Abstract

We present our latest research that focuses on the most ancient of social rituals: “storytelling” – exemplified through telling a geovisual analytics 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 in this interactive process and help advancing research critical to the dissemination of official statistics by means of web-enabled reporting tools. Geovisual analytics is a technique that can help illustrating high-dimensional statistical temporal data which for the eye are hard perceive or interpret. We introduce visual “storytelling” means for the author to import large spatio-temporal statistical data, explore and discern trends, create a story with snapshots and metadata and finally publish understandings and knowledge embedded in web page. The story guides the reader in the directions of both context and discovery through a highly engaging intuitive visual interface “vislet” based on cognitive principles and at the same time follow the analyst’s way of logical reasoning. Value no longer relies solely on the content but also on the ability to access this information.

Keywords: [statistics visualization, storytelling, publish understanding and knowledge]

1. Introduction

Statistical data have great potential to generate knowledge and serve as basis for decisions taken by many actors in society. In this paper, we introduce an innovative web-enabled geovisual analytics platform (Andrienko, Jern 2010) for exploring geographic statistics data and subsequently publishing the analytics understanding and knowledge. Our solution is based on richer and more dynamic visual user interfaces facilitating methods that can uncover hidden structures and relations and let the analyst present her findings and reasoning to a broader audience. The combination of a geovisual analytics explorer tool (explore, gain insight and understanding) and a visual publisher tool facilitating an authoring process that integrates statistics visual analysis, collaboration and publication based on the most ancient of social rituals “storytelling”. Telling a story
about, for example, sustainable development over time and in addition let the reader dynamically participate through blogs.

Understanding the variety in regional economic structures and performance is essential knowledge for initiating development which could improve regional competitiveness and in turn enhance national growth (OECD 2010). The results from our research make these variations more visible and understandable, providing region-by-region indicators in the form of motion graphs and maps that could lead to better identification of areas that are outperforming or lagging behind. Patterns of growth and the persistence of inequalities are analyzed over time, highlighting the factors responsible for them. How can such significant knowledge about these statistical facts be collaborated and published to analysts and citizens?

We exploit two tools Statistics eXplorer and Publisher for an integrated statistical analysis, collaboration and publication process facilitating storytelling aimed at unveiling the essential messages and educate the citizens in the statistics. The author can simply press a button to publish gained knowledge from a visual interactive discovery process and place into 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.

Figure 1: Euro eXplorer – ageing population in Europe 1995-2007 for NUTS2 regions. Regions with a low percentage of population aged 65+ i.e. 16.1% (the 50th percentile) are coloured in blue. The motion bubble chart represents four indicators age 0-14, age 65+, total population (bubble size) and 65+ bubble colour. Five regions are highlighted in the bubble and time graphs: Inner London, Slaskie, Latvija, Galicia and Dresden. We see from the trails that London has a negative ageing trend compared to the other selections. The time graph shows the trends for selected regions. The story providing the analytics reasoning and understanding of the data is shown in the right panel.
An innovative storytelling mechanism is introduced (figure 2) for the author (domain expert) to: 1) import statistical data; 2) explore and make discoveries through trends and patterns, derive insight and discovery of the unexpected 3) orchestrate snapshots and describe metadata that is the foundation for 4) creating a story that can be 5) shared with colleagues and reach consensus and trust and 6) finally publishes (html) “tell-a-story” using a “Vislet” that is embedded in educational blogs or web pages providing readers with an interactive learning experience.

2. System Implementation

Statistics eXplorer is based on inhouse developed component toolkit adapted for Adobe’s Flash basic graphics and does not require installation of any other software and will run anywhere. The toolkit facilitates innovative methods from information and geographical Visualization such as the choropleth map, dynamic histogram, table lens, parallel axes plot “profile plot”, scatter plot, time graph, and pie and time glyphs, flow map (trade and migration) applied and customized for 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. Data are normally preloaded with a set of basic indicators such as demographics, economic indicators, education statistics etc. but the user can also insert external (own) data to be mixed with preloaded data based on a standard unicode spreadsheet format. A
new version will open for optional data base interfaces such as SDMX, PC-AXIS or other API solutions.

Figure 3: Statistics eXplorer provides an open data architecture for flexibility

In order to detect complex patterns it is convenient to view statistics data through a number of different visual representations simultaneously (figure 1), each of which is best suited to highlight specific features. Any filtering, highlighting or colouring made in one of the linked views is transmitted to all the others.

The conceptual data model for our Statistics eXplorer platform can be seen as a data cube with three dimensions: space, time and indicators. The spatial dimension is represented by the regions and the indicators are various indicators (GDP growth, elderly dependency rate, etc). Time is the period or point in time to which the data refer. The general method for finding a value in the cube is by its position (space; time; indicator) and fast access time is essential for motion graphs. Space-time-indicator awareness means that the data cube can be analysed and visualized across all three dimensions simultaneously.

Figure 4: Statistics eXplorer conceptual data model based on (when, what, when) events

Statistics eXplorer performs this task by integrating and time-linking all its motion graphs (figure 5): choropleth map, scatter plot, dynamic histogram, flow map, table lens, data table, glyphs, time graph, parallel axes plot and regional distribution plot.
3. Storytelling and Publishing

Collaboration is achieved through a mechanism in the toolkit that supports the storage of interactive events in an analytical reasoning process through “memorized interactive visualization views” or “snapshots” that can be captured at any time during an explorative data analysis process and becomes an important task of the storytelling authoring analytical reasoning process.

3.1 Snapshot

When exploring and making sense of comprehensive statistics data, we need a coherent cognitive workspace on which to attach our discoveries for organizing and navigating our thoughts. The GAV Flash toolkit includes such means by capturing, saving and packaging the results of a Statistics eXplorer “gain insight” process in a series of “snapshots” that could help the analyst to highlight views of particular interest and subsequently guide other analysts to follow important discoveries (figure 6). The snapshot tool creates a single or a continuous series (story) of visualization captures during the exploration process. In a typical scenario the analyst selects relevant attributes for a snapshot, e.g. time step, highlighted regions for comparisons, class values for colour legend, filter inquiry conditions for selected attributes and finally highlights the “discoveries” in the map view from a certain angle.

The analyst requests a snapshot with the Capture function in the Story Editor (figure 1) that results in a snapshot class operation scanning through all active connected views for properties to be captured. Each of these properties will then be parsed into XML and written to a file that also contains details on which data (indicators, time and GIS regions) was used and a unique name for each component. When a snapshot is activated, the saved state of the snapshot class will be read from the XML file and parse its nodes back into component properties again. The previously marked properties will then be applied and set the state of the application (figure 6).
3.2 Storytelling

Storytelling, in our context, is about telling a story on a subject, based on statistics, and related analytics reasoning about how gained knowledge was achieved. Storytelling within this participative web context, could more engagingly draw the user into exciting reflections and sometimes change a perspective altogether. The story is placed in the hands of those who need it, e.g. policy and decision makers, teachers but also the informed citizens. Dynamic visual storytelling is a way of telling stories through interactive web-enabled visualization. Our proposed novel storytelling technology could advance research critical to collaboration and dissemination of digital media and enable a leap in understanding by the audience, so as to enhance informed decision making.
### 3.3 Publisher and Vislets

A Vislet is a standalone Flash application (widget) assembled from our low-level toolkit represented by, for example, a single map view or a composite time-linked map and scatter plot view (figure 9). A Vislet facilitates the transition of selected statistics into communicative sense-making news entities with integrated metadata and interactive time animated visualization that could engage the user. Publisher is the application tool that imports a story and generates the HTML code that represents the Vislet and metadata (figure 7). First, the user selects appropriate visual representation for the Vislet e.g. map, scatter plot, histogram, time graph etc. The size of the Vislet window with metadata is set and Publisher generates the HTML code. This code is manually copied and finally manually (copy/paste) embedded into a web page. The Vislet can now be opened in the reader’s Web browser and dynamically communicate the story. A Publisher server maintains the Vislet flash (swf) files together with a story repository, statistical data and regional shape maps. The Vislets run locally in the client’s Flash Player and can thus achieve dynamic interactive performance.

### 4. The use of Vislets in Statistics Denmark and Statistics Sweden

The paper includes some interesting examples from two national statistical institutes (NSI), where the use of Vislets and storytelling is becoming a new way to broaden the understanding of the large amount of regional statistics available in the statistical databases. The use of Statistics eXplorer also gives statisticians within the NSI's a tool to analyse and get new insight in their statistics and the findings may be published as Vislets with narrative text.

#### 4.1 The case of Denmark

Statistics Denmark has chosen Statistics eXplorer as preferred tool to serve two different purposes: 1) analysing and presenting the regular official statistics published by the office, and 2) creating added value to paying customers who buy special data tailored to their needs.

The major part of Statistics Denmark’s work is fulfilling the role as the nation’s infrastructure for official statistics: Providing high quality data, objective and impartial, for everyone to use as a basis for decision making and information. Much of the statistics produced have geography as an essential dimension,

---

**Figure 8**: Example of a Vislet produced with Denmark eXplorer and embedded into a web page
and it is important to allow users to understand the impact of geographical differences and structures. Until recently the dissemination of such data has been rather dull and less attractive to many categories of users. For this reason we are convinced that our statistics are grossly under-utilised, as compared to its great potential. So we started in 2010 using Statistics eXplorer for communicating the essentials of official statistics to a broad range of users on Statistics Denmark’s web site. This has so far been done only for a few selected statistics, notably the Income Statistics. The tool has been used by statistical experts to analyse in depth the geographic structures and correlations, and as a result, Vislets with good stories have been presented on the web site. This has proved to be an excellent way of catching the attention of many users, including the media. Behind the Vislets, the full eXplorer application with its full functionality has also been made accessible to be enjoyed by users with a wish to go more in depth and make their own analysis.

The official statistics, which are freely accessible to all users, are of course far from satisfying all needs of all decision makers. But the basic data behind the statistics have great potential for being recombined, especially in a statistical system as the Danish one, based to a large extent on rich administrative sources that can be linked across domains using identifiers and geographic coordinates. There are innumerable opportunities for creating new information and analysing it, and this process has a cost that must be borne by the users. Consequently, Statistics Denmark offers a lot of possibilities for helping customers who buy data specifically adapted to their needs. These customers need more than data: they need a tool that can help them analyse their special data and make their findings clear to their own community. For instance, the planning division of a municipality may acquire a set of statistics broken down by small areas, such as school districts; but for the municipal politicians to use this material and make wise decisions, it must be presented in a way so as to reveal the stories they contain. Here Statistics eXplorer comes in handy. So Statistics Denmark offers to host the data together with the tool and Vislets, as an add-on. These pages can be linked from the customer’s own web site or intranet.

4.2 The use of Vislets in Statistics Sweden

By the end of January 2011 Statistics Sweden will launch both an updated Statistics eXplorer and a collection of Vislets. This is the outcome of a project that has been running since October 2010 and proves that eXplorer has been successfully implemented at Statistics Sweden. The possibilities with Vislets have been crucial for the decision which tool to implement, as it is a necessary bridge from the statisticians expert knowledge to the broad range of users of scb.se - many of whom don’t have any practice in using statistics.

The starting point is to provide a Statistics eXplorer that consists of a large number of statistical indicators on municipality level, organised by subject areas. As a first step six areas were identified: population, housing, income, public finances, labour market and education. The project contacted the statisticians responsible for these subject areas and introduced the eXplorer tool. The statisticians were also asked to create a Vislet after using eXplorer to find some interesting links between different indicators.
Figure 9: Example of a dynamic document with integrated Vislet based on public SCB indicators, educational text, map, motion chart, snapshots and time series.

The Vislets for those six subject areas are presented on the web as a Regional Statistics Atlas, together with the full Statistics eXplorer. The idea is to let this Atlas grow with more subject areas during 2011 and also to schedule the updates and have support and expert knowledge available within Statistics Sweden to help statisticians who want to use the tool. A collaboration between GIS-experts, web-experts and IT-experts has been established during the project and will continue to have important roles to support the use of Vislets in the organisation.

The Vislets are integrated in the web-environment at Statistics Sweden and a special template has been created in EPiServer to allow the Vislet to be shown using as much as possible of the web page. Tests on using Vislets in the predefined template for press releases lead on to the next step: how to integrate Vislets in the publication process with release simultaneously with our other publishing channels (press releases, statistical databases, tables and graphs).

5. Conclusion

The technique introduced in this paper allows the analyst (author) to communicate with interested readers through visual discoveries captured into snapshots together with descriptive text. Selected indicators and visual representations can be published together with their metadata, thus facilitating the comprehension of statistical information by non-expert readers. We believe that this advanced storytelling technology can be very useful
for media as some examples of using Statistics eXplorer to tell a story have already showed. At the same time, the Vislet technique can help developing agile on-line publications, which draw the attention on recent trends and inequalities. Reviews and evaluations performed by several statistical institutes highlight the following features:

- Statistics eXplorer can easily be customized by a statistics organisation - requires only regional boundaries (shape file) and associated indicator data;
- Statistics eXplorer is a comprehensive tool for advanced users – the Vislet approach is regarded as a painless and more attractive alternative for the general public;
- Encourage collaboration between statistics analysts and users of statistics;
- Possibility to capture, save and open discoveries (snapshots) with attached analytics reasoning metadata, e.g. storytelling;
- IT expertise is not required to publish interactive visualization embedded in blogs or web pages;
- Possible strategic tool for news media to publish statistics news on the web;
- It is easy to import external statistical data into eXplorer;
- Ability to have dynamic time-link views and see the multi-dimensionality of regional development;
- Increased expectations in terms of user experience;
- Statistics eXplorer will encourage more educational use of official statistics.


References


