

A journey towards a data visualisation strategy

Anu Ots (Statistics Estonia)

Anu.Ots@stat.ee

Abstract

We are living in a world where the amount of data can easily overwhelm our users. Good data visualization is more important than ever. There are many different options to choose and consider when one journeys towards a data visualisation strategy. We want to share our ideas and plans, including some samples what we have done so far and how our users have reacted in response. Hopefully we can start a discussion, get some feedback and harmonize the views between different countries.

A journey towards a data visualisation strategy

Anu Ots (Statistics Estonia)

Anu.Ots@stat.ee

Abstract

We are living in a world where the amount of data can easily overwhelm our users. Good data visualisation is more important than ever. There are many different options to choose and consider when one journeys towards a data visualisation strategy. We want to share our ideas and plans, including some examples of what we have done so far and how our users have reacted in response. Hopefully we can start a discussion, get some feedback and harmonise the views between different countries.

Introduction

In the European Statistical System, the visualisation and interactive presentation of statistics is a very popular topic at the moment. To introduce statistics, several countries have made a short video, launched an interactive map application or developed an app for smart devices. With regard to the use of interactive tools, the aim is to present statistics in a manner that is as simple as possible but also illustrative and interesting. Statistical publications continue to be published as printed material, but in most countries electronic dissemination has moved from PDF-based dissemination to the introduction of a smart app. There is a wide selection of graphics editing software and tools for making statistics interactive, so the options are endless. It seems that the main difficulty lies in orientating oneself in this abundance of possibilities, and to make the most suitable choice for one's organisation.

While at the 2010 United Nations meeting on the dissemination of statistics social media was a relatively new topic and countries were discussing whether they should use social media, in 2016 most countries use one or another social media channel. When talking about information campaigns and statistics dissemination, it is said that social media is an effective and free channel, which facilitates the dissemination of statistics, promotes statistical literacy and increases the use of statistics. Several countries have started to actively engage pupils and students in order to steer them towards using official statistics. Both presenting statistics in an interactive manner and disseminating data in social media play an important role in stimulating young people's interest in statistics. Statistical authorities need to move with the times and be where statistics aficionados and potential users are – in social media and on smart devices.

As there is a great number of suppliers on the information market, most national statistical institutes are facing the question how to reach their existing and new users by offering reliable information in an entertaining manner.

I will be describing the best practices that Statistics Estonia (SE) has been using in visualising data and statistics, and explaining what has led us to the idea that there is actually the need to write out a visualisation strategy for the organisation.

Users of statistics

Most statistical institutes divide their users into groups based on the frequency of using statistics and/or on the importance of the user.

SE's segmentation of users

Users of official statistics include all legal and natural persons who use official statistics. Different users need different types of statistics. The following user groups can be identified based on the frequency of use of statistics and/or the user's importance (participation in the statistics production process; dissemination of statistics to the general public; use of statistics for important decisions concerning the society):

1. **Infrequent users** use statistics infrequently or randomly and are mostly interested in basic data;
2. **Main users** use statistics regularly or (actively) in specific periods; mostly use statistics for analyses, decision-making, further dissemination etc.;

3. **Key users** – mostly use very detailed statistics, incl. micro-data; use the data collected for official statistics to create new value for the society and/or participate in the preparation of the statistical programme and the development of statistics.

Pyramid of user needs

Key users

Mostly use very detailed statistics, incl. micro-data; use the data collected for official statistics to create new value for the society and/or participate in the preparation of the statistical programme and the development of statistics.

- ✓ Plan future developments
- ✓ Make models and projections
- ✓ Do research

Main users

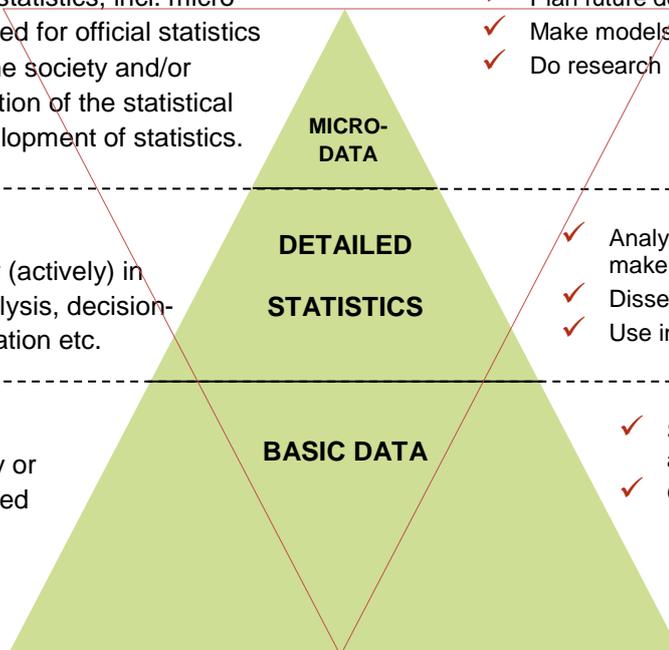
Use statistics regularly or (actively) in specific periods – for analysis, decision-making, further dissemination etc.

- ✓ Analyse the current situation and make the necessary decisions
- ✓ Disseminate statistical data
- ✓ Use in studies (e.g. students)

Infrequent users

Use statistics infrequently or randomly, mostly interested in basic data.

- ✓ Study the general economic and social background
- ✓ Get information



Visualisation helps to gain more users of statistics among infrequent users because, although they use statistics randomly and are mostly interested in basic data, such users are great in number. They do not think about statistics very often, they like information they can relate to, they do not want to spend too much time on statistics, they will share the information they like. They can benefit from statistics but they do not really care about the source of the data (and there are a lot of competitors for NSI-s).

SE has developed a range of products specifically designed for infrequent users. In order to explain statistics and to provide better an overview of the available data, these products also feature data visualisation (interactive graphs, thematic maps, etc.).

SE's products for infrequent users

Pre-defined tables (<http://www.stat.ee/pre-defined-tables>)

Pre-defined tables contain basic data on various subject areas. The tables are based on tables published in the Statistical Database and feature the unique code of the corresponding database table. The pre-defined tables are updated automatically whenever the data in the Statistical Database are updated.

News releases (<http://www.stat.ee/news-releases>)

A news release is a brief, news-like piece of information presenting a selection of recently produced official statistics. News releases notify the public of the completion and availability of up-to-date data in a certain subject area of statistics. Separate news releases are published to notify the public of conferences organised by SE or of other major events. News releases are published on the front page of SE's website and sent to journalists by e-mail. Other users can sign up for a notification of news releases – subscribers will receive a notification by e-mail when a news release is published.

Statistics blog (<https://statistikaamet.wordpress.com/>)

The statistics blog is a tool for promoting official statistics, for improving statistical literacy and for establishing the contributors as opinion leaders. The blog is also a place for information that cannot be published in SE's traditional channels (news releases, publications). The blog can be used to draw wider attention to important issues in the European Statistical System or in the Estonian society, to highlight recent trends or to present new indicators that are already in use or planned to be implemented.

"Estonian statistics" app for smart devices (<http://www.stat.ee/statistics-on-smart-devices>)

The app offers convenient access to official statistics on smart devices. It includes the main indicators of most subject areas. Statistical information can be displayed as figures, thematic maps and tables. The latest news releases, blog posts and statistical publications can also be accessed through the app.

Calculators (<http://www.stat.ee/calculators>)

The consumer price index calculator allows users to calculate the monthly, quarterly and annual change of the consumer price index.

The purchasing power calculator allows users to compare the purchasing power of a sum of money (income, wages, savings etc.) in different years.

Population pyramid (<http://www.stat.ee/population-pyramid>)

The interactive population pyramid shows the population change in Estonia since 1923 and population projections up to the year 2050. In the case of projections, users can select one of two options, each based on a different assumption. By navigating on the pyramid, it is possible to view the population by age and year of birth, or the population of selected age groups.

Thematic maps (<http://www.stat.ee/thematic-maps>)

Thematic maps are graphic presentations of statistics. Thematic maps show the value of a statistical indicator in different parts of the country, using various colours, patterns and raster graphics. On SE's website, there are thematic maps about different subject areas. The maps allow users to compare the indicators in different counties and rural municipalities.

New visualised products created by SE in recent years and their Facebook reach

How to measure the cost-effectiveness and/or impact of data visualisation?

There are 1.3 million people living in Estonia. Facebook is the most important social media channel, which is used by about 600,000 inhabitants, thus we can make a broad generalisation that every other person living in Estonia uses Facebook. Statistics Estonia has a Facebook profile since 2010. Just like most statistical authorities, SE uses Facebook as a channel that is ideal for disseminating statistics on social media. For the following examples of data visualisation, I will be indicating the number of people reached on Facebook.

During the time that SE has been using Facebook, we have observed that posts that contain an image (photo, graph, map, infographic) have a greater Facebook reach (FBR). At the same time, it also depends greatly on the statistics domain in question, on how interesting people find it. Social and demographic statistics are more popular than statistics on the economy, for example, and posts concerning the former are shared more actively on Facebook and thus reach a greater number of people.

Considering the population figure of Estonia and the fact that Statistics Estonia's posts on Facebook are in Estonian (although we also share Eurostat's and our closest neighbours' news releases published in English, and the product information issued by Eurostat), we have developed the following scale based on our experience in terms of the number of people reached:

FBR 500: common;
FBR over 1,000: good;
FBR over 2,000: popular.

Product: Statistics of names (<http://www.stat.ee/public/apps/nimed/>)



In May 2016, we launched a new product that was meant to present statistics in a more personal way. It is basically a statistics application of first names describing in detail the popularity of a specific name in Estonia and visually comparing the popularity of the name by different age groups, counties and birth months. Additionally, there are top 50 lists of the most popular names by gender, year of birth and county. Information on the name day celebrated on the current date and on all the days in a year is also available.

Prior to launching this product, we imagined that it would gain some popularity, but on the day the application was launched it went completely viral. It spread on social media and was talked about on all the biggest national news portals where it was among the first news due to the great popularity of the product. The application set new records for daily and hourly user statistics on our website, even surpassing previous records from 2012, when the online population and housing census was conducted. The traffic to our website was so unpredictably heavy that we were contacted by network monitoring specialists asking if we were experiencing some kind of a cyber-attack (on the launch day there were 1.4 million page views – remember that there are 1.3 million people in Estonia). Fortunately, the product was so small and optimised that despite the high traffic we did not suffer any failures or slowdowns on our website. Even now, months later, the product still accounts for nearly a fifth of the page views on our website.

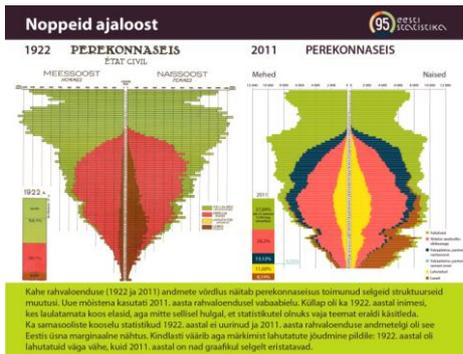
In order to introduce the new product to our users, we also published a blog post, and the product was so popular that the blog post received 10,900 views on the day the application of names was launched, and the corresponding post on Facebook reached 47,600 people.

Infographics

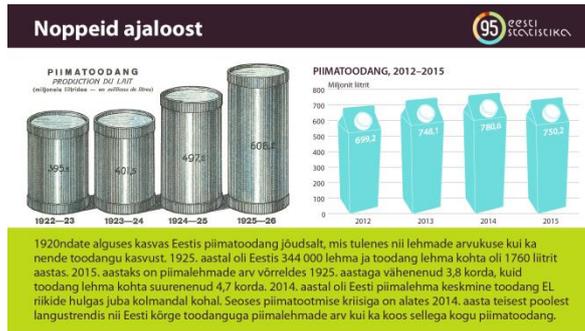
Using infographics to boost the popularity of statistics is a relatively new topic for Statistics Estonia. The plan to start publishing infographics was discussed for a long time, but it was only in 2015 that we invited our colleagues from the Digital Content Team of the Office of National Statistics of the United Kingdom to share their experience and give us a brief training on the topic. In 2016, we have been practising what we heard and learnt. We have made infographics mainly for anniversaries (SE's anniversary, Independence Day of the Republic of Estonia, Accountants Day) and have also monitored their Facebook reach. We plan to also start compiling infographics for the publication of regular statistics as well, but in order to do that we would have to sell the idea to statisticians and establish a rapid response team.

The year 2016 sees the 95th anniversary of the Estonian Statistical System. To celebrate the anniversary, we have also published a monthly infographic "Bits of history" on our statistical blog. In these infographics we have used interesting visualisations from printed publications published in the 1920s and compared these with similar visualisations that were made using the latest data.

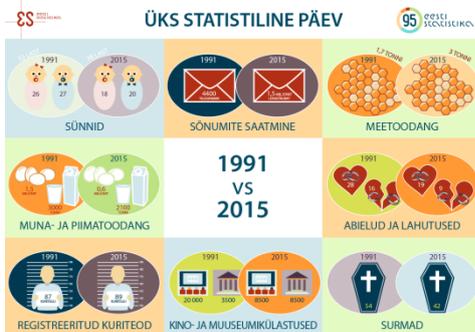
FBR – 3,500



FBR – 1,100



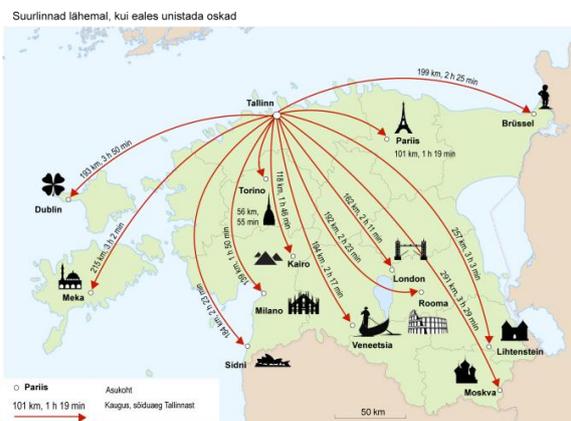
As for other infographics which were popular based on their Facebook reach, this year, the Republic of Estonia celebrated the 25th anniversary of the restoration of independence and Statistics Estonia published an infographic to mark the occasion. The infographic was titled “One Statistical Day” and it featured a comparison of indicators from 1991 and 2015, based on the data published on various subject areas for these two years. The infographic reached 15,900 people on Facebook.



Based on SE’s experience, infographics presented as maps reach the greatest number of people on Facebook. One of the blog posts published this summer gave entertaining advice on how to spend one’s holiday in Estonia and suggested visiting Estonian villages that bear the same name as various cities in the world (the infographic on the left shows the location of the village and its distance from the capital).

What was also popular on Facebook was a thematic map which advertised the application of names by showing the most popular names of children born in 2011–2015 by county (see the map on the right).

FBR – 41,700



FBR – 47,600

Aastatel 2011–2015 sündinud laste populaarseimad nimed maakonniti



With regard to the Facebook reach of infographics, we have also noticed that if an infographic has been posted on Facebook to advertise a blog post and if there is a link that directs the user to the analysis posted on the blog, it often happens that people who have shared and liked the infographic do not end up reading the blog post itself. At the same time, infographics help the introduction of the blog post to reach a greater number of people on Facebook.

Statistics Estonia has not taken advantage of the paid options offered by Facebook to increase the reach of posts. Perhaps we should give it more thought.

Videos (<https://www.youtube.com/user/Statistikaamet>)

Our videos, which are meant to serve as marketing material, have been outsourced by Statistics Estonia. Since the end of 2015, we have a video introducing the 2020 register-based population and housing census ([The first register-based population and housing census in Estonia](#)) and another video introducing SE's products and services ([Official statistics – the primary source of reliable data](#)). We have shared these videos on YouTube and used them as introductory and marketing material on information days and at conferences.

Estonia's largest news portal (Delfi) has produced animated videos (fact-toons) based on SE's survey data: [Does having a kid make a man poor?](#) and [Estonian Family Portrait](#).

Since making videos is not a very difficult task these days, Statistics Estonia is planning to start producing our own fact-toons and help videos for our products (e.g. a help video on how to use SE's [statistics map application](#)).

Challenge

1) Budget

In Statistics Estonia we have a situation where the general budget of our organisation (excluding the budget for population censuses) has been the same for almost a decade.

On the other hand, there is growing demand for statistics both on the national and international level; and there is growing demand for innovative and engaging statistical products.

The respondents, too, would like to submit less and less data and expect government organisations to share the necessary information via registries. If they must submit data they expect there to be the most convenient options for doing so.

All this means we have constant development and innovation needs which are often without funding.

In this situation we are trying to find sources of saving and one possibility is to identify products and statistical indicators or even statistical actions which our users do not really need and which are not required by any national or international agreements.

In order to do this, we have done extensive reviews of our statistical actions and indicators in co-operation with our main customers, and we have also analysed the cost and usage statistics of our statistical products.

As a result, we have found some products we can discontinue but unfortunately these tend to be the smaller and cheaper ones. So we are still faced with the challenge of saving money.

2) Selection of design software and tools

There is a wide selection of paid and free of charge design software and tools that can be used for making the presentation of statistics interactive, and therefore the list of possibilities is endless. The main difficulty lies in orientating oneself in this abundance of options and in making a choice that is best for the organisation.

In visualising its products, SE uses the paid software offered by Adobe. Using this software requires technical skills and training, and making one infographic, graph or map can take up to three working days. What is the most important, of course, is the idea – the story that we want to tell with the help of data visualisation.

Programs that are actually required for this need to build something professional fast (using, for example, pre-made building blocks, pre-applied themes, connection with a data source for easy updating) and be easy to use after a brief training.

There is a number of free-of-charge software on the market, you only need to define your requirements and choose a tool that is best suited to them (<https://infogr.am/> – Charts & Infographics the Easy Way; <https://www.easel.ly/> – Create and share visual ideas; <https://piktochart.com/> – Easy-to-Use Infographic Maker; <http://visual.ly/> – Visual Content That Works; <http://www.dipity.com/> – Digital timeline).

3) **Visualisation team**

SE’s visualisation team consists of three graphic designers and two GIS experts who design our products. At the same time, coming up with the design requires close co-operation with statisticians, marketing specialists and designers in order to set rules regarding what we should respond to and how. We need to have a quick response team and tell a story, not just present figures.

Summary: Visualisation strategy for SE?

