Explore, gain insight, create Stories with Snapshots and Publish

Select statistics indicator data – explore, inquire, animate in Statistics eXplorer - make discoveries through trends and patterns and derive insight – important visual discoveries are captured into snapshots together with your descriptive and analytics reasoning text - gained knowledge is then the foundation for creating and exporting your story that can be shared with colleagues and reach consensus and trust for your discoveries and knowledge. Get feedback from colleagues and adopt the story - finally use Statistics Publisher to deploy your story based on interactive “Vislets” loaded with your preferred indicators and snapshots –copy the automatically created HTML code with links included to geographic and data sources - publish by embedding the HTML code into blogs or Web pages accessible to thousands of people that can interactively follow your logic reasoning.

Statistics data have great potential to generate knowledge and serve as basis for decisions taken by many actors in society. NComVA introduces innovative tools web-enabled statistics visualization software for exploring, presenting and publishing regional statistics data for a single year or animated time series based on richer and more dynamic visual user interfaces. Statistics eXplorer facilitates methods for exploring geographic statistics that can uncover hidden structures and relations and let the analyst present her findings through storytelling to a broad audience. The combination of Statistics eXplorer (explore, gain insight and knowledge) and Statistics Publisher represents the tool for an integrated statistics analysis, collaboration and publication process facilitating storytelling aimed at producing statistical news content in support of an automatic authoring process. The analyst can press a button in Statistics Publisher that creates the HTML from a visual discovery process in Statistics eXplorer and can then be deployed in any blog or web site. This process focuses on the most ancient of social rituals “Storytelling” - telling a story about a region’s development over time and shape the measure of economic growth and well-being. Discoveries that more engagingly draw us into reflections about the knowledge on how life is lived - and can be improved – from region to region and in addition let the reader dynamically participate. Statistics eXplorer and Publisher represent the platform for dissemination of embedded dynamic statistics data visualization with the analytics sense-making metadata (story) based on snapshots, associate text or sound joined together and publishable in any web pages such as blogs, wikis or web pages. Publishing official statistics through assisted content creation with emphasis on visualization and metadata represents a key advantage of our storytelling and has the potential to change the terms and structures for learning and sharing knowledge.
Statistics eXplorer is based on our own developed solid object-oriented component platform GAV Flash framework adapted for Adobe’s Flash basic graphics and Flex and does not require installation of any other software and will run anywhere.

GAV Flash includes a collection of suitable geo-visualization and information visualization methods (see figure) specially adapted and customized to the behaviour and requirement of large multidimensional and spatio-temporal statistics data. Interactive features that support a spatial analytical reasoning process are applied such as tooltips, brushing, highlight, visual inquiry and conditioned statistics filter mechanisms that help detecting outliers.

GAV Flash is built upon Adobe Flex, Statistics eXplorer and takes advantage of the Flex user interface functionalities. Buttons, panels and sliders are integrated with visual representations from GAV Flash and represent a thoroughly tested solid and platform. Its open architecture allows new or enhanced methods to be incorporated as demonstrated in new releases.

All GAV Flash components rest on different basis such as visualizations, managers or data providers and interact with each other, without knowing the exact nature of the others. By separating the data structure from the visual representations, Statistics eXplorer can work regardless of input. Data are normally preloaded with most basic indicators such as demographics, economic, education etc. but the user can also insert external (own) data to be mixed with preloaded data based on standard spreadsheet format. Coming version eXplorer V4 opens for optional data base interfaces such as SDMX, PC-AXIS or customized solutions.

In order to detect complex patterns it is convenient to view statistics data through a number of different visual representations simultaneously, each of which is best suited to highlight different features. The linked views are coordinated using the GAV Flash data linking method based on the data cube model and colouring scheme. Any filtering, highlighting or colouring made in one of the linked views is transmitted to all the others.
Statistics eXplorer opens default with three frequently used views: choropleth map, scatter plot and histogram:

These multiple linked views are separated by interactive splitters, allowing the user to scale the sizes of individual views focusing screen area on important discoveries. The scatter plot share views with the data grid and table lens while the histogram shares with the time graph and parallel axes chart (“Profile Plot”). Space-time-indicator awareness means that data are analysed and visualized across all three dimensions simultaneously. The scatter (bubble) plot is an effective time animation method and here integrated with a map and the histogram but could also be the time graph (see below) or a profile plot and makes motion visualization even more efficient. The bubbles are displayed at the intersection values for two chosen indicators on the X-axis and Y-axis. The plot also shows two extra dimensions through the use of colour which corresponds to that on the map and through the size of each bubble, so the number of indicators simultaneously analysed is increased to four.

Efficient animation simultaneously in all views, requires a data model optimized for handling spatio temporal and multidimensional data. Our GAV Flash data model can be seen as a data cube with three dimensions: space, time and indicators. The spatial dimension is represented by the regions, the indicators (GDP growth, elderly dependency rate, etc) and time is represented by the yearly values of the indicators. The general method for finding a value in the cube is by its position (where, when, what).

Our state-of-the-art colour legend facilitates statistical methods for class value calculations and scientifically tested perceptual acceptance of colour scales. Statistics eXplorer’s standard approach of the visualization scales the colour map according to value density based on percentile calculations, which means that value ranges with a high amount of regions receive a larger range of colours to help distinguish them. The dynamic handlers in the colour legend can be moved to assign preferred class values.

A filter handle moves along the colour legend and removes regions outside the range value.
A diverging color scale is used to explore net migration flow for European TL3/NUTS3 regions. 0.0 is assigned to the centre point (where blue change to yellow/red).

**Map layer architecture** - overlay several types of maps such as colored statistical regions, country boundaries, background maps e.g. Google or Bing Maps, and pie chart objects – control transparency level for each layer – a feature typical for an advanced GIS application. Switch layers at speed of light.

Five map layers are applied in the Flow Map application.

In the pie chart, the arc length (size of pie chart) of each sector can be set proportional to the quantity (population) it represents. The position of the pie chart within a region is calculated by a sophisticated method that search for a representative point. Two sliders control the size of pie chart (arc length) and the opacity (see through the pie chart). A special pie chart color legend is also displayed in the Map view.

A pie chart layer (age group 0-14 and 65+) is superimposed on this NUTS2 map of Europe displaying GDP Growth rate. We see ageing population coincident with low GDP Growth rate (blue).

Dynamic histogram exploring and comparing world population ages 0-14 and 65+ ordered after 65+. All countries are viewed together.

The dynamic Histogram chart can be a significant visualization complement providing actual ordered values, coordinated and linked with the map and scatter plot. Histogram can be displayed in both vertical and horizontal mode. Histogram is based on a focus & context method using fisheye technique meaning that more bars than available space can be displayed.

The histogram also displays multiple indicators (up to four) as stacked bars or single bars. Here we have three indicators showing age population for 0-14, 15-64 and 65+ that should add up to 100%. We can see the difference between using stacked or single bars for comparison. The panel is an example of a hidden GUI interface that here controls attributes for the histogram. All GUI panels are hidden in Statistics eXplorer to maximize the size of available layout for the interactive views.
The distribution plot is a special implementation of the histogram and shows regional differences in the share of elderly population that are high within countries. For example in Southern Australia the share of elderly population is three times higher than in the Northern Territory. Similar wide differences are observed in Italy, Spain, Turkey and United States.

The Parallel Axes Chart provides the analyser with a superior exploration technique that enables visual comparison of selected regions for 4 -10 indicators by representing each region as a profile passing through the parallel axes. Each axis represents a single indicator. A profile forms a visual representation of the characteristics for the region. Differences between selected regions can be found by visually comparing the profiles representing them.

Another advantage with this exploration technique is the capability to dynamically make visual inquiries and remove data. Filtering data is a critical step in the process of statistics data analysis. Filter out uninteresting regions; reduce the data set to a smaller, more manageable size. Each indicator axis has a pair of range sliders which define the bottom and top range for the query area. The range of an indicator can be specified by “dragging” the handles on the top and bottom of the corresponding range slider (see figure). Regions with values for a selected indicator, that fall outside of the specified range, are filtered out. A combination of range slider movements can be used to dynamically formulate a more complex visual inquiry. These visual conditions and constraints will immediately reflect the visual contents in all linked views. An example of a query using the sliders is shown below. After a dynamic query operation, applied to the indicator “Labour Productivity”, regional values below the mean (green line) was removed. A tooltips shows the exact value for the slider position along the indicator axis. A second condition is then given for indicator “Unemployment rate”, where regions with higher rate (above mean value) are removed.

Two filter conditions are performed: 1) Keep regions with high “Labour Productivity” (above mean value) AND 2) Keep regions with low “Unemployment Rate” (below mean value). The map to the left shows the result after removing the regions that do not fulfil these two conditions.

Percentile filter operation applied on population age group 65+. Red regions represents regions with high (95th percentile) and blues outliers with low percentage of elderly people (5th percentile).

Table Lens is a method to dynamically explore large amounts of tabular data and displays regional data values in columns and rows without scrollbars and without obscuring any data and filling the cells with scaled and coloured horizontal small bars. Each row in the Table Lens represents a single region and the columns represent a specific indicator. The power of the Table Lens is that it has the combined abilities to sort records, focus “zoom” in on interesting areas in the data (to reveal exact numerical information) and remove unwanted records.

Table lens with map and highlighted regions for comparison
Statics eXplorer Wizard is a step-by-step semi-automatic tool assisting the user to create a new customized Statistics eXplorer application e.g. Denmark or Italy eXplorer with county and municipality regions, Manhattan eXplorer with local regions. Geographical shape files are here merged with corresponding indicator data. These two data sources must have the same unique naming convention. Indicators are linked to corresponding regions for statistics analyzing. The Wizard can also be used to edit an already existing Statistics eXplorer application such as adding and replacing indicator data or pre-load new Stories. The outcome of a Wizard process will be an eXplorerFiles.zip file, containing the eXplorer.swf file and necessary data files for the new Statistics eXplorer application to run on your client computer or be installed on a server for public use. Just unzip the eXplorerFiles.zip and launch index.html in the root folder.

Use the Wizard to customize your Statistics eXplorer with preferred geographic regions defined through the ESRI shape file (.shp) format and DBF (.dbf) files, where the shape file has to be in a polygon shape. Associate indicators are organized for a single year or time series and by category (demographics, economy, education, migration etc.).

Multilingual support includes English, Italian, Spanish, Swedish, Danish, Norwegian and a mechanism for implementing your own language is available. Add your preferred logo and then with a final click - your customized Statistics eXplorer application is completed and ready to empower analysers and public users with rich, interactive visualizations with your preferred indicators.
Facilitate the use of statistics, through easier-friendy visualization tools

Summary of features and functionalities for Customizable Statistics eXplorer

- Web 2.0 compliant explorative, communicative and collaborative statistical visualization;
- Discern trends or patterns - derive insight and understanding – communicate knowledge effectively through Storytelling – publish for action;
- Identify what’s meaningful in the data in ways that your eyes can see and brain can understand;
- Keep focused on what's important and not be distracted by visual content that doesn't add to your understanding of the data;
- Interact with the data that you frequently need (filtering, sorting, etc.);
- Easily navigate through the analytical process, from view to view, interaction to interaction, overview to detail, riding the wave of thought smoothly throughout the process;
- Support large spatial-temporal and multidimensional regional data (>1,000 regions);
- Innovative methods from Information and Geographical Visualization (choropleth map, scatter plot (bubbles), time graph, dynamic histogram, table lens, parallel axes “profile plot”, scatter matrix frequency plot) applied and customized for statistics data;
- Web compliant – visualization tools are optimized using Adobe’s ActionScript for Flash 10 and Flex 3;
- Regional statistical data is simultaneously explored through multiple-linked and coordinated views e.g. map, scatter plot, scatter matrix, table lens, table, profile plot, pie chart, histogram etc;
- Map layer architecture - overlay several types of maps such as coloured statistical regions, country boundaries, background maps e.g. Google or Bing Maps, flow and pie chart – control transparency level for each layer;
- Dynamic state-of-the-art colour legend facilitating statistical methods for class value calculations and scientifically tested perceptual acceptance of colour scales;
- Mechanism for Storytelling for creating dynamic Web documents;
- Analytical reasoning facilitated by interactive visual interfaces and metadata;
- Visual inquiry and conditioned statistics filter mechanisms that help detecting outliers;
- Pick interesting regions for multiple indicator profile and comparison;
- Interactive time animation based on an innovative multi-dimensional data cube architecture;
- Dynamic motion (time animation) visualization using map, scatter-plot, scatter matrix, table lens, data grid, profile plot, histogram, pie chart and flow map;
- Screen space usage is optimized for visualization – no unnecessary visible GUI panels;
- Modular multiple-view design - scale the individual views and allocate more space to the visual representation that is most important;
- Support for categorical data and visual classification of, for example, urban, rural and intermediate regions.
- Integrated mechanism Statistics Publisher for creating dynamic Blogs and HTML documents with embedded interactive visualization “vislets” based on eXplorer storytelling;
- Thoroughly tested – eXplorer has been in production at OECD, Statistics Sweden and Denmark, EU Commission and several others since 2008, where errors and inappropriate properties have been settled;
- User can insert external (own) data and mix with preloaded data based on standard spreadsheet format;
- Optional SDMX and PC-AXIX data interface formats and customizable data supplier;
- Multilingual, so you can switch between languages (English, Italian, Spanish, Swedish, Danish) and mechanism for implementing your own language;
- Generic eXplorer – easy to introduce new indicators and map - all eXplorer applications use the same compiled code. The only difference is the configuration and data behind it;
- Statistics eXplorer is build on our in-house developed solid and proven object-oriented component architecture and framework GAV Flash platform for easy add-ons or replacement customizable Statistics eXplorer applications;

Statistics eXplorer is built on a solid object-oriented layered component platform GAV Flash that provides high flexibility for further optimization, improvements and future extension. This architecture has allowed innovative and special user requested improvements since the first release in late 2008 at the OECD web site. Components that don’t perform acceptable can easily be replaced by our developers.