



stats sa

Department:
Statistics South Africa
REPUBLIC OF SOUTH AFRICA

Quality assurance and assessment of statistics from new and administrative data sources Statistics South Africa

8 December 2022



statssa.gov.za



twitter.com/StatsSA



facebook.com/StatsSA

IMPROVING LIVES THROUGH DATA ECOSYSTEMS



stats sa
Department:
Statistics South Africa
REPUBLIC OF SOUTH AFRICA



Content

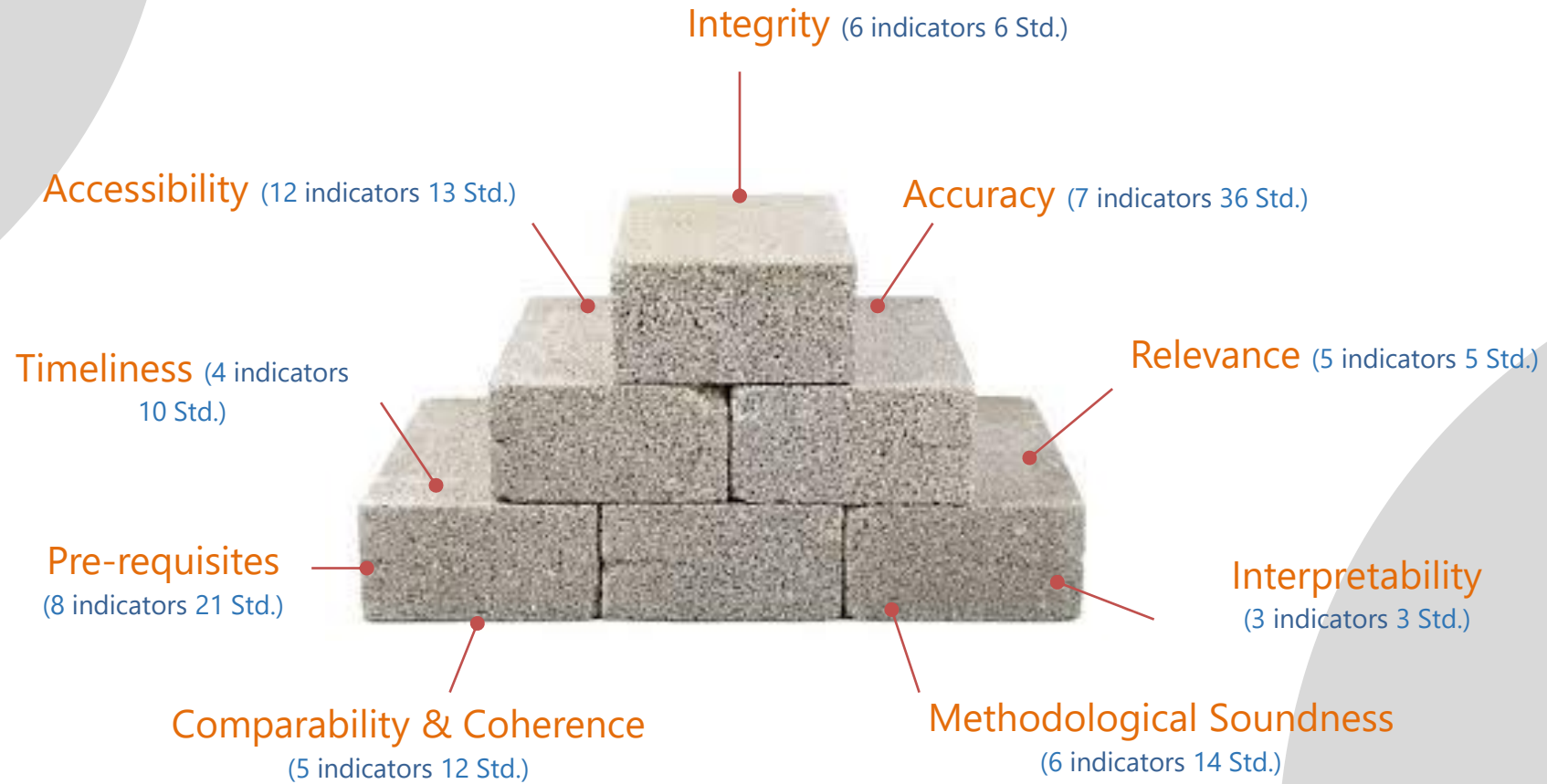
- Background
- Example of the SVC/GSBPM mapped to SASQAF indicators
- The application of SASQAF for quality assurance and assessment of admin data sources
- Case study based on other Data Quality frameworks
- Conclusion

Background: Quality Framework (SASQAF)

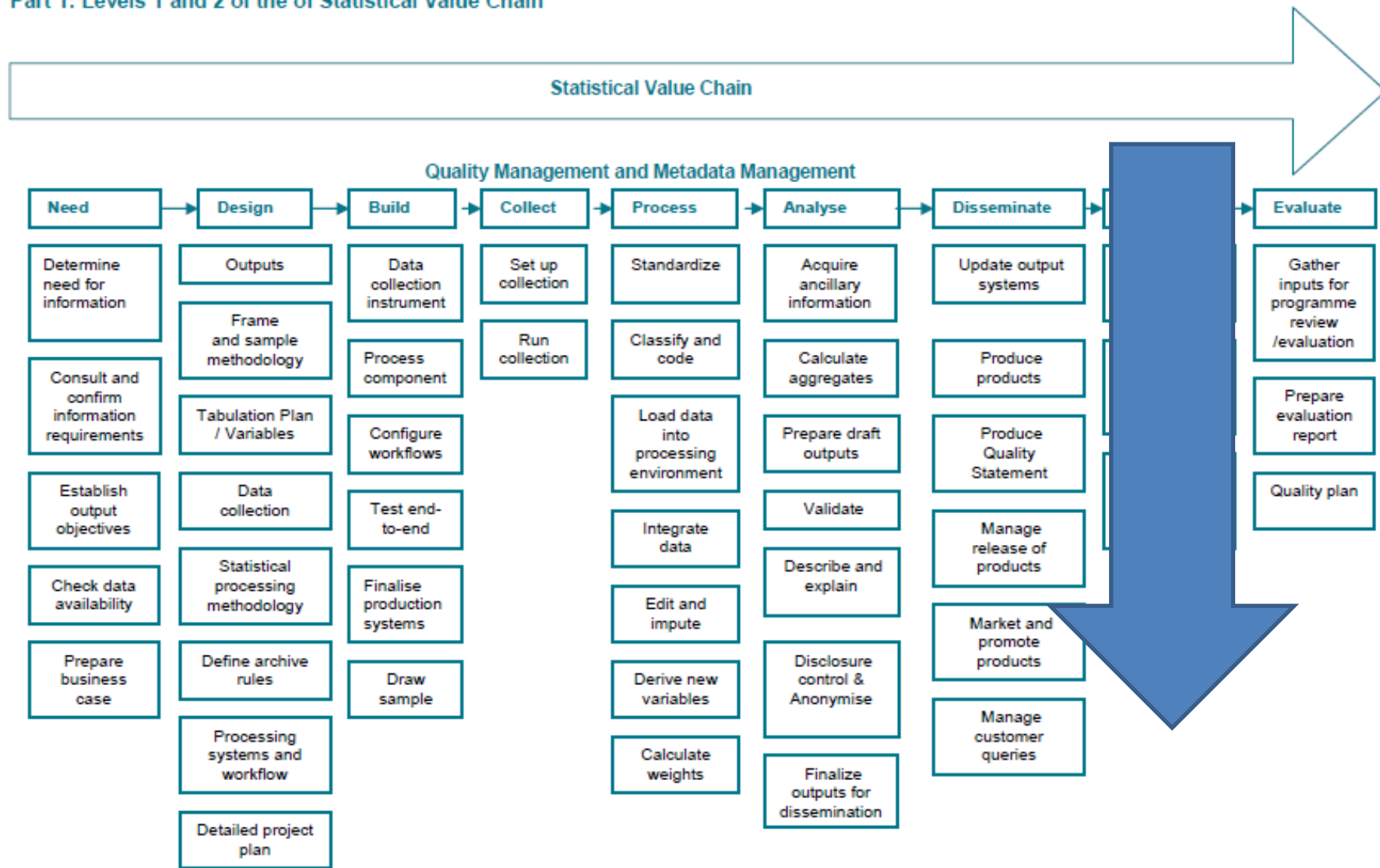
- In pursuance of S14, Stats SA developed a quality framework for assurance and assessment of statistical products within Stats SA and in the NSS community
- The main purpose of the framework is to provide a flexible structure for the assessment of statistical products
- The SG uses the framework to evaluate the quality of statistics produced in the NSS
- SASQAF is also be used as tool to improve quality of statistical products
 - *self-assessments by data producers*
 - *Used for Independent Quality assessment (DQAT)*



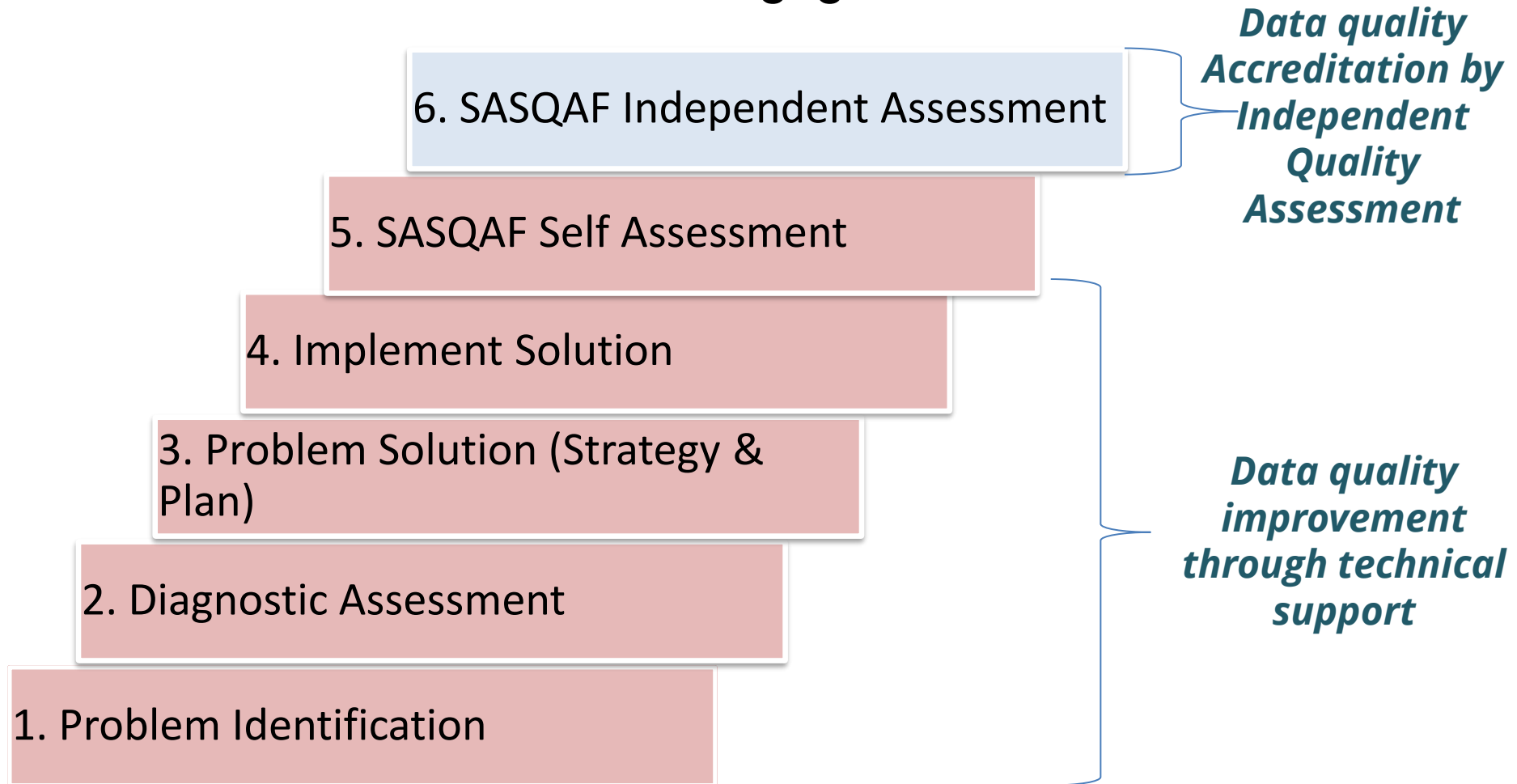
SASQAF Dimensions – Indicators – Standards



Part 1: Levels 1 and 2 of the of Statistical Value Chain



Phases of engagement



Background cont... SVC/GSBPM mapped to SASQAF indicators for Quality Assurance

Activities of the statistical value chain		Quality dimensions and indicators	
Phases	Sub-processes	Quality dimension	Quality indicator
			<ul style="list-style-type: none"> • piloting • data collection • editing and imputation of data • analysis of data • revision data
Design	Tabulation plan/variables Identify concepts	Methodological soundness	8.1 Concepts, definitions, and classifications used follow accepted standards, guidelines or good practices (national, international, peer-agreed)
Design	Statistical processing methodology	Methodological soundness	8.3 Methodologies used follow accepted standards, guidelines or good practices (national, international, peer-agreed), viz.: <ul style="list-style-type: none"> • questionnaire design • sampling • sampling frame design • frame maintenance • piloting
Process	Integrate data	Comparability and coherence	7.4 A common set of identifiers (for the purpose of record matching) exist and have been agreed upon by data producers.
Analyse	Validate	Accuracy	3.1 Measures of sampling errors for key variables are calculated. Amongst others these are: <ul style="list-style-type: none"> • standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) • design effect (DEFF)



South African Statistical Quality Assessment Framework (SASQAF)



Second edition

The application of SASQAF for quality assessment of outputs from admin data sources

Quality indicators and standards that should be selected and reported		CPI	SNAP	FCM	GHS	QLFS	SAT-DTS	Electricity	FSHE	MACOD	SAPS	SAPAD
List of indicators	List of Standards											
Accuracy												
1.1 Measures of sampling errors for key variables are calculated. Amongst others these are: <ul style="list-style-type: none"> • standard error • coefficient of variation (CV) • confidence interval (CI) • mean square error (MSE) and • design effect (DEFF). 	1.1.1 Measures of sampling errors must be calculated for the main variables. They must be available for other variables on request.	N	N	N	Y	Y	Y	Y	N	N	N	N
	1.1.2 Measures of sampling errors must fall within acceptable standards. At a minimum the following must be calculated: standard error, coefficient of variation, confidence interval, mean square error. The low accuracy of variables (if these exist), is explained. Metrics: $SE = \sqrt{\text{Var}(\hat{\theta})}$ $CV = \frac{\sqrt{\text{Var}(\hat{\theta})}}{E(\hat{\theta})}$ $MSE(\hat{\theta}) = \text{Var}(\hat{\theta}) + B^2(\hat{\theta})$ Where $\hat{\theta}$ is an estimator of a parameter of interest	N	N	N	Y	Y	Y	Y	N	N	N	N
	1.1.3 Scientific sampling techniques must be used. Metrics: a = Design effect (θ) = $1 + \delta(n-1)$	N	N	N	Y	Y	Y	Y	N	N	N	N
1.2 Measures of non-sampling errors are calculated, viz.:	1.2.1 The extent of measures of non-sampling errors must be kept to an acceptable level. Metrics:	Y	N	N	Y	Y	Y	Y	N	N	N	N

The application of SASQAF for quality assurance and assessment of outputs from admin data sources

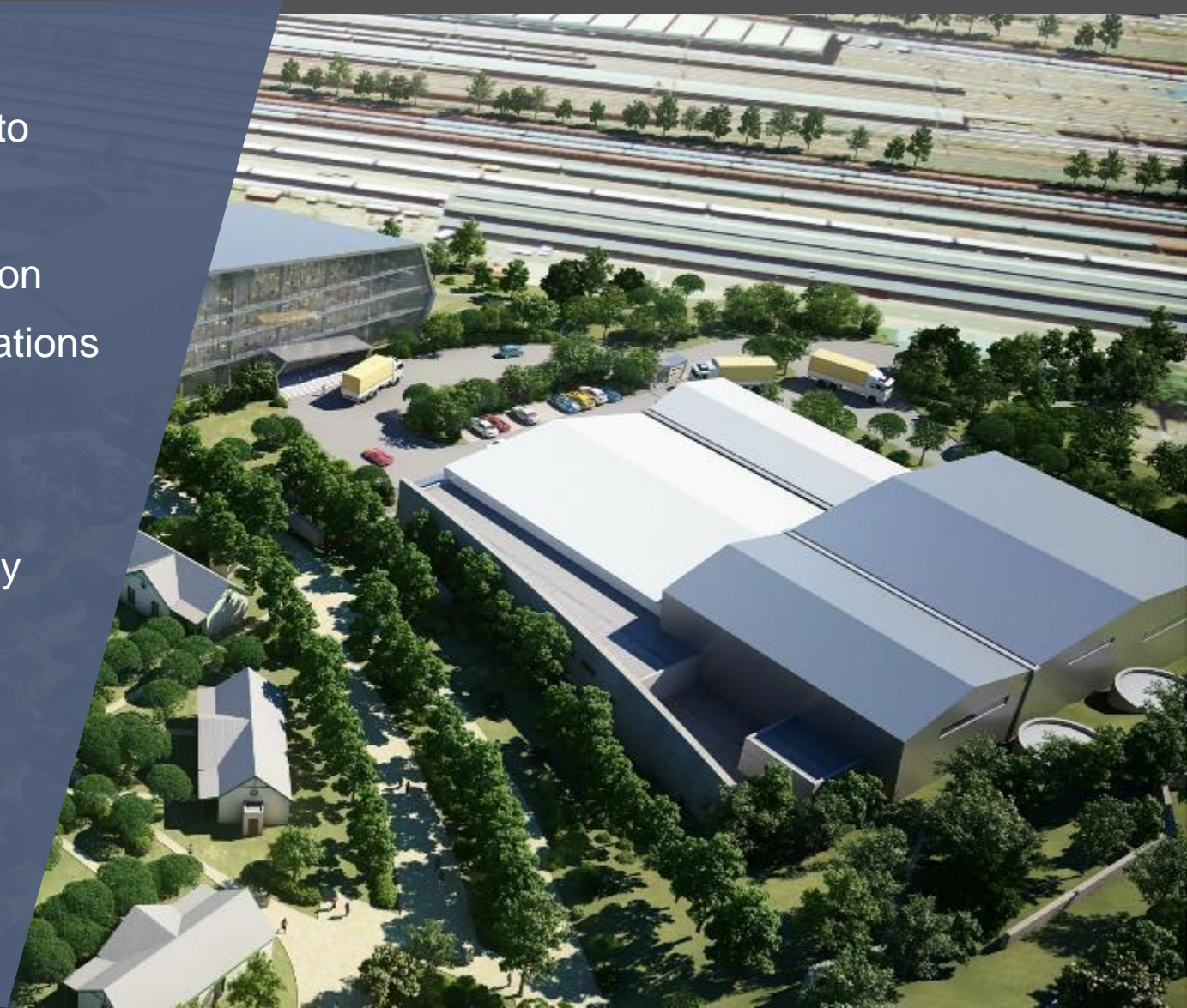
Examples of Administrative data sources

SASQAF

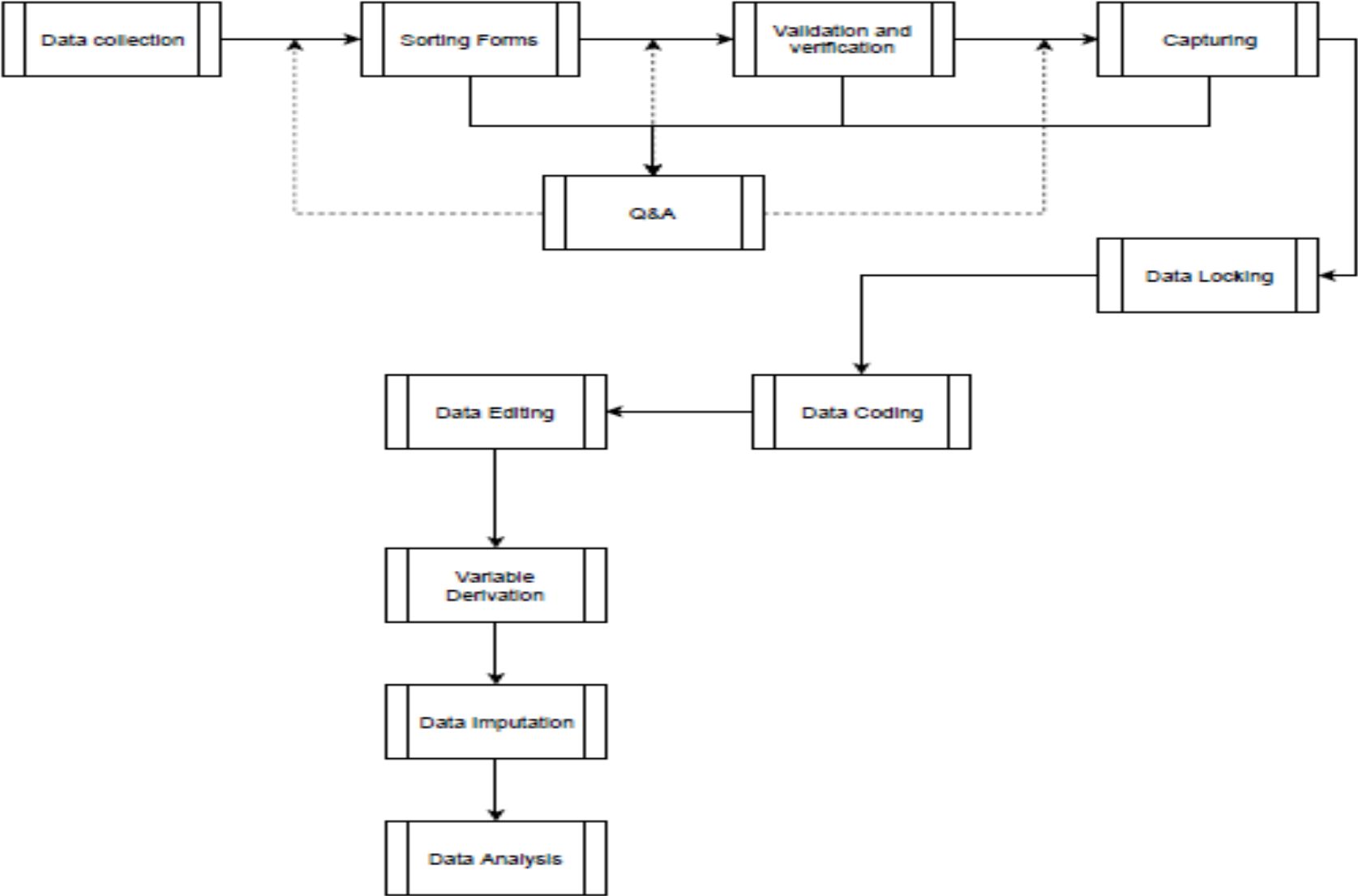
- *South African Police Services*
 - *Quality Standards to guide the collection and processing of data*
- *South African Reserve Bank*
 - *SARB – currently undertaking an online self assessment.*
- *Department of Home Affairs/Department of Health*
 - *Death Notification Forms for Mortality and causes of Death. Stats SA provide inputs into the updating the DNF Form*

Case Study: Quality Assurance for administrative data using SASQAF

- Approach
 - Selection of indicators that are applicable to administrative data sources
 - Technical support giving advice on collection methods, methodology as well as specifications on what variables to include.
- Current developments
 - Research on different administrative quality assurance frameworks (other countries)
 - Consultation (National statistical system)



Case Study: Quality Assurance for administrative data using SASQAF (SVC – Implementation phase)



Case Study: Quality Assurance for administrative data using SASQAF

Morocco	Australia	Canada	New Zealand	Proposed
Institutional environment	Institutional environment			Prerequisites of quality
Relevance	Relevance	Relevance	Relevance	Relevance
Accuracy	Accuracy	Accuracy	Accuracy	Accuracy
Timeliness	Timeliness	Timeliness	Timeliness	Timeliness
Accessibility	Accessibility		Accessibility	Accessibility
Coherence	Coherence	Coherence	Coherence	Coherence
	Interpretability		Interpretability	
Coordination and Corporation				

Conclusion

Application of SASQAF when assessing statistics from the administrative data sources

- Status quo
 - SASQAF 2nd edition
 - Currently developing a separate checklist/framework for quality assessment of administrative data sources based on the General SASQAF.
- Lesson learnt:
 - Some of the indicators/standards are not applicable to administrative data
- Challenges
 - Takes time to identify the indicators that are not applicable
- Way forward/Recommendation
 - develop a separate SASQAF for administrative data sources

Quality Framework

Thank You



statssa.gov.za



twitter.com/StatsSA



facebook.com/StatsSA

IMPROVING LIVES THROUGH DATA ECOSYSTEMS



stats sa
Department:
Statistics South Africa
REPUBLIC OF SOUTH AFRICA

