Building a Dedicated Team for Innovation in Data Visualisation - SCAD's Experience

IAOS Conference 2016 Abu Dhabi, UAE

Authors: Badria Abdulla Obaid; Gregory Pole
Statistics Centre – Abu Dhabi
P.O. Box: 6036 Abu Dhabi, UAE
T: +971.2.81.00.000

Abu Dhabi, August 2016

Abstract:

Statistics Centre – Abu Dhabi (SCAD) is motivated by innovation. In particular – innovation in dissemination.

In 2013, SCAD established a dedicated Data Visualisation team. This team has since developed many innovative and forward-looking data visualisation tools and outputs. Some examples include: infographics, animated visualisations, thematic maps, reports generators, statistical dashboards (eSCAD), D3, quizzes, mobile apps, videos, and more. These tools have been applied to a wide range of data sets.

The client uptake of these tools has been impressive. The number of interactive tools users increased by 520 per cent between 2014 and 2015.

SCAD's Data Visualisation Team has a varied skill set - including a combination of statistical, information technology, and graphical design expertise. This team also has a strong research role. That is, the ability to locate and understand new technologies (both hardware and software), and how these might apply to statistical dissemination.

This paper aims to highlight the benefits of establishing a dedicated data visualisation team within a statistical organisation; the outputs; the main challenges faced along the way; and the skill sets required. In addition, the paper will share some of the unique outputs generated by SCAD's Data Visualisation team.

Key words: *data visualisation, interactive tools, statistics, design, user experience, internet, innovation, technology.*

1. Introduction

1.1 Abu Dhabi

Abu Dhabi is the federal capital of the United Arab Emirates (UAE) and the largest of the seven emirates. Geographically, Abu Dhabi lies on the borders with the Kingdom of Saudi Arabia, the Sultanate of Oman, and the Arabian Gulf. Over the past 40 years, Abu Dhabi has experienced significant population growth and economic development.

To manage the growth and prosperity of the Emirate, the Government of Abu Dhabi required an official agency that could provide statistics for decision-making and policy setting.

1.2 Statistics Centre – Abu Dhabi

Statistics Centre – Abu Dhabi (SCAD) was established in accordance with Law #7 for the year 2008. SCAD is responsible for the collection, classification, storage, analysis and dissemination of official statistics covering social, demographic, economic, environmental and cultural indicators.

As a young statistical office, SCAD is in the fortunate position of being able to implement best practices from international bodies and leading National Statistical Organisations (NSOs). SCAD is aiming to be a world leader in innovative and efficient methods for data collection, analysis and dissemination.

2. Data Visualisation

2.1 What is data visualisation?

Data visualization is a term that describes any effort to help people understand the significance of data by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based data can be exposed and recognized easier with data visualization software. [1]

2.2 Examples of data visualisation

Some common data visualisation outputs include: infographics, interactive web visualisations (e.g. D3, SAS, etc), animated videos, charts (bar, pie, box plot, histogram, scatter plot, etc), dashboards, tree maps, thematic maps, etcetera.

2.3 What are the benefits of data visualisation?

2.3.1 Quicker understanding of data (pre-attentive processing)

Presenting statistics in a graphical representation can aid users to see large amounts of data in clear, cohesive ways – and draw conclusions from that information quickly. You could say it's a case of '*a picture is worth a thousand data records'*

A human can distinguish differences in line length, shape, and colour readily without significant processing effort - these are referred to as 'pre-attentive attributes'. For example, it may require significant time and effort (attentive processing) to identify the number of times the digit '5' appears in a series of numbers; but if that digit is different in size, orientation, or colour, instances of the digit can be noted quickly through pre-attentive processing.[2] See Figures 1 and 2.

2.3.2 Reveal relationships, patterns, and trends

Edward Tufte wrote, "*Graphics reveal data. Indeed graphics can be more precise and revealing than conventional statistical computations*". [3]

In that respect, consider Charles Minard's graphic that presents the loss of troops in Napoleon's invasion of Russia (see Figure 3). The visualisation uses six variables: the size of the army, its location on a two-dimensional surface (x and y), time, direction of movement, and temperature. The line width illustrates a comparison (size of the army at points in time) while the temperature axis suggests a cause of the change in army size. This multivariate display on a two dimensional surface tells a story that can be grasped immediately. [4]

Compared to textual data, data visualisations make it easier to discover:

- Trends (time-series) – e.g. a line chart showing unemployment rate over a 10-year period
- Rankings – e.g. bar chart showing exports by in descending order.
- Proportions (part-to-whole) – e.g. pie chart or bar chart can show the comparison of ratios
- Deviations – e.g. an infographic showing population growth from a base year.
- Frequency distributions – e.g. a histogram displaying population by travel time to work
- Correlations - e.g. a scatter plot comparing unemployment (X) and inflation (Y) for a sample of months.
- Geographic or geospatial – e.g. a choropleth map comparing import value by countries.

2.3.3 Interact with data

Unlike publications and tables, data visualization tools enable users to interact with data in an engaging and customisable way. Common data visualization features include: drill-downs, sliders, data selectors, variable drag and drop, filtering, axis swapping, colour and range selectors, etc.

With data visualisation tools users are free to investigate the data in ways that correspond to their needs. Similarly, the statistician (creator) is not obliged to have to predict what the user requires. The statistician can provide the data and a data visualisation tool and give the user the flexibility and freedom to design their own outputs. Several SCAD examples of interactive data visualization tools are provided in Section 4.

3. Creating a Data Visualisation Team

3.1 Multi-disciplinary team

It is SCAD's belief that a data visualisation team requires a mixture of skills and expertise from several disciplines. To make great data visualisations, input and consultation is required from the graphic design, IT, and statistical disciplines. SCAD's Data Visualisation team includes the following roles:

Graphic artists – Consult with statisticians to understand the stories in the data and reflect these in infographics. Coordinate design work for tool interfaces, colour palettes, icons, logos, etc. Has professional understanding of visual perception and layout.

Technology developers: Produce the interactive tools, including access to datasets, coding, testing, etc. Works with clients/statisticians to understand requirements. Conduct research on new and emerging technologies.

Statisticians: Subject matter experts who are often the 'client' during development. Provide advice on the theme/story in the underlying data. Provide quality review of both the accuracy of the presentation data and the functionality of the tool. Approve the release the data visualisation.

Manager: Coordinates multiple data visualisation projects. Liaison point between executive management, statisticians, and clients. Encourages research and development within the team.

4. SCAD's Data Visualisation Outputs

4.1 Introduction

The following list provides an overview of some recent SCAD data visualization initiatives. These initiatives are officially released projects. SCAD also has completed some 'experimental' developments, such as augmented reality app, which are not included below.

It is SCAD's hope, that by sharing these visualisations, other statistical offices will be encouraged to progress data visualization within their dissemination suite.

4.2 Infographics

An infographic is a visual representation of data using charts, icons, or pictures intended to present information quickly and clearly. SCAD creates both single infographics and storyboard infographics.

See Figure 4, or link→https://www.scad.ae/en/pages/infographics.aspx

4.3 Animated Visualisations (time-series)

One of the best methods for presenting time-series data is by employing animated visualisations. Most animations of this type have a 'play' and 'stop' buttons that allows the user to see a progression of changes over time. For example, changes in Abu Dhabi's population size over 40 years.

See Figure 5, or link→https://www.scad.ae/SCADDocuments/AD_DevStatistics_Timeseries_EN.html

4.4 Mobile Applications

SCAD's Data Visualisation team has developed several mobile apps. These include:

- *Key Statistics* – three taps to get time-series data.
- *SCAD mServices* – a range of SCAD services and innovative tools available on mobile.
- *Statistical Calculators* – personal inflation and purchasing power calculators.

- *AD Price Watch* – uses CPI data to track average price changes of items.
- *AD in Figures* – uses picture icons of statistics to assist younger users access data.

See Figure 6, or link→https://www.scad.ae/en/pages/mobileapp.aspx

4.5 Statistical Quizzes

SCAD has developed two interactive and fun quizzes that support learning about Abu Dhabi Emirate and how it compares with OECD countries.

See Figure 7, or link→https://www.scad.ae/en/pages/datavisualizations.aspx

4.6 Thematic Maps

Using thematic (choropleth) maps is a great way to show geographic distribution of statistics.

See Figure 8, or link→https://www.scad.ae/en/pages/datavisualizations.aspx

4.7 Videos (animated)

SCAD has developed several animated educational videos that explain statistical themes and provide related data.

See Figure 9, or link→https://www.scad.ae/en/pages/datavisualizations.aspx

4.8 Automated Country Reports

This foreign trade tool uses interactive maps to allow users to select a country, or countries to generate a comprehensive report in seconds.

See Figure 10, or link→https://www.scad.ae/en/pages/datavisualizations.aspx

4.9 Dashboards (eSCAD)

eSCAD is an on-line dashboarding tool that is interactive and can combine several data sources into a single platform. This popular dissemination tool can be customised by filtering variables (e.g. region, citizenship, gender, etc.). Additionally, the user can "drill down" through an indicator to another more detailed dashboard.

See Figure 11, or link→https://dv.scad.ae/dashboards/viewer/?guestuser=Public&dashID=142&c=0

4.10 Interactive Yearbook

This online tool provides an interactive spider diagram as an alternative way of allowing users to access the Statistical Yearbook tables.

See Figure 12, or link→https://www.scad.ae/en/pages/datavisualizations.aspx

4.11 D3 Charter

Data-Driven Documents, or D3, is a JavaScript library for producing dynamic, interactive data visualizations in web browsers. It makes use of the widely implemented SVG, HTML5, and CSS standards. In contrast to many other libraries, D3.js allows great control over the final visual result. [5] SCAD has built several D3 applications, including a charting tool that allows the user to load data and create beautiful charts.

See Figure 13, or link→https://www.scad.ae/SCADDocuments/AD_DevStatistics_Timeseries_EN.html

5. Conclusion

SCAD recognizes its responsibility in expanding and improving the quality, timeliness, and accessibility of statistics produced in the Emirate of Abu Dhabi. The suite of new and innovative dissemination outputs outlined in this paper is evidence of SCAD's commitment to this cause. This project has raised expectations of the types of output SCAD can produce and has set a positive benchmark for other future statistical outputs. SCAD hopes that this paper will provide other NSOs with valuable information when considering developing similar innovative outputs.

6. References

[1] TechTarget, http://searchbusinessanalytics.techtarget.com/definition/data-visualization

[2] Steven Few, Perceptual Edge, Tapping the Power of Visual Perception, September 2004

[3] Tufte, Edward, The Visual Display of Quantitative Information, Cheshire, Connecticut: Graphics Press. ISBN 0-9613921-4-2, 1983.

[4] Wikipedia, https://en.wikipedia.org/wiki/Data_visualization

[5] Wikipedia, https://en.wikipedia.org/wiki/D3.js

7. Figures

Figure 1: Attentive processing (identify the digit 5)

Figure 2: Pre-attentive processing (identify the digit 5)

Figure 3: Napoleon's invasion of Russia (Charles Minard)

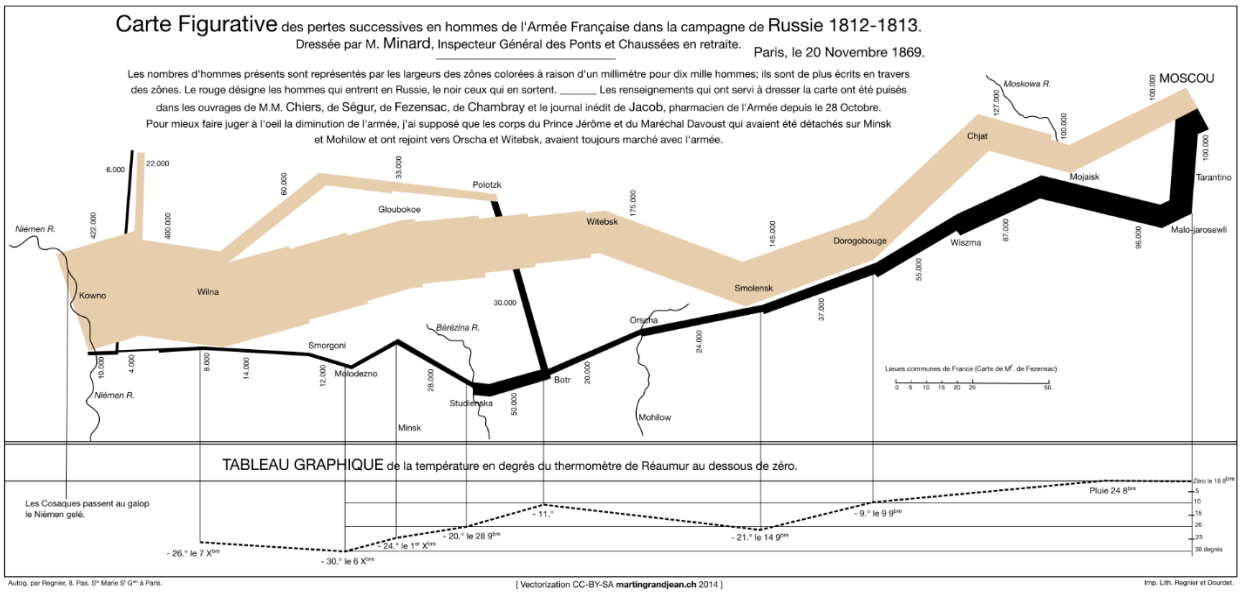


Figure 4: Infographics

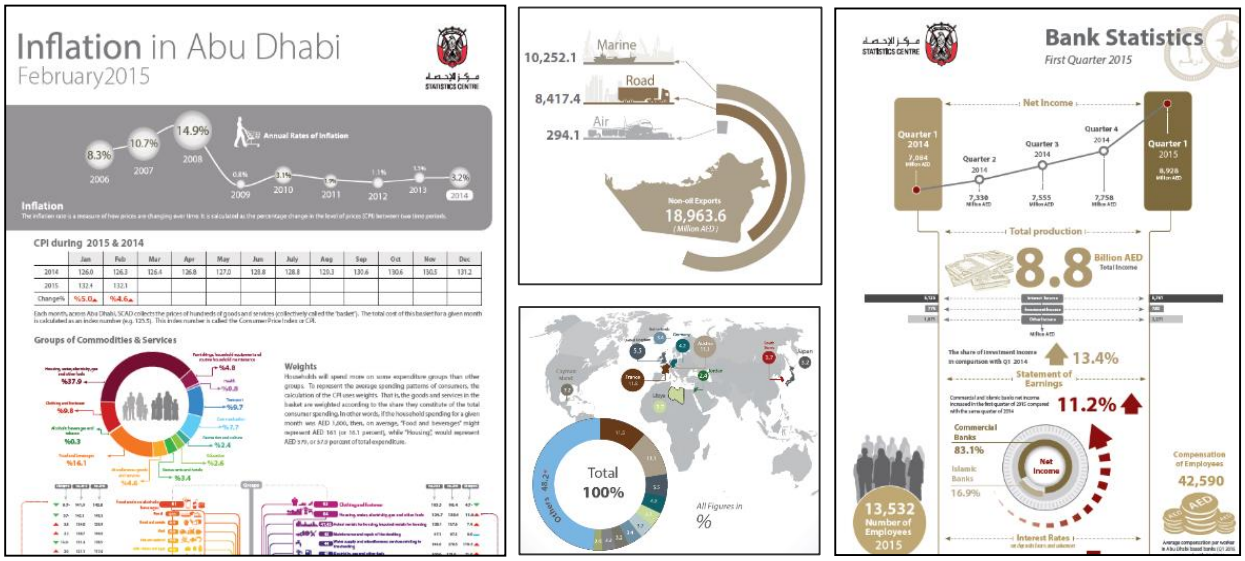


Figure 5: Animated Visualisations (time-series)

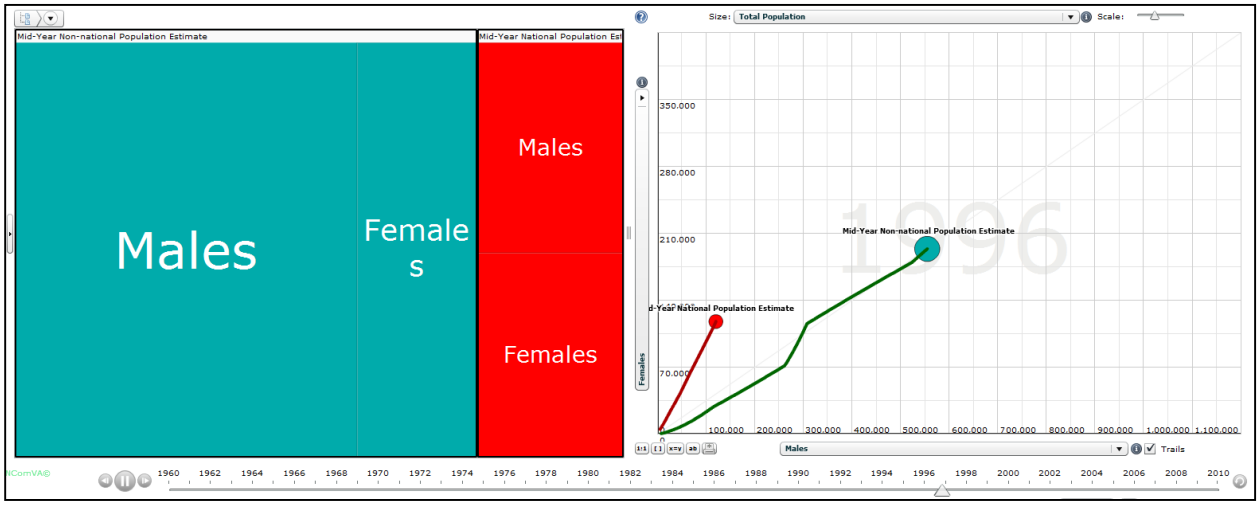
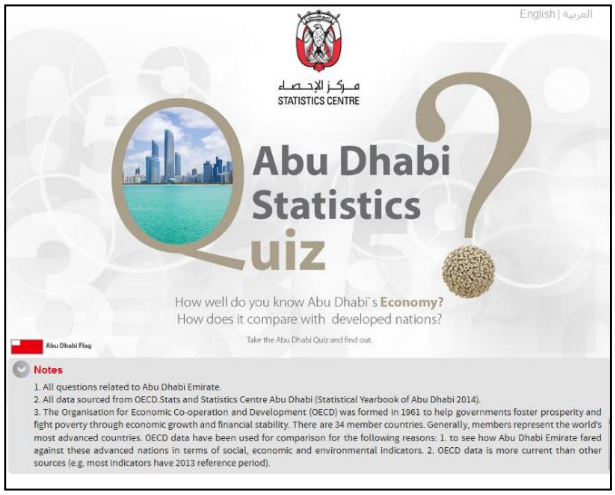


Figure 6: Mobile Apps

Figure 7: Statistical Quizzes



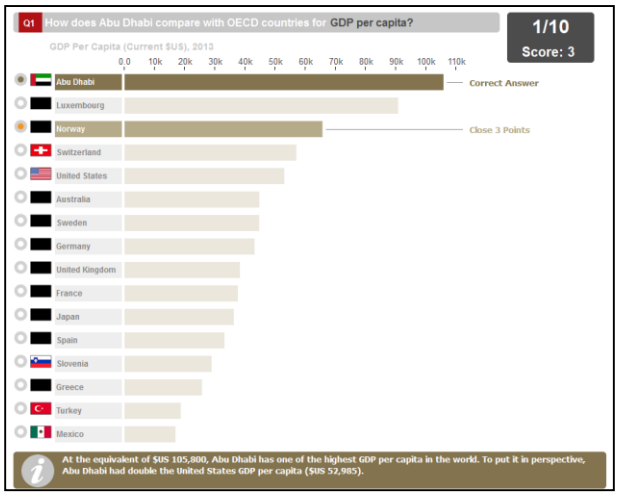
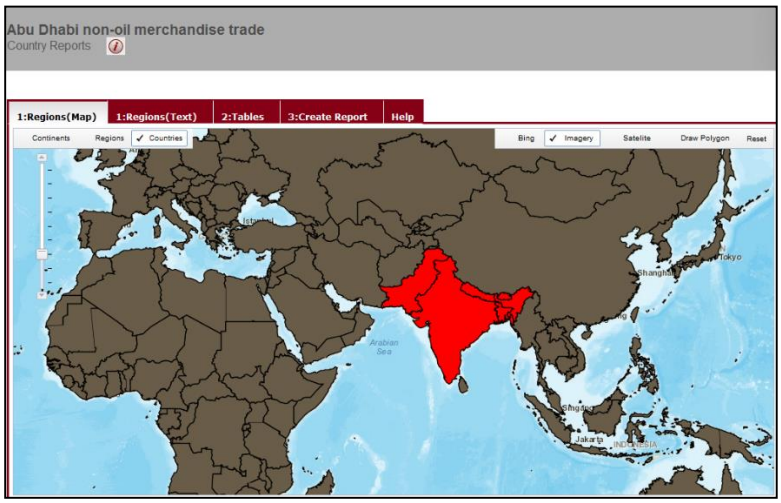


Figure 8: Thematic Maps

Figure 9: Videos (animated)

Figure 10: Automated Country Reports



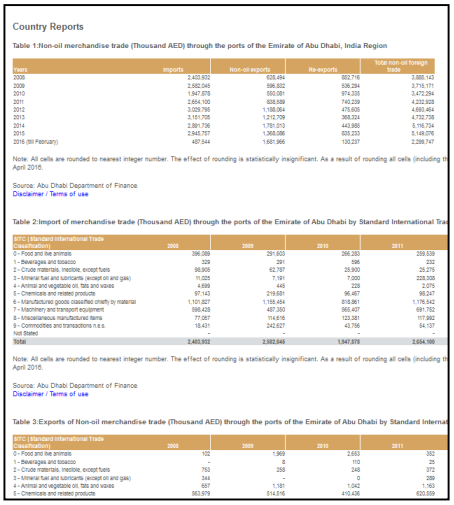


Figure 11: Dashboards

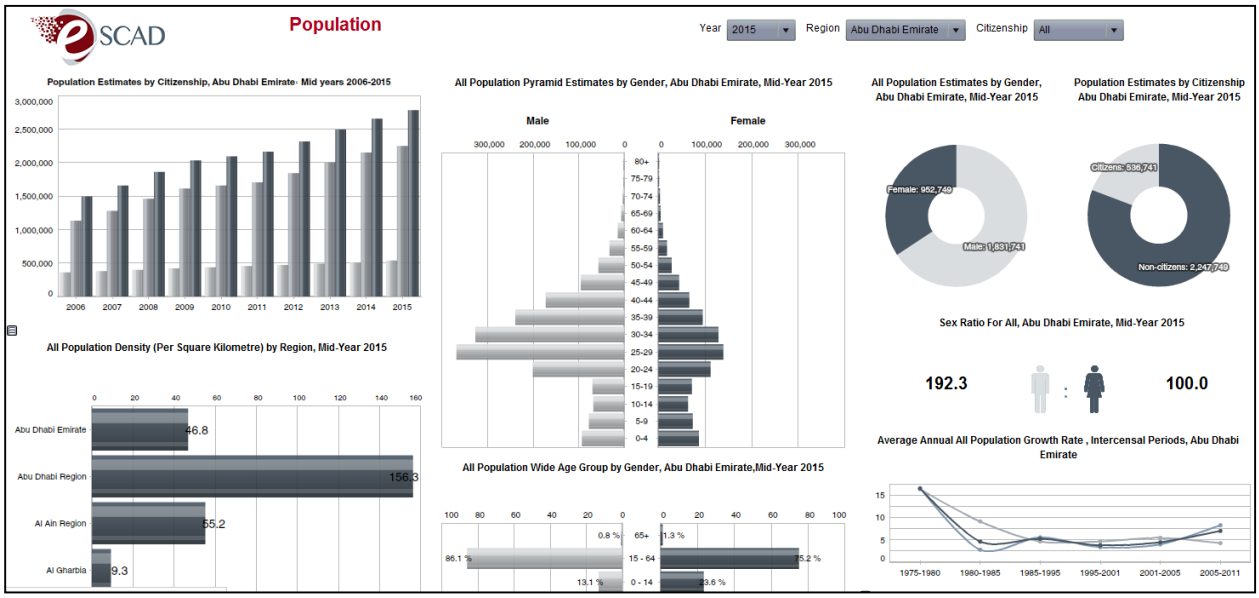
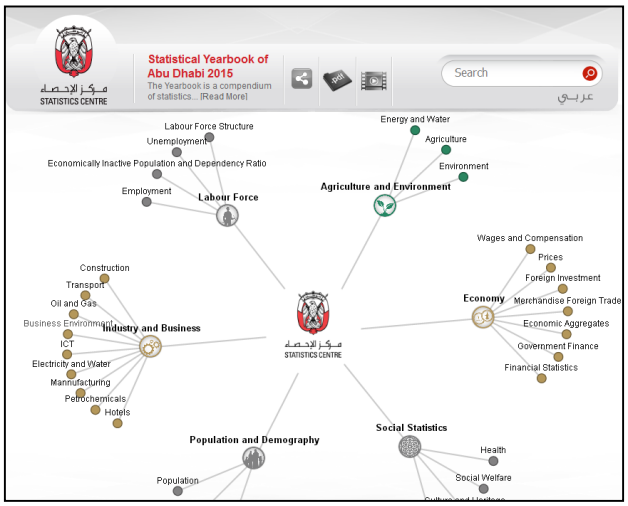


Figure 12: Interactive Yearbook



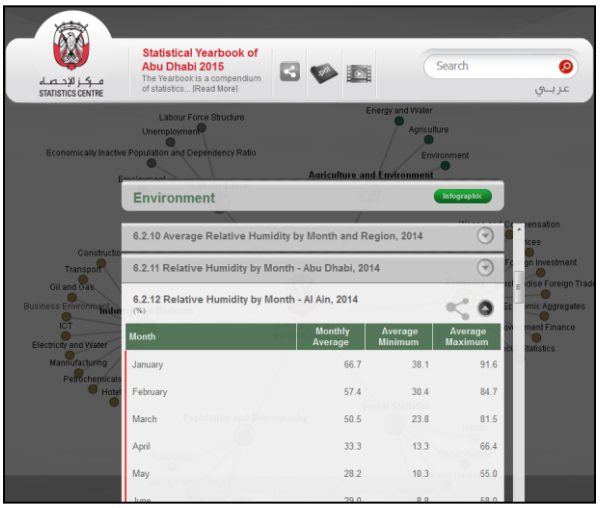
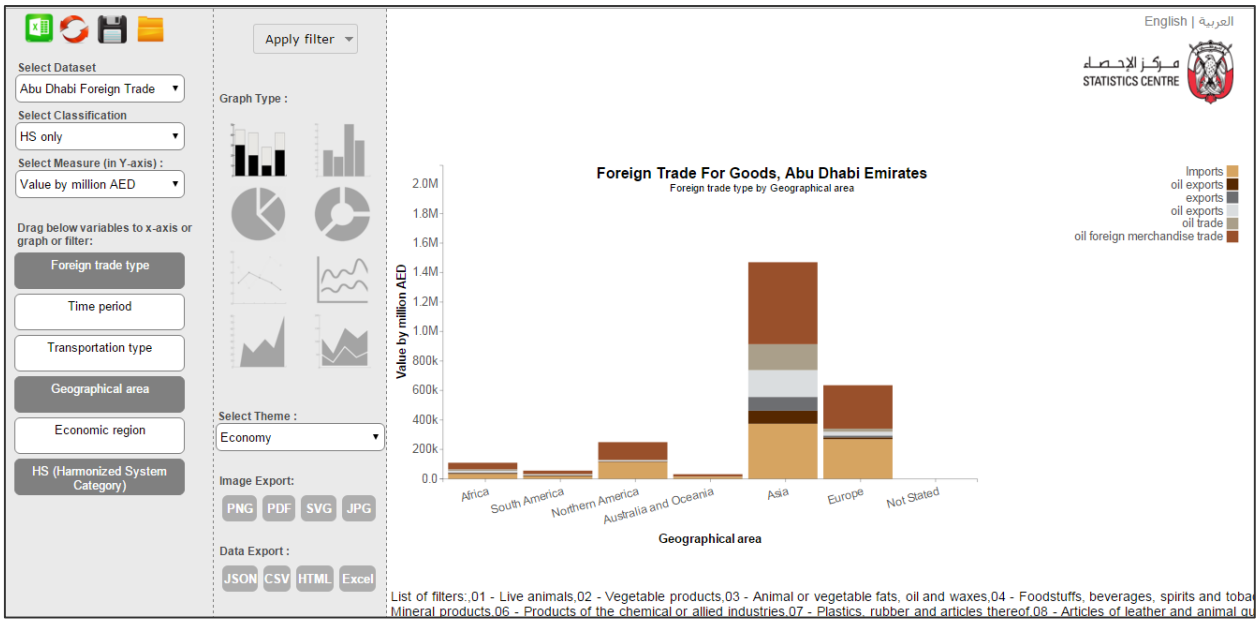



Figure 13: D3 Charter