestände zugreift. Daraus und aus der laufenden Kommunikation mit den Benutzerinnen und Benutzern der Datenbank, z.B. über Funktionalität, Usability und inhaltliche Vollständigkeit, wird ersichtlich, wo Bedarf an zusätzlichem statistischen Inhalten oder technischer Aufrüstung besteht.

Richtlinien:

- Ziel ist, jeden seitens Statistik Austria publizierten fachstatistischen Datenbestand in STATcube einzulagern und diesen den Nutzerinnen und Nutzern auf anwenderfreundliche Art und Weise weiterverarbeitbar zur Verfügung zu stellen.
- Werden Daten in STATcube eingelagert, muss geplant werden, welche Merkmale in welcher Gliederungstiefe zur Verfügung gestellt werden. Jedenfalls sollen Merkmale vorhanden sein, die eine räumliche und zeitliche Eingrenzung der betreffenden statistischen Masse erlauben.
- Bevor Datenbestände eingelagert werden, muss geprüft werden, ob Geheimhaltungsprobleme auftreten können. Es ist zu definieren, wie mit Zellen umgegangen wird, deren Werte aufgrund der Verletzung von Geheimhaltungsregeln nicht in originärer Form ausgewiesen werden können. Dies betrifft auch Maßnahmen eventueller Gegenlöschungen. Sämtliche Richtlinien der statistischen Geheimhaltung (siehe Kapitel 3.6.6) haben auch für die Publikation in STATcube ihre Gültigkeit. In diesem Zusammenhang sei auf die im Internet publizierten [Richtlinien zur statistischen Geheimhaltung](#) verwiesen.
- Daten, die erstmals eingelagert werden, müssen getestet werden, ehe sie für die Nutzung freigegeben werden. Insbesondere ist zu überprüfen, ob die Werte vor und nach dem technischen Konversionsprozess, der der Einlagerung vorangeht, identisch sind. Datennutzerinnen und -nutzer sollen darüber informiert werden, dass die Daten nunmehr in der Datenbank verfügbar sind.
- Eingelagerte Datenbestände sind regelmäßig zu warten. Die Publikation der Ergebnisse und die Einlagerung in die Datenbank sollen zeitgleich erfolgen.
- Mitarbeiterinnen und Mitarbeiter sollen, wenn notwendig, Benutzerinnen und Benutzer von STATcube unterstützen, falls sich online verfügbare Hilfsfunktionen als nicht ausreichend erweisen. In diesem Zusammenhang wird auf die auf der Webseite von Statistik Austria verfügbaren [Unterlagen](#) verwiesen.

3.6.4 Datenübermittlung an Eurostat

Ein wesentlicher Anteil der Produkte von Statistik Austria wird auf Grundlage europäischer Anforderungen erstellt. Dies betrifft vor allem Statistiken, bei denen ein starkes Bestreben zur Harmonisierung der Konzepte auf europäischer Ebene besteht. In diesem Zusammenhang ist **Eurostat** ein wichtiger Partner für die Bundesanstalt. Die Zusammenarbeit ist zum einen in der Planungsphase gegeben, wo in - von Eurostat organisierten - Arbeitsgruppensitzungen Konzepte, Definitionen, Methoden und Lieferprogramme zwischen Eurostat und den Mitgliedstaaten der Europäischen Union diskutiert werden. Zum anderen ist Eurostat auch jene Stelle, an die Statistik Austria gemäß einem **vereinbarten Lieferprogramm** verpflichtet ist, Daten zu übermitteln. Der Umfang, die Termine, und die Inhalte der Lieferungen werden in entsprechenden EU-Verordnungen festgelegt.

Prinzipiell können derartige Übermittlungen folgendes umfassen:

- I. Übermittlung eines Tabellenprogramms
- II. Übermittlung von Einzeldaten
- III. Übermittlung von Metadaten in Form von Qualitätsberichten bzw. detaillierten Methodenbeschreibungen

Die Lieferung von Tabellen an Eurostat ist zumeist eine Teilmenge des nationalen Publikationsprogramms. Man spricht auch von der Lieferung ausgewählter Ergebnisse. Nachdem die nationalen Daten zur Schätzung europäischer Aggregate dienen, ist die termingerechte Übermittlung auch unter dem Blickwinkel der aktuellen Verfügbarkeit europäischer Ergebnisse zu sehen. Daher ist die Einhaltung dieser Liefertermine als besonders prioritär zu erachten. Eurostat selbst stellt die europäischen Aggregate und Statistiken wiederum über eigene Publikationsschienen (z.B. Pressemitteilungen, Datenbanken, Statistics Explained, Jahrbücher) den Nutzerinnen und Nutzern zur Verfügung.

Hinsichtlich der Übermittlung von Einzeldaten an Eurostat hat die Bundesanstalt dafür zu sorgen, dass diese in entsprechend anonymisierter Form erfolgt, sodass kein Rückschluss auf einzelne Personen oder Unternehmen erfolgen kann.

Qualitätsberichte und/oder detaillierte Methodenbeschreibungen der Mitgliedstaaten dienen Eurostat als Nachweis dafür, dass die zuvor in den Arbeitsgruppensitzungen vereinbarten Empfehlungen und Qualitätsstandards eingehalten wurden. Obwohl die Formate und Inhalte der von Eurostat eingeforderten Qualitätsberichte abhängig von der jeweiligen Materie variieren, orientieren sie sich doch alle an der Definition der Qualität in der Statistik (vgl. Kapitel 1). Diese Berichte sind in elektronischer Form gemäß den in SDMX implementierten Standards SIMS (Single Integrated Metadata Structure) elektronisch an Eurostat über entsprechend implementierte Plattformen zu übermitteln.

Im Sinne der Europäischen Statistik Verordnung agiert Statistik Austria als „Single point of Contact“ für Datenlieferungen an die Europäische Kommission (EUROSTAT). Dies gilt insbesondere auch für Daten anderer nationaler Produzenten – „Other National Authorities (ONA)“. In diesem Sinne strebt Statistik Austria an, gegenseitige Vereinbarungen mit nationalen Produzenten zu treffen, auf deren Basis eine zeitgerechte und den Qualitätskriterien der Amtlichen Statistik entsprechende Datenlieferung garantiert wird. Diese Vereinbarungen sollen auch gegenüber der Kommission kommuniziert werden.

Richtlinien:

- Es ist jedwede Anstrengung seitens Statistik Austria zu unternehmen, um die Anforderungen von Eurostat zu erfüllen. Dies betrifft sowohl die Inhalte und den Umfang des Lieferprogramms als auch die terminlichen Erfordernisse, die Art der Datenübermittlung und ggf. vorliegende Datenvalidierungserfordernisse.
- Die Mitarbeiterinnen und Mitarbeiter, die an den Arbeitsgruppen teilnehmen, sollen dort Positionen vertreten, die hausintern im fachstatistischen Bereich abgestimmt sind. Nur auf diese Weise kann eine konsistente Vertretung österreichischer und hausinterner Interessen über alle hierarchischen Gremien des ESS sichergestellt werden.
- Die Lieferung von Daten bzw. Metadaten an Eurostat muss dokumentiert werden. Insbesondere müssen alle diesbezüglichen Vorgänge in der Standard-Dokumentation beschrieben werden.
- Einzeldaten sind auch an Eurostat nur in anonymisierter Form zu übermitteln. Die Lieferung soll gemäß einem vorher vereinbarten Satzaufbau erfolgen.
- Die Mitarbeiterinnen und Mitarbeiter von Statistik Austria sollen in der Zusammenarbeit mit Eurostat darauf hinwirken, dass die Inhalte und Formate übermittelter Metadaten über die Zeit stabil bleiben, damit ein hohes Maß an zeitlicher Vergleichbarkeit der Qualitätssinformation gegeben ist.
- Verpflichtende Qualitätsberichterstattung innerhalb des Europäischen statistischen Systems im Sinne Artikel 12 der EU-Statistik Verordnung ist, entsprechend den Europäischen Metadaten Standards (SIMS, ESMS, ESQRS) an Eurostat zu übermitteln. Voraussetzung dafür ist jedoch die entsprechende Implementierung seitens Eurostat.
- Es ist seitens Statistik Austria dafür Sorge zu tragen, dass auch die Datenlieferungen der anderen nationalen Produzenten zeitgerecht und den Qualitätskriterien der Amtlichen Statistik entsprechend an Eurostat erfolgen.

3.6.5 Internationale Anfragen

Statistik Austria agiert im internationalen Umfeld und ist Teil des Europäischen Statistischen Systems (ESS). Die Kommunikation zwischen NSI's und anderen nationalen und internationalen Organisationen und Institutionen, die mit der Entwicklung, Erstellung und Verbreitung von Statistiken befasst sind, ist für den Wissensaustausch von eminenter Bedeutung. Diese Vernetzung findet nicht nur in Arbeitsgruppen, Ausschüssen und auf internationalen Veranstaltungen, wie z.B. Tagungen und Konferenzen statt, sondern erfolgt auch durch Anfragen, die direkt an ein NSI gerichtet werden. Statistik Austria ist bemüht, **internationale Anfragen** zufriedenstellend zu beantworten. Oftmals werden derartige Anfragen von NSIs und internationalen Organisationen, wie der UN oder der OECD, in Form von Fragebögen z. B. zur Befüllung von Datenbanken/ Publikationen bzw. im Rahmen von Forschungsprojekten gestellt, im Rahmen dessen ein NSI oder eine internationale Organisation federführend beteiligt ist. Die Beantwortung liegt im Interesse der Bundesanstalt. Unter anderem können die Ergebnisse solcher Projekte, die häufig im Rahmen des ESS durchgeführt werden, für die Amtliche Statistik von Nutzen sein. Anfragen von Eurostat oder durch beigeordnete Gremien, wie das Europäische Beratungsgremium für die Statistische Governance (ESGAB), oder den Europäischen Beratenden Ausschuss für die Statistik (ESAC), stehen oft im Zusammenhang mit dem Verhaltenskodex für Europäische Statistiken und dessen Einhaltung, zu der sich Statistik Austria bekennt.

Anfragen können globaler Natur sein und allgemeine, die Statistik Austria betreffende Aspekte zum Inhalt haben (Management-Themen, Arbeitsprogramm, budgetäre bzw. personalspezifische Gegebenheiten etc.). Es kann sich diesbezüglich aber auch um die methodische bzw. praktische Handhabung gewisser Themengebiete handeln (z. B. Metadatenmanagement, Zeitreihenbereinigung, Imputation etc.). Zudem kann die Anfrage ein statistisches Fachgebiet direkt betreffen, indem an Daten oder/ und an speziellen methodischen, oder konzeptuellen Gegebenheiten Interesse besteht. So können zum Beispiel spezielle Ergebnisse, die nicht Teil der standardmäßigen Publikation sind, gewünscht werden, was mit einer Sonderauswertung vergleichbar ist (vgl. Kapitel 3.5.9).

Die Beantwortung jeder Anfrage sollte - wenn nicht zwingende Gründe dagegen sprechen (z. B. Geheimhaltung) - jedenfalls vollständig erfolgen. Sämtliche fachlich betroffenen Organisationseinheiten von Statistik Austria sollten einbezogen werden. Die Stabsstelle „Internationale Beziehungen“ (IB) muss, falls sie nicht selber die Beantwortung der Anfrage vornimmt im Zuge der Beantwortung durch die verantwortliche Organisationseinheit informiert werden. Diese Informationspflicht betrifft nur Anfragen, die über Routinetätigkeiten hinausgehen (wie z.B. die Beantwortung von Fragebögen). Datenanfragen sowie Routineanfragen wie z.B. durch Eurostat-Verordnungen geregelte) sind jedenfalls ausgenommen

Im Sinne einer besseren Verwaltung und Koordinierung solcher Anfragen wird von IB eine zentrale Erfassung vorgenommen. Bei Dir./Abt. übergreifenden Aussendungen werden Koordinierung und Erfassung von IB übernommen, ansonsten lediglich die von den Direktoren/Abteilungen/Stabsstellen erhaltenen Informationen zur Dokumentation.

Richtlinien:

- Nach Einlangen einer internationalen Anfrage soll die Stabsstelle Internationale Beziehungen informiert werden. Diese nimmt entweder nur einen Vermerk in der zentralen Archivierungsliste zur Dokumentation vor (sollte die Anfrage bereits an die fachlich kompetente Mitarbeiterin oder den fachlich kompetenten Mitarbeiter gerichtet worden sein) oder sie leitet die Anfrage - falls erforderlich - an die zuständigen Expertinnen und Experten bei Statistik Austria weiter. Die Informationspflicht hinsichtlich der Anfragen umfasst insbesondere Anfragen folgender Organisationen:
 - Die Europäische Kommission insbesondere Eurostat, ESGAB und ESAC
 - Andere der Europäischen Kommission zuzurechnende Institutionen
 - UNO (inkl. UN/ECE)
 - OECD
 - EZB
 - Andere NSIs (Falls sie z.B. sie als Koordinatoren wichtiger Projekte (wie auftreten)
- Die Stabsstelle Internationale Beziehungen erfasst die beantworteten Anfragen. Sollte es zu einer Anfrage oder zu Teilen davon bereits ähnliche oder idente Anfragen gegeben haben, muss die Antwort auf die aktuelle Anfrage mit den damals gegebenen Antworten abgestimmt werden. Dazu kann auch die Dokumentation der Stabsstelle Internationale Beziehungen herangezogen werden.
- Die Beantwortung einer internationalen Anfrage soll erst nach einem vollständig abgeschlossenen Abstimmungsprozess unter sämtlichen mit der Anfrage befassten Organisationseinheiten erfolgen. Über die Freigabe der Antwort entscheiden, falls nur eine Direktion/ Abteilung betroffen ist, die jeweiligen Direktorinnen und Direktoren bzw. Abteilungsleiterinnen und Abteilungsleiter. Für direktionsübergreifende bzw. Statistik Austria im Allgemeinen betreffende Anfragen liegt die Koordination bei der Stabsstelle Internationale Beziehungen. Die Freigabe der Antwort erfolgt in diesem Fall seitens der Generaldirektion. Die Verantwortung für die einzelnen Bestandteile obliegt jeweils den zuständigen Direktorinnen und Direktoren bzw. Abteilungsleiterinnen und Abteilungsleitern.
- Internationale Anfragen müssen möglichst vollständig und fristgerecht beantwortet werden.

3.6.6 Revisionen

Von der Amtlichen Statistik publizierte Ergebnisse werden von Nutzerinnen und Nutzern oftmals als Grundlage für weiterführende Analysen verwendet oder dienen Verantwortungsträgerinnen und Verantwortungsträgern als Entscheidungsgrundlage. Daher ist das Vertrauen in publizierte Zahlen seitens der Nutzerinnen und Nutzer eine unbedingte Voraussetzung für die Relevanz der Produkte statistischer Ämter.

Ein Problem für diese Wechselwirkung tritt auf, wenn bereits veröffentlichte Daten, die für einen bestimmten Berichtszeitraum (oder auch Stichtag) gültig sind, ex-post verändert werden. In diesem Zusammenhang wird von einer **Revision** gesprochen. Der generelle Umgang mit Revisionen wird in der Revisionspolitik von Statistik Austria geregelt. Dort findet sich im Anhang auch eine Liste all jener Statistiken, für die laufende Revisionen durchgeführt werden.

Der Begriff der Revision wird in der Amtlichen Statistik auch in einem breiteren Zusammenhang verwendet, nämlich dann, wenn es zu Änderungen bei umfassenden, weite Teile der Amtlichen Statistik betreffenden Grundlagenthematiken (z. B. Klassifikationen) kommt. In diesem Kapitel werden nicht diese sondern ausschließlich Datenrevisionen thematisiert.

Laufende Revisionen sind solche, die standardmäßig gemäß einer a priori fixierten Publikationsprozedur folgend erfolgen. Zu fix in vorgegebenen Terminen werden die bis dato veröffentlichten Ergebnisse durch andere ersetzt. Im einfachsten Fall gibt es zwei Publikations-schritte. Dabei werden **vorläufige Ergebnisse** publiziert, die dann durch **endgültige Ergebnisse** ersetzt werden. In Kapitel 3.5.8 wurde bereits die Thematik der Flash-Estimates angesprochen, die die zeitlich gesehen extremste Form der Veröffentlichung vorläufiger Ergebnisse darstellt. Manchmal kann es der Revisionsplan auch vorsehen, dass es mehrere Generationen vorläufiger Ergebnisse gibt („Vintages“). Resultate, die durch laufende Revisionen ersetzt werden, gelten nicht als endgültig. Erst jene Ergebnisse, die nach der letzten laufenden Revision publiziert werden, haben finalen Status. Laufende Revisionen korrigieren nicht Fehler innerhalb des Produktionsprozesses, sondern basieren zumeist darauf, dass die neu zu publizierenden Zahlen auf einer umfangreicheren Datenbasis beruhen. Substantiell für Datennutzerinnen und -nutzer ist das Wissen, dass es sich bei einem publizierten Datum um ein vorläufiges Ergebnis handelt, also die Nutzung und Weiterverarbeitung ein gewisses Risiko birgt. Darüber hinausgehend ist die Beurteilung des zu erwartenden Ausmaßes der Abweichung und ob die vorzeichenmäßige Richtung der Differenz systematisch ist.

Bei manchen Produkten, die aus langen Zeitreihen bestehen, kann es zu umfangreichen Revisionen kommen, die die Abänderung der Werte lang zurückreichender Zeitpunkte zur Folge haben. Solche „großen Revisionen“ können auch mehrere Produkte involvieren, da vor allem bei synthetischen Produkten (z. B. VGR) Änderungen in als Datenquellen genutzten Statistiken Berücksichtigung finden.

Als **anlassbezogene Revisionen** werden solche bezeichnet, die nicht im Vorhinein geplant gewesen sind und dadurch zustande kommen, dass **bereits als endgültig** gekennzeichnete Ergebnisse aufgrund neuer Erkenntnisse geändert werden müssen. Die Charakteristik derartiger Revisionen liegt in der Tatsache, dass sie unvorhergesehen erfolgen und daher für etwaige Nutzerinnen und Nutzer, die die bereits publizierten Zahlen für weiterführende Analysen verwendet haben, problematisch sind. Die Gründe für ein derartiges Vorgehen können vielfältig sein. Aufgrund der bereits angesprochenen Problematik der möglichen Datenverwertung durch Nutzerinnen und Nutzer muss für jedes Produkt individuell überlegt werden, ob Ergebnisse revidiert werden sollen oder nicht. Jedoch gibt es Szenarien, die dies mehr oder minder zwingend erfordern, z. B.

- **Änderungen der Datenquellen:** Nach der Veröffentlichung der Ergebnisse stehen eine oder mehrere neue Datenquellen zur Verfügung, oder eine verwendete Datenquelle hat sich fundamental geändert.
- **Neue methodische Erkenntnisse:** Die Bundesanstalt ist bestrebt bei der Erstellung statistischer Ergebnisse immer die neuesten Methoden („current best practices“) anzuwenden. Sollte sich nach Veröffentlichung von Resultaten herausstellen, dass ein neueres Verfahren/eine neuere Methode dem/der zuvor angewendeten vorzuziehen ist, sollte eine Revision ins Auge gefasst werden.
- **Korrektur von Fehlern:** Passiert im Zuge des Produktionsprozesses trotz der angewendeten Qualitätssicherungsprozesse ein nach profunder Analyse hinsichtlich der damit verbundenen Ergebnisabweichungen als schwerwiegend eingeschätzter Fehler, müssen jene Resultate, die davon betroffen sind, revidiert werden.

Der Umgang mit Revisionen, vor allem die Art und Weise, wie die Informationsverpflichtung über Revisionen gegenüber Nutzerinnen und Nutzern wahrgenommen wird, findet sich in der im Internet publizierten [Revisionspolitik von Statistik Austria](#). Im Anhang dieses Dokuments finden sich auch Informationen über die Revisionspläne der Produkte von Statistik Austria.

Richtlinien:

- Wenn vorläufige oder nicht endgültige Ergebnisse, für die Revisionen zu erwarten sind, publiziert werden, sollen diese eindeutig als solche gekennzeichnet werden. Dies betrifft sowohl die Ankündigung der Veröffentlichungstermine im [Veröffentlichungskalender](#) von Statistik Austria als auch die Kennzeichnung innerhalb der publizierten Ergebnisse.
- Im Falle laufender Revisionen sollen die entsprechenden Revisionspläne transparent und öffentlich zugänglich sein. In diesem Sinne muss der Anhang der im Internet verfügbaren Revisionspolitik von Statistik Austria laufend gewartet werden.
- Es muss dafür Sorge getragen werden, dass etwaige durch eine Revision hervorgerufene Änderungen in Ergebnissen auch in anderen, mit diesen Ergebnissen verbundenen statistischen Produkten nachgezogen und auch dort dokumentiert werden.
- Laufende Revisionen sind in der Standard-Dokumentation zu beschreiben. Dies betrifft sowohl die zeitliche Dimension der Veröffentlichung der einzelnen Revisionschritte, als auch quantitative Informationen über die Abweichung einzelner Datengenerationen von den endgültigen Ergebnissen.
- Werden anlassbezogene Revisionen durchgeführt, müssen Benutzerinnen und Benutzer umgehend darüber informiert werden. Dabei soll die Information in der gleichen Publikationsschiene erfolgen, wie die Veröffentlichung der ursprünglichen Ergebnisse. Die Informationspflicht bei anlassbezogenen Revisionen ist in der Revisionspolitik von Statistik Austria umfassend geregelt.

3.6.7 Geheimhaltung

Amtliche Statistik zielt immer auf die Beschreibung von Massenerscheinungen ab. In diesem Sinne sind zumeist Aggregate von Interesse, hinter deren Ermittlung eine gewisse Anzahl von Einzeldatensätzen steht. Das Einzeldatum an sich oder welche physische oder juristische Person sich dahinter verbirgt, ist für statistische Zwecke nicht von Belang.

Wer Statistik Austria Auskünfte erteilt, kann darauf vertrauen, dass seine persönlichen Angaben ausschließlich statistischen Zwecken dienen. Die Angaben der Befragten werden geheim gehalten und dürfen ohne rechtliche Grundlage nicht an andere Institutionen oder Stellen weitergegeben werden, weder an staatliche Institutionen noch an private oder sonstige Stellen. Das BStatG 2000 enthält entsprechende Maßnahmen zum Schutz der Geheimhaltungsinteressen von Personen und Unternehmen und zur Sicherung der Vertraulichkeit von Einzeldaten, darunter das Löschen von Namen und Adressen zum frühestmöglichen Zeitpunkt oder die Verschwiegenheitspflicht der mit Aufgaben der Amtlichen Statistik betrauten Personen. Die Vertraulichkeit persönlicher Angaben ist daher für Statistik Austria mehr als eine gesetzliche Pflicht. Sie hat ein ureigenes Interesse daran, die statistische Geheimhaltung durch technische und organisatorische Maßnahmen zu garantieren.

Eine hausinterne [„Datensicherheitsvorschrift“](#) regelt den Umgang mit personenbezogenen Daten. Die Mitarbeiterinnen und Mitarbeiter müssen sich zur Einhaltung des Datengeheimnisses [schriftlich](#) verpflichten.

Bei der Auswertung kleiner Zellen kann es zu der Situation kommen, dass durch Kombination verschiedener Auswertungsdimensionen Einzeldatensätze identifizierbar werden. Aus diesem Grund muss darauf geachtet werden, dass die statistische Geheimhaltung gewahrt bleibt und das Risiko der Identifizierung („Disclosure risk“) so minimal ist, dass eine Zuordnung von Merkmalswerten zu Einzelindividuen de facto ausgeschlossen werden kann. In solchen Fällen ist es notwendig, in Tabellen bzw. in den Mikrodaten manipulativ einzugreifen, um jene Werte zu verändern, die eine Identifizierung ermöglichen.

Riskante Zellwerte, also solche, anhand derer man Einzeldatensätze identifizieren könnte, können entweder in Tabellen unterdrückt werden ("primäre Geheimhaltung") – hier muss es aber auch zu Gegenlöschungen kommen ("sekundäre Geheimhaltung"), da ansonsten die Identifizierbarkeit einzelner statistischer Einheiten und damit auch das Erlangen von einheitenspezifischer Information durch Rückrechnungen möglich ist.

Eine andere Methode der primären Geheimhaltung ist es, solche wie oben erwähnten riskanten Werte zu verschmutzen/verfälschen. Dazu eignet sich besonders die Methode des (Target) Record Swappings, bei der ein Teil der Merkmale einer statistischen Einheit mit denen einer ähnlichen Einheit getauscht werden. Das wird auf einen bestimmten Prozentsatz der Daten angewandt. Ein großer Vorteil dieser Methode ist, dass es keiner sekundären Geheimhaltung bedarf, da Rückrechnungen hier unmöglich sind. Außerdem wird das Verfahren direkt auf die Mikrodaten angewandt, es ist daher nicht nötig, in jede einzelne Auswertungstabelle einzugreifen, und die Ergebnisse sind immer konsistent. Weitere Informationen sind in dem Dokument ["Target Record Swapping"](#) zu finden.

Nachdem es durch derartige Eingriffe zu Informationsverlusten für Datennutzerinnen und -nutzer kommt, gilt es mathematisch gesehen ein Optimierungsproblem zu lösen, damit das Ausmaß der Veränderungen minimiert wird. Seit einigen Jahren gibt es auch in Kooperation von NSI's mit der wissenschaftlichen Welt rege Forschungsaktivitäten, die sich mit der Entwicklung von Geheimhaltungsverfahren und der Implementierung von Software zur Umsetzung dieser Verfahren beschäftigen. In der hausinternen Richtlinie [„Statistische Geheimhal-](#)

[tung in Publikationen und bei Weitergabe von Daten](#)“ finden sich unter anderem jene Verfahren und Programme, die bei Statistik Austria zur Anwendung kommen.

Neben der typischen Publikation und Weitergabe von Daten in aggregierter Form kann es unter bestimmten Voraussetzungen zur Weitergabe von Mikrodaten kommen. Für die sichere Abgabe von Einzeldaten muss zunächst überlegt werden, welche Merkmale (Schlüsselmerkmale) ein eventuell an Datenmissbrauch Interessierter zur Verfügung haben könnte, die ihm die Identifizierung einzelner Personen/Einheiten ermöglichen. Merkmale in diesem Sinne können beispielsweise aus öffentlich abrufbaren Registern stammen oder branchenspezifisches Insiderwissen sein, das ein Angreifer zur Verfügung hat. Wenn die Kombination dieser Merkmalsausprägungen zu Zellen mit geringer Besetzung führt, dann ist das Risiko einer Identifizierung von Einheiten in solchen Zellen groß. Um das Risiko der Identifizierung von Einzeldaten so gering wie möglich zu halten, gibt es zwei Arten von Maßnahmen: Datenreduktion (Stichproben, Vergrößerung von Merkmalsausprägungen, Löschung von Variablen oder Datensätzen) und Datenverschmutzung (Verschmutzung der Daten, „Data/Record Swapping“, Mikroaggregation, Löschung und Imputation). Bei Anwendung aller dieser Maßnahmen muss jedoch immer darauf geachtet werden, dass durch die Datensicherungsmaßnahmen die Datenqualität nicht zu stark leidet.

Richtlinien:

- Bevor Ergebnisse - in welcher Form auch immer - publiziert werden, muss überprüft werden, ob es zu Geheimhaltungsproblemen kommen kann. Dies gilt sowohl für die Publikation und Weitergabe von Daten in aggregierter Form (Häufigkeitstabellen aber auch Indizes), als auch insbesondere für die Abgabe von Mikrodaten.
- Namen und Adressen sollen immer frühestmöglich von den Einzeldatensätzen gelöscht werden. Die sichere Aufbewahrung von Einzeldatensätzen unter Wahrung sämtlicher Grundsätze der Datensicherheitsvorschrift ist für alle Mitarbeiterinnen und Mitarbeiter von Statistik Austria verpflichtend.
- Werden Respondentinnen und Respondenten im Zuge der Datensammlung angesprochen, sollen sie darauf aufmerksam gemacht werden, dass sich Statistik Austria im vollen Umfang zum Datenschutz und zur statistischen Geheimhaltung bekennt, die Daten nur zu statistischen Zwecken verwendet werden und keinesfalls persönliche Daten weitergegeben werden.
- Ein Ziel der statistischen Geheimhaltung ist die Vermeidung der Veröffentlichung von schutzwürdigen Zellen. In den Veröffentlichungen von Statistik Austria wird eine Zelle als schutzwürdig angesehen, wenn eine zu geringe Anzahl an statistischen Einheiten zu ihr beitragen (meist müssen mehr als 2 Einheiten zu einer Zelle beitragen).
- Die Wahl der Methode sowie die technische Umsetzung, die zur Wahrung der statistischen Geheimhaltung herangezogen wird, sind von der statistischen Materie und von der jeweiligen Problemstellung abhängig. Jedenfalls ist bei allen verfahrenstechnischen Fragen, die mit Geheimhaltung verbunden sind, der Bereich Methodik mit einzubeziehen.
- Über diese in diesem Dokument gegebenen Richtlinien hinaus gilt die im Internet publizierte Richtlinie zur „Statistische Geheimhaltung in Publikationen und bei Weitergabe von Daten“ in vollem Umfang.
- Datennutzerinnen und -nutzer sind über die zur Geheimhaltung herangezogenen Verfahren zu informieren. Gegebenenfalls soll auf die zu erwartenden Restriktionen und Informationsverluste hingewiesen werden. Jedenfalls ist in der Standard-Dokumentation unter dem Punkt „Behandlung vertraulicher Daten“ darauf einzugehen.

3.6.8 Dokumentation und Qualitätsberichterstattung

Ein statistisches Produkt liefert Ergebnisse und Daten, mit dem Ziel, diese zu interpretieren und Erkenntnisse zu gewinnen. Um die Anwendung der bei der Datenerstellung verwendeten Methoden für die Nutzerinnen und Nutzer statistischer Produkte nachvollziehbar zu machen, müssen ergänzend ausreichend **Metadaten** zur Verfügung gestellt werden.

In statistischen Datenbanken werden diejenigen Daten als Metadaten bezeichnet, die nicht direkt den Inhalt einer Statistik darstellen, für das Verständnis des statistischen Produkts bzw. für dessen Herstellungsprozess jedoch von essentieller Bedeutung sind (z. B. Branchen- oder Berufsbezeichnungen, Gemeindeverzeichnisse und andere Kataloge). Zu den statistischen Metadaten zählen auch Beschreibungen der Datenfelder in Umfrageformularen, Variablendefinitionen und -abgrenzungen, Erläuterungen sowie unter Umständen auch komplette Formularbeschreibungen. Die eigentlichen statistischen Daten bezeichnet man in Abgrenzung zu den Metadaten als Mikrodaten und Makrodaten. Ohne ausreichende Metadaten entsteht bei Nutzerinnen und Nutzern möglicherweise ein Wissensdefizit, das zu Verständnisproblemen, Fehlinterpretationen und in weiterer Folge zu Fehlentscheidungen führen kann.

Aus diesem Grund kommt der **Dokumentation** bei der Erstellung statistischer Produkte besondere Bedeutung zu. Sie dient nicht nur internen Zwecken, sondern beinhaltet eben jene Metadaten, die für die sinnvolle Nutzung des Produkts notwendig sind. Wie bereits in Kapitel 1 beschrieben, ist der Qualitätsbegriff für die Amtliche Statistik auf ein solides definitorisches Fundament gestellt worden. Eine damit verbundene Frage ist naturgemäß, in wie weit die Qualität eines statistischen Produkts den gegebenen Qualitätskriterien entspricht. Reine quantitative Maße sind selten verfügbar und somit sind **Qualitätsberichte** in textueller Form ein geeignetes Instrument, um die Güte eines statistischen Produktes anhand der durch die Definition gegebenen Dimensionen zu beschreiben.

Dokumentation, Bereitstellung von Metainformation und Qualitätsberichterstattung stehen in einem unmittelbaren Zusammenhang. Daher ergibt sich für die Produzentinnen und Produzenten von Statistiken ein synergetischer Effekt, der es erlaubt, Dokumentationsarbeit und Qualitätsberichterstattung miteinander zu verbinden. In diesem Sinne gibt es bei Statistik Austria ein standardisiertes System der Beschreibung statistischer Produkte. Dieses Berichtschema der [Standard-Dokumentationen](#) stellt den Benutzerinnen und Benutzern in standardisierter Form Informationen über die zugrunde liegenden Konzepte, Definitionen und Erläuterungen samt Angaben zu den verwendeten Methoden und zur Qualität der Statistiken zur Verfügung (Metainformationen).

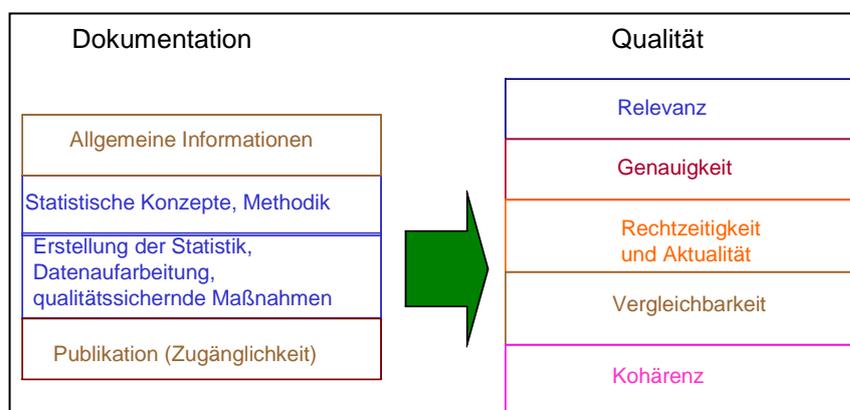


Abbildung 15: Beschreibende Inhalte der Dokumentation werden zur Qualitätsbewertung herangezogen

Wie in *Abbildung 15* zu sehen, ist es die Intention, anhand zuvor gegebener produktrelevanter Beschreibungen, Qualitätsbewertungen für die definierten Qualitätsdimensionen zu liefern.

Alle Standard-Dokumentationen von Statistik Austria sind nach einem fixen Schema strukturiert, dessen Hauptüberschriften in *Abbildung 16* dargestellt sind.

Executive Summary
1. Allgemeine Informationen
2. Konzeption und Erstellung
2.1 Statistische Konzepte, Methodik
2.2 Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen
2.3 Publikation (Zugänglichkeit)
3. Qualität
4. Ausblick
Glossar
Abkürzungsverzeichnis

Abbildung 16: Struktur der Standard Dokumentationen

Die Kapitel 1 bis 3 sind ihrerseits relativ stark untergliedert und es kann der Fall sein, dass einzelne Subkapitel – vor allem im Kapitel 2 – für manche Produkte nicht von Relevanz sind, wohingegen andere Unterüberschriften benötigt werden können. Hierbei haben die Autorinnen und Autoren der Dokumentationen durchaus Gestaltungsspielraum.

Mit dem Konzept der Standard-Dokumentationen verfolgt Statistik Austria grundsätzlich zwei Ziele. Einerseits soll einer eiligen Leserin bzw. einem eiligen Leser durch die vorangestellten zusammenfassenden Teile ein rascher Überblick über den Zweck, Inhalt und die wichtigsten Konzepte geliefert werden. Andererseits soll intensiven Nutzerinnen und Nutzern, die an detaillierteren Informationen interessiert sind, die Möglichkeit gegeben werden, ihr Wissen über das Produkt durch die in den Kapiteln 2 und 3 vorhandenen Informationen zu vertiefen.

In Kooperation mit dem Statistikrat-Ausschuss „Qualitätssicherung“ werden im Rahmen des Qualitätsmanagements bei Statistik Austria laufend „Feedback-Gespräche“ zur Qualität der verschiedenen statistischen Produkte auf Basis von „Standard-Dokumentationen“ durchgeführt. Inhalt und Ziele der „Feedback-Gespräche“ sind die kritische Auseinandersetzung mit den Qualitätsaspekten der Statistiken unter besonderer Berücksichtigung der verwendeten Methoden und Verfahren, die Identifikation von Qualitätsverbesserungspotentialen und Erarbeitung von Empfehlungen für Verbesserungsmaßnahmen sowie die Verbesserung der Standard-Dokumentation der jeweils besprochenen Statistiken, wobei insbesondere die Sicht von Nutzerinnen und Nutzern und externer Expertinnen und Experten einfließen soll.

Neben Standard Dokumentationen können auch andere Berichtspflichten für statistische Produkte existieren. Als hauptfall ist hier die Qualitätsberichterstattung an Eurostat zu nennen (dazu siehe Kapitel 3.6.3). Weitere Berichte, die zumeist in Kooperation und auf Wunsch des Auftraggebers erstellt werden, enthalten oftmals sehr umfangreiche methodische Beschreibungen. In solchen Fällen spricht man von einem Methodenbericht, der in Inhalt und Substanz zumeist weit über eine Standard-Dokumentation hinausgeht.

Richtlinien:

- Zu jedem statistischen Produkt, das von Statistik Austria erstellt wird, muss eine Standard-Dokumentation erstellt werden. Standard-Dokumentationen werden im Internet unter dem Zweig „Dokumentationen“ veröffentlicht, der nach den Fachgebieten der Statistiken strukturiert ist.
- Standard-Dokumentationen müssen nach der den Mitarbeiterinnen und Mitarbeitern im Intranet zur Verfügung stehenden, vorgegebenen Gliederung strukturiert sein. Sämtliche Kapitelüberschriften auf 2-Steller Ebene sind verpflichtend einzuhalten. Für darunterliegende Gliederungsebenen sollen nur jene Punkte Eingang finden, die für das Produkt von Relevanz sind. Auch können ab der Ebene von 3-Stellern neue Überschriften einge-zogen werden, falls dies zweckmäßig erscheint.
- Jede Standard-Dokumentation muss ein Executive Summary beinhalten, dessen textuelle Länge den Umfang von 1,5 Seiten möglichst nicht übersteigt. Das Executive Summary soll die wichtigsten inhaltlichen Charakteristika der Statistik zusammenfassen und bedeutsame Hinweise zur Datennutzung beinhalten. Im Anschluss an den Text des Executive Summary ist eine Tabelle mit den wichtigsten Eckpunkten der behandelten Statistik anzuführen. Das Executive Summary soll eine grafische Darstellung zur Veranschaulichung der erstellten Statistik bzw. zu deren Einbettung in ein Gesamtsystem beinhalten.
- Bei der Erstellung der Standard-Dokumentation soll immer eine möglichst klare Ausdruckweise gewählt werden. Wiederholungen von bereits einmal verwendeten Textpas-sagen sollen vermieden werden. Fachbegriffe sollen bei ihrem ersten Auftreten erklärt werden und Abkürzungen bei erstmaliger Verwendung ausgeschriebenen werden. Überdies sollen Fachbegriffe und Abkürzungen in einem ergänzenden Glossar bzw. Abkürzungsverzeichnis erklärt werden. Tabellen und Abbildungen müssen beschriftet und ausreichend erklärt werden. Wird im Text auf andere externe Dokumente hingewiesen, die im Internet verfügbar sind, soll mit einem Hyperlink darauf verwiesen werden. Die in der Standard-Dokumentation vorhandenen Hyperlinks sollen vierteljährlich überprüft und gegebenenfalls ersetzt werden.
- Für Produkte mit mehrjähriger Periodizität ist für jeden Berichtszeitraum/Stichtag eine eigene Standard-Dokumentation zu erstellen. Bei jährlicher oder unterjähriger Periodizität ist die vorhandene Standard-Dokumentation einmal pro Jahr einer Überarbeitung zu unterziehen. Ändern sich jedoch bei einem derartigen Produkt fundamentale Konzepte, muss auch dort eine neue Standard-Dokumentation erstellt werden.
- Verantwortlich für die Erstellung einer Standard-Dokumentation ist jeweils die verantwortliche Projektleitung. Bevor eine Standard-Dokumentation im Internet publiziert wird, muss sie einer redaktionellen Überarbeitung unterzogen werden. Jedenfalls muss sie durch den zuständigen Direktor/Abteilungsleiter bzw. die zuständige Direktorin/Abteilungsleiterin freigegeben werden. Im Anschluss ist sie der Stabsstelle Qualitätsmanagement und Methodik zu übermitteln, wo sie inhaltlich und formal überprüft wird. Erst nach der Freigabe durch das Qualitätsmanagement wird sie ins Internet gestellt.

- Jene Verbesserungsvorschläge, die sich als Ergebnis des Feedback-Gesprächs ergeben, müssen in die Standard-Dokumentation eingearbeitet werden, wobei insbesondere die Sicht von Nutzerinnen und Nutzern und externen Expertinnen und Experten Beachtung finden soll.
- Mitarbeiterinnen und Mitarbeiter von Statistik Austria, die in Arbeitsgruppen an der Entstehung von Verordnungen über Qualitätsberichterstattung mitwirken, sollen proaktiv darauf einwirken, dass die Qualitätsberichte aussagekräftig, sinnvoll interpretierbar und über die Zeit stabil sind.
- Projektleiterinnen und Projektleiter bei Statistik Austria müssen dafür sorgen, dass auch die interne Dokumentation sämtlicher Arbeiten erfolgt. Die Nachvollziehbarkeit der wesentlichen Arbeitsschritte und deren Reproduzierbarkeit müssen gegeben sein.

4. Maßnahmen zur Qualitätssicherung

Die im Kapitel 3 erläuterten Prinzipien und Standards geben einen Einblick in die Komplexität der Erstellung Amtlicher Statistiken. Damit aus niedergeschriebenen Prinzipien gelebte Praxis entsteht, bedarf es entsprechender Maßnahmen, die die Einhaltung geforderter Standards gewährleisten. Derartige Maßnahmen, die sich unter dem Oberbegriff Qualitätssicherung zusammenfassen lassen, müssen im Falle der Amtlichen Statistik alle Dimensionen, die Bestandteil der Qualitätsdefinition (siehe Kapitel 1) sind, betreffen. Somit bedarf es nicht nur interner Maßnahmen, die natürlich eine zentrale Rolle spielen, sondern auch die Einbeziehung der Nutzerinnen und Nutzer ist ein entscheidender Faktor zur laufenden Überprüfung der Relevanz, Zugänglichkeit und Klarheit sowie der Verwendbarkeit des zur Verfügung stehenden Datenmaterials.

4.1 Interne Maßnahmen

Naturgemäß unterliegen sämtliche Bearbeitungsschritte bei der Erstellung der statistischen Produkte einer **laufenden Kontrolle**. Entscheidend dabei ist, dass die Überprüfung des aus einem Arbeitsschritt resultierenden Ergebnisses mit größtmöglicher Sorgfalt vorgenommen wird. Inhaltlich soll die Kontrolle sich an den in der Planungsphase erarbeiteten Konzepten orientieren. Das bedeutet, dass mit der Durchführung eines Arbeitsschrittes eine gewisse Erwartungshaltung hinsichtlich des zu erzielenden Ergebnisses verbunden ist und dann Handlungsbedarf besteht, wenn das realisierte Ergebnis deutlich von dem zu erwartenden abweicht.

Die zuvor angesprochene Kontrolle des Outputs verschiedener Produktionsschritte ist auch dadurch gegeben, dass für etliche Arbeitsvorgänge die **Einbeziehung spezieller Organisationseinheiten** zumindest dringend angeraten wird. Dies betrifft die IT-Abteilung ebenso wie die in der Abteilung Register, Klassifikationen und Methodik befindlichen Einheiten, die die Fachdirektionen im Zuge der Erstellung von Statistiken bei Arbeitsschritten, die ein bereichsübergreifendes Know-how erfordern, unterstützen.

Die Durchführung der qualitätssichernden Maßnahmen erfolgt in den jeweiligen Fachbereichen. Unterstützt werden die Fachdirektionen dabei durch die Stabsstelle Qualitätsmanagement, deren Aufgaben unter anderem auch die Wahrnehmung zentraler Aufgaben wie die Wartung der Standard-Dokumentationen und des Veröffentlichungskalenders vorsehen. Ein wichtiger Partner bei der Überwachung der Einhaltung der Qualitätsgrundsätze, die im BStatG 2000 niedergeschrieben sind, ist der **Statistikrat** und hier vor allem der **Qualitätsausschuss**. Der Qualitätsausschuss ist nicht nur bei jedem Feedback-Gespräch vertreten, sondern spielt auch eine proaktive Rolle bei wichtigen Qualitätsfragen, die in regelmäßig stattfindenden gemeinsamen Sitzungen diskutiert werden.

Die Komplexität, die bei der Produktion Amtlicher Statistiken vorherrscht, erfordert eine umfassende **Dokumentation** aller Arbeitsschritte. Diese Notwendigkeit ist einerseits wegen der Nachvollzieh- und Reproduzierbarkeit der angewandten Methoden gegeben. Andererseits dient eine ausführliche Dokumentation auch als Qualitätsnachweis dafür, dass bestimmte Arbeitsschritte, deren Durchführung als erforderlich („Lege Artis“) angesehen werden, erfolgt sind. Auch was die Fundiertheit der eingesetzten Methoden (siehe auch Kapitel 2) angeht, dient die Dokumentation als entsprechender Nachweis.

Die Überprüfung der Einhaltung der Qualitätsrichtlinien, die dieses Dokument vorgibt, ist auch von der internen Sicht her von entscheidender Bedeutung. Seit 2016 werden aus diesem Grund bei Statistik Austria **interne Qualitätsaudits** durchgeführt. Dabei wird basierend auf einer Selbstbewertung eine Einschätzung der Einhaltung der Qualitätsrichtlinien für die

dem Audit unterzogenen Statistiken erarbeitet. Durchgeführt werden die Audits hauptverantwortlich durch die Stabsstelle Qualitätsmanagement und Methodik unter Beteiligung des Risikomanagements.

Neben den laufenden Kontrollen und den bisher besprochenen Maßnahmen zur Qualitätssicherung gibt es bei der Produktion von Statistik auch die Idee der **Ex-Post Qualitätssicherung**. Diese Maßnahmen sind zumeist sehr projektspezifisch und deren Durchführung hängt auch von den jeweiligen Ressourcen ab. Beispiele für Ex-Post Maßnahmen sind unter anderem:

- Ex-post Kontrolle durch Nacherhebungen
- Ex-post Kontrolle der Datenerfassung, Signierung und der Mikro- und Makro-Plauschkorrekturen
- Analyse von Antwortausfällen
- Zusammenführung mit anderen Beständen (Anreicherung um zus. Merkmale) zum Zwecke der ex-post Qualitätssicherung
- Spezielle Qualitätssicherungsprojekte

Neben den detaillierten Dokumentationen, die internen Zwecken dienen, ergibt sich auch die Notwendigkeit, Nutzerinnen und Nutzern die Verwendbarkeit der statistischen Ergebnisse durch Bereitstellung geeigneter Dokumentationen zu erleichtern. Mit dem **System der Standard-Dokumentationen** stellt Statistik Austria ihren Kundinnen und Kunden in standardisierter Form Informationen über die zugrunde liegenden Konzepte, Definitionen und Erläuterungen samt Angaben zu den verwendeten Methoden und zur Qualität der Statistiken zur Verfügung (Metainformationen). Durch die Metainformationen werden die Nutzerinnen und Nutzer in die Lage versetzt, die angebotenen statistischen Informationen richtig zu interpretieren, deren Informationsgehalt auszuschöpfen und ihre Qualität einzuschätzen. Die Erstellung einer Standard-Dokumentation sollte nach Möglichkeit für jedes Produkt von Statistik Austria erfolgen.

4.2 Einbindung von Nutzerinnen und Nutzern

Wie bereits einleitend erwähnt, erwarten Nutzerinnen und Nutzer der Ergebnisse von Statistik Austria eine hohe Qualität des zur Verfügung stehenden Datenmaterials. Deshalb ist es wichtig, dass sowohl im Zuge des Planungsprozesses als auch bei der Qualitätsbewertung wichtige Nutzergruppen integriert sind. Daher ist das System der **Fachbeiräte**, die einen wichtigen Baustein bei der Planung neuer und Erweiterung bzw. Modifikation bestehender Statistiken bilden, eine wichtige Plattform, wo Schlüsselnutzer („Power User“) sich einbringen können. Die Diskussion in den Fachbeiräten ist auch insofern wichtig, als dort ein Kompromiss zwischen Wünschen der Nutzerinnen und Nutzer einerseits und dem, was durch Statistik Austria machbar erscheint, erzielt werden kann.

Die Meinung der Nutzerinnen und Nutzer hinsichtlich der Qualität der erstellten Produkte ist für Statistik Austria natürlich auch nach Vorliegen des Endproduktes von Wichtigkeit. Um diesem Anliegen permanent Rechnung zu tragen, werden seit 2003 in Kooperation mit dem Statistiker-Ausschuss „Qualitätssicherung“ im Rahmen des Qualitätsmanagements bei Statistik Austria laufend **Feedback-Gespräche** zur Qualität der verschiedenen statistischen Produkte auf Basis der bereits erwähnten Standard-Dokumentationen durchgeführt. Inhalt und Ziele der Feedback-Gespräche sind die kritische Auseinandersetzung mit den Qualitätsaspekten der Statistiken unter besonderer Berücksichtigung der verwendeten Methoden und Verfahren, die Identifikation von Qualitätsverbesserungspotentialen und Erarbeitung von Empfehlungen für Verbesserungsmaßnahmen sowie die Verbesserung der Standard-Dokumentation der jeweils besprochenen Statistiken, wobei insbesondere die Sicht von Nutzerinnen und Nutzern und externen Expertinnen und Experten einfließen soll.

Neben den Feedback-Gesprächen führt Statistik Austria alle zwei Jahre eine **Nutzerinnen- und Nutzerbefragung** durch. Der Verhaltenskodex für Europäische Statistiken benennt im Grundsatz 11 - Relevanz explizit Nutzerbefragungen als ein maßgebliches Instrument zur Überwachung dieser Qualitätsdimension. Die Ergebnisse dieser Befragung werden jeweils intern diskutiert und im Internet publiziert.

Aktualität und Rechtzeitigkeit sind immer bedeutender werdende Qualitätsdimensionen. Das Bekenntnis von Statistik Austria zu hoher Aktualität und Termintreue spiegelt sich im Bestreben wider, Projektpläne nach Möglichkeit so zu gestalten, dass Veröffentlichungstermine nahe bei der betrachteten Referenzperiode (bzw. Stichtag) liegen. Der im Internet verfügbare **Veröffentlichungskalender** von Statistik Austria beinhaltet eine umfassende Vorschau auf geplante Veröffentlichungstermine von Hauptergebnissen auf der Webseite von Statistik Austria. Alle Projektleiterinnen und Projektleiter sind verpflichtet, Termine zu nennen, die dann verbindlich für die Publikation der jeweiligen Ergebnisse sind und im Veröffentlichungskalender eingetragen werden. Die Überprüfung der Einhaltung der Termine und die Wartung des Kalenders werden zentral durch die Stabsstelle Qualitätsmanagement erledigt.

Anhang 1: Europäische Statistiken: Verhaltenskodex

Präambel

Vision des Europäischen Statistischen Systems

„Das Europäische Statistische System wird im Bereich der statistischen Informationsdienste weltweit führend sein und zum wichtigsten Informationslieferanten für die Europäische Union und ihre Mitgliedstaaten werden. Gestützt auf wissenschaftliche Grundsätze und Methoden wird das Europäische Statistische System ein Programm harmonisierter europäischer Statistiken anbieten und laufend verbessern. Dies bildet eine wesentliche Grundlage für demokratische Prozesse und Fortschritte in der Gesellschaft.“

Der Auftrag des Europäischen Statistischen Systems

„Wir stellen der Europäischen Union, der Welt und der Öffentlichkeit unabhängige und qualitativ hochwertige Informationen über Wirtschaft und Gesellschaft auf europäischer, nationaler und regionaler Ebene zur Verfügung und machen diese Informationen für die Zwecke der Entscheidungsfindung, Forschung und Diskussion jedermann zugänglich.“

Um diesen Auftrag zu erfüllen und die Vision umzusetzen, streben die Mitglieder des Europäischen Statistischen Systems nach Zusammenarbeit und kontinuierlichem Austausch mit den Nutzerinnen und Nutzern. Dabei wird den Grundsätzen des Verhaltenskodex für europäische Statistiken und den allgemeinen Qualitätsmanagementgrundsätzen, einschließlich Führungsverpflichtung, Partnerschaft, Mitarbeiterzufriedenheit und kontinuierlichen Verbesserungen, sowie der Integration und Harmonisierung Rechnung getragen.

Verhaltenskodex für europäische Statistiken

Der Verhaltenskodex für europäische Statistiken basiert auf 15 Grundsätzen für den institutionellen Rahmen, die statistischen Prozesse und die statistischen Produkte. Für jeden Grundsatz bietet ein Satz von Indikatoren vorbildlicher Praktiken eine Referenz für die Überprüfung der Umsetzung des Kodex. Die Qualitätskriterien für europäische Statistiken sind im europäischen Statistikrecht festgelegt. Statistische Stellen, darunter die Kommission (Eurostat), nationale statistische Ämter und andere einzelstaatliche Stellen, die für die Entwicklung, Erstellung und Verbreitung europäischer Statistiken verantwortlich sind, sowie Regierungen, Ministerien und der Europäische Rat verpflichten sich zur Einhaltung des Kodex. Die Grundsätze des Verhaltenskodex bilden zusammen mit den allgemeinen Qualitätsmanagementgrundsätzen einen gemeinsamen Qualitätsrahmen für das Europäische Statistische System.

Institutioneller Rahmen

Institutionelle und organisatorische Faktoren wirken sich maßgeblich auf den Wirkungsgrad und die Glaubwürdigkeit einer Stelle aus, die europäische Statistiken entwickelt, erstellt und verbreitet. Die relevanten Aspekte in diesem Zusammenhang sind fachliche Unabhängigkeit, das Mandat zur Datenerhebung, angemessene Ressourcen, die Verpflichtung zur Qualität, statistische Geheimhaltung, Unparteilichkeit und Objektivität.

GRUNDSATZ 1: FACHLICHE UNABHÄNGIGKEIT

Die fachliche Unabhängigkeit der statistischen Stellen gegenüber anderen politischen, Regulierungs- oder Verwaltungsstellen sowie gegenüber den Akteuren des Privatsektors ist der Garant für die Glaubwürdigkeit der europäischen Statistiken.

Indikatoren

INDIKATOREN

1.1: Die Unabhängigkeit der nationalen statistischen Ämter und von Eurostat von politischer und anderer externer Einflussnahme bei der Entwicklung, Erstellung und Verbreitung von Statistiken ist gesetzlich festgelegt und für andere statistische Stellen gewährleistet.

1.2: Die Leiterinnen und Leiter der nationalen statistischen Ämter und von Eurostat und gegebenenfalls die Leiterinnen und Leiter anderer statistischer Stellen sind auf einer hierarchischen Ebene angesiedelt, die den Zugang zu hochrangigen politischen Entscheidungsträgern und Verwaltungsstellen gewährleistet. Die Leiterinnen und Leiter verfügen über die höchstmöglichen fachlichen Qualifikationen.

1.3: Die Leiterinnen und Leiter der nationalen statistischen Ämter und von Eurostat und gegebenenfalls die Leiterinnen und Leiter anderer statistischer Stellen sind dafür verantwortlich, dass die Entwicklung, Erstellung und Verbreitung der Statistiken in unabhängiger Weise erfolgt.

1.4: Die Leiterinnen und Leiter der nationalen statistischen Ämter und von Eurostat und gegebenenfalls die Leiterinnen und Leiter anderer statistischer Stellen tragen die alleinige Verantwortung für die Festlegung der statistischen Methoden, Standards und Verfahren sowie des Inhalts und des Zeitplans der statistischen Veröffentlichungen.

1.5: Die statistischen Arbeitsprogramme werden veröffentlicht, und über den Stand der Arbeiten wird regelmäßig Bericht erstattet.

1.6: Statistische Veröffentlichungen sind klar als solche erkennbar und werden getrennt von politischen bzw. Grundsatzserklärungen veröffentlicht.

1.7: Soweit angebracht, nehmen die nationalen statistischen Ämter und Eurostat und gegebenenfalls andere statistische Stellen öffentlich Stellung zu statistischen Fragen, auch zu Kritik an amtlichen Statistiken und zu deren Missbrauch.

1.8: Die Ernennung der Leiterinnen und Leiter der nationalen statistischen Ämter und von Eurostat und gegebenenfalls anderer statistischer Stellen beruht allein auf deren fachlicher Eignung. Die Gründe für die Beendigung der Amtszeit sind gesetzlich festgelegt. Darunter fallen nicht solche Gründe, die die fachliche oder wissenschaftliche Unabhängigkeit beeinträchtigen könnten.

GRUNDSATZ 2: MANDAT ZUR DATENERHEBUNG

Die statistischen Stellen haben ein eindeutiges gesetzliches Mandat zur Erhebung von Angaben für die Zwecke europäischer Statistiken. Verwaltungen, Unternehmen und private Haushalte sowie die Öffentlichkeit im weiteren Sinne können gesetzlich dazu verpflichtet werden, auf Anforderung statistischer Stellen für die Zwecke europäischer Statistiken, den Zugriff auf Daten zu gewähren oder Daten zu liefern.

INDIKATOREN

2.1: Das Mandat der statistischen Stellen zur Erhebung von Angaben für die Entwicklung, Erstellung und Verbreitung europäischer Statistiken ist gesetzlich festgelegt.

2.2: Die statistischen Stellen sind gesetzlich dazu befugt, Verwaltungsdaten zu statistischen Zwecken zu verwenden.

2.3: Die statistischen Stellen können die Beantwortung bei statistischen Erhebungen auf der Basis einer Rechtsgrundlage verbindlich vorschreiben.

GRUNDSATZ 3: ANGEMESSENE RESSOURCEN

Die den statistischen Stellen zur Verfügung stehenden Ressourcen reichen aus, um den aktuellen statistischen Erfordernissen Europas zu entsprechen.

INDIKATOREN

3.1: Es sind qualitativ angemessene und ausreichende Personal-, Finanz- und Datenverarbeitungs-Ressourcen vorhanden, um dem aktuellen statistischen Bedarf zu entsprechen.

3.2: Umfang, Gliederungstiefe und Kosten der Statistiken entsprechen dem Bedarf.

3.3: Es gibt Verfahren, mit denen Forderungen nach neuen Statistiken gegenüber den Kosten dieser Statistiken abgewogen und gerechtfertigt werden können.

3.4: Es gibt Verfahren, mit denen beurteilt werden kann, ob sämtliche Statistiken weiterhin benötigt werden oder, ob die Erstellung eines Teils von ihnen eingestellt oder eingeschränkt werden kann, um Ressourcen freizusetzen.

GRUNDSATZ 4: VERPFLICHTUNG ZUR QUALITÄT

Die statistischen Stellen sind zur Qualität verpflichtet. Sie ermitteln systematisch und regelmäßig Stärken und Schwächen mit dem Ziel der kontinuierlichen Verbesserung der Prozess- und Produktqualität.

INDIKATOREN

4.1: Die Qualitätspolitik ist festgelegt und öffentlich zugänglich. Für das Qualitätsmanagement sind Organisationsstruktur und -instrumente vorhanden.

4.2: Verfahren zur Planung und Überwachung der Qualität des statistischen Produktionsprozesses sind vorhanden.

4.3: Die Produktqualität wird regelmäßig überwacht und im Hinblick auf mögliche Zielkonflikte beurteilt. Die Qualitätsberichterstattung erfolgt gemäß den Qualitätskriterien für europäische Statistiken.

4.4: Die wichtigsten statistischen Produkte werden, falls angemessen, auch unter Hinzuziehung externer Sachverständiger, regelmäßig gründlich überprüft.

GRUNDSATZ 5: STATISTISCHE GEHEIMHALTUNG

Die Anonymität der Datenlieferanten (private Haushalte, Unternehmen, Verwaltungen und andere Auskunftgebende (Respondenten)), die Geheimhaltung ihrer Angaben und deren ausschließliche Verwendung für statistische Zwecke sind unter allen Umständen gewährleistet.

INDIKATOREN

5.1: Die statistische Geheimhaltung ist gesetzlich vorgeschrieben.

5.2: Die Mitarbeiterinnen und Mitarbeiter unterzeichnen bei ihrer Einstellung rechtlich verbindliche Geheimhaltungsverpflichtungen.

5.3: Die vorsätzliche Verletzung des Statistikgeheimnisses wird geahndet.

5.4: Die Mitarbeiterinnen und Mitarbeiter erhalten Leitlinien und Anweisungen für die Wahrung des Statistikgeheimnisses bei der Erstellung und Verbreitung von Statistiken. Die Geheimhaltungspolitik wird der Öffentlichkeit kommuniziert.

5.5: Physische, technische und organisatorische Vorkehrungen zum Schutz der Sicherheit und Integrität statistischer Datenbanken sind getroffen.

5.6: Für externe Nutzerinnen und Nutzer, die auf statistische Mikrodaten zu Forschungszwecken zugreifen möchten, gelten strenge Vorschriften.

GRUNDSATZ 6: UNPARTEILICHKEIT UND OBJEKTIVITÄT

Die statistischen Stellen entwickeln, erstellen und verbreiten europäische Statistiken unter Wahrung der wissenschaftlichen Unabhängigkeit und in objektiver, professioneller und transparenter Weise, wobei alle Nutzerinnen und Nutzer gleich zu behandeln sind.

INDIKATOREN

- 6.1:** Die Statistiken werden auf einer von statistischen Überlegungen getragenen objektiven Grundlage erstellt.
- 6.2:** Die Wahl der Quellen und der statistischen Methoden sowie alle Entscheidungen bezüglich der Verbreitung von Statistiken erfolgt aufgrund von statistischen Überlegungen.
- 6.3:** Werden in veröffentlichten Statistiken Fehler festgestellt, so werden sie zum frühest möglichen Zeitpunkt berichtigt, und die Öffentlichkeit wird davon in Kenntnis gesetzt.
- 6.4:** Informationen zu den verwendeten Methoden und Verfahren sind öffentlich zugänglich.
- 6.5:** Das Datum und der Zeitpunkt, zu dem statistische Daten veröffentlicht werden, werden vorab mitgeteilt.
- 6.6:** Alle größeren Revisionen und Änderungen der Methoden werden vorab angekündigt.
- 6.7:** Alle Nutzerinnen und Nutzer haben gleichzeitigen und gleichberechtigten Zugang zu statistischen Daten. Jeglicher bevorzugte Vorabzugang externer Nutzerinnen und Nutzer ist beschränkt, kontrolliert und wird öffentlich bekanntgegeben. Falls Daten unberechtigterweise an die Öffentlichkeit gelangen, werden die Modalitäten der Vorabfreigabe so überarbeitet, dass die Unparteilichkeit gewährleistet ist.
- 6.8:** Die Veröffentlichung statistischer Ergebnisse und entsprechende Erklärungen auf Pressekongressen erfolgen objektiv und unparteilich.

Statistische Prozesse

Bei der Organisation, Erhebung, Verarbeitung und Verbreitung europäischer Statistiken werden internationale Standards, Leitlinien und vorbildliche Praktiken von den statistischen Stellen uneingeschränkt eingehalten. Wenn die statistischen Stellen für ihr solides Management und ihre Effizienz bekannt sind, kommt dies der Glaubwürdigkeit der Statistiken zugute. Die relevanten Aspekte in diesem Zusammenhang sind eine solide Methodik, geeignete statistische Verfahren, die Vermeidung einer übermäßigen Belastung der Auskunftgebenden (Respondenten) und Wirtschaftlichkeit.

GRUNDSATZ 7: SOLIDE METHODIK

Qualitativ hochwertige Statistiken basieren auf einer soliden Methodik. Diese erfordert geeignete Instrumente und Verfahren sowie ein entsprechendes Know-how.

INDIKATOREN

- 7.1:** Der für europäische Statistiken verwendete allgemeine methodische Rahmen trägt europäischen und anderen internationalen Standards, Leitlinien und vorbildlichen Praktiken Rechnung.
- 7.2:** Es gibt Verfahren, die gewährleisten, dass Standardkonzepte, -definitionen und -klassifikationen in der gesamten statistischen Stelle einheitlich verwendet werden.
- 7.3:** Um eine hohe Qualität zu gewährleisten, werden das Unternehmensregister und die Erhebungsgrundlagen für Bevölkerungserhebungen regelmäßig evaluiert und erforderlichenfalls angepasst.
- 7.4:** Zwischen den nationalen und den europäischen Klassifikationssystemen besteht eine enge Übereinstimmung.
- 7.5:** Es werden Absolventen der einschlägigen Studiengänge eingestellt.

7.6: Die statistischen Stellen verfolgen eine Politik der kontinuierlichen beruflichen Weiterbildung ihrer Mitarbeiterinnen und Mitarbeiter.

7.7: Zur Verbesserung der Methodik sowie der Wirksamkeit angewandter Methoden und, sofern möglich, zur Förderung besserer Instrumente werden Maßnahmen in Zusammenarbeit mit der Wissenschaft durchgeführt.

GRUNDSATZ 8: GEEIGNETE STATISTISCHE VERFAHREN

Geeignete statistische Verfahren – von der Erhebung bis zur Validierung der Daten – bilden die Grundlage für qualitativ hochwertige Statistiken.

INDIKATOREN

8.1: Falls europäische Statistiken auf Verwaltungsdaten basieren, werden die für administrative Zwecke verwendeten Definitionen und Konzepte den Erfordernissen der Statistik soweit wie möglich angepasst.

8.2: Die Fragebogen für statistische Erhebungen werden vor der Erhebung der Daten systematisch getestet.

8.3: Die Erhebungspläne sowie die Stichprobenziehung und Schätzverfahren basieren auf soliden Grundlagen und werden regelmäßig überprüft und erforderlichenfalls überarbeitet.

8.4: Die Datengewinnung sowie die Eingabe und Kodierung der Daten werden regelmäßig kontrolliert und erforderlichenfalls angepasst.

8.5: Für das Editieren und Imputationen werden geeignete Verfahren eingesetzt, die regelmäßig überprüft und erforderlichenfalls überarbeitet oder aktualisiert werden.

8.6: Revisionen erfolgen nach standardisierten, bewährten und transparenten Verfahren.

8.7: Die statistischen Stellen sind an der Gestaltung von Verwaltungsdaten beteiligt, um deren Eignung für statistische Zwecke zu erhöhen.

8.8: Es werden Vereinbarungen mit den Eignern von Verwaltungsdaten getroffen, in denen die gemeinsame Verpflichtung zur Nutzung dieser Daten für statistische Zwecke bekräftigt wird.

8.9: Die statistischen Stellen arbeiten mit den Eignern von Verwaltungsdaten zusammen, um die Datenqualität zu gewährleisten.

GRUNDSATZ 9: VERMEIDUNG EINER ÜBERMÄSSIGEN BELASTUNG DER AUSKUNFTGEBENDEN (RESPONDENTEN)

Der Beantwortungsaufwand steht in einem angemessenen Verhältnis zum Bedarf der Nutzerinnen und Nutzer und ist für die Auskunftgebenden (Respondenten) nicht übermäßig hoch. Die statistischen Stellen überwachen den Beantwortungsaufwand und legen Ziele für dessen schrittweise Verringerung fest.

INDIKATOREN

9.1: Der Bedarf an Angaben für europäische Statistiken wird in Bezug auf Umfang und Gliederungstiefe auf das absolut erforderliche Maß begrenzt.

9.2: Der Beantwortungsaufwand wird so gleichmäßig wie möglich auf die Erhebungspopulationen verteilt.

9.3: Die von den Unternehmen verlangten Angaben werden soweit möglich direkt aus deren Buchhaltung entnommen, und im Interesse der leichteren Übermittlung dieser Angaben werden möglichst elektronische Hilfsmittel eingesetzt.

9.4: Administrative Datenquellen werden – wann immer möglich – herangezogen, um doppelte Datenanforderungen zu vermeiden.

9.5: Innerhalb der statistischen Stellen erfolgt generell eine gemeinsame Datennutzung, um eine Vervielfachung der Erhebungen zu vermeiden.

9.6: Die statistischen Stellen fördern Maßnahmen, die die Verknüpfung von Datenquellen ermöglichen, um den Beantwortungsaufwand zu reduzieren.

GRUNDSATZ 10: WIRTSCHAFTLICHKEIT

Die Ressourcen werden effektiv eingesetzt

INDIKATOREN

10.1: Durch interne und unabhängige externe Maßnahmen wird der Ressourceneinsatz der statistischen Stelle überwacht.

10.2: Das Produktivitätspotenzial der Informations- und Kommunikationstechnologie wird bei der Datenerhebung, -verarbeitung und -verbreitung soweit als möglich ausgeschöpft.

10.3: Zur Vergrößerung des statistischen Potenzials von Verwaltungsdaten und zur Begrenzung des Zurückgreifens auf direkte Erhebungen werden proaktive Anstrengungen unternommen.

10.4: Zur Steigerung der Effizienz und Wirksamkeit fördern und realisieren die statistischen Stellen standardisierte Lösungen.

Statistische Produkte

Die angebotenen Statistiken entsprechen dem Nutzerbedarf. Die Statistiken stehen in Einklang mit europäischen Qualitätsstandards und decken den Bedarf der europäischen Institutionen, Regierungen, Forschungseinrichtungen und Unternehmen sowie der Öffentlichkeit im Allgemeinen. Wichtige Aspekte in diesem Zusammenhang sind Relevanz, Genauigkeit und Zuverlässigkeit, Aktualität, Kohärenz, Vergleichbarkeit zwischen Regionen und Ländern sowie leichte Zugänglichkeit für die Nutzerinnen und Nutzer.

GRUNDSATZ 11: RELEVANZ

Die europäischen Statistiken entsprechen dem Bedarf der Nutzerinnen und Nutzer.

INDIKATOREN

11.1: Es gibt Verfahren zur Konsultation der Nutzerinnen und Nutzer, zur Überwachung der Relevanz bestehender Statistiken und des Ausmaßes, in dem sie den Bedarf der Nutzerinnen und Nutzer tatsächlich decken sowie zur Einbeziehung des neu entstehenden Bedarfs und der neu entstehenden Prioritäten der Nutzerinnen und Nutzer.

11.2: Prioritäre Anforderungen werden erfüllt und im Arbeitsprogramm abgebildet.

11.3: Die Zufriedenheit der Nutzerinnen und Nutzer wird regelmäßig überprüft und systematisch verfolgt.

GRUNDSATZ 12: GENAUIGKEIT UND ZUVERLÄSSIGKEIT

Die europäischen Statistiken spiegeln die Realität genau und zuverlässig wider.

INDIKATOREN

12.1: Die Basisdaten, die vorläufigen Ergebnisse und die statistischen Produkte werden regelmäßig evaluiert und validiert.

12.2: Stichprobenfehler und Nicht-Stichprobenfehler werden gemessen und systematisch gemäß den europäischen Standards dokumentiert.

12.3: Zur Verbesserung statistischer Prozesse werden Datenrevisionen regelmäßig analysiert.

GRUNDSATZ 13: AKTUALITÄT UND PÜNKTLICHKEIT

Die europäischen Statistiken sind aktuell und werden pünktlich veröffentlicht.

INDIKATOREN

13.1: Die Aktualität erfüllt europäische und andere internationale Veröffentlichungsstandards.

13.2: Für die Veröffentlichung der Statistiken wird ein täglicher Standardzeitpunkt bekanntgegeben.

13.3: Die Periodizität der Statistiken trägt dem Nutzerbedarf weitestmöglich Rechnung.

13.4: Abweichungen vom Veröffentlichungskalender werden vorab bekanntgegeben und erläutert, und ein neuer Veröffentlichungszeitpunkt wird festgesetzt.

13.5: Vorläufige Ergebnisse von akzeptabler Gesamtgenauigkeit können veröffentlicht werden, wenn dies für nützlich erachtet wird.

GRUNDSATZ 14: KOHÄRENZ UND VERGLEICHBARKEIT

Die europäischen Statistiken sind untereinander und im Zeitablauf konsistent und zwischen Regionen und Ländern vergleichbar; es ist möglich, miteinander in Beziehung stehende Daten aus unterschiedlichen Quellen zu kombinieren und gemeinsam zu verwenden.

INDIKATOREN

14.1: Die Statistiken sind in sich kohärent und konsistent (d. h. die rechnerischen und buchungstechnischen Identitätsbeziehungen bleiben gewahrt).

14.2: Die Statistiken sind über einen ausreichenden Zeitraum betrachtet vergleichbar.

14.3: Die Erstellung der Statistiken erfolgt auf der Grundlage von einheitlichen Standards in Bezug auf den Geltungsbereich, die Definitionen, die Einheiten und die Klassifikationen, die für die verschiedenen Erhebungen und Quellen gelten.

14.4: Die Statistiken aus den verschiedenen Quellen und von verschiedener Periodizität werden verglichen und miteinander in Einklang gebracht.

14.5: Die Vergleichbarkeit der Daten verschiedener Länder wird innerhalb des Europäischen Statistischen Systems durch regelmäßige Kontakte zwischen dem Europäischen Statistischen System und anderen statistischen Systemen gewährleistet. Methodische Untersuchungen werden in enger Zusammenarbeit zwischen den Mitgliedstaaten und Eurostat durchgeführt.

GRUNDSATZ 15: ZUGÄNGLICHKEIT UND KLARHEIT

Die europäischen Statistiken werden klar und verständlich präsentiert, in geeigneter und benutzerfreundlicher Weise veröffentlicht und sind zusammen mit einschlägigen Metadaten und Erläuterungen entsprechend dem Grundsatz der Unparteilichkeit verfügbar und zugänglich.

INDIKATOREN

15.1: Die Statistiken und die entsprechenden Metadaten werden in einer Weise präsentiert und archiviert, die eine korrekte Interpretation und aussagekräftige Vergleiche erleichtert.

15.2: Die Verbreitung erfolgt mit Hilfe moderner Informations- und Kommunikationstechnologie sowie, falls angemessen, durch gedruckte Veröffentlichungen.

15.3: Maßgeschneiderte Analysen werden, wenn dies möglich ist, bereitgestellt und die Öffentlichkeit wird davon in Kenntnis gesetzt.

15.4: Der Zugang zu Mikrodaten ist zu Forschungszwecken gestattet und unterliegt besonderen Regeln oder Vorschriften.

15.5: Die Metadaten sind im Einklang mit standardisierten Metadaten-Systemen dokumentiert.

15.6: Die Nutzerinnen und Nutzer werden fortlaufend über die Methodik der statistischen Prozesse, einschließlich der Verwendung von Verwaltungsdaten, informiert.

15.7: Die Nutzerinnen und Nutzer werden fortlaufend über die Qualität der statistischen Produkte in Bezug auf die Qualitätskriterien für europäische Statistiken informiert.

Anhang 2: Fundamentale Prinzipien der Amtlichen Statistik der Vereinten Nationen

Grundsatz 1: Die amtliche Statistik ist ein unverzichtbarer Teil des Informationssystems einer demokratischen Gesellschaft, sie stellt Regierungen, der Wirtschaft und der Öffentlichkeit Daten über die wirtschaftliche, demographische, soziale und Umweltsituation zur Verfügung. Zur Erfüllung von erwiesenen Informationsbedürfnissen dieser Art werden auf unparteiischer Grundlage durch staatliche statistische Institutionen amtliche Statistiken erstellt und zugänglich gemacht, um dem Recht der Bürger und Bürgerinnen auf öffentliche Information zu entsprechen.

Grundsatz 2: Um das Vertrauen in die amtliche Statistik zu erhalten ist es erforderlich, dass die statistischen Institutionen ihre Entscheidungen über Methoden und Verfahren für die Erhebung, Aufarbeitung, Speicherung und Präsentation statistischer Daten ausschließlich auf Grund von sachbezogenen Überlegungen, d.h. nach wissenschaftlichen Prinzipien und im Einklang mit der Berufsethik, treffen.

Grundsatz 3: Um die korrekte Interpretation von Daten zu erleichtern legen die statistischen Institutionen die verwendeten statistischen Quellen, Methoden und Verfahren gemäß wissenschaftlichen Standards offen.

Grundsatz 4: Die statistischen Institutionen sind berechtigt, auf unrichtige Interpretationen und den Missbrauch von Statistiken zu reagieren.

Grundsatz 5: Daten für statistische Zwecke können aus den verschiedensten Quellen stammen, sowohl aus statistischen Erhebungen wie auch aus Verwaltungsdaten. Die statistischen Institutionen wählen die Quellen unter Berücksichtigung der Kriterien Qualität, zeitliche Verfügbarkeit, Kosten und Respondentenbelastung aus.

Grundsatz 6: Personenbezogene Daten, die von statistischen Institutionen für die Erstellung von amtlichen Statistiken beschafft werden, und zwar gleichgültig, ob sie sich auf natürliche oder auf juristische Personen beziehen, sollen streng vertraulich behandelt und ausschließlich für statistische Zwecke verwendet werden.

Grundsatz 7: Die Gesetze, Verordnungen und Maßnahmen, die die Grundlage der statistischen Systeme bilden, sollen öffentlich zugänglich sein.

Grundsatz 8: Die Koordination zwischen den statistischen Institutionen innerhalb eines Landes ist von großer Bedeutung für die Konsistenz und Leistungsfähigkeit des statistischen Systems.

Grundsatz 9: Die Verwendung internationaler Konzepte, Klassifikationen und Methoden durch die statistischen Institutionen jedes einzelnen Landes fördert die Vergleichbarkeit und die Leistungsfähigkeit der statistischen Systeme auf allen innerstaatlichen und internationalen Ebenen.

Grundsatz 10: Bilaterale und multilaterale statistische Kooperation trägt zur Verbesserung der Systeme amtlicher Statistik in allen Ländern bei.

Anhang 3: Task Force zur Überarbeitung der Qualitätsrichtlinien

Thomas Burg	Qualitätsmanagement und Methodik
Johanna Einfalt	Register, Klassifikationen und Geoinformation
Reinhard Fiedler	Register, Klassifikationen und Geoinformation
Markus Fröhlich	Qualitätsmanagement und Methodik
Thomas Glaser	Direktion Bevölkerung
Martin Hirsch	Direktion Unternehmen
Thomas Karner	Direktion Unternehmen
Petra Koll	Direktion Bevölkerung
Alexander Kowarik	Qualitätsmanagement und Methodik
Kathrin Lais	Direktion Volkswirtschaft
Eva Milota	Direktion Raumwirtschaft
Franz Neumann	Direktion Raumwirtschaft
Alexandra Schindlar	Direktion Volkswirtschaft
Magdalena Six	Qualitätsmanagement und Methodik
Beatrix Tomaschek	Medien- und Informationspolitik
Christoph Waldner	Register, Klassifikationen und Geoinformation

Glossar

Anlassbezogene Revision

Anlassbezogene Revisionen sind solche, die nicht im Vorhinein geplant gewesen sind und dadurch zustande kommen, dass bereits als endgültig gekennzeichnete Ergebnisse aufgrund neuer Erkenntnisse geändert werden müssen.

Ausreißer (Outlier)

Ausreißer bei quantitativen Merkmalen weichen von der Majorität der Daten ab und können großen Einfluss auf klassische Schätzungen haben, welche durch Ausreißer stark verfälscht werden können, was in weiterer Folge zu inkorrekten Ergebnissen von Statistiken führt. Dies betrifft vor allem quantitative Merkmale. Einerseits kann es sich um Messfehler handeln (z.B. Fehler bei Angabe des Einkommens) oder um valide aber extreme Datenwerte handeln (z.B. eine reiche Person).

Auswahlrahmen

Ein Auswahlrahmen ist eine IT-mäßig gespeicherte Menge von Einheiten, die die **Grundgesamtheit** möglichst gut abbilden soll. Die wichtigsten Auswahlrahmen bei Statistik Austria sind das **Zentrale Melderegister** und das **Unternehmensregister für statistische Zwecke**.

Authentischer Datenbestand

Der Authentische Datenbestand ist der finale Datenbestand, von dem Publikations- bzw. Ergebnistabellen erstellt werden. Die im authentischen Datenbestand enthaltenen (Mikro-)Daten werden geplanter Weise in Ihren Merkmalswerten nicht mehr verändert oder um weitere Merkmale angereichert.

Backcasting

Unter Backcasting versteht man die Anwendung von statistischen Vorhersagemethoden zur Schätzung von Werten, die sich auf in der Zeit zurückliegende Perioden beziehen. Dies wird z.B. im Fall von Revisionen von Klassifikationen für die Schätzung von den neuen Klassifikationsausprägungen zugeordneten Kennzahlen angewendet.

Bereichsspezifisches Personenkennzeichen Amtliche Statistik (bPK-AS)

Das bereichsspezifische Personenkennzeichen (bPK-AS) ersetzt und pseudonymisiert bei Personenbefragungen die Identitätsdaten. Dadurch wird eine Nutzung der Einzeldatensätze, beispielsweise durch eine Verknüpfung mit anderen Einzeldaten, ermöglicht, ohne die Geheimhaltungsvorschriften zu verletzen.

Bruttoinlandsprodukt

Das Bruttoinlandsprodukt ist die zentrale Kenngröße der **volkswirtschaftlichen Gesamtrechnungen** und misst die Wirtschaftsleistung eines Landes. Das Bruttoinlandsprodukt kann sowohl über die Entstehungsseite als auch über Verwendungs- und Vermögensseite berechnet werden.

Datenbank, Datenbanksystem

Eine Datenbank, auch Datenbanksystem genannt, ist ein System zur elektronischen Datenverwaltung. Die wesentliche Aufgabe einer Datenbank ist es, große Datenmengen effizient, widerspruchsfrei und dauerhaft zu speichern und benötigte Teilmengen in unterschiedlichen, bedarfsgerechten Darstellungsformen für Benutzer und Anwendungsprogramme bereitzustellen

Datenbereinigung (Data Cleaning)

Datenbereinigung subsumiert den Vorgang der **Plausibilitätsprüfung** und den Vorgang der **Imputation**. Es handelt sich dabei um die Hauptarbeitsschritte, die notwendig sind um aus Rohdaten statistisch verwertbare Daten zu kreieren.

Datenerfassung

Datenerfassung beschreibt den Vorgang, zuvor gewonnene Rohdaten in ein elektronisches System einzupflegen. Im Falle einer elektronischen bzw. computerunterstützten Erhebung bzw. bei der Übermittlung administrativer Datenbestände ist dies bereits implizit gegeben.

Datenwürfel

Ein Datenwürfel ist eine Form der logischen Darstellung von Daten innerhalb der bei Statistik Austria in Verwendung befindlichen Datenbank STATcube. Daten werden dabei als Elemente eines mehrdimensionalen Würfels betrachtet.

Diagramm

Ein Diagramm ist eine grafische Darstellung von Daten, Sachverhalten oder Informationen. Je nach der Zielsetzung des Diagramms werden höchst unterschiedliche Typen eingesetzt.

ESAC (Europäischer Beratender Ausschuss für Statistik)

ESAC hat 24 Mitglieder, welche Nutzerinnen und Nutzer, Auskunftgebende und andere Akteure im Bereich der Gemeinschaftsstatistiken repräsentieren. Die Aufgabe des ESAC ist es, dafür zu sorgen, dass bei der Entwicklung der Statistikprogramme die Erfordernisse der Nutzer berücksichtigt und die Informationslieferanten und Datenproduzenten nicht über Gebühr belastet werden. Der Ausschuss gibt Stellungnahmen zum Statistischen Mehrjahresprogramm ab, insbesondere hinsichtlich dessen Relevanz für die Erfordernisse der europäischen Integration und nimmt auch zur Ausgewogenheit der Verteilung von Prioritäten und Ressourcen auf verschiedene Bereiche des Statistischen Mehrjahresprogramms sowie zum jährlichen statistischen Arbeitsprogramm der Kommission Stellung.

ESGAB (European Statistical Governance Advisory Board - Europäisches Beratungsgremium für Statistische Governance)

ESGAB wurde für die unabhängige Überwachung des **Europäischen Statistischen Systems (ESS)**, genauer gesagt die Umsetzung des **Verhaltenskodex für europäische Statistiken** eingesetzt. Das Gremium berät die Europäische Kommission (**Eurostat**), wenn es um geeignete Maßnahmen zur erleichterten Durchführung des Kodex, um einschlägige Informationen für Nutzerinnen und Nutzer und Datenlieferanten sowie um die Aktualisierung des Kodex geht.

ESMS (Euro-SDMX Metadata Structure)

ESMS wurde bis 2009 als vorgegebene Struktur für Qualitätsberichterstattung wie beispielsweise die Übermittlung von **Metadaten** verwendet

ESQRS (ESS Standard for Quality Reports Structure)

ESQRS ist eine vorgegebene Struktur für Qualitätsberichtserstattung wie beispielsweise die Übermittlung von **Metadaten**.

ESS (European Statistical System – Europäisches Statistisches System)

Das Europäische Statistische System ist eine Partnerschaft zwischen der statistischen Stelle der Gemeinschaft, d. h. der Kommission (**Eurostat**), den nationalen statistischen Ämtern (NSIs-National Statistical Institutes) und anderen einzelstaatlichen Stellen, die in den einzelnen Mitgliedstaaten für die Entwicklung, Erstellung und Verbreitung europäischer Statistiken zuständig sind. Diese Partnerschaft umfasst auch die EWR- und EFTA-Länder.

Eurostat

Eurostat ist das statistische Amt der Europäischen Union mit Sitz in Luxemburg. Es hat den Auftrag, die Union mit europäischen Statistiken zu versorgen, die Vergleiche zwischen Ländern und Regionen ermöglichen. Eurostat hat den Rang einer Generaldirektion der Europäischen Kommission und ist dem Kommissar für Wirtschaft und Währung zugeordnet.

Feedback-Gespräche

Feedback-Gespräche werden regelmäßige Treffen externer Nutzer und Nutzerinnen, Mitgliedern des Statistikrat-Ausschusses „Qualitätssicherung“, und MitarbeiterInnen und Mitar-

beitern von Statistik Austria genannt, um auf Basis der **Standard-Dokumentation** die Qualität eines statistischen Produkts zu diskutieren. Inhalt und Ziele der "Feedback-Gespräche" sind die kritische Auseinandersetzung mit den Qualitätsaspekten der Statistiken unter besonderer Berücksichtigung der verwendeten Methoden und Verfahren, die Identifikation von Qualitätsverbesserungspotentialen und Erarbeitung von Empfehlungen für Verbesserungsmaßnahmen sowie die Verbesserung der **Standard-Dokumentation** der jeweils besprochenen Statistiken, wobei insbesondere die Sicht von Nutzerinnen und Nutzern und externer Expertinnen und Experten einfließen soll.

Flash Estimates

Flash Estimates bezeichnen Schätzwerte, die so früh wie möglich nach Ende der Referenzperiode produziert und veröffentlicht werden. Sie basieren dabei auf einer – im Vergleich zu herkömmlichen Schätzungen – reduzierten Informationsmenge und werden, sobald die Datenbasis vollständig bzw. zumindest vollständig zur Verfügung steht, revidiert.

Gebundene Hochrechnung

Werden die hochgerechneten Ergebnisse aus der Stichprobe an externe Informationen angepasst, so spricht man von einer gebundenen Hochrechnung. Ein häufiges Beispiel für eine derartige Anpassung ist die Kalibrierung der hochgerechneten Ergebnisse, sodass die geschätzte Anzahl und die aus externen Quellen bekannte Anzahl an Frauen bzw. Männern oder die bekannte Anzahl an Personen pro Bundesland übereinstimmen.

Geheimhaltung

Durch die Anwendung von Geheimhaltungsmethoden soll verhindert werden, dass aus veröffentlichten Daten von Statistik Austria Rückschlüsse auf individuelle und sensible Informationen gezogen werden können. Man unterscheidet grundsätzlich zwischen **Geheimhaltung von Mikrodaten** und **Geheimhaltung tabellarischer Daten**.

Geheimhaltung von Mikrodaten

Vor der Veröffentlichung von Einzeldaten etwa in Form von Datensätzen für die Anwendung in Forschung und Lehre wird die Einhaltung des Schutzes persönlicher Information von Respondenten durch Anwendung von Geheimhaltungsmethoden bestmöglich garantiert. Einige häufig angewendete Geheimhaltungsmethoden von Mikrodaten sind etwa Vergrößerung von Merkmalen, Lokale Unterdrückung, Mikroaggregation und Postrandomisierung. Gleichzeitig wird bei der Anwendung von Geheimhaltungsverfahren darauf geachtet, den Informationsverlust zu minimieren.

Geheimhaltung tabellarischer Daten

Vor der Veröffentlichung von Daten in tabellarischer oder in interaktiver Form (wie etwa in **STATcube**) muss durch Anwendung von Geheimhaltungsmethoden der notwendige Datenschutz gewährleistet werden. Eine häufig angewendete Methode sind Datenunterdrückungen (**Primäre und Sekundäre Geheimhaltung**)

Geographisches Informationssystem (GIS)

Ein Geoinformationssystem (GIS) ist ein rechnergestütztes System, das aus Hardware, Software und Daten besteht. Mit ihm können raumbezogene Daten erfasst und bearbeitet, gespeichert und verwaltet, analysiert und recherchiert sowie dargestellt werden. Geographische Informationssysteme ermöglichen es, statistische Daten nicht nur auf Basis vorgegebener räumlicher Einheiten (wie der Verwaltungsgliederung oder der Zählsprenkel) darzustellen, sondern für beliebig definierte räumliche Gebilde.

Gesamtrechnungen

Gesamtrechnungen verknüpfen und ergänzen statistische Basisdaten in systematischer Weise, so dass sie einem vorgegebenen Kontenplan entsprechen. Zumeist werden die Kontenpositionen in monetären Bewertungen angegeben. .

Geschichtete Zufallsstichprobe

Bei geschichteten Zufallsstichproben ist die Auswahlwahrscheinlichkeit nur innerhalb einer Gruppe von Respondenten (einer „Schicht“) mit denselben Eigenschaften (beispielsweise Geschlecht, Bundesland, etc.) gleich groß. Durch eine Schichtung kann garantiert werden, dass aus jeder Schicht ausreichend Einheiten gezogen werden, bzw. erhält man bei günstiger Schichtung genauere Ergebnisse.

Grundgesamtheit

Menge aller Einheiten, die man beschreiben will, beispielsweise österreichische Bevölkerung, Unternehmen in Österreich, Haushalte in Österreich. Eine Grundgesamtheit wird auch oftmals statistische Population genannt.

Hochrechnung

Die Hochrechnung ist eine geschätzte Extrapolation der Ergebnisse aus der Stichprobe auf die Gesamtpopulation. Dies geschieht durch die Zuweisung eines Hochrechnungsgewichts zu jeder Stichprobeneinheit, weshalb auch von Gewichtung gesprochen wird.

Imputation

Imputation ist jener Vorgang, der fehlende, ungültige bzw. inkonsistente Werte, die während der **Plausibilitätsprüfung** entdeckt wurden, ersetzt. Imputation soll vor allem dem Zweck dienen, die durch Antwortverweigerung entstandene Verzerrung (**Non-Response Bias**) auszugleichen. Die Ersetzung durch plausible und intern kohärente Werte erfolgt durch wissenschaftliche und erprobte Methoden.

Index

Vor allem für die wirtschaftliche Entwicklung sind Indizes wichtige Konjunkturindikatoren, die meist in kurzen Zeitabständen veröffentlicht werden. Ein Index misst nicht den Absolutwert eines Merkmals, sondern die jeweilige Veränderung des Merkmals in Bezug auf die zuletzt gemessene Periode oder in Bezug auf eine Basisperiode. Je nachdem ob dabei die Veränderung Preisen oder von Mengen oder gemessen wird, unterscheidet man zwischen Preis und Mengen- bzw. Volumenindizes. Oftmals beschreiben Indizes die Entwicklung mehrerer, sachlich zusammenhängender Größen (z.B. der Warenkorb beim Verbraucherpreisindex). Die Aggregation der einzelnen Größen erfolgt dann nicht rein summarisch sondern basiert auf einem Gewichtungsschema.

Indikator

Im sozio-ökonomischen Kontext dienen Indikatoren oftmals zur Messung bzw. Operationalisierung von Konzepten, durch welche auch theoretische bzw. abstrakte Begriffe messbar gemacht werden, selbst wenn diese nicht direkt ermittelbar sind. Bekannte sozio-ökonomische Indikatoren sind beispielsweise die Indikatoren zu Armut und sozialer Eingliederung oder die Lebenserwartung.

Im ökonomischen Kontext misst ein volkswirtschaftlicher Indikator eine Messgröße, welche Aussagen über die konjunkturelle Entwicklung oder die wirtschaftliche Situation einer Volkswirtschaft erlaubt.

Input Output System

Im Projekt Input-Output-Statistik werden Güterströme in einer Volkswirtschaft betrachtet und in zwei Darstellungsformen – Aufkommens- und Verwendungstabellen bzw. Input-Output-Tabellen - präsentiert: Aufkommens- und Verwendungstabellen stellen das Güteraufkommen und die Güterverwendung in einer detaillierten Gliederung gegenüber. Durch die Herstellung von Aufkommens- und Verwendungsgleichheit für jedes Gut werden die Ergebnisse des entstehungsseitigen und des verwendungsseitigen Ansatzes der Bruttoinlandsproduktsberechnung auf Konsistenz geprüft.

Intervallskala

Bei intervallskalierten Merkmalen lassen sich zusätzlich zu den Eigenschaften der **Ordinalskala** die Abstände zwischen den verschiedenen Merkmalsausprägungen exakt bestimm-

men, allerdings existiert kein natürlicher Nullpunkt für die Skala. Beispiele sind Temperaturangaben oder Zeitpunkte. Eine Intervallskala misst ein quantitatives Merkmal.

Irreguläre Komponente einer Zeitreihe

Die irreguläre Komponente ist diejenige Komponente der **Zeitreihe**, die Einflüsse wiedergibt, die weder der **Trend- und Konjunkturkomponente** noch der **Saisonkomponente einer Zeitreihe** zugeordnet werden können.

Item Non Response (partieller Antwortausfall)

Item Non Response beschreibt den Ausfall einzelner Untersuchungsmerkmale. Bei einer Erhebung können beispielsweise nur bestimmte Fragen beantwortet worden sein, während Antworten zu anderen fehlen (z.:B. weil die Antwort verweigert wurde).

Kartogramm

Kartogramme sind thematische Karten, in denen quantitative Aussagen auf festgelegten Bezugsflächen dargestellt werden. Die jeweilige Angabe gilt für die gesamte Fläche, ohne interne Variationen zu beachten. Häufig werden statistische Angaben, die auf einer Ebene der Verwaltungsgliederung erhoben wurden, als Kartogramm dargestellt. Viele gesammelte Kartogramme finden sich im [OnlineAtlas](#) von Statistik Austria.

Klassifikation

Eine Klassifikation ist eine Einteilung von Beobachtungen in diskrete, vollständige und einander ausschließende Kategorien (Klassen). Die einzelnen Klassen werden in der Regel mittels Klassifizierung – das heißt durch die Einteilungen von Objekten anhand bestimmter übereinstimmender Merkmale – gewonnen. Klassen gewährleisten die vollständige Erfassung der Grundgesamtheit. Die im statistischen Produktionsprozess zur Verwendung kommenden Klassifikationen sind in den meisten Fällen international standardisiert. Wichtige Klassifikationen sind beispielsweise die **NACE** oder die **CPA**.

Klassifikationsdatenbank

Die Klassifikationsdatenbank von Statistik Austria beinhaltet eine große Anzahl nationaler und internationaler Klassifikationen und Systematiken. Sie ist unentgeltlich im Internet abrufbar.

Laspeyres-Index

Das Indexkonzept nach Laspeyres zeichnet sich dadurch aus, dass die Gewichtung des Warenkorbs für die Basisperiode erfasst und für die einzelnen Berichtsperioden konstant gehalten wird und die Information über den Indexgegenstand (zumeist Preise) beobachtet wird.

Laufende Revision

Laufende Revisionen sind solche, die standardmäßig gemäß einer a priori fixierten Publikationsprozedur erfolgen. Zu fix vorgegeben Terminen werden die bis dato veröffentlichten, vorläufigen Ergebnisse durch andere ersetzt. Nach der letzten planmäßigen Revision spricht man von endgültigen Ergebnissen.

Lokale Unterdrückung

Lokale Unterdrückung bezeichnet das gezielte Sperren einzelner Ausprägungen in ausgewählten Variablen. Häufig wird Lokale Unterdrückung angewendet, um sicherzustellen, dass jede Ausprägungskombination von einem Set von Variablen - den **Schlüsselvariablen** – zumindest k Einheiten im Datensatz zugeordnet werden kann. Dies entspricht k-Anonymität.

Makroplaus (Macro Editing)

Makroplaus bezeichnet die Überprüfung von Aggregaten und Verteilungen hinsichtlich inhaltlicher Glaubwürdigkeit. Dies geschieht zumeist durch den Vergleich zu Vorperioden, bzw. den Vergleich mit anderen Datenquellen.

Mikroaggregation

Mikroaggregation bedeutet grundsätzlich, dass möglichst ähnliche Objekte in einem ersten Schritt gruppiert werden. In einem zweiten Schritt werden schließlich die Ausprägungen von numerischen Variablen der gewählten Einheiten durch eine Statistik - üblicherweise dem Mittelwert ersetzt. Durch die Mikroaggregation numerischer Variablen wird sichergestellt, dass jede einzelne Ausprägung der numerischen Variablen mehrfach im Datensatz auftritt.

Mikroplaus (Micro Editing)

Mikroplaus bezeichnet die Überprüfung von Einzeldatensätzen mithilfe von Plausibilitätsregeln. Zielrichtung dabei ist es, dass die Einzeldatensätze fehlerfrei, logisch konsistent und inhaltlich plausibel sind. Besonderes Augenmerk kommt dabei dem Erkennen von Ausreißern zu („Outlier Detection“), da diese die Ergebnisse stark beeinflussen können.

Nicht-Zufallsstichprobe

Bei einer Nicht-Zufallsstichprobe gibt es Einheiten der Grundgesamtheit, die zwar im Auswahlrahmen erfasst sind, jedoch eine Auswahlwahrscheinlichkeit von Null haben. Ein häufiges Beispiel sind sogenannte Konzentrationsstichproben, bei der nur Einheiten ab einem gewissen Schwellenwert ausgewählt werden (bei Unternehmenserhebungen beispielsweise der Umsatz oder die Anzahl der Beschäftigten). Nicht-Zufallsstichproben lassen zwar keine unmittelbaren Rückschlüsse auf die Grundgesamtheit mittels Hochrechnung zu, allerdings ist eine Ergebnisdarstellung für die Grundgesamtheit mithilfe verfügbarer Sekundärdaten in Kombination mit Modellrechnungen doch möglich.

Nominalskala

Ein Merkmal ist nominal skaliert, wenn seine möglichen Ausprägungen zwar unterschieden sein können, aber keine natürliche Rangfolge aufweisen, typisches Beispiel ist das Geschlecht. Eine Nominalskala beschreibt ein qualitatives Merkmal.

Non Response (Antwortausfall)

Non Response steht für das Nichtreagieren auf Fragen bei Befragungen. Nicht zufällige, ergo systematische Antwortausfälle können zu Verzerrungen der Ergebnisse (Non Response Bias) führen. Man unterscheidet zwischen **Item Non Response** und **Unit Non Response**.

Open Data (Offene Daten)

Open Data sind sämtliche Datenbestände, die im Interesse der Allgemeinheit den Open Data Prinzipien gemäß ohne jedwede Einschränkung zur freien Nutzung, zur maschinenlesbaren Weiterverbreitung und zur freien Weiterverwendung frei zugänglich gemacht werden.

Ordinalskala

Die Ordinalskala dient in der Statistik der Charakterisierung von Ausprägungen, zwischen denen eine natürliche Rangordnung besteht, allerdings ist es nicht möglich, Aussagen über die Abstände zwischen den Rangplätzen zu treffen. Beispiele sind Selbsteinstufungen, Zufriedenheit mit Einkommen, Bewertung von Gesundheit. Eine Ordinalskala beschreibt ein qualitatives Merkmal.

Paarsche-Index

Der Paarsche-Index zeichnet sich dadurch aus, dass die Gewichtung des Warenkorb für jede Berichtsperiode gesondert vorgenommen wird und somit der Wert des Warenkorb aus aktuellen Verbrauchszahlen ermittelt wird.

Persönlicher Inflationsrechner

Der persönliche Inflationsrechner von Statistik Austria erlaubt es allen Nutzerinnen und Nutzern, durch eigene Angaben zu den haushaltsspezifischen Ausgaben eine für den jeweiligen Haushalt aussagekräftige Inflationsrate zu berechnen und diese mit der nationalen Inflationsrate zu vergleichen. Der persönliche Inflationsrechner ist gratis auf der Homepage von Statistik Austria aufrufbar.

Plausibilitätsprüfung (Statistical Editing)

Unter Plausibilitätsprüfung („Statistical Editing“) wird das Aufspüren von Fehlern in Daten verstanden. Dabei sollen Situationen erkannt werden, die fehlerhaft, logisch inkonsistent sind bzw. inhaltlich implausibel erscheinen.

Postrandomisierung

Um Variablen, die für die Datenanalyse absolut notwendig sind und aus diesem Grunde nicht aus dem Datensatz entfernt oder vergrößert werden können zu schützen, besteht die Möglichkeit, ein Postrandomisierungsverfahren anzuwenden. Die grundlegende Idee bei diesem Verfahren ist es, dass grundsätzlich jede Ausprägung in einer Variable entweder unverändert bleibt oder – mit gewissen Wahrscheinlichkeiten - in eine andere Kategorie wechseln kann.

Primäre und Sekundäre Geheimhaltung tabellarischer Daten

Primäre Geheimhaltung (Primärsperren) bezeichnet das Unterdrücken von Zellwerten, anhand derer man Einzeldatensätze identifizieren könnte. Da es durch Rückrechnungen dennoch zu einer Identifizierbarkeit einzelner statistischer Einheiten kommen könnte, sind auch Gegenlösungen notwendig, diese werden als sekundäre Geheimhaltung (Sekundärsperren) bezeichnet.

Primärstatistische Daten

Primärstatistischen Daten sind solche, die für einen zuvor definierten Zweck als Input zur statistischen Produktion erstmalig gewonnen werden. Man spricht auch von einer primärstatistischen Erhebung.

Reine Zufallsstichprobe („simple random sample“)

Bei einer reinen Zufallsstichprobe hat jede Einheit der Grundgesamtheit, welche auch im **Auswahlrahmen** erfasst ist, dieselbe von Null verschiedene Auswahlwahrscheinlichkeit.

Revision

Eine Revision ist eine ex-post Veränderung von statistischen Ergebnissen, die bereits veröffentlicht wurden. Man unterscheidet zwischen **laufenden Revisionen** und **anlassbezogenen Revisionen**.

Rohdaten

Rohdaten sind jene Daten, die bei einer Datenerhebung unmittelbar gewonnen worden sind. Sie zeichnen sich dadurch aus, dass noch keinerlei Datenbearbeitung stattgefunden hat. Dadurch sind Rohdaten oftmals lücken- und fehlerhaft.

Saisonale Bereinigung von Zeitreihen („Saisonbereinigung“)

Unter Saisonbereinigung für Zeitreihen werden jene Verfahren verstanden, die die saisonalen Einflüsse aus einer Zeitreihe herausfiltern und somit eine exakter interpretierbare Darstellung der Zeitreihe ermöglichen.

Saisonkomponente einer Zeitreihe

Die Saisonkomponente ist diejenige Komponente der **Zeitreihe**, welche die periodischen Schwingungen mit der Periode eines Jahres beschreibt. Sie entspricht dem Zyklus der Jahreszeiten.

Satellitenkonten

Satellitenkonten als Gesamtrechnungssysteme decken Gebiete ab, worüber die **volkswirtschaftliche Gesamtrechnung** nur unzureichend Informationen liefert. Auch die Satellitenkonten beruhen auf europaweit akkordierten Konzepten, Definitionen und Klassifikationen. Beispiele umfassen die Umwelt-Gesamtrechnungen oder das System der Gesundheitsausgaben.

Schlüsselvariable

Indirekte Identifikationsvariablen, deren Ausprägungskombinationen ein Angreifer verwenden könnte, um eine eindeutige Identifikation einer Person im Datensatz vorzunehmen, werden

allgemein als Schlüsselvariablen bezeichnet. Beispiele sind das Wohnsitzbundesland, Alter oder die Schulbildung einer Person.

SDMX (Statistical Data and Metadata Exchange)

SDMX ist eine internationale Initiative, welche sich zum Ziel gesetzt hat, die Mechanismen und Prozesse für den Austausch von statistischer Information und zugehöriger **Metadaten** zwischen internationalen Organisationen und ihren Mitgliedsstaaten zu standardisieren und zu vereinheitlichen. Die Informationsübermittlung zwischen Statistik Austria und **Eurostat** basiert auf SDMX.

Sektorkonten

Sektorkonten sind eine Darstellungsform in der Volkswirtschaftlichen Gesamtrechnung, welche für die einzelnen institutionellen Sektoren (beispielsweise Kapitalgesellschaften, Staat, Private Haushalte) eine Beschreibung der verschiedenen Phasen des Wirtschaftskreislaufs liefern. Die Phasen des Wirtschaftskreislaufs umfassen die Produktion, die Einkommensentstehung, -verteilung, -umverteilung und –verwendung von finanziellem und nicht-finanziellem Vermögen.

Sekundärstatistische Daten

Sekundärstatistische Daten sind solche, die bereits zu einem anderen Zweck gesammelt wurden. Dabei kann es sich sowohl um einen der statistischen Produktion dienlichen Zweck handeln oder aber ein gänzliche anderer z.B. der Verwaltung dienend Zweck handeln. Werden sekundärstatistische Daten für die Produktion von Statistik verwendet spricht man von sekundärstatistischer Datennutzung.

Signierung

Signierung bezeichnet das Zuweisen von - gemäß einer definierten lexikalischen Klassifikation -alphanumerischen Schlüsseln (Codes) zu einer zumeist textuellen Antwort. Der Signiervorgang erfolgt meist automatisch bzw. automationsunterstützt; nur bei speziellen Anwendungen ist eine manuelle Signierung erforderlich.

SIMS (Single Integrated Metadata Structure)

SIMS wurde von der Task Force on Quality Reporting entwickelt und legt einen SDMX basierten Rahmen für eine einheitliche Beschreibung und Dokumentierung der Metadaten von statistischen Produkten fest. Durch SIMS soll eine einheitliche und konsistente Qualitätsberichtserstattung über statistische Produkte einzelner nationaler statistischer Institute an Eurostat erreicht werden.

Standard-Dokumentation

Die Standard-Dokumentation bei Statistik Austria ist ein standardisiertes Berichtsschema, welches Nutzerinnen und Nutzern Informationen über die den statistischen Produkten zugrunde liegenden Konzepte, Definitionen und außerdem Angaben zu verwendeten Methoden und zur Qualität der statistischen Produkte zur Verfügung stellt. Die Standard-Dokumentation folgt einem fixen Berichtsschema. Standard Dokumentationen existieren zu jedem statistischen Produkt von Statistik Austria. Sie werden in regelmäßigen Abständen upgedated und mit den Nutzern und Nutzerinnen und den Mitgliedern des Statistikrat-Ausschusses „Qualitätssicherung“ in Form von **Feedback-Gesprächen** diskutiert.

STATcube – die statistische Datenbank

STATcube ist das statistische Datenbanksystem von Statistik Austria und steht unter www.statcube.at online zur Verfügung. Damit ist es Nutzerinnen und Nutzern möglich, Auswertung nach ihren Bedürfnissen online erstellen und in verschiedenen Formaten ausgeben zu lassen. Der Name basiert auf dem Datenbankelementen in STATcube, die man als **Datenwürfel** („cube“) bezeichnet.

Statistische Maßzahl

Eine statistische Maßzahl ist das Ergebnis einer Funktion, die auf die Daten angewandt wird,

um die Verteilung eines Merkmals innerhalb einer statistischen Masse zu charakterisieren. Typische statistische Maßzahlen sind der Mittelwert und der Median.

Statistische Metadaten

Statistische Metadaten sind Daten, welche andere Daten beschreiben. Sie stellen nicht direkt den Inhalt einer Statistik dar, sind aber für das Verständnis des statistischen Produkts bzw. für dessen Herstellungsprozess von essentieller Bedeutung. Zu Metadaten zählen beispielsweise Beschreibungen der Datenfelder in Umfragen, Variablendefinitionen und – abgrenzungen, Erläuterungen sowie komplette Formularbeschreibungen. Für jedes statistische Produkt werden Metadaten bei Statistik Austria in einer **Standard-Dokumentation** gemäß einem vorgegebenen standardisierten Schema dargestellt.

Statistische Register

Ein Register ist ein Verzeichnis, welches in geordneter Form Einheiten und die ihnen zugeordneten Informationen beinhaltet. Statistische Register umfassen beispielsweise Personen, Unternehmen oder Gebäude. Sie sollen die Grundgesamtheit abbilden und dienen oftmals als **Auswahlrahmen** für Stichprobenziehungen. Wichtige Register sind beispielsweise das **Zentrale Melderegister (ZMR)** und das **Unternehmensregister für statistische Zwecke(URS)**.

Stichprobenerhebung

Bei einer Stichprobenerhebung werden nicht Daten aller Einheiten der Grundgesamtheit einer Population erhoben, sondern lediglich eines Teils. Die betrachtete Teilmenge der Population (die Stichprobe), wird aus dem **Auswahlrahmen** gezogen und ist zumeist von der Quantität weitaus kleiner als die Grundgesamtheit. Man unterscheidet zwischen **Zufallsstichproben und Nicht-Zufallsstichproben**. Nur bei ersteren können die stichprobentheoretischen Verfahren für Hoch- und Fehlerrechnung zur Schätzung von Parametern der Grundgesamtheit zur Anwendung gebracht werden.

Stichprobenfehler

Der Stichprobenfehler entsteht durch die Verwendung einer Stichprobe statt der gesamten Grundgesamtheit. Er misst die Ungenauigkeit des geschätzten Parameters und ist definiert als die Standardabweichung der Schätzfunktion für diesen Parameter.

Target Record Swapping

Target Record Swapping ist eine Geheimhaltungsmethode, bei der ein Teil der Merkmale einer statistischen Einheit mit denen einer ähnlichen Einheit getauscht werden. Das wird auf einen bestimmten Prozentsatz (1-15%) der Daten angewandt. Wenn nach dem Record Swapping Auswertungen von Einzelmerkmalen über die gesamte Menge der Einheiten durchführt werden, sind die Ergebnisse identisch mit jenen vor dem Swapping.

Treemap

Eine Treemap dient der Visualisierung hierarchischer Strukturen, die hierbei durch ineinander verschachtelte Rechtecke dargestellt werden. Damit können anschaulich Größenverhältnisse dargestellt werden, indem die Fläche der Rechtecke proportional zur Größe der darzustellenden Dateneinheit gewählt wird. Bei Statistik Austria kommen Treemaps beim [persönlichen Inflationsrechner](#) zur Veranschaulichung des nach eigenen Angaben gewichteten Warenkorbs zur Anwendung.

Trend- und Konjunkturkomponente einer Zeitreihe

Die Trend- und Konjunkturkomponente ist diejenige Komponente der **Zeitreihe**, welche die mittel- und langfristige Entwicklungsrichtung der Zeitreihe angibt

Unit Non Response (vollständiger Antwortausfall)

Unit Non Response beschreibt den kompletten Ausfall einer Erhebungseinheit durch Verweigerung der Teilnahme an einer Erhebung.

Unternehmensregister für statistische Zwecke(URS)

Das Unternehmensregister für statistische Zwecke stellt die Evidenz der in Österreich ansässigen und wirtschaftlich tätigen Unternehmen, ihrer Betriebe und Arbeitsstätten sowie der Einrichtungen des Staates und der Non-Profit-Organisationen dar. Es umfasst mit Ausnahme der Land- und Forstwirtschaft und der Fischerei, sowie der Privatzimmervermietung und der Vermietung von Immobilien durch Private Haushalte sämtliche andere Wirtschaftsbereiche.

Verhaltenskodex für europäische Statistiken („Code of Practice“)

Der Verhaltenskodex für europäische Statistiken basiert auf 15 Grundsätzen für den institutionellen Rahmen, die statistischen Prozesse und die statistischen Produkte. Für jeden Grundsatz bietet ein Satz von Indikatoren vorbildlicher Praktiken eine Referenz für die Überprüfung der Umsetzung des Kodex. Die Qualitätskriterien für europäische Statistiken sind im europäischen Statistikrecht festgelegt. Statistische Stellen, darunter die Kommission (**Eurostat**), nationale statistische Ämter und andere einzelstaatliche Stellen, die für die Entwicklung, Erstellung und Verbreitung europäischer Statistiken verantwortlich sind, sowie Regierungen, Ministerien und der Europäische Rat verpflichten sich zur Einhaltung des Kodex. Die Grundsätze des Verhaltenskodex bilden zusammen mit den allgemeinen Qualitätsmanagementgrundsätzen einen gemeinsamen Qualitätsrahmen für das **Europäische Statistische System**.

Verhältnisskala/Rationalskala

Bei Merkmalen, die auf einer Verhältnisskala messbar sind, existiert zusätzlich zu den Eigenschaften der Intervallskala ein absoluter Nullpunkt. Beispiele sind das Längenmaß oder Altersangaben. Eine Verhältnisskala misst ein quantitatives Merkmal.

Verwaltungsdaten

Verwaltungsdaten sind Daten, die bei Stellen in Wahrnehmung von bundes- oder landesgesetzlich übertragenen Aufgaben oder in Vollziehung unmittelbar anwendbarer gemeinschaftsrechtlicher Vorschriften angefallen sind. Es sind also Daten, die ursprünglich zu nicht statistischen Zwecken gesammelt wurden, aber als **sekundärstatistische** Quelle genutzt werden können.

Volkswirtschaftliche Gesamtrechnungen (VGR)

Die Volkswirtschaftlichen Gesamtrechnungen sind ein umfassendes Gesamtrechnungssystem, welches die statistische Darstellung der Gesamtwirtschaft in einem Kontenschema zum Gegenstand hat. Die zentrale Rechnungsgrundlage, das Europäische System volkswirtschaftlicher Gesamtrechnungen (ESVG 2010), umfasst europaweit akkordierte Konzepte, Definitionen und Klassifikationen. Die zwei Hauptdarstellungsformen des ESGV sind **Sektorkonten** und die Tabellen nach Wirtschaftsbereichen und das **Input-Output System**

Zeitreihe

Eine Zeitreihe ist eine periodische, sequentielle Betrachtung einer Variablen über die Zeit. In der Zeitreihenanalyse wird davon ausgegangen, die ursprüngliche Zeitreihe in drei Komponenten zerlegen zu können: die **Trend- und Konjunkturkomponente**, die **Saisonkomponente** und die **irreguläre Komponente**.

Zentrales Melderegister (ZMR)

Das Zentrale Melderegister ist ein durch das Bundesministerium für Inneres geführtes amtliches Verzeichnis, in dem der ständige oder vorübergehende Aufenthalt von Personen erfasst wird, soweit er der Meldepflicht unterliegt. Das zentrale Melderegister dient bei Statistik Austria als Grundlage für die Konstruktion von **Auswahlrahmen** für Personen- und Haushaltserhebungen.

Zufallsstichprobe (Random Sample)

Bei einer Zufallsstichprobe wird jede gezogene Einheit gemäß einem vorher festgelegten Stichprobenplan zufällig gezogen, wobei jede Einheit, welche im **Auswahlrahmen** erfasst ist, eine eindeutige, von Null verschiedene Auswahlwahrscheinlichkeit besitzt. Man unter-

scheidet zwischen reinen Zufallsstichproben und geschichteten Zufallsstichproben. Wichtig ist, dass keine Stichprobeneinheit bewusst oder willkürlich (also nicht auf Grundlage eines Zufallsexperiments) ausgewählt wird.

Bibliographie

Biemer, P; Lyberg, L (2014). Total Survey Error Paradigm. Short Course Materials within the Quality in Official Statistics Conference 2014. Vienna.

Brancato, G., Macchina, S., Murgia, M., Signore, M., Simeoni, G., Blanke, K., Körner, T., Nimmergut, A., Lima, P., Paulino, R., Hoffmeyer-Zlotnik, J.H.P., (2006). Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System, European Statistical System.

Dillman, D; Smyth, J; Christian, L (2009). Internet, Mail and Mixed-Mode Surveys. The Tailored Design Method.

Groves, R; Fowler, F; Couper, M; Lepkowski, J; Singer, E; Tourangeau, R (2009). Survey Methodology.

Presser, Stanley et al. (2004). Methods for Testing and Evaluating Survey Questions. In Public Opinion Quarterly, Vol. 68 No.1 Pp 109-130.

Allen, R.g.G.D 1975. *Index Numbers in Theory and Practice*, London, MacMillan

Archer, D. 1995 . *Maintenance of business registers Business Survey Methods*

Biemer, P., L. Lyberg. 2003. *Introduction to Survey Quality*. New York. Wiley.

Box, G.E.P. and G.M. Jenkins. 1976. *Time Series Analysis, Forecasting and Control*. San Francisco. Holden Day.

Bundesstatistikgesetz 2000 *Bundesgesetz über die Bundesstatistik*

http://www.statistik.at/web_de/ueber_uns/aufgaben_und_grundsaeetze/bundesstatistikgesetz/index.html

Brunauer Monika 2010 *Backcasting Methods at Statistics Austria*

[Paper presented at the Quality Conference Q2010, Helsinki](#)

Burg Thomas 2002. *Indicators for the Edit & Imputation process in the Austrian Quality Report System*

[Paper presented at the UN/ECE Work-Session on Data Editing 2002 Helsinki](#)

Burg Thomas 2003 *Procedures to improve the data-cleaning process based on quality information*

[Paper presented at the UN/ECE Work-Session on Data Editing 2003 Madrid](#)

Burg Thomas 2004 *Quality reports at Statistics Austria*

Paper presented at the Quality Conference Q2004, Mainz

Burg Thomas, Hackl Peter, Holzer Werner 2008. Quality at a Glance, Documentation of Quality Indicators at Statistics Austria

[Paper presented at the Quality Conference Q2008, Rom](#)

Burg Thomas 2009, *Consequences of a fully developed Quality Report System based on a 5-years Experience at Statistics Austria*

[Paper presented at the 57th session of the ISI 2009, Durban](#)

Burg Thomas 2010. *Are the Standard Documentations really Quality reports*

[Paper presented at the Quality Conference Q2010, Helsinki](#)

Cochran, William G. 1977,. *Sampling Techniques* New York, Wiley.

EUROSTAT 1999 *Handbook on quarterly national accounts*

<http://ec.europa.eu/eurostat/documents/3859598/5854053/CA-22-99-781-EN.PDF/faab8e70-d414-4bf0-8db1-a05b6e94e4aa?version=1.0>

EUROSTAT 2001 *Handbook on Price and Volume Measures in National Accounts*

<http://www.cbs.nl/NR/rdonlyres/54E6E10F-1B8E-4161-A9A7-9CB1B9D1C013/0/handbookpricevolume.pdf>

- EUROSTAT 2002 *Variance estimation methods in the European Union*
http://ec.europa.eu/eurostat/documents/64157/4374310/14-Monographs_off-stat-variance-estimation-methods-EU-2002.pdf/af98a8bc-688d-4bba-9748-43d648c723f4
- EUROSTAT 2002 *Handbook on improving quality by analysis of process variables*
<http://ec.europa.eu/eurostat/documents/64157/4373903/06-Handbook-on-improving-quality-by-analysis-of-process-variables.pdf/b0006e09-1708-4f8e-97e0-6a54d840b92b>
- EUROSTAT 2003 *Edimbus: Recommended Practices for editing and imputation of cross sectional business statistics*
<http://ec.europa.eu/eurostat/documents/64157/4374310/30-Recommended+Practices-for-editing-and-imputation-in-cross-sectional-business-surveys-2008.pdf>
- EUROSTAT 2003 *Quality Reporting on Flash Estimates*
Working Group Assessment of Quality in Statistics, Luxemburg
- EUROSTAT 2003 *Definition of Quality in Statistics*
[Working Group Assessment of Quality in Statistics](#), Luxemburg
- EUROSTAT 2004 *Handbook of Recommended Practices for Questionnaire Development and Testing in the European Statistical System*
<https://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10364.aspx>
- EUROSTAT 2005 *Europäische Statistiken Verhaltenskodex für die nationalen und gemeinschaftlichen Statistik Stellen*
<http://ec.europa.eu/eurostat/de/web/quality/european-statistics-code-of-practice>
- EUROSTAT 2005 *Handbook on data quality assessment methods and tools*
<http://ec.europa.eu/eurostat/documents/64157/4373903/05-Handbook-on-data-quality-assessment-methods-and-tools.pdf/c8bbb146-4d59-4a69-b7c4-218c43952214>
- EUROSTAT 2005 *Benchmarking through calibration of weights for microdata*
<http://ec.europa.eu/eurostat/documents/3888793/5836321/KS-DT-05-007-EN.PDF/07247935-0d03-47e2-90f2-b9239e7bbf86>
- EUROSTAT 2007 *Guidelines for finding a balance between accuracy and delays in the statistical surveys*
<http://ec.europa.eu/eurostat/documents/64157/4374310/28-GUIDELINES-FOR-BALANCE-BETWEEN-ACCURACY-AND-DELAYS-2007.pdf/cad273c7-8534-4c4d-8662-10adfd411b0f>
- EUROSTAT 2008 *Handbook for Monitoring and Evaluating Business Survey Response Burdens*
<http://ec.europa.eu/eurostat/documents/64157/4374310/12-HANDBOOK-FOR-MONITORING-AND-EVALUATING-BUSINESS-SURVEY-RESPONSE-BURDEN.pdf>
- EUROSTAT 2008 *Eurostat sampling reference guidelines: Introduction to sample design and estimation techniques*
<http://ec.europa.eu/eurostat/documents/3859598/5901961/KS-RA-08-003-EN.PDF/833f7740-0589-47e1-99a5-c14878a2c1a8>
- EUROSTAT 2009 *Towards an ESS structure for Good Practices*
Working Group Assessment of Quality in Statistics, Luxemburg
- EUROSTAT 2009 *ESS guidelines on Seasonal Adjustments*
<http://ec.europa.eu/eurostat/documents/3859598/5910549/KS-RA-09-006-EN.PDF>
- EUROSTAT 2009 *ESS Handbook for Quality Reports*
<https://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10243.aspx>
- Fiedler R.,Schwerer E., Berka C., Moser M., Humer S. 2010 *Quality Assessment of register-based Statistics - A Quality Framework*
[Paper presented at the Quality Conference Q2010, Helsinki](#)
- Findl Peter 2004 *Improving Efficiency and Process Quality - Statistics Austria's Experience*
Paper presented at the Quality Conference Q2004, Mainz
- Fuller, W.A. 1993. *Measurement Error Models*. New York: Wiley-Interscience
- Groves, R.M., D.A. Dillman, J.L. Eltinge and R.J.A. Little, 2001. *Survey Non-response*. New York. Wiley-Interscience.
- Hackl Peter 2004 *The Quality Assurance Committee at Statistics Austria*
Paper presented at the Quality Conference Q2004, Mainz
- Hackl Peter, Holzer Werner 2007 *Quality Management: A Pragmatic Approach, Experiences and Results*
Paper presented at 56th session of the ISI, Lisbon

Qualitätsrichtlinien der Statistik Austria

- Hackl Peter 2009 Cross national Comparability: Relevance and reality for an NSO
[Paper presented at the 57th session of the ISI 2009, Durban](#)
- Hahn, M, Lindén, H. 2006: The European Statistics Code of Practice for a High Quality European Statistical System.
Paper presented at the Quality Conference Q2006, Cardiff
- Hartmann Helmut 2002 *Technik der Statistik*,
Lernbehelf zur Grundausbildung der Statistik Austria
- Hartmann Helmut 2002 *Theorie der Statistik*,
Lernbehelf zur Grundausbildung der Statistik Austria
- Haslinger Alois 1997 *Automatic Coding and Text Processing using N-grams*.
Conference of European Statisticians. Statistical Standards and Studies –
No. 48, Genva
- Haslinger Alois 2004 *Data Matching for the Maintenance of the Business Register of Statistics Austria*
Austrian Journal of Statistics Volume 33 (2004), Number 1&2, 55-67
- Holzer Werner 2004 *Planning and Carrying out User/Customer Satisfaction Surveys by NSIs in EU and EEA Countries*
Paper presented at the Quality Conference Q2004, Mainz
- IMF 2004 *Revisions Policy for Official Statistics: A Matter of Governance*
<http://www.imf.org/external/pubs/ft/wp/2004/wp0487.pdf>
- Kutzenbergwer Ewald, Holzer Werner, Rainer Norbert 2001
Introducing Total Quality Management (TQM) in Statistics Austria
Paper presented at the Quality Conference Q2001, Stockholm
- Lessler, J.T. and W.D. Kalsbeek. 1992. *Nonsampling Errors in Surveys*. New York. Wiley
- Lyberg, L., P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz and D. Trewin (eds.) 1997.
Survey Measurement and Process Quality. New York. Wiley-Interscience
- Lyberg Lars at al. 2001 *Summary report from the Leadership group (LEG) on Quality*
Paper presented at the Quality Conference Q2001, Stockholm
- OECD 2008 *Handbook on Constructing Composite Indicators: Methodology and User Guide*
<http://www.oecdbookshop.org/oecd/display.asp?CID=&LANG=en&SF1=DI&ST1=5KZN79PVDJ5J>
- Rainer Norbert 2004 *Response Burden Barometer of Statistics Austria*
Paper presented at the Quality Conference Q2004, Mainz
- Rainer Norbert 2004 *Measuring response burden in a system of Total Quality Management*,
[International Conference: Statistics - Investment in the Future 2004 Prague](#)
- Rainer Norbert 2008 *Quality Reporting and Quality Indicators for Statistical Business Registers*
[Paper presented at the Quality Conference Q2008, Rome](#)
- Rainer Norbert 2008 *Quality of administrative data – A challenge for the maintenance of the Statistical Business Register*
[Paper presented at the UN/ECE Work-Session on Data Editing 2008 Vienna](#)
- Rainer Norbert 2009 *Business register as a tool for linking various administrative data sources*,
[International Conference: Statistics - Investment in the Future, Prague](#)
- Rainer Norbert 2010 *Quality aspects and quality criteria of a classification revision and its implementation*
[Paper presented at the Quality Conference Q2010, Helsinki](#)
- Statistics Canada 2009 *Quality Guidelines 5th Edition*
<http://www.statcan.gc.ca/pub/12-539-x/12-539-x2009001-eng.pdf>
- Statistics Finland 2007 *Quality Guidelines for Official Statistics 2nd Revised Edition*
http://www.stat.fi/org/periaatteet/qg_2ed_en.pdf
- Statistik Austria 2010 *Das Leitbild der Statistik Austria*
http://www.statistik.at/web_de/static/das_leitbild_der_statistik_austria_022224.pdf
- Statistik Austria 2010 *Ergebnisse des Belastungsbarometers*
http://www.statistik.at/web_de/statistiken/unternehmen_arbeitsstaetten/respondentenbelastung/index.html
- Templ Mathias 2008 *Visualization of Missing Values before Imputation*
[Paper presented at the Quality Conference Q2008, Rome](#)

Qualitätsrichtlinien der Statistik Austria

- UN/ECE 1994 *Statistical data Editing Volume No 1 Methods and Techniques 1*
<http://www.unece.org/stats/publications/editing/SDE1.htm>
- UN/ECE 1995 *Guidelines for the modeling of statistical data and metadata*
<http://www.unece.org/stats/publications/metadatamodeling.pdf>
- UN/ECE 1997 *Statistical data Editing Volume No 2 Methods and Techniques 2*
<http://www.unece.org/stats/publications/editing/SDE2.htm>
- UN/ECE 2000 *Guidelines for statistical metadata on the internet*
<http://www.unece.org/stats/publications/metadata.pdf>
- UN/ECE 2006 *Statistical data Editing Volume No 3 Impact on Data Quality*
<http://www.unece.org/stats/publications/editing/SDE3htm>
- UN/ECE 2007. *Managing Statistical Confidentiality and Microdata, Access – Principles and Guidelines of Good Practice*. Geneva
<http://unstats.un.org/unsd/EconStatKB/Attachment194.aspx>
- UN/ECE 2009 *Making Data meaningful*
<http://www.unece.org/stats/documents/writing/Welcome.html>
- United Nations 1994 *Fundamentale Prinzipien der amtlichen Statistik der Vereinten Nationen*
http://www.statistik.at/web_de/ueber_uns/aufgaben_und_grundsaeetze/fundamentale_prinzipien_der_amtlichen_statistik_un/index.html
- United Nations 2003 *Handbook of Statistical Organization 3rd Edition*
http://unstats.un.org/unsd/publication/SeriesF/SeriesF_88E.pdf
- Wallgren, A. and B. Wallgren. 2007. *Register-based Statistics: Administrative Data for Statistical Purposes*. New York. John Wiley and Sons.
- Wolter, K. 2007. *Introduction to Variance Estimation*. 2nd edition New York. Springer-Verlag..