ESA/STAT/AC.202/S8.2

International Seminar on Early Warning and Business Cycle Indicators

14 to 16 December 2009
Scheveningen, The Netherlands

Creating statistical information; visualisations and coherent indicator sets

Floris van Ruth
Statistics Netherlands

"Creating statistical information; visualisations and coherent indicator sets"

Paper for the International Seminar on Early Warning and Business Cycle Indicators 14-16 December 2009, Scheveningen, The Netherlands

Floris van Ruth and Symon Algera

Introduction

Data, or statistics, by themselves are not information. For numbers to gain meaning, something needs to be added This can take the form of analysis or interpretation. But this is both labour intensive and potentially problematic for statistical agencies as it might introduce a subjective element. Another approach to giving meaning to statistics is to provide context by structuring statistical information. A very basic, and standard, way to do this is by reporting statistics together with previous periods' estimates. Adding other economic indicators relating to the same time period is a slightly more advanced variant. Most analysis is in the end based upon placing realisations into context by relating them to other (statistical) factors or earlier developments. A fruitful approach is therefore to exploit the structure and relationships inherently present in statistical data. The aim is to find a way to present these as objectively as possible, and in a data-centred fashion.

Things rarely develop in isolation, and there is a lot of knowledge present in statistical agencies and the academic community on connections and structure in the economy. Often, users are not only interested in a statistic per se, but also in the wider phenomenon it represents or is connected to. The art (science?) of constructing coherent indicator sets aims to help the user by structuring statistical information in such a way that relevant connections become visible, preferably yielding a comprehensive picture of the central theme. An important additional function of indicator sets is that they can act as a kind of guide in the sometimes bewildering supply of statistics. Most non-expert users tend to find it hard to interpret raw statistical information, and certainly to find the indicators most relevant for their information need. Combining statistics adds value to the individual indicator as it results in more reliable interpretations and shows their broader relevancy. This approach has also been described as statistical storytelling. Its essence is selecting and structuring statistical information, thus making connections visible and yielding a comprehensive picture of the central theme.

This paper aims to give a broad sketch of the nature of coherent indicator sets, and how they can be used to transform data into statistical information. It also gives an outline of the methodology developed at Statistics Netherlands for constructing coherent indicator sets and visualisations. This will be illustrated using examples from Statistics Netherlands business cycle reporting. But the next section will first explain how coherent indicator sets can help meet the demand for more and faster information on current developments.

Coherent indicator sets and the need for faster information

This approach also ties in with the current need of quick and reliable information to monitor the current crisis and with which future crises can be signalled earlier [OECD 2009]. Many users consider the Quarterly Accounts to slow for this purpose. On the other hand the various monthly indicators are considered to be too fragmented. It is often unclear which monthly indicators should be monitored, and what the overall situation is. In order to monitor the crisis, Statistics Netherlands has developed an intermediate product, which combines the timeliness of the monthly indicators with the summarizing properties of the Quarterly Accounts, this is the Business Cycle Factsheet (BCF) . The BCF and its components can be considered as coherent indicator sets. Coherent indicator sets are a new element of the statistical program, which can be illustrated by the place of the Business Cycle Factsheet in the system of statistics.

The Dutch system of statistics can roughly be described by a matrix with the rows containing the time dimension and the columns the degree of integration [Algera 2005]. The matrix shows in a glance the main characteristics of the various statistics and the interrelationships between them .

Statistical matrix (with some relevant examples)

Time	*Degree of integration*		
	Single	*Combined*	*Integrated*
Future Month	Tendency surveys Short-term statistics	*Business Cycle Factsheet and coherent indicator sets*	
Quarter			Quarterly national accounts
Year	Structural business statistics		National accounts

The time column distinguishes statistics relating to future, monthly, quarterly and annual data.

Under degree of integration:

- the column headed single contains statistics which are obtained from a single survey of statistical units. The data are simply the survey findings;
- in the combined column the statistics are the outcome of combining different statistics (and different surveys), but are not yet integrated;
- the integrated column contains data that are the outcome of an integration process involving detailed checks of all available information. The national accounts and the quarterly national accounts are examples of integrated statistics.

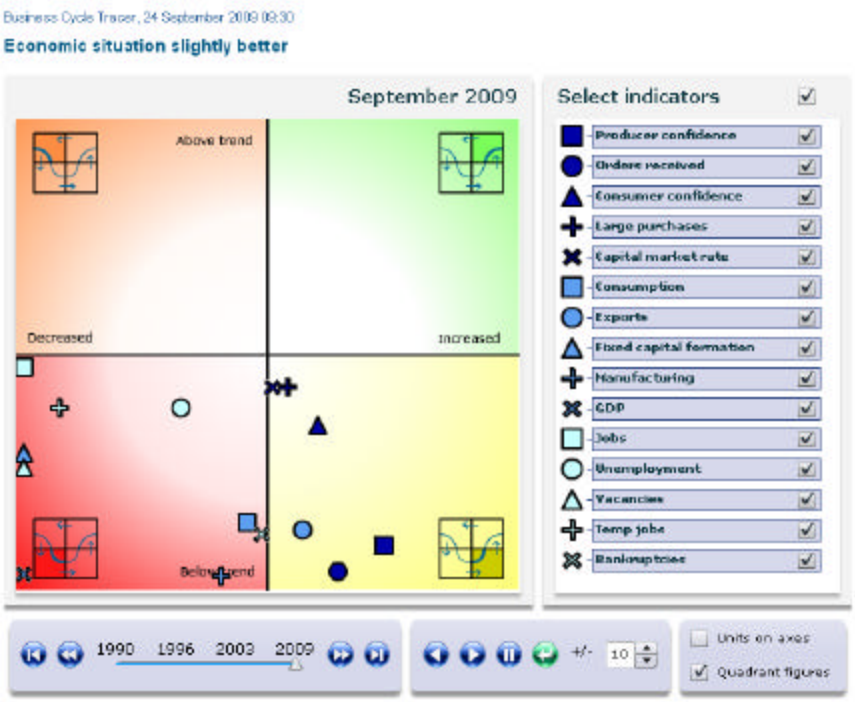
Such a matrix can be used to explain some relevant aspects of coherence. Reading from top to bottom of each column, the reported results become more reliable and detailed, but take longer to publish. Reading from left to right, the data become more comprehensive and reliable, but again as a rule are published later. Moreover, as a consequence, the quality of statistics can be assessed in two ways, one in terms of predictive power (the columns of the statistical matrix), the other in terms of consistency (the rows of the statistical matrix).
The Business Cycle Factsheet and its components (Business Cycle Tracer, BCT indicator, Business Cycle Dashboard and the Exports Radar) are based on tendency surveys and monthly statistics and on compound indicators/infographics. So, within the matrix above, the Factsheet is located on the rows *future* and *month* and in the column *combined*, in the shaded area.
Though not being fully integrated, the Factsheet provides a first insight into the development of variables in the integrated statistics. Structuring indicator sets, the intermediate stage of integration, is a way to achieve some of the advantages of National Accounts style integration, but with greater flexibility and timeliness. It is also possible to construct indicator sets for different economic phenomena, not just national production.

It should be stressed here that the coherent indicator sets, at least those developed by Statistics Netherlands, do not aim to model or predict the associated phenomena. They are solely meant as a tool for monitoring and analysis. The indicators are not selected based on their leading character or forecasting power. Therefore, they are not built for early warning. But, and this is an important but, they do show what factors are potentially relevant for the development of key economic quantities. If one knows which indicators influence realisations, and how economic variables are connected to general economic conditions, then future realisations and developments should come as less of a surprise.

The working of coherent indicator sets

The amount of information embedded in a coherent indicator set and how it can be used to assess current and coming developments can be illustrated by the Statistics Netherlands Business Cycle Tracer. This is a tool which has been especially developed for the monitoring of the Dutch Business Cycle [Van Ruth et al. 2005]. It consists of a set of fifteen carefully selected and filtered macro-economic indicators, which are placed in a diagram according to their medium-term development (above or below trend) and their short-term development (increasing or decreasing).

The Statistics Netherlands Business Cycle tracer for September 2009; a transition from the recession quadrant (bottom left) to the upswing quadrant (bottom right) can be seen.

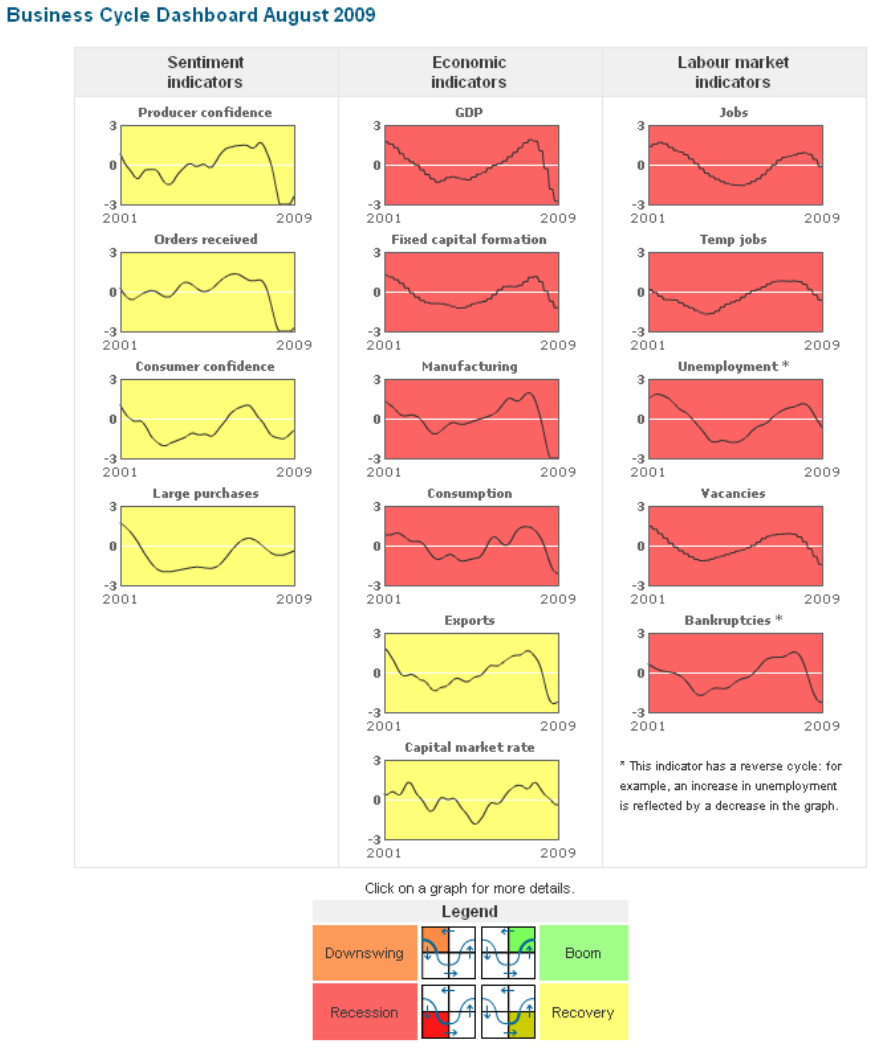


The whole is calibrated to be coincident with the Dutch business cycle, but the indicator set is a (balanced) mix of lagging, coincident, and leading indicators. There are two aspects of the composition of this indicator set which are especially relevant for communicating statistical information. For a start, the Tracer is a concretization of the concept of the business cycle. It shows the business cycle as a phenomenon that is present in all or most major economic indicators, and it shows which indicators are relevant for monitoring the business cycle. The other important piece of information it communicates is that economic indicators have a lagging, coincident or leading character compared to the general business cycle. Thus, by watching the leading indicators one has some warning concerning future business cycle developments. On the other hand, it shows that it may take some time for business cycle developments to manifest themselves in lagging indicators, giving an indication of what to expect there. The time function and interactive features of the Business Cycle Tracer application allow users to explore these relations and dynamics for themselves. Also, this graphical representation and visual interpretation is often easier and quicker to understand than a table or even a textual analysis.

Role of visual applications

The coherent indicator set approach to reporting statistical information is greatly boosted by the possibilities offered by web-based applications. There is currently a shift in emphasis from reporting numbers to offering alternative presentations and analytical tools. This has been made feasible by the possibilities the internet offers for constructing interactive and dynamic applications. These are two mutually reinforcing developments; coherent indicator sets give meaning to interactive applications and the applications allow for new methods of presentation. Dynamic and interactive options allow the user to explore the phenomenon and connections himself. One should never underestimate the power of visual representations. For most users, these are much easier and faster to interpret than data in tabular form. Perhaps even more important, skilfully constructed visualisations can, in a glance, make visible the overall situation. A powerful example of this property of visualisations is the Business Cycle Dashboard, which shows the individual graphs of the component indicators of the Statistics Netherlands Business Cycle Tracer.

Business Cycle Dashboard; the cyclical development of the component indicators of the Business Cycle Tracer. The colour of the graph corresponds to the colour code of the relevant business cycle phase from the Tracer diagram. Indicators are grouped into sentiment indicators, economic indicators and labour market indicators.

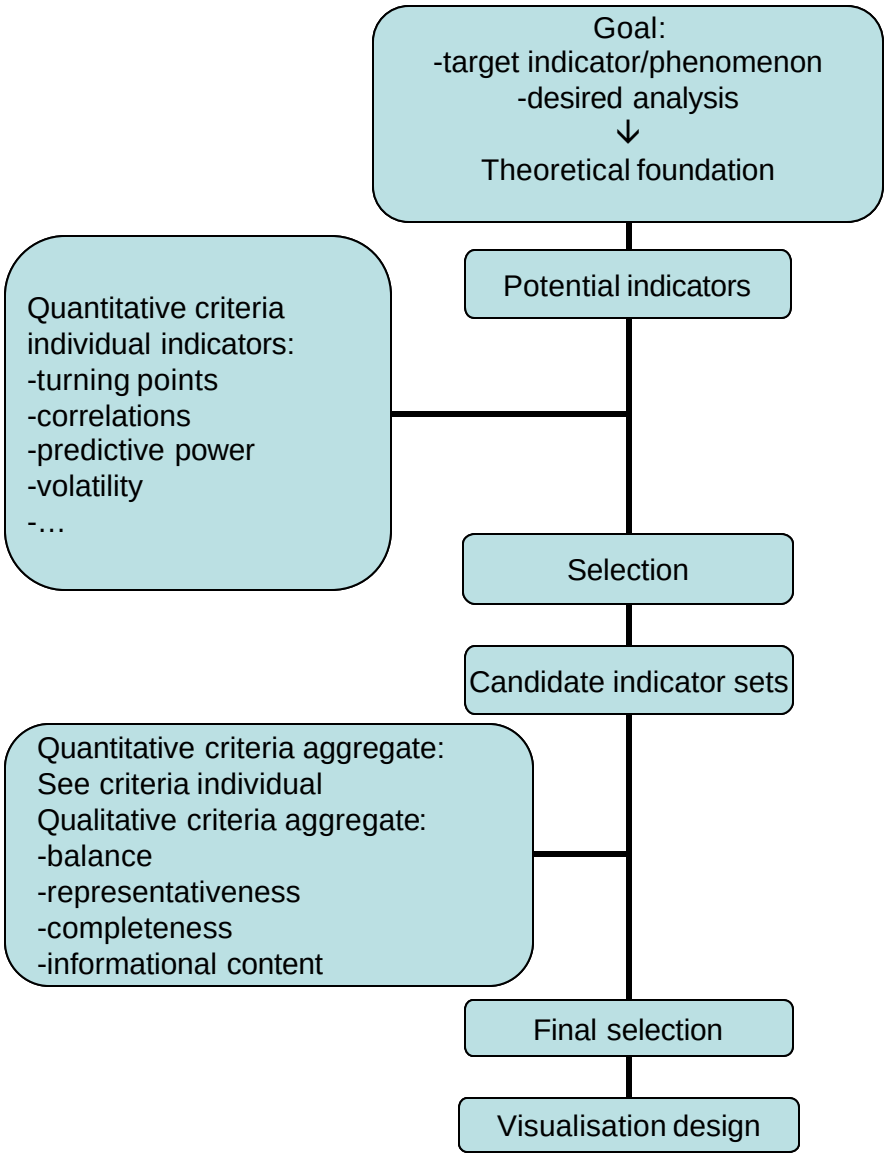


Each graph is given a colour, corresponding to the business cycle phase of the component indicator. In a glance, the overall distribution of the individual indicators is visible, and also the resulting overall situation. The colours are a very powerful tool for characterising and communicating the development of the individual indicators.
At Statistics Netherlands a program is underway to provide access to important statistics via interactive and graphical applications. >>**link CBS interactief**<< These range from dynamic maps, via customisable graphs to somewhat more novel applications.

Methodology

The value and credibility of an indicator set and visualisation depend critically on the selection of the indicators and the design of the monitoring tool. Fancy graphics alone won't make a tool useful, and a confused concept or sloppy indicator selection will destroy its credibility. The development process is part science, part art. For the tool to be considered objective and reliable, the development process needs to be based as much as possible on theoretic and quantitative considerations. But to end up with a useful and comprehensible tool, a certain amount of judgement is inevitable. As far as the actual visualisation is concerned, certain aspects are generally desirable, such as a time function, user customisation options, clarity and use of colour codes. In general however, each concept will lead to a different form of visualisation, which should be determined by the phenomenon concerned.

At Statistics Netherlands , a general methodology has been developed for the construction of indicator sets. It is a chain of several steps, which act as a kind of sieve where, based on a conceptual framework and ensuing conditions, potential indicators are selected or rejected. This can easily represented in a flow-chart, see the diagram below. This however, does not do justice to the paramount importance of the first step, deciding exactly what it is one wishes to monitor. Another way of formulating this is defining which questions does one want to help answer with the monitoring tool. A clearly defined and well-thought out concept will make the development process go much smoother, and result in a much more focused and therefore useful end product. The relevant considerations and criteria needed in the following steps of the selection and design process will also follow more naturally from a well thought-out concept.
The next phase is using results from theory and empirics in determining which general factors determine or are connected to the selected phenomenon. The outcomes of this analysis will guide the next step, the selection of candidate indicators. This is actually the first step of the true selection process. The following steps have a more quantitative character, as the potential indicators are benchmarked against the quantitative selection criteria. From this follows a list of well performing indicators, which can be grouped into different candidate indicator sets. These can be scored on aggregate performance according to the quantitative and qualitative criteria formulated in the beginning. This will most likely be an iterative process, as indicators are dropped and added in trying to achieve an optimal composition.

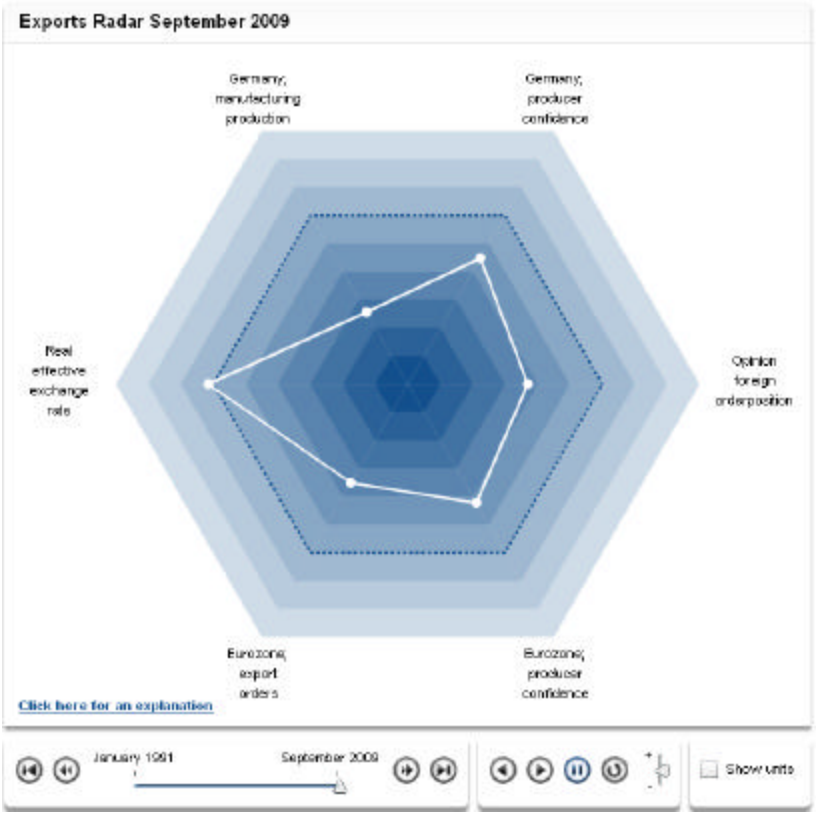


The final stage of the development process is the design of the visualisation itself. This is an altogether separate process from the indicator selection. For a certain indicator set, a variety of visualisations will in general be possible, though some might be more optimal for the chosen aim than others. It can be advantageous to extend the development process by presenting the finished product to strategic partners and potential users. While not really part of the development per se, this can result in useful feedback, but also builds credibility and support. It is more or less common practice at Statistics Netherlands to ask a wide variety of relations for feedback, which can result in important new insights.

An illustration; the Export Radar and the Economic Dashboard

An interesting illustration of how this process works in practice is the development of the Export Radar [Van Ruth 2009b]. This is a visual tool for analysing export conditions, part of the Business Cycle Fact Sheet. It consists of six economic indicators, all relevant for Dutch exports. Together, they show whether conditions are favourable or unfavourable for Dutch exports. Using the time function, it is also possible to see whether conditions have improved or deteriorated compared to the previous month or any earlier time period.

The Export Radar; a graphical representation of a structured indicator set reflecting conditions for Dutch exports.

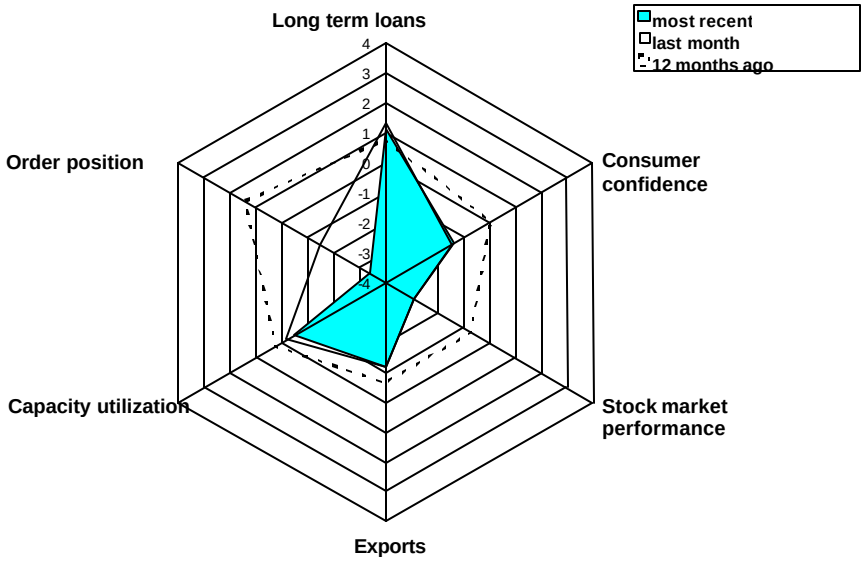


The origin was a desire to find a way to place the realisations of key economic variables into context. The aim was to sketch the environment of influences which drive export development. Ideally, the tool should be able to assist in answering questions like “why have exports grown/declined (this much)?”. Having defined the general objective, the next step was to make the concepts “environment” and “context” more concrete. Fundamental to the concept of the Export Radar concept is the identification of factors which determine the development of exports. For exports this relatively simple, competitiveness and developments in the main markets, Germany and the rest of the Euro-zone. An important qualification is that the emphasis here was on the short- and medium-term, as we were developing a tool for business cycle analysis. A variant focussing on the longer term development would probably

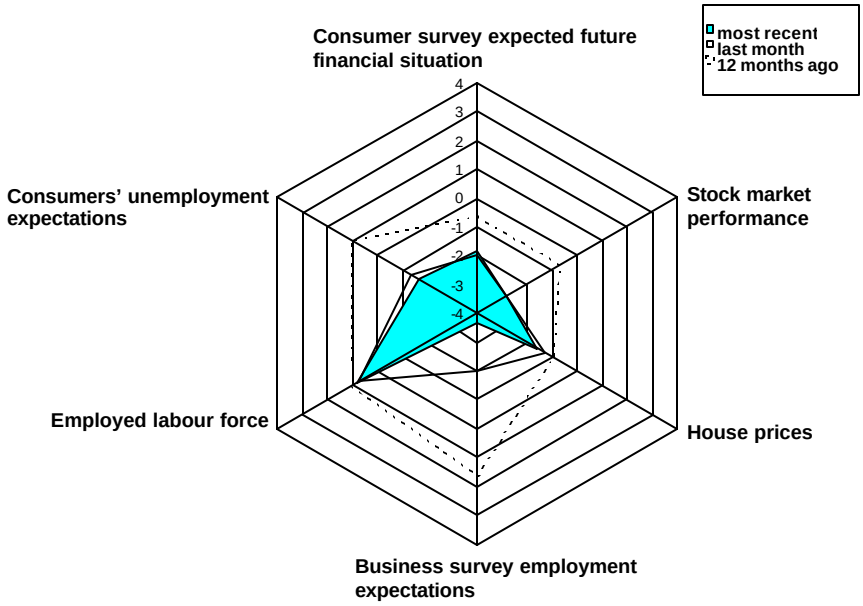
identify different factors and indicators. As the aim was to develop a system identifying drivers of export growth, correlation analysis and predictive testing were natural selection tools to identify the strongest potential indicators. Things were somewhat less straightforward in the construction of the optimal indicator set.There were quantitative criteria for evaluating the aggregate, such as correlation with export growth, turning point correspondence and communality of the indicator set. But these had to be balanced by more qualitative factors such as representativeness of the indicators for the underlying factors, the fact that more smoothly developing indicators are more informative, and inherent informative value of the indicators. Thus, there is a certain amount of judgement involved in the selection of the final indicator set. When the selection phase of the development has been completed, it is time to start thinking about the presentation. Remember, the aim was both to characterize the conditions for exports, and to show how the different aspects of export conditions are influencing these. The refore, a visualisation was needed that gives both an overall picture, and also shows how the component indicators are developing. That is why the choice fell on the radar (or spider) diagram, where the surface area gives an indication of the overall state, but the component indicators are all individually depicted. Note that this diagram does not attempt to quantify either the overall conditions or a corresponding export realisation, nor does it attempt to do an pronouncement on the exact strength of the influence of the different indicators. It is purely an analytical tool which structures statistical information. But by doing so, it shows relationships and gives information not readily available before.

The Export Radar is itself however only a component of a larger indicator set. Statistics Netherlands plans to introduce an Investment Radar and a Consumption Radar as well [Van Ruth 2009a,c]. These will complement the Export Radar, and together they will form an Economic Dashboard, which gives a complete overview, in the expenditure approach, of conditions in the private sector of the economy. This is a structured and visual variant of the classic table with core economic data.

The Investment Radar; a graphical representation of a structured indicator set reflecting conditions for business investment in The Netherlands.



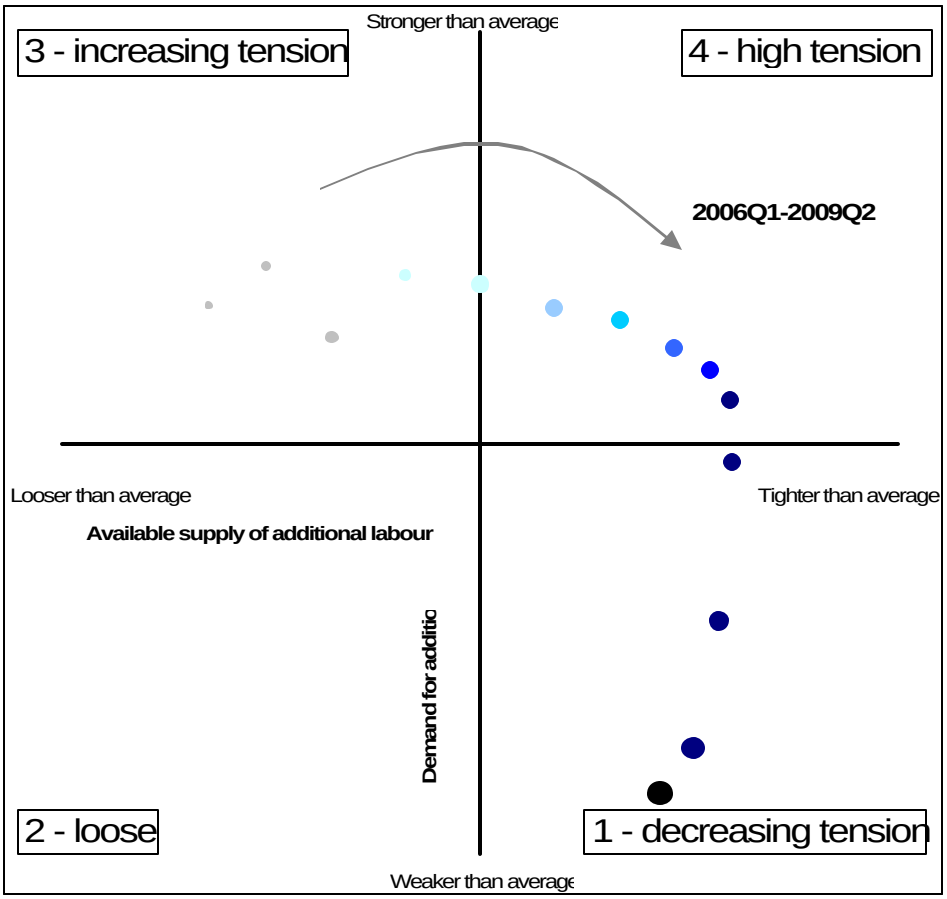
The Consumption Radar; a graphical representation of a structured indicator set reflecting conditions for Dutch household consumption.



A few more examples

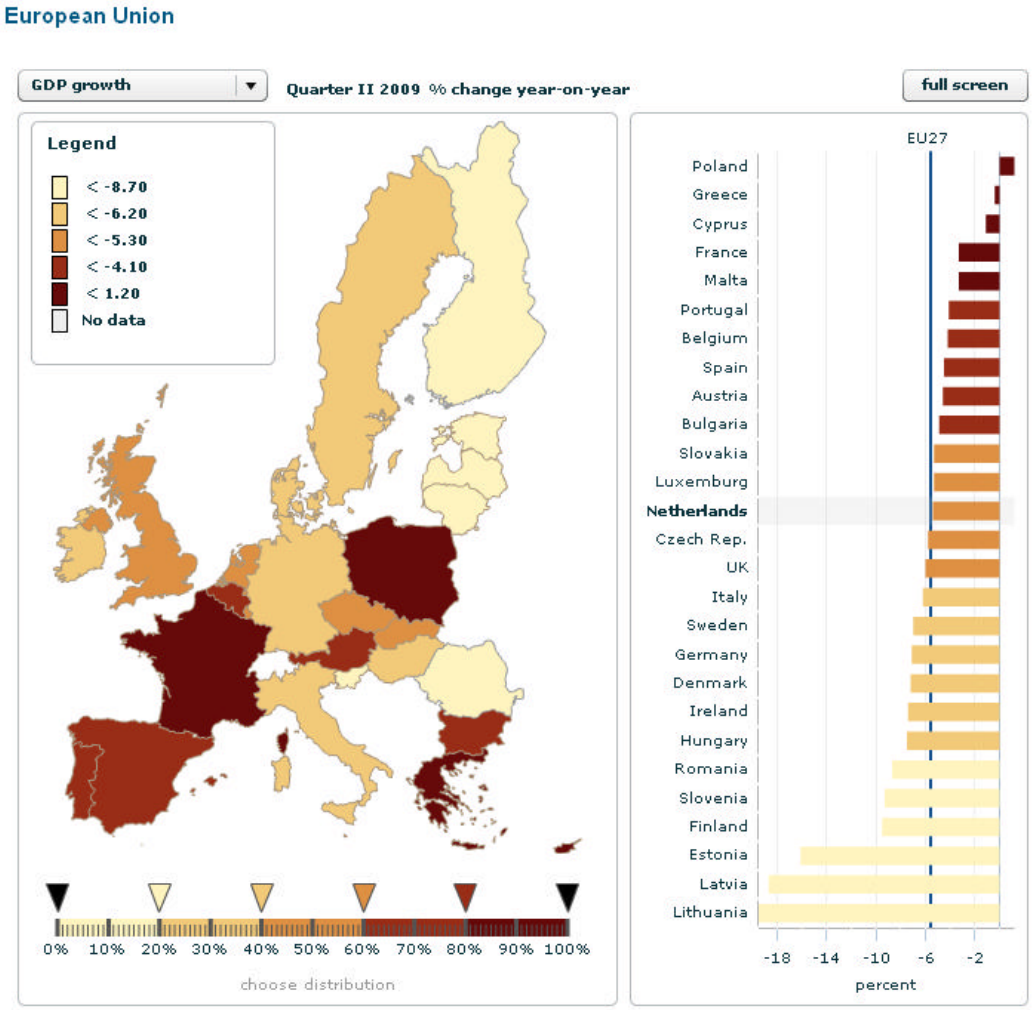
To complete this paper two more examples of visualisations of economic data will be shown, to illustrate the variety in potential approaches. The Labour Market Tension Gauge is still under development and very experimental. It takes the concept of indicator sets and statistical visualisations one step further. It is a tool for analysing labour market developments, deriving global characterisation of the state of the labour market from selected relevant statistics [Van Ruth 2009d]. It is an example of how showing related statistical indicators in a structured fashion can yield information not present in the individual indicators. A graphical representation does this much clearer and is easier to interpret than a table containing the same statistics. The basic idea is that to fully characterise the situation on the labour market, both supply and demand indicators are needed. The vertical axis shows the intensity of demand for additional labour, the horizontal axis the relative tightness of labour supply. Each quadrant of the diagram represents one of four possible states of the labour market. For a non-expert user, this offers a relatively easy to understand characterisation of the labour market, and it shows that several types of statistics, and more importantly which ones, are needed for a meaningful analysis of the state of the labour market.

The Labour Market Tension Gauge; A graphic method for characterising the state of the Dutch labour market, based on confronting indicators of supply and demand for additional labour.



A more conventional visualisation is the comparison of key economic indicators in the EU from the Statistics Netherlands Economic Monitor. It shows a chart of the EU, with the option of choosing one of six economic indicators. The countries of the EU are then colour coded according to the value of their realisation of the chosen statistic, while the graph to the right of the chart the actual values are shown, ordered by size. This is a particularly effective way to meet the demand for between country comparisons of economic performance. The country shapes on the map are immediately recognisable, while the intensity-based colour code gives an instant impression of the relative development. Thus, this visualisation is much easier and faster to interpret than a traditional table containing the same data.

A graphical presentation of the comparison of different key economic indicators in the European Union.



Source: Statistics Netherlands/ Eurostat

Conclusions

This paper has aimed to show that coherent indicators sets and visualisations are powerful ways to extract and make widely available the information present in statistical data. Structuring statistical indicators is a way to add value to individual indicators and to better meet the information needs of users. Explicitly showing connections between economic indicators does not only communicate important knowledge, but also means that new developments come less as a surprise. Coherent indicator sets and the visual and interactive possibilities offered by the internet are mutually reinforcing phenomena. Without the interactive and visual possibilities, a large fraction of the information present in the indicator sets would remain hidden for most users. But it has also hopefully been shown that without a well-thought out concept and a rigorous development process, illustrated here, a visualisation will never reach its full potential.

References

Algera, S.B., 2005, "Statistical Benchmarking", paper presented at the OECD/Eurostat workshop. April 2005, Luxemburg.
OECD, 2009, "Implications of the crisis for statistics: towards an OECD work plan", OECD, Paris.
Ruth, F.J. van, 2009a, "A conditions monitor for household consumption. Statistics Netherlands Discussion Paper 09006, The Hague/Heerlen.
Ruth, F.J. van, 2009b, A conditions monitor for exports.", Statistics Netherlands Discussion Paper 09007, The Hague/Heerlen.
Ruth, F.J. van, 2009c, "A conditions monitor for fixed capital formation.", Statistics Netherlands Discussion Paper 09008, The Hague/Heerlen.
Ruth, F.J. van, 2009d, "Analysing the labour market: a tension gauge.", Second draft. Statistics Netherlands.
Ruth, F.J. van, J.G. Schouten & R.F. Wekker, 2005, "The Statistics Netherlands' Business Cycle Tracer. Methodological aspects; concept, cycle computation and indicator selection.", Second draft. Statistics Netherlands. (via www.cbs.nl)
Vlag, P., Algera, S.B., and Van Cleef, B., 2009, "Measuring business cycles with rapid estimates", Statistics Netherlands Discussion Paper 09038