Promoting Civic Engagement via Exploration of Evidence:
Challenges for Statistics Education

Co-funded by the Erasmus+ Programme of the European Union

Dynamic Visualisation Tools

Jim Ridgway*, James Nicholson*, Pedro Campos** and Sonia Teixeira**

* University of Durham, UK

** University of Porto, Portugal

Abstract

This document has three components: (1) *Preamble* - an explanation of what we proposed to do, and what we have achieved in the PCS project. This reproduces text in the Final Report of PCS to ERASMUS+, but is included here for clarity and completeness; (2) a review of *Tools for Visualisation* - notably websites that support data visualisation and data analysis. We begin with examples, then present a spreadsheet where (hyperlinked) tools are categorized in terms of their properties; (3) *examples of short activities with visualisations* – visualisations and activities we have used in teaching episodes and workshops.

Suggested citation: Ridgway, J., Nicholson, J., Campos, P., & Teixeira, S. (2018). *Dynamic Visualisation Tools: a review.* A product of the ProCivicStat project. Retrieved [Date] from http://IASE-web.org/ISLP/PCS.

For more information, extensive teaching resources, supporting papers, datasets, contacts, and our Call for Action and Recommendations: See the ProCivicStat website under the International Association for Statistics Education (IASE) website here: http://iase-web.org/islp/pcs. You can also visit our original website at www.procivicstat.org, though it will not be updated after Fall 2018.

Acknowledgments: The authors thank the ERASMUS+ program of the European Commission for its support of the *ProCivicStat* project. However, the opinions expressed in this document reflect the authors' own views and not necessarily those of the sponsoring agency or the EC.

Preamble

This Output has made a number of significant contributions, notably in terms of innovation and transferability.

Innovation: Above and beyond our proposal for Dynamic Visualisation Tools (DVT), we created CivicStatMap, a searchable tool which offers direct access to teaching materials that incorporate DVT. This innovative resource allows users to access resources in a variety of languages, using a number of filters.

Transferability: We have conducted an extensive review of DVT, which we extended to include tools that both display data and have facilities for data analysis. This is a free-standing resource which is inherently valuable to anyone planning curriculum activities in statistics education.

Impact: We have run workshops based on these materials at Multiplier Events in and Berlin and Szeged, and at international conferences in Rabat and Kyoto. Materials will be presented (by invitation) at the 2018 South African Annual Statistics Conference. The materials have also been the basis for a number of academic papers.

Differences between what was proposed and what was done

The proposal was written in early 2015, and the ideas proposed for dynamic visualization tools (DVT) were ambitious, and reflected 'state of the art' thinking in 2015. Our original concept was to modify some existing DVT so as to make it easy to import new data sets. However, DVT have subsequently become far more 'mainstream'; many of the developments we proposed are now routine features of open source DV packages. The project as a whole spent fewer days on developing DVT than was originally planned (204 not 270 days), and devoted less time to repurposing existing DVTs, and more time to writing tutorial advice on uploading and analyzing data to new packages of DVT (including iNZight, CODAP and Tableau). As promised, we produced a comprehensive review of DVTs. Here, we show how we have achieved all and more than we aimed to do in the original proposal.

O2-A1: Tool Identification and Assembly

We conducted an extensive review of existing DVT, many of which are embedded in packages that facilitate data analysis. This review is available on the PCS website, and has been the basis for academic papers. A version is available on the website that provides links to relevant resources. There, tools are described in terms of their availability (open source, free, free for educational use, pay etc.), the sorts of data that can be analysed and displayed (e.g. macrodata, microdata), and the sorts of displays available (graphs, charts, maps, relational diagrams etc.).

The candidate visualisations set out in the proposal were:

A pan-European version of the Constituency Explorer – an analogous resource wherein users can chose variables and examine their distribution across states and regions has been

implemented by NCVA at Linkoping University - see
https://stats.oecd.org/OECDregionalstatistics/#story=0

An interactive population pyramid – these are available from NSOs (e.g. The UK, Hungary, Germany) and are incorporated into some of our teaching materials – see https://www.ksh.hu/interactive_agepyramids and Task 13 in Output O2.

A population flow diagram – there are many examples in the public domain – see http://atlas.cid.harvard.edu/

Gapminder – we have used this in our teaching resources.

An interactive display of multivariate count data – a number of these have been adapted from SMART Centre resources (see https://www.dur.ac.uk/smart.centre/freeware/), for example, an activity based on educational attainment in the UK presented at the ICOTS10 workshop (see Task 10 in Output O2).

A quiz that runs on smartphones – a generic shell has been implemented to test local knowledge for UK constituencies (e.g. *For every 100 people, how many were born in the UK?).*

O2 - A2: Creating Interactive Teaching Tools

And

O2 – A3: Validated Classroom Tools

The stated aim was to produce robust functioning data visualisations. These can be viewed in our teaching materials.

O2 – A4: Uploaded Data Visualisations

Free standing examples can be found on the website accessible from the document reviewing DVTs; more importantly, many are embedded in teaching materials.

O2 – A5: Publication

Again, free standing examples can be found in the document reviewing DVTs; again, more importantly, many are embedded in teaching materials.

Going beyond the Proposal:

We have created CivicStatMap, a resource that provides a simple way to access teaching materials, indexed by language, educational level, statistical topic, and visualisation tool/analysis package.

We have included resources relevant to data science, notably CODAP, and Decision Trees

Tools for Data Visualization: Some Examples

This section provides an overview of a large number of software tools that use data visualization as a way to communicate information clearly and efficiently via statistical graphics, charts and plots. Many of them have facilities for data analysis. Our primary aim was to review tools which are free to use; a number of commercial tools are free to use for educational purposes; tools which are not free, were tested on their trial version.

At the end of this Section, we provide a Table with a systematic comparison of the tools, using descriptors such as: **Ways to access the tool** (Open source / Trial / Free / Online / Cloud / Educational / Payable), **Easy of use**, **Types of charts available** (Lines, Tables, Scatter, Graphs, Bar Charts, Maps, Combined Charts, Relational Diagrams, Boxplot) **Upload data/existing data**, **MicroMacro data, Connect with apps,** etc.

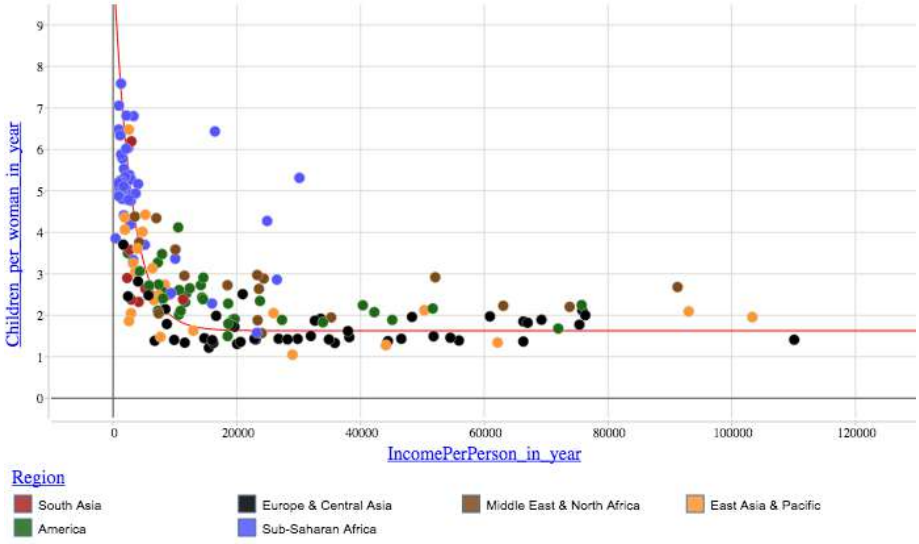
For the most part, we used the same microdata set (Nigerian Migration data from the World Data Bank) to test the main features of the software.

1. CODAP (Common Online Data Analysis Platform)

https://codap.concord.org

CODAP is free educational software for data analysis. This web-based, open source, client-based data science tool is designed as a platform for developers and as an application for students in grades 6-14. CODAP is document-based (like a word processor or spreadsheet), makes it easy to import data, has a drag and drop interface and allows (with some restrictions) web-scraping. Multiple data representations are dynamically linked. Designed as a tool for data science education, CODAP supports transformation and restructuring of hierarchically ordered data sets. The software is available in English, Chinese, German, Hebrew, Spanish and Turkish.

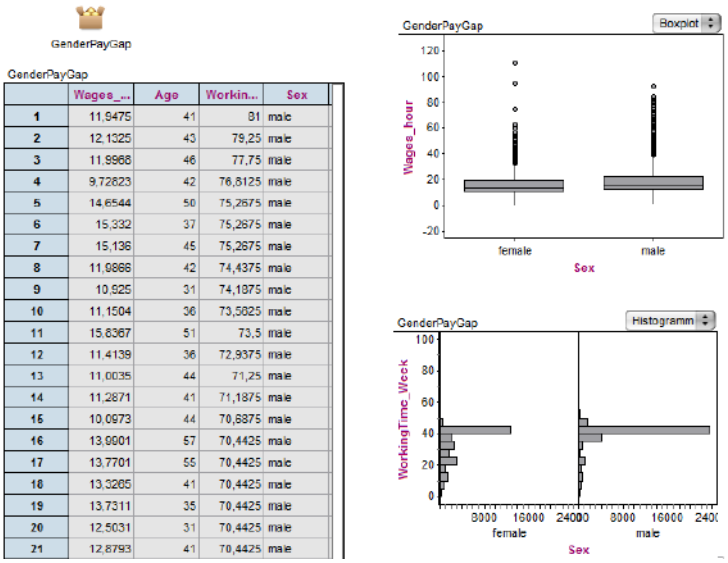
The figure below shows the average number of children per woman versus the average income per person in 2008, together with a fitted exponential curve. World regions are represented in different colours.



2. Fathom

Fathom is a data visualization and modeling tool developed for use by middle school students through to university students. The software can be used to learn and teach subjects like mathematics, statistics, social science, or physical or biological science content, basically in any course in which data are relevant. The tool offers a variety of conventional data displays like histograms, boxplots, bar charts, and scatter plots to display the distribution of categorical and numerical variables and to explore relationships between them

The figure below shows different explorations on the German Gender Pay Gap data.



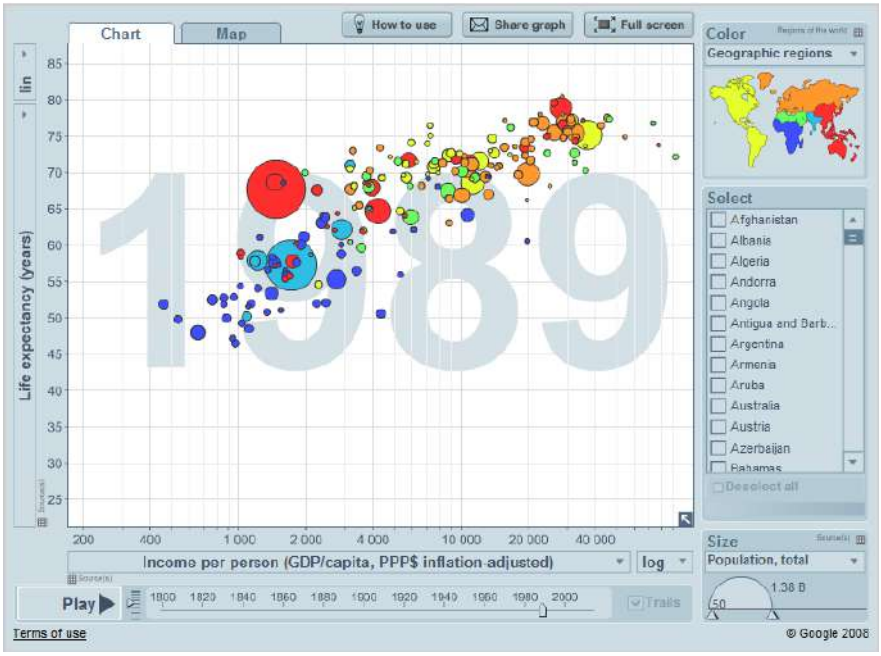
English and German versions of the software are available.

3. Gapminder

https://www.gapminder.org/

Gapminder is a project available online which promotes the understanding of statistics and other information about global development. The project includes data on topics such as Health, Population, Work, Environment, Education, Economy, Society, Infrastructure and Energy.

The Gapminder tool is very intuitive and extremely easy to use. In addition, it has the possibility of understanding the evolution of a question during a period of time. The tool supports combined and relational type of graphs. It is not possible to upload your own data. The following graph shows information about how long people live and how much money they earn.



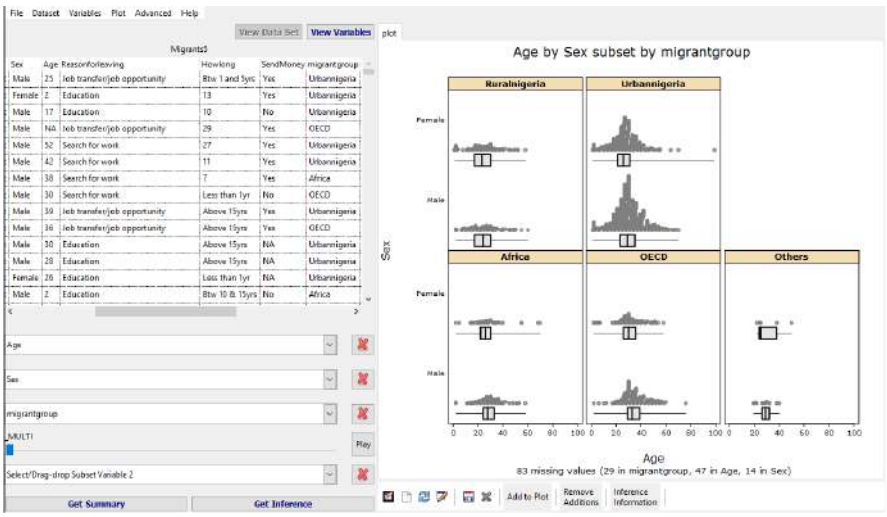
4. iNZight

https://www.stat.auckland.ac.nz/~wild/iNZight/index.php

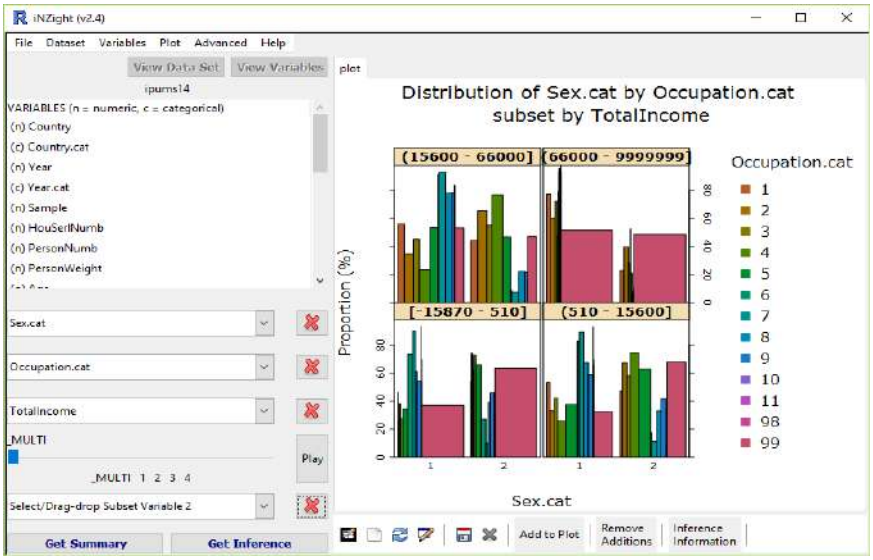
iNZight is a student-driven project, specifically written to expose rich statistical toolsets to users who would otherwise find it difficult. This freely available tool has a strong emphasis on visual representation of data and results.

The tool automatically identifies the type of variable, allows users to combine the variables and presents charts according to their type. It is necessary to be careful with some combinations which may not make sense. iNZight allows users to import their own data, even microdata.

The figure below shows information about age by sex and by region (data source: Migration data from the World data bank).



The figure below shows the distribution of sex by occupation and total income (differences of gender related with work - data source: GenderME data from IPUMS).

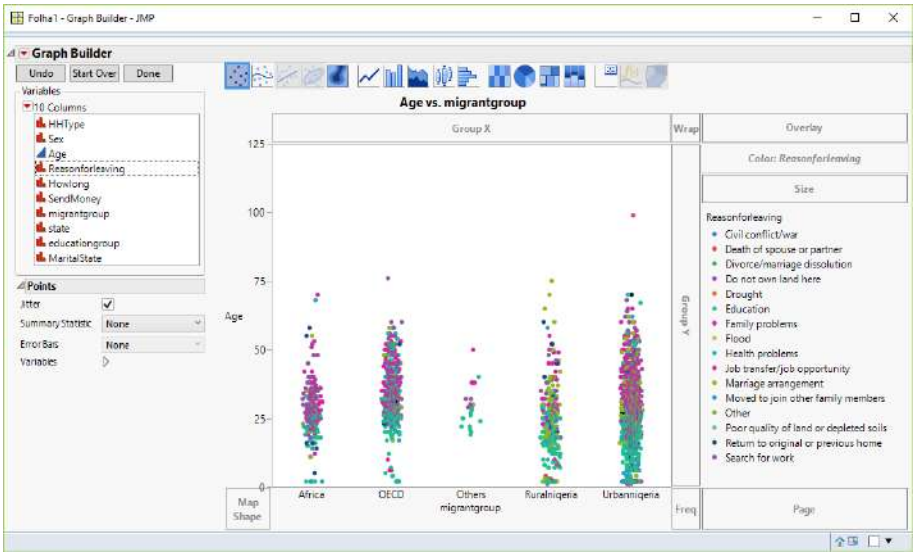


5. JMP

http://www.jmp.com/en_us/home.html

JMP is an interactive exploratory data and statistical analysis tool which automatically displays graphs with statistics, enabling users to visualize and uncover data patterns. This tool is easy to use. While it may not be as intuitive as others in this document, you create a graphic just by dragging and dropping the desired variables in place you want (axis, size, or other), it has several types of charts such as lines, combined, scatter, boxplots and bar. JMP allows users to import their own data, even microdata.

The figure below shows the age by region where migrants go and the reason why the migrants leave the place where they live (data source: Migration data from the World data bank).

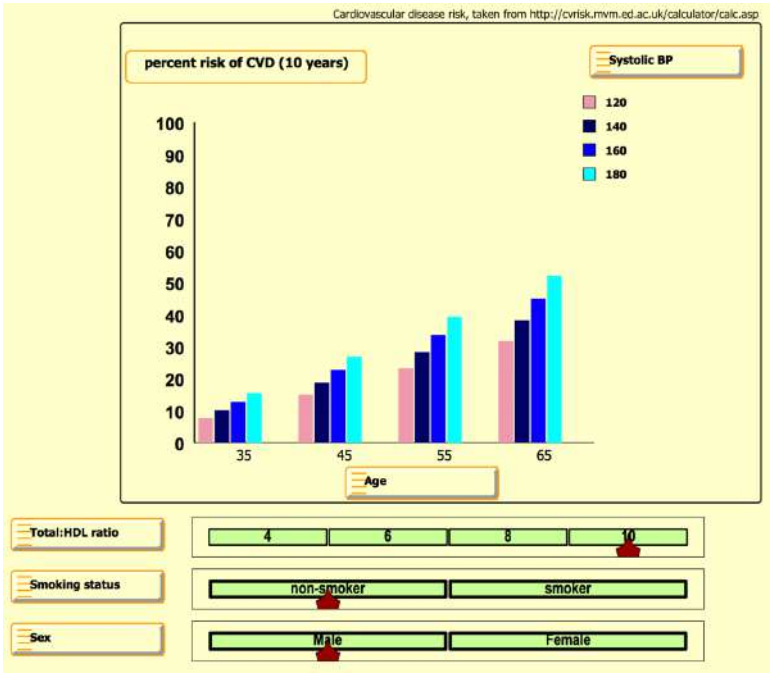


6. SMART Centre visualization tool

https://www.dur.ac.uk/smart.centre/

The SMART Centre focuses on envisioning data and reasoning from evidence. In particular, creating displays to present data in ways that can be understood, and to stimulate debate about important issues, grounded on evidence. The SMART Centre tool is an online tool and is very easy to use. Up to 6 variables can be displayed and explored. It is necessary for the data to be prepared, and it does not allow microdata to be displayed.

The following graph shows the risk of a coronary vascular disease as a function of blood pressure, age, cholesterol, smoking status and sex.



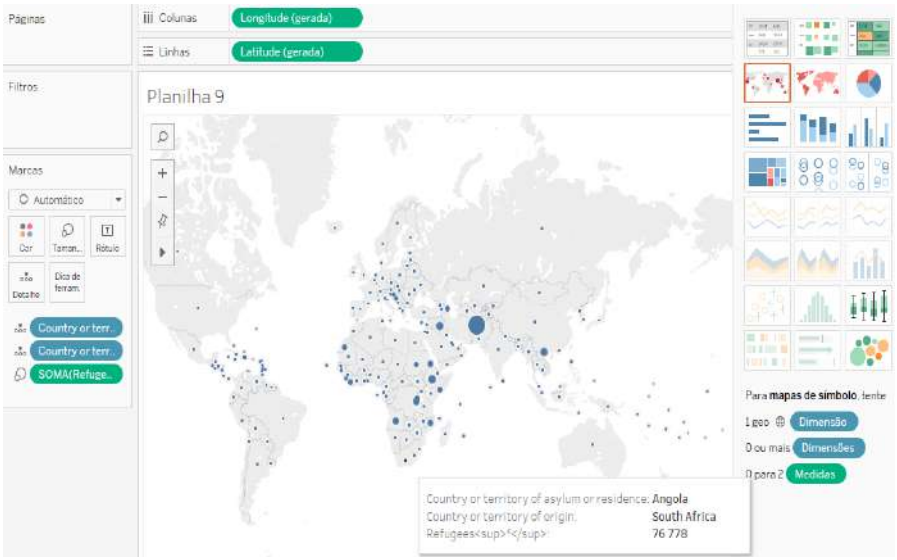
7. Tableau

http://www.tableau.com/

Tableau is mainly a data visualization tool. Users can drag and drop data into the system and watch it update in real-time. It is easy to use. The tool supports a wide variety of charts such as maps, lines, tables, bar, scatter, combined, relational and boxplots. The charts can be easily embedded in any web page. It is possible for users to import their own data.

The figure below is an example of the interface of Tableau with the Natural Disasters dataset. The graph shows the total of deaths by occurrence for region and the total of deaths by disaster type for region.

The figure below is an example of the interface of Tableau with the Refugees dataset. The graph shows the countries with more refugees, and where they came from (country of origin).



8. TinkerPlots

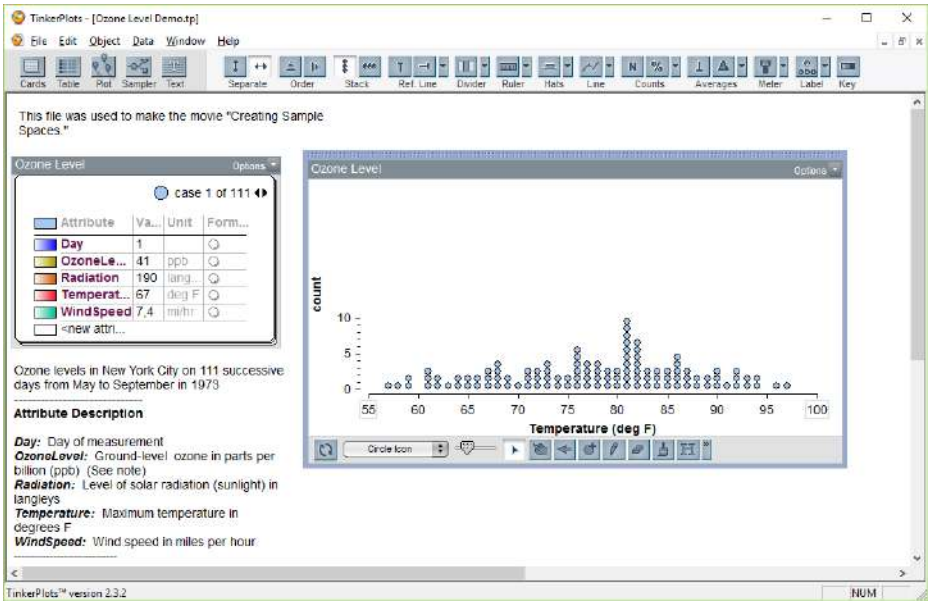
http://www.tinkerplots.com/

TinkerPlots is a data visualization and modeling tool developed for use by primary and middle school students through to university students. This educational tool with its constructive philosophy can be used to learn and to teach subjects like mathematics, statistics, social science, or physical or biological science content, basically in any course in which data are relevant. Fundamental structure of the data analysis tool is that data are stored as cards and

plots can be created by three basic operations like stack, separate, and order. Furthermore conventional statistical features such as median, mean, boxplot, histogram, and so on can be accessed easily.

The figure below shows the Ozone level by Temperature, using the Ozone data set from TinkerPlots.

English and German versions of the software are available.



9. Online population pyramids

With the help of online population pyramids we can visualize the age distribution of a society and show population projection.

http://www.worldlifeexpectancy.com/world-population-pyramid tool was developed by LeDuc Media. It shows population pyramids of different countries between 1950 and 2050. The website contains other map visualizations for instance a country health profile.
https://populationpyramid.net was started in 2011, the interactive pyramids and a line chart show data from all over the world between 1950 and 2100. The website contains other visualization tools on demographic indicators.
https://www.census.gov/data-tools/demo/idb/informationGateway.php was developed by US Census Bureau, which contains different demographic indicators around world. The website also can generate population pyramids.

Several national statistical agencies also have interactive population pyramids. For instance
https://www.ksh.hu/interactive_agepyramids is the visualization of Hungarian Statistical Office. It is available on country and on regional level.

A visualization of Germany is available at https://www.destatis.de/EN/Service/InteractiveVisual/InteractiveVisualised.html

UK population projections which allow different migration flow effects to be explored by selecting different scenarios from the drop down menus at the top of the screenshot below – see https://www.ons.gov.uk/visualisations/dvc418/pyramids_projections/index.html#20/0/3/66/98/false/true/2016/0

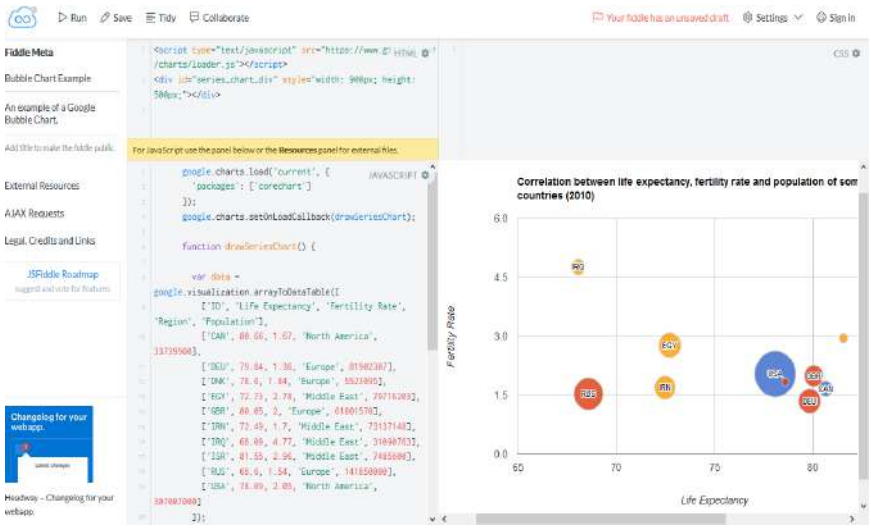
10.Gephi

Gephi is an open source and free graph visualization tool used to visualize and explore social and economic networks. The software contains multiple built-in functions to filter, cluster and customize networks along several characteristics, which makes it the proper tool to uncover patterns, trends and outliers within the data. With the help of further plugins also freely available user experience can be further customized to meet the needs of users. Gephi is publicly available for download under a General Public License at https://gephi.org/users/download/
A quick start guide, visualization tutorial, videos and explanations on how to get started with Gephi and assistance for beginner users are available at the following link: https://gephi.org/users/

11.Google Chart

Google Chart Tools (see https://developers.google.com/chart/) provide several ways to add charts to any web page. Charts may be static or interactive. The tool offers several types of charts which cover the most commonly used chart types like bar, tables, area, scatter, maps, lines and combined. The format of the data and the number of variables required for each type of chart are very important in this tool. It is necessary to preprocess/prepare the data in advance, before uploading it to the tool.

The graph below shows the relation between life expectancy, fertility rate and population of some countries.



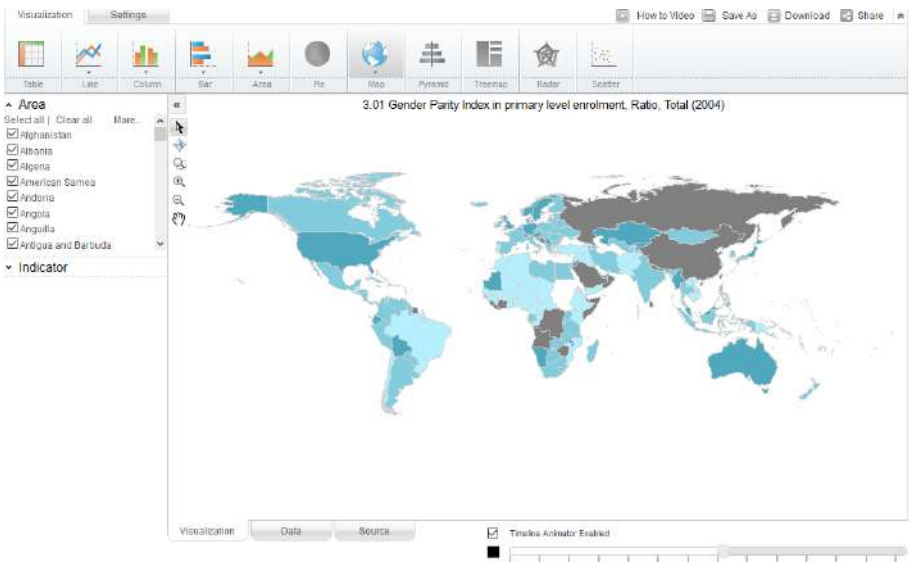
12.DevInfo

http://www.devinfo.org/libraries/aspx/Home.aspx

DevInfo is used to compile and disseminate data on human development. The DevInfo project is an interagency initiative managed by UNICEF on behalf of the United Nations (UN) System. The project includes data on topics like Nutrition, Health, Environment, Women, Education, Economy, Demography and information and Communication.

The tool is available online, and it is very intuitive and extremely easy to use. At the beginning and after choosing the data files, it is possible to see the data and visualize it. DevInfo makes suggestions about the charts available. It is also possible for users to upload their own data.

The simulation shows the gender parity index in primary level enrolment on world for 1991 and between 1999 and 2011 (data available in *Topic/Education/Gender equity* of DevInfo). Ratio for each color is available on *Settings*.



13.Infogr.am

https://infogr.am/

Infogr.am is a data visualization tool that creates and publishes data visualizations. The tool is easy to use. It has an educational version, and the charts are fast to share and easy to understand. Being easy to use, this tool considers the whole file and produces the chart by asking which type of chart the user wishes to produce (it may cause some errors, if you are not careful with the data). Infogr.am reads microdata but does not deal with particular files (such as the Migration data from the World data bank).

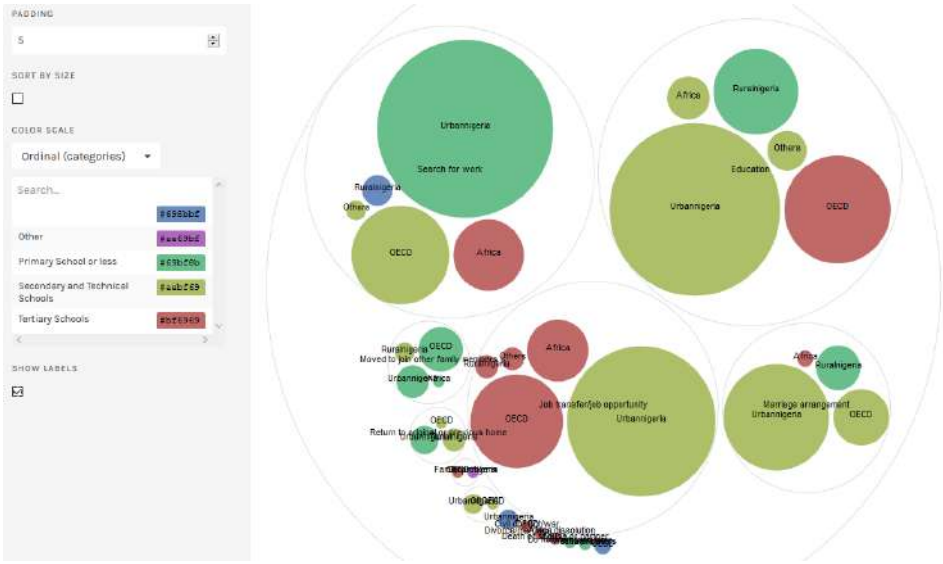
The following pie chart shows the results of PISA levels for Boys (BL1B, L1B, L2B,…) and Girls (BL1G, L1G,…) in each country. The country selected is Vietnam.

14.RAW

http://raw.densitydesign.org/

Raw is an online, free and open source tool that allows users to simply paste the data and create graphs in few simple steps. It is extremely easy to use and intuitive. It is necessary for users to be careful to show labels and legends on visualisations. Allows the use and upload of microdata. Raw also allows those with sufficient programming skills to add new types of chart.

The graph below shows region where the migrants go and educational group by reason why migrants leave the place where they live. The reason why they leave the place where they live is in the middle of the circumference and in the middle of each circle we have the region where they go. The education group is represented by the colour (data source: Migration data from the World data bank).

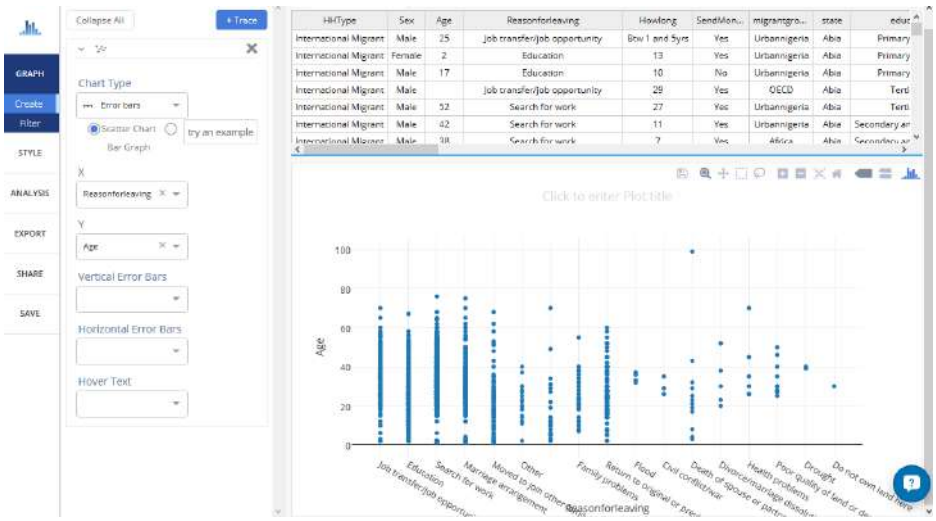


15.Plotly

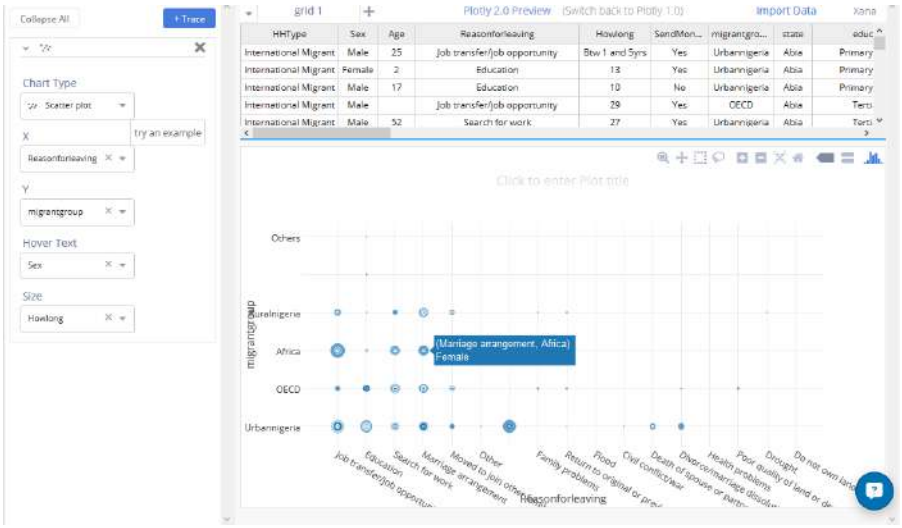
https://plot.ly/

Plotly is an analytics and data visualization tool which provides online graphing, analytics and statistics tools. The Plotly tool is easy to use and has several types of charts such as lines, matrix, combined, scatter, boxplots, maps and bar. It allows users to import their own data and deals with microdata.

The figure bellow shows the age by the reason why the migrants leave the place where they live (data source: Migration data from the World data bank).



The graph below shows the region where migrants go by reason why they leave, sex and how long migrants stay away (data source: Migration data from the World data bank).



16.Visualizefree

http://visualizefree.com/visualizations.jsp

Visualize Free is a free visual analysis tool, useful for sifting through multi-dimensional data to spot trends. It is extremely easy to use, intuitive and is available online.

This figure shows an example of visualization included on the page of the project, which is possible to explore, analysis of cause of death for major reasons in 2005 by race and sex.

17.NationMaster

http://www.nationmaster.com/

NationMaster is a vast compilation of data from hundreds of sources. It allows you to get maps and graphs on several kinds of statistics.

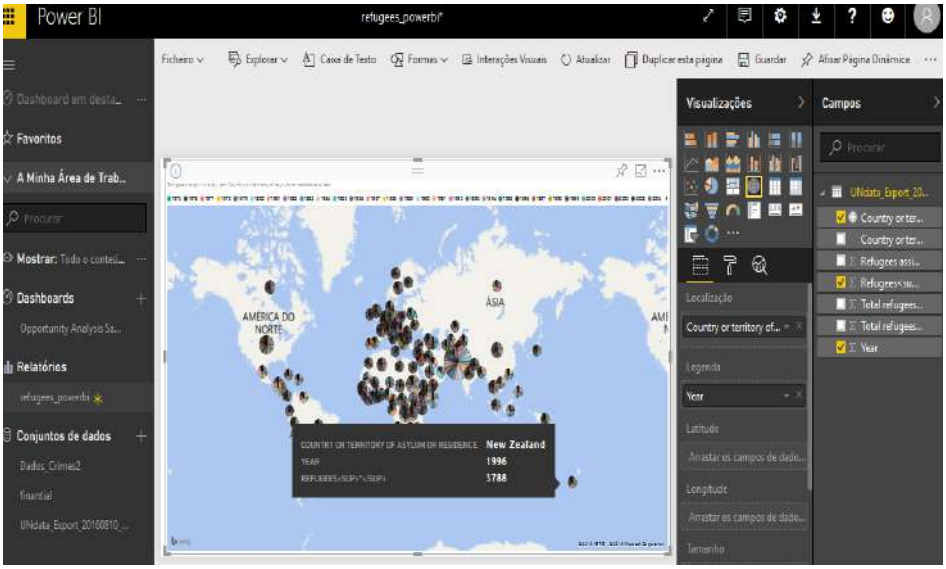
The map above represents a comparison between countries on number of human trafficking, from category crime.

18.Power BI

https://powerbi.microsoft.com/en-us/

Power BI is a suite of business analytics tools to analyze data and share insights. Monitor your business and get answers quickly with rich dashboards available on every device.
The tool is easy to use and has several types of charts (lines, tables, combined, relational, maps and bar). The Power BI has data to use, but is possible to import our data file and use microdata.

The figure below shows the number of refugees, by year, for each asylum country (data source: Refugees data from the United Nations).



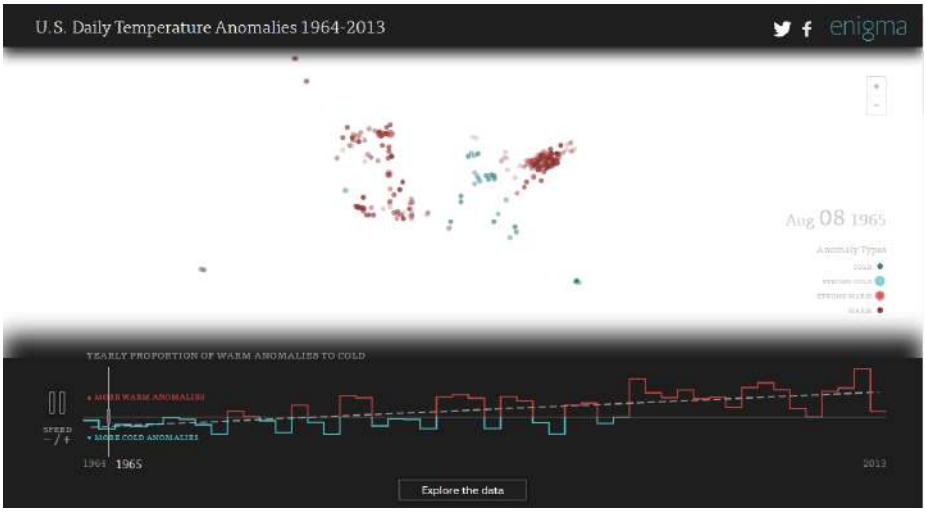
19.Other Projects:

- **Refugees**

The Refugee Project is a narrative, temporal map of refugee migrations since 1975. United Nations data is complemented by original histories of the major refugee crises of the last four decades, situated in their individual contexts.

- **Enigma**

At Enigma, the priority is to enable the exploration of climate change. Weather signals alone can be powerful indicators, but it is by placing them in relation to the economy and society that the impact of climate change becomes meaningful.



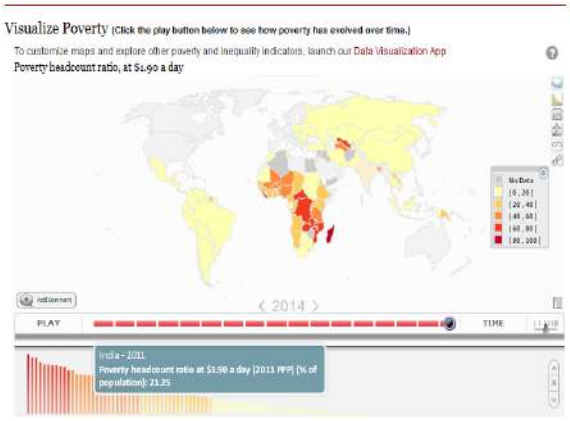
- **Women in Science**

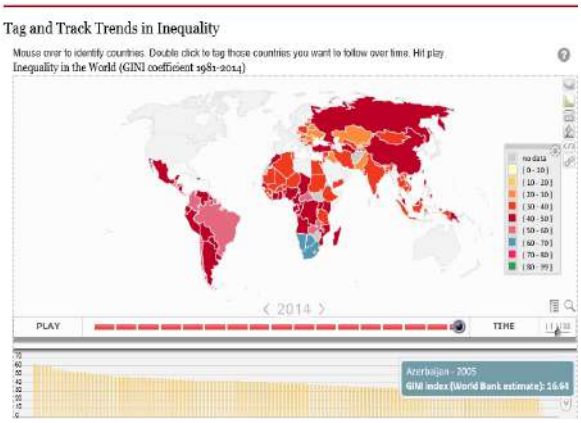
Less than 30% of the world's researchers are women. Women in Science is an interactive tool which let explore and visualize gender gaps in the pipeline leading to a research career, from the decision to get a doctorate degree to the fields of research women pursue and the sectors in which they work.

- **PovcalNet/ Poverty and Equity (World Bank)**

PovcalNet is an interactive computational tool which allows users to replicate the calculations made by the Bank's researchers and also allows calculating the numbers under different assumptions. A wider range of poverty and inequality measures than those found in the background paper are given by this tool.

Poverty indicators include the poverty headcount ratio, poverty gap, and number of poor at both international and national poverty lines. Inequality indicators include the Gini index and income or consumption distributions. The database includes national, regional and global estimates.





Google Earth Pro

https://www.google.com/earth/

Google Earth offers high resolution images of the Earth Surface, seafloor, along with information about specific locations, current global temperatures, wind speeds, precipitation, sea temperatures. Specific topics are addressed such as travel, nature, sports and history, and there is a section on education with topics that range from animal diets.

Tools for Visualisation Table

Below we present a table where (hyperlinked) tools, including those outlined above, are categorized in terms of their properties.

Pro Civic Stat

Promoting Civic Engagement via Exploration of Evidence: Challenges for Statistics Education

Co-funded by the Erasmus+ Programme of the European Union

Tool	Open source/ Trial/Free/ Online/Cloud/ Educational/ Pay	Ease of use	Lines	Tables	Scatter Graphs	Bar Charts	Maps	Combined Charts	Relational Diagrams	Boxplot	Others (not mentioned before)	Data: UD=upload/ D=existing data	Micro/ Macro data	Connect with apps
BIRT	OS/T		✓	✓	✓	✓					Pie/Stock/Bubble	UD	Ma	✓
Chart.js	OS/F		✓			✓					Pie/Polar/Radar/ Bubbles	UD	Mi/Ma	
Chartblocks	F/O		✓		✓	✓					Pie	UD	Ma	✓
ChartFX	F/T		✓		✓	✓	✓	✓		✓	Pie/Bubble/ Candlestick	UD	Mi/Ma	✓
Chartist.js	OS/F		✓			✓					Pie	UD	Mi/Ma	
CODAP	OS/F		✓	✓	✓	✓	✓		✓	✓	simulation, restruct-ering data , extensible	D/ DU	Mi/Ma	
D3.js	OS/F		✓	✓	✓	✓	✓	✓	✓	✓	Many other types	UD	Mi/Ma	✓
datamatic	T/F		✓		✓	✓	✓	✓			Pie/Rose/Bubbles/ Punch	UD	Ma	✓
datavisual	T		✓		✓	✓	✓	✓			Pie/Bubble	UD	Ma	✓
Datawrapper	F/O		✓	✓	✓	✓	✓				Pie/Donnut	UD	Ma	✓
DevInfo	F/O		✓	✓	✓	✓	✓		✓	♟	Pie/Radar	D/UD	Ma	
Dygraphs	OS		✓									UD	Mi/Ma	
Fathom	E / P		✓	✓	✓	✓	♟	♟	✓	✓	Simulations	D/ UD	Mi/ Ma	
Flare	OS/F		✓		✓	✓					Pie/Bubbles/Donnut	UD	Mi/Ma	✓
Flot	OS/F		✓		✓	✓		✓			Pie	UD	Mi/Ma	
FusionCharts	OS/T/F		✓	✓	✓	✓	✓	✓		✓	Pie/Radar/ Candlestick	UD	Mi/Ma	✓
Gapminder	O							✓	✓		Bubble	D	Ma	
Gephi														
GGobi	F				✓	✓		✓		♟		UD	Ma	
Google Charts	F/O		✓	✓	✓	✓	✓	✓			Pie/Bubble/Donnut/Can dlestick	UD	Mi/Ma	✓
Highcharts	T/F/C		✓	✓	✓	✓	✓	✓		✓	Pie/Bubble/Funnel	UD	Ma	✓

Co-funded by the Erasmus+ Programme of the European Union

Tool	Open source/ Trial/Free/ Online/Cloud/ Educational/ Pay	Ease of use	Lines	Tables	Scatter Graphs	Bar Charts	Maps	Combined Charts	Relational Diagrams	Boxplot	Others (not mentioned before)	Data: UD=upload/ D=existing data	Mi=Micro/ Ma=Macro data	Connect with apps
InetSoft	T		✓		✓	✓	✓	✓	✓			UD	Mi/Ma	✓
Infogr.am	T/E		✓		✓	✓		✓			Pie	UD	Ma	✓
Infovis	F/T					✓					Pie/Graphs/ SpaceTree	UD	Ma	
Inzight	F				✓	✓		✓		✓	Dot/Grid density	UD	Mi/Ma	
JMP	T/E		✓		✓	✓		✓		✓	Pie/Bubble/Clusters	UD	Mi/Ma	✓
jQuery Sparklines	F		✓			✓		✓		✓	Pie	UD	Mi/Ma	
NationMaster	O		✓				✓				Pie	D	Ma	
NVD3.js	OS/F		✓		✓	✓		✓			Pie	UD		✓
Online Charts	F/O		✓		✓	✓		✓			Pie/Bubbles/Radar	UD	Ma	
Plotly	T/E		✓	✓	✓	✓	✓	✓		✓	Funnel/Pie/Bubbles	UD	Mi/Ma	✓
Power BI	F/O		✓	✓		✓	✓	✓	✓		Funnel/Pie	D/UD	Mi/Ma	✓
Raw	OS/F/O		✓		✓	✓		✓		✓	Many other types	UD	Mi/Ma	✓
Slemma	T/O			✓	✓	✓	✓	✓			Pie	UD	Mi/Ma	
Smart Center	O					✓						UD	Ma	
Tableau	T/O/E		✓	✓	✓	✓	✓	✓	✓	✓	Many other types	D/UD	Mi/Ma	✓
Timeline											Timeline	UD	Mi/Ma	
TinkerPlot	E/ P		✓	✓	✓	✓			✓	✓	Pie/ Simulation tool	D/UD	Mi/Ma	
Tuva														
Visualizefree	F/O		✓		✓	✓	✓	✓	✓		Bubble	D/UD	Mi/Ma	
ZingChart	T		✓		✓	✓	✓	✓		✓	Pie/Rank flow/Chord	UD	Mi/Ma	✓

KEY:		
		Easy to use
		Not so dificult as the orange and not so easy as the green. In general is easy to use it. Don't need programming skills to make analysis.
		Is possible to use an interface to facilitate/skip the programming part. The first contact can not be very easy. Not so dificult as the red.
		Need some programming skills (libraries)

Short Activities with Data Visualisations

Data visualisations are encountered every day in print and other media. The data revolution had led to the creation of a wide range of data visualisations (DVs), and these include DVs where the user has control over some key elements of the display, such as the choice of variables, and where data can be displayed dynamically; new DVs continue to be invented. DVs facilitate the display of complex information in an accessible way; however, some user skills are required for proper interpretation. This Section shows how we have used DVs in presentations and workshops to (a) provide examples of short activities that can foster the development of PCS skills, and can form the basis for assessing PCS skills and (b) develop critical skills in interpreting visual displays.

Examples of Short Activities that can Foster the Development of PCS Skills

Task 1: Exploring Sexually Transmitted Infections

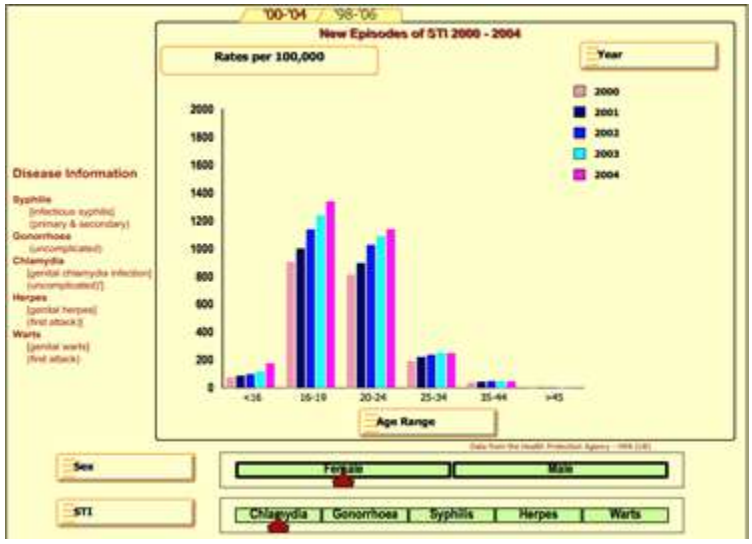
Background: This is one of a number of displays created by the SMART centre at Durham University, designed for (free) educational use. Here, the associations between 4 factors and new incidents of STI can be displayed and explored by dragging variable names in and out of the display, and by using sliders.

Data are available on new incidences of sexually transmitted diseases from every sexual health clinic in the UK. Here, we show data for the period 2000 to 2004.

Instructions:

1. Find the SMART visualization on the web:
https://www.dur.ac.uk/resources/smart.centre/Freeware/STI_GUM_update1.swf

Explore it. Then continue with the Instructions below.



Instructions (cont.):

2. Analysis questions:

   Q1: What trends can you see in the incidence of Chlamydia over time?

   Q2: What trends can you see in the incidence of genital warts over time?

   Q3: Explain any differences you see.

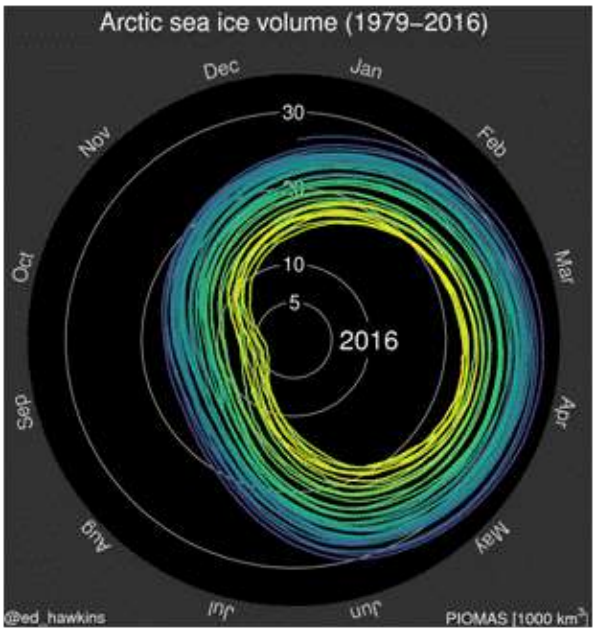
Task 2: Arctic Sea Ice Volume

Background: This graphic comes from the Polar Science Center at the University of Washington. There are no precise measurements of total ice volume. The data shown are based on a model that is validated via submarine and satellite data; the observations are collected at the Unified Sea Ice Thickness Climate Data Record.

Instructions:

1. Search for "Arctic sea climate lab", or use this URL:
   http://www.climate-lab-book.ac.uk/files/2016/06/icevol.gif

   Under "climate spirals", run the visualization and explore it. Then continue with the Instructions below.



Instructions (cont.):

2. Answer the questions below:

   Q1: What variables are being visualised?

   Q2: What trends can you see within each year?

   Q3: What trends can you see across years?

   Q4: Is our planet getting warmer?

Task 3: When Will You Die?

Background: This graphic comes from Nathan Yau's Flowingdata website. The original data is taken from the Actuarial Life Table of the Social Security Administration (USA).

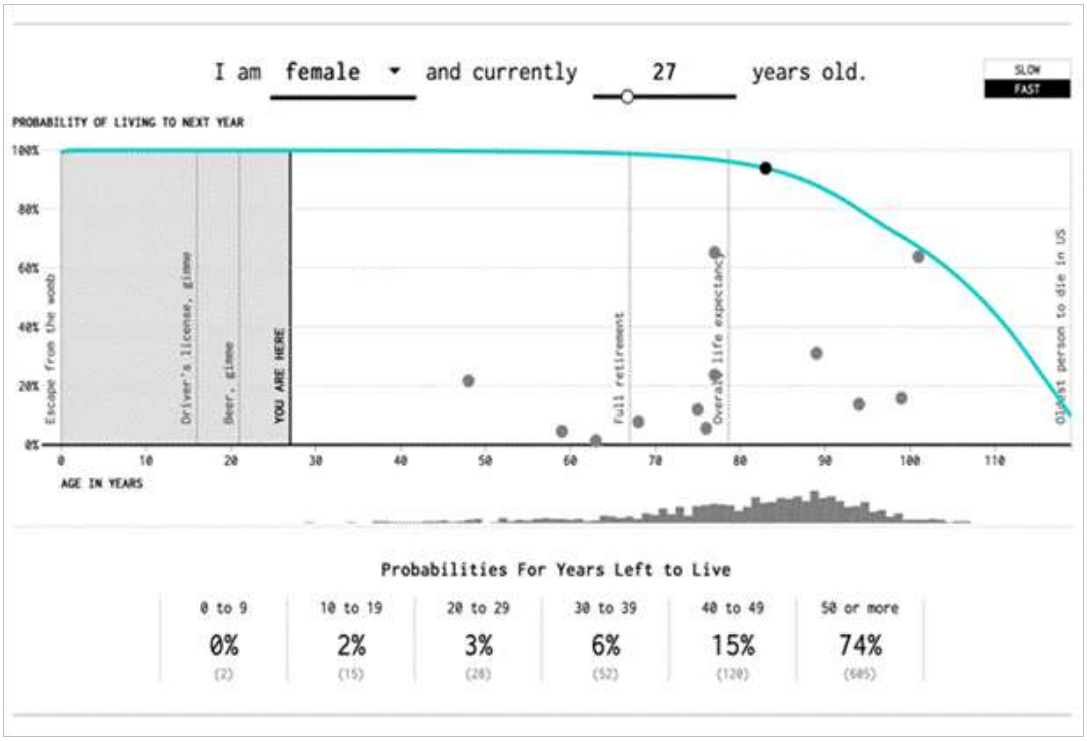
Source: https://www.ssa.gov/oact/STATS/table4c6.html

Instructions:

1. Find the Flowingdata visualization on the web:

http://flowingdata.com/2015/09/23/years-you-have-left-to-live-probably

2. Run it (scroll down, read some text). Then continue with the Instructions below.



Instructions (cont.):

3. Answer the questions below:

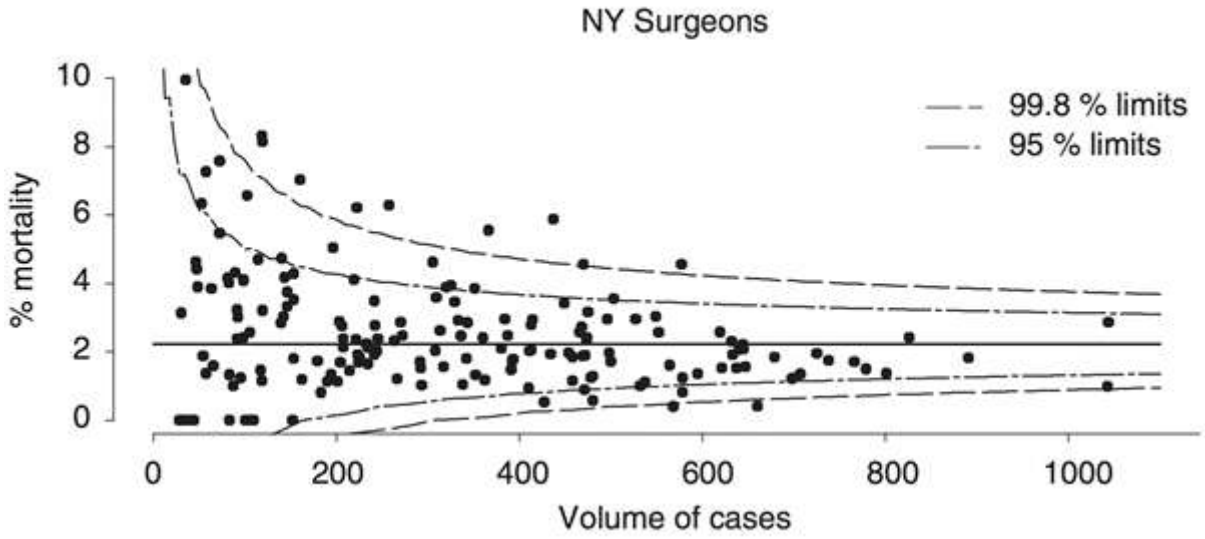
Q1: What does '74%' mean?

Q2: When will ***you*** die?

Q3: Under what assumptions is the life expectancy of a 30 year old lower than the life expectancy of a 70 year old?

Task 4: Funnel Plots

Background: The Figure shows risk-adjusted mortality rates following coronary artery bypass grafts in New York State 1997-99 for 175 surgeons who conducted at least 25 operations. It is taken from Spiegelhalter, D. (2005). Statistics in Medicine 2005(14): 1185-1202. DOI: 110.1002/sim.1970



Instructions:

1. Examine the display.
   Q1: How is the number of operations performed by a surgeon related to patient mortality?

   Q2: Explain why it can be more sensible for someone to choose a surgeon with a higher mortality rate than one with a lower mortality rate for their bypass graft.

2. Discuss your impressions and thoughts with your group (add more notes if appropriate).

Task 5: Infant Development

Background: Growth curves are used in many countries to track the development of infants in order to identify abnormal patterns of development. The chart shows length-for-age and weight-for-age for neonatal girls in the USA. There is obvious non-linearity. Users are expected to map development over time, and to look for abnormal patterns of development. These growth curves have been taken from the Centers for Disease Control and Prevention in the USA – see

https://www.cdc.gov/growthcharts/clinical_charts.htm

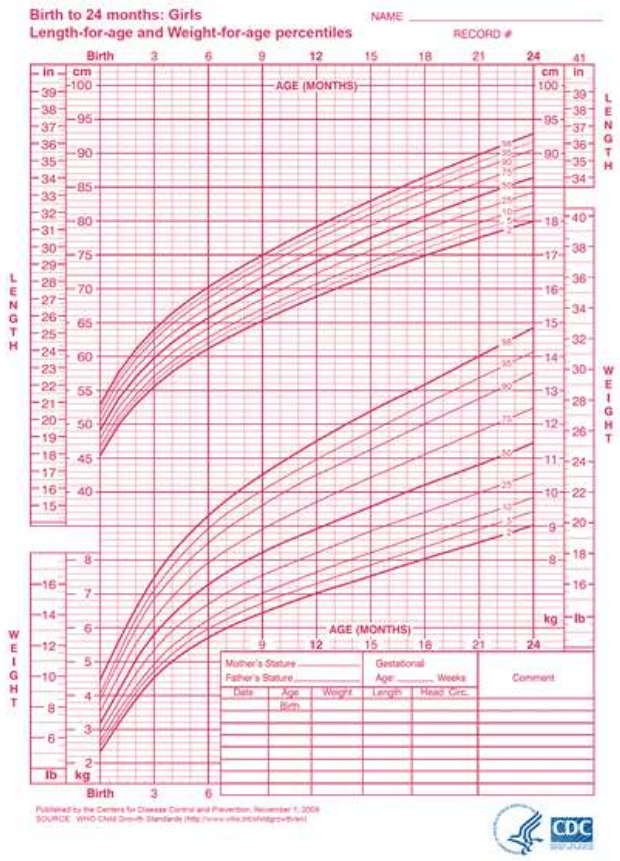
Instructions:

1, Complete the tasks, then answer the question below. Then continue with the Instructions afterwards.

T1: Draw the trajectory of a girl who is normal at birth, then fails to thrive.

T2: Draw the trajectory of a girl who shows signs of becoming obese.

Q1: When might you NOT use these charts to plot developmental progress?



Task 6: OECD Better Life Index

Background: The OECD Better Life Index focuses on aspects of life that matter to people and that influence the quality of their lives. This chart shows different aspects of areas of well-being in Germany, based on a ranking of all OECD countries. Longer lines show areas of relative strength, shorter lines show areas of relative weakness.

Details: www.oecd.org/statistics/Better-Life-Initiative-2016- country-notes-data.xlsx.

Instructions:

1. Examine the display below. Then continue with the Instructions afterwards.

Instructions (cont.):

2. Analysis questions:

Q1: What statistical knowledge do you need to make sense of this graphic?

Q2: What statistical questions do you have about the validity of this approach?

Task 7: Indices – using OECD's Better Life Index

Background: The BLI invites users to input their own criteria for 'quality of life' and creates the user's own index. This is then used to rate the Quality of Life in OECD countries.

Instructions:

Go to http://www.oecdbetterlifeindex.org/#/11111111111

Rate the importance of different topics to you.

Use the data available to examine sex differences in BLI.

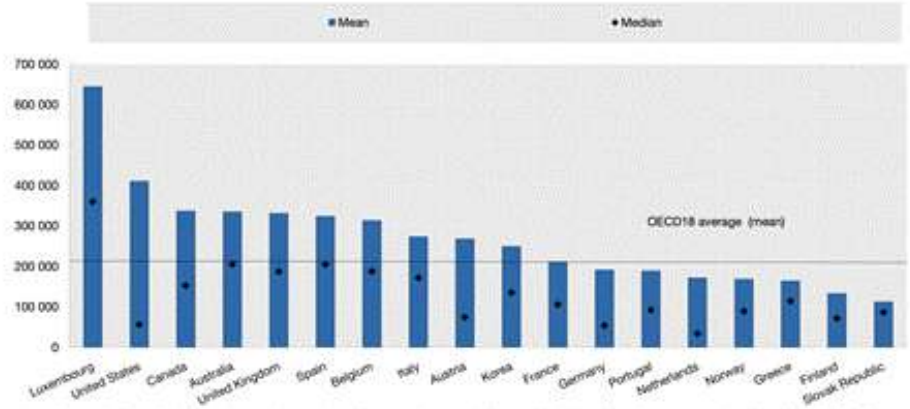
Change your index to make YOUR country ad desirable as possible, then as undesirable as possible.

Write down some important features of indicator systems.

Task 8: Conjectures about Electoral Success

Use the figure below to provide one explanation of the election of President Trump. Using this analysis, which other countries are most likely to produce 'surprising' election results?

Figure 1. **Mean and median net wealth per household in selected OECD countries**
2010 or latest available year, values in 2005 USD PPPs



Note: Countries are ranked from left to right, in decreasing order of mean household wealth. Wealth values are ranked in 2005 USD: first, wealth values in different years are expressed in prices of the same year (2005) though consumer prices indexes; second, national values are converted into a common currency through the use of purchasing power parities for household consumption.
Source: OECD Wealth Distribution Database.

Task 9: Population and Politics

This graphic is a screenshot taken from gapminder.org

The graph shows the population of China and the USA from 1800 to 2015.

Q1: Should the USA insist that China reinstates its 'one-child' policy? Justify your answer.

Task 10: Explaining Educational Attainment

What factors are associated with educational attainment?

Here we present population data from England on educational attainment in successive years, disaggregated by sex, ethnicity, region, and entitlement to free school meals (a measure of family poverty) in an interactive display for you to explore, explain, and offer advice about what to do!

Context and Data: In England, almost all children take high-stakes tests at the end of compulsory schooling (aged 16 years). These tests are subject based (e.g. in mathematics, geography, science). Grades A* to C are viewed as a 'good pass'.,Entry to many careers (e.g. the police, civil service, nursing) requires a 'good pass' in at least 5 subjects.mSo *5 good passes* is a reasonable measure of educational attainment. Another measure only counts *5 good passes* if they include passes in maths and English.

Before we start to explore…What do you expect to see?

Please circle one of **yes ?? no**

S1: Girls will do better than boys **yes ?? no**

S2: Some ethnic groups will do better than others **yes ?? no**

S3: Children from richer homes will do better than children from poorer homes **yes ?? no**

S4: Scores will get better over time **yes ?? no**

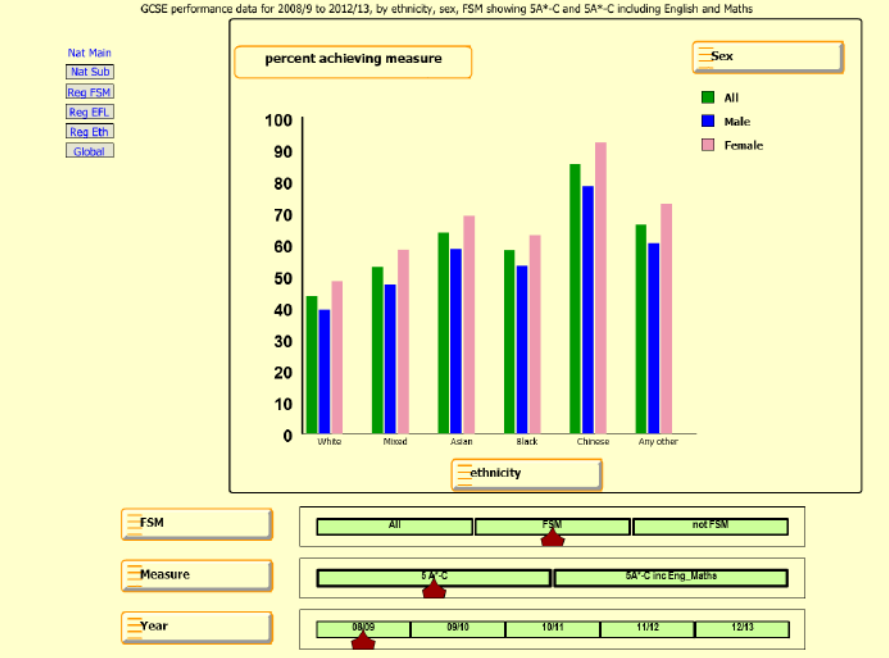
S5: Scores in some regions will be consistently higher than in other regions **yes ?? no**

Strength of Association

Put these factors in order of strength of association with educational attainment ('1' for the factor with the strongest association, '2' for the next strongest association...)

Sex	Ethnicity	Poverty	Year of Test	Region

Exploring the Data Visualisation -we will guide you through ways to explore the DataViz



Check your answers to the earlier questions! Write down some things you discovered.

Role Play!

Congratulations! – you have just been appointed Minister for Education in England.

You want to raise educational attainment.

Q1: What might you do, based on the data here?

Q2: What other information would you want to see, to help your policy making?

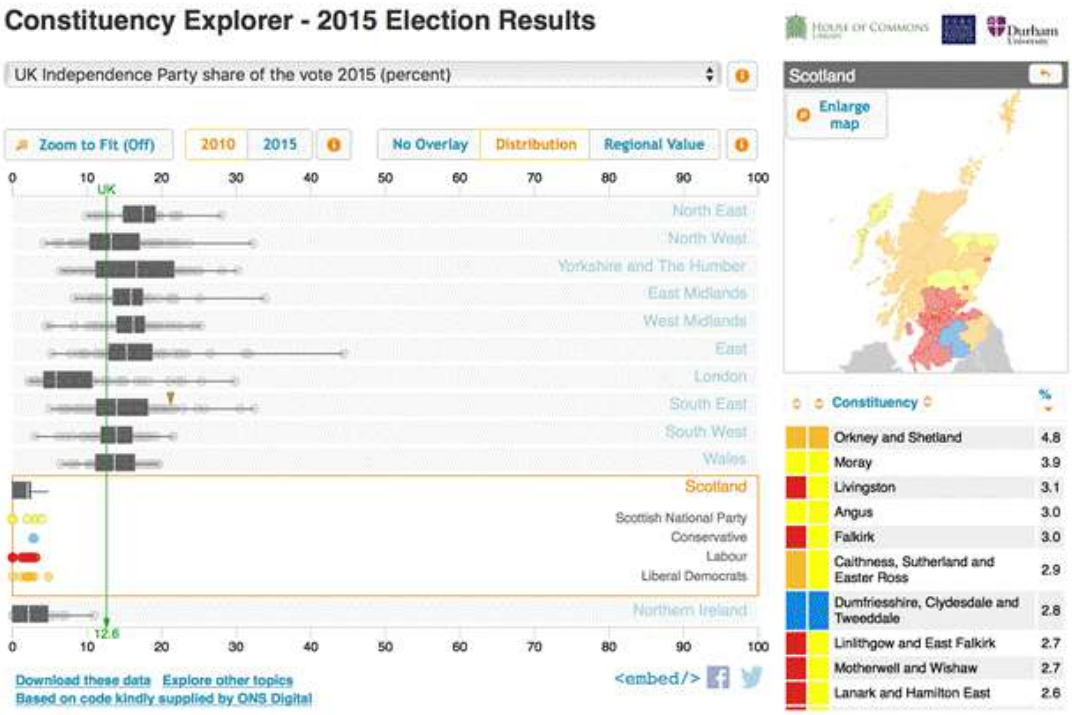
Task 11: Understanding Brexit Voting Patterns

Background: The Constituency Explorer was created by Durham University and the House of Commons Library to provide data on 150+ variables for every constituency in the UK, ahead of the 2015 general election. It now includes the 2015 election results, and users can view the successes and failures of political parties, dynamically.

Instructions:

1. Find The Constituency Explorer on the web:
http://www.constituencyexplorer.org.uk/explore/2015_election_results

Start to explore it. Then continue with the Instructions below.



Instructions (cont.):

2. Analysis questions:

   Q1: What changes in winning party can be seen in Scotland between 2010 and 2015?

   Q2: The United Kingdom Independence Party (UKIP) strongly supported Brexit. Were there more UKIP supporters in Labour (red) or Conservative (blue) constituencies?

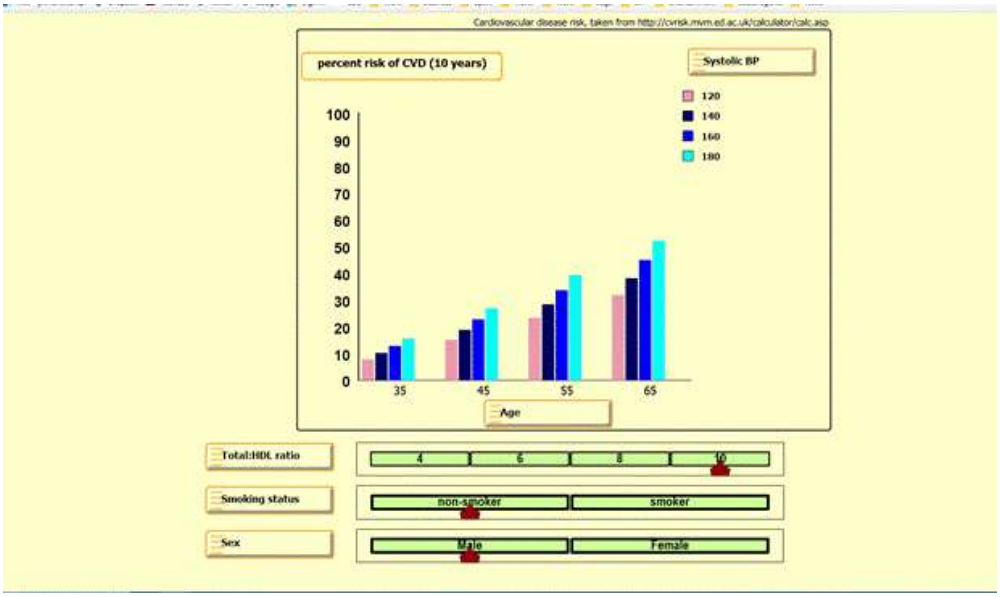
Task 12 – Exploring Risks Associated with Cardiovascular Disease

Instructions:

1. Go to the SMART Center website at:
https://www.dur.ac.uk/resources/smart.centre/Freeware/6dCVAsex.swf

2. Press START

3. Explore the visualization: try to move sliders, click on things, drag & switch the variables names on X and Y axes, etc. Study the changing data displays. What can you conclude?



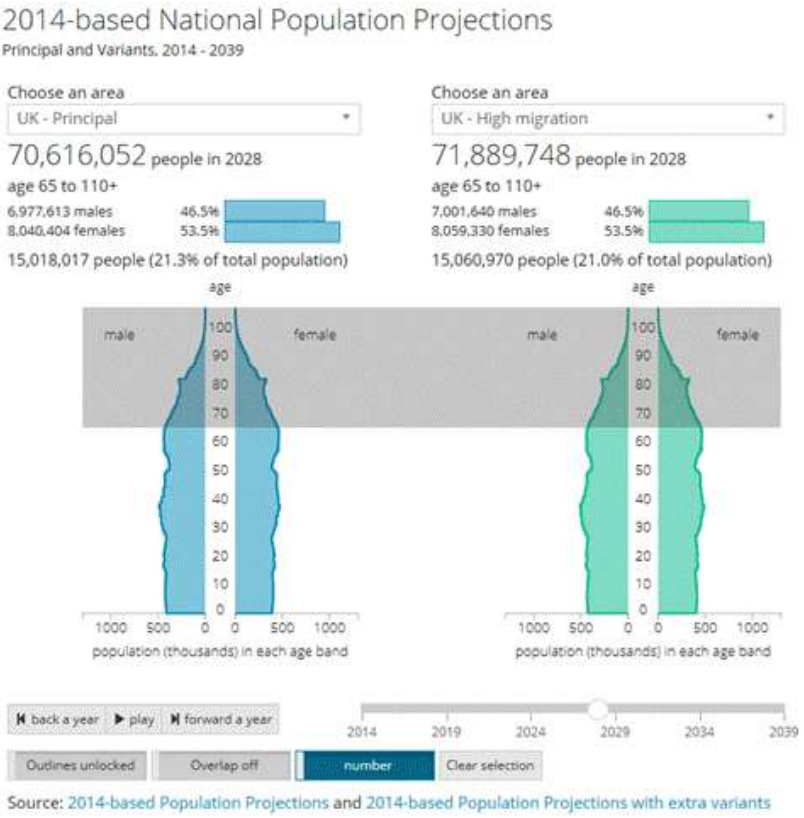
Task 13 - Dynamic Population Pyramids

This is an example of a dynamic population pyramid produced by the Office for National Statistics that allows the user to explore the effects of changing some of the modelling assumptions.

Instructions:

1. Go to https://www.ons.gov.uk/visualisations/nesscontent/dvc219/pyramids/index.html

The screenshot below shows the comparison of the UK principal projection for 2028 (in blue on the left) with the projection assuming high migration (in green on the right), highlighting the group of people aged 65 or more (in grey across the top).



Source: 2014-based Population Projections and 2014-based Population Projections with extra variants

Instructions (cont.):

2. Explore how different the population projections are for the different factors ONS uses in their projection models, such as fertility, migration, life expectancy and many others. You can choose high and low levels of a factor to display side by side. Note that when you change a new 'area' to display the year will return to 2014 (the base year) so the two displays will be identical until you move the projection time forward.

3. Explore how the differences in projected values behave as you project further into the future.

Task 14 - Identifying 'user beware' Conceptual Errors in Graphical Displays

A large number of studies have analysed misleading graphics presented in print media, and have offered heuristics to enable users to avoid being misled. It is important that students learn to read and critique dynamic graphical displays; however, we feel that emphasising flaws in some displays might distract students from seeing the great benefits that interactive displays can offer. Here, we show just 2 displays with design flaws.

Example 1: Eurostat Satisfaction with Personal Relationships

Background: The Eurostat widget is an application designed to be embedded in any website. It shows a selection of Eurostat data in a variety of ways. A number of different displays can be viewed. Here, we present a single sample chart

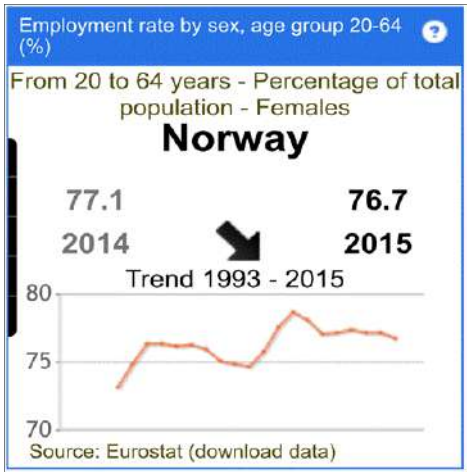
Source: http://ec.europa.eu/eurostat/cache/infographs/qol/index_en.html

Instructions:

Identify 2 ways in which this display is misleading

Example 2: Eurostat Charts: Employment of Females in Norway

Source: http://ec.europa.eu/eurostat/en/web/main/wtg



Co-funded by the
Erasmus+ Programme
of the European Union

Instructions:

1. Examine the display above.
2. Answer the questions below:

Q1: How is the 'Employment rate' measured?

Q2: How would you change this graph?

Q3: Explain your answer.

Q4: What can you conclude from this graph?