# Creating Interactive Visualisations Using Tableau Desktop

### *CS5044 – Practical 1*

Jamie Munro, Student Number: 210027910

**Word Count:** 989

## Questions

1. Which regions make the largest contributions to globally famous biographies?
2. Which domains produce the most famous figures and how does this vary by region?
3. How do the contributions of female authors change over time?

## Visualisation

*Table 1 Visual mapping*

| Attribute | View | Attribute Type | Visual Variable | Expressive |
|---|---|---|---|---|
| Birth Country | A | Categorical | Position (map) | Yes |
| Total views (country) | A | Ordinal | Size (radius) | Yes |
| Modal Domain (country) | A | Categorical | Colour | Yes |
| Birth Country | A* | Categorical | Text | No |
| Total Books (country) | A* | Ordinal | Text | No |
| Total views (country) | A* | Ordinal | Text | No |
| Total English views | A* | Ordinal | Text | No |
| Total Non-English views | A* | Ordinal | Text | No |
| Average number of languages | A* | Ordinal | Text | No |
| Modal Domain (country) | A* | Categorical | Text | No |
| Birth City | B | Categorical | Position (map) | Yes |
| Total page views (city) | B | Ordinal | Size (radius) | Yes |
| Modal Domain (city) | B | Categorical | Colour | Yes |
| Birth City | B* | Categorical | Text | No |
| Total Books (city) | B* | Ordinal | Text | No |
| Total views (city) | B* | Ordinal | Text | No |
| Total English views | B* | Ordinal | Text | No |
| Total Non-English views | B* | Ordinal | Text | No |
| Average number of languages | B* | Ordinal | Text | No |
| Modal Domain (country) | B* | Categorical | Text | No |
| Total Views (author) | C | Ordinal | Length (height) | Maybe |
| English Views (author) | C | Ordinal | Length (height) / colour | Maybe |
| Non-English Views (author) | C | Ordinal | Length (height) / colour | Maybe |
| Author Name | C/C* | Ordinal | Text | Maybe |
| Total Views (author) | C* | Ordinal | Text | No |
| English Views (author) | C* | Ordinal | Text | No |
| Non-English Views (author) | C* | Ordinal | Text | No |
| Gender Ratio | D | Ordinal | Angle / colour | Yes |

| Number of authors per gender | D* | Ordinal | Text | No |
|---|---|---|---|---|
| Domain Ratios | E | Ordinal | Angle / colour | Yes |
| Number of authors per domain | E* | Ordinal | Text | No |
| Books per birth year per domain | F | Ordinal | Size (height) / colour | Maybe |
| Books per birth year per domain | F* | Ordinal | Size (height) / colour | No |
| Author name | G | Categorical | Text | No |
| Author field/domain | G | Categorical | Colour | No |
| Author gender | G | Categorical | Symbol | No |
| Author name | G* | Categorical | Text | No |
| Author field/domain | G* | Categorical | Text | No |
| Author gender | G* | Categorical | Text | No |
| Author birth city | G* | Categorical | Text | No |
| Author birth year | G* | Ordinal | Text | No |
| Historical Popularity Index | G* | Ordinal | Text | No |
| Total page views | G* | Ordinal | Text | No |
| English page views | G* | Ordinal | Text | No |
| Non-English page views | G* | Ordinal | Text | No |
| Number of languages | G* | Ordinal | Text | No |

\* Tooltip

Table 1 shows the mapping between attribute and visual. Figure 1 is an annotated screenshot showing the layout of the visualisation. Table 2 provides a description of each of the layout labels.
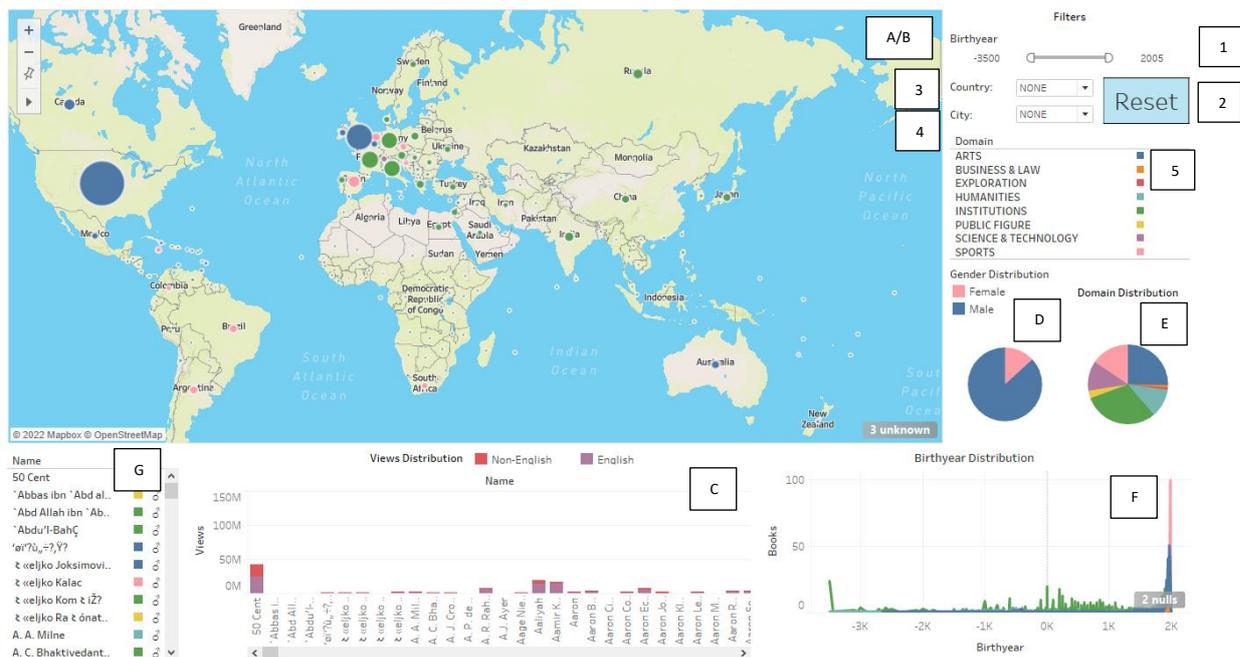


*Figure 1 Annotated Screenshot of visualisation*

*Table 2 Figure 1 label annotations*

| Label | Description |
|---|---|
| A | World map |
| B | Country map (not shown) |
| C | View bar chart |
| D | Gender pie chart |
| E | Domain pie chart |

| F | Birth year line graph |
| G | Author details |
| 1 | Birth year filter |
| 2 | Reset filter button |
| 3 | Country dropdown |
| 4 | City dropdown |
| 5 | Domain legend/filter |

The main view of the visualisation (A) is a map overlaid with a circle for each country containing a book from the dataset. The circle is sized according to the total number of views of books from that country. This is an effective encoding as it clearly shows on the map which countries make the greatest contributions and allows comparison between countries. The circle is coloured according to the modal domain within that country. This provides the viewer with an understanding of which domains are most popular in each country. Further information, as shown in table 1 is displayed in a tooltip when a user hovers over a circle.

If the user clicks on a country, the main view changes (B) to a map centred on the country, showing similar information to before, but with individual cities plotted instead of countries, as shown in figure 2.  Users can also select a country using the dropdown menu (3). Users can select a city by clicking on it, or selecting it from the city dropdown (4).



*Figure 2 Country Map (B)*

The domain legend (5) shows which colours are used to represent each of the domains in the charts and can also be used to filter to one or more domain. The birth year slider (1) can be used to filter by birth year. The remaining charts are filtered according to the filters (including country and city).

The gender (D) and domain pie charts (E) show the distributions of gender and domain across the selected data respectively. These are effective encodings as the user won't need to compare between gender and domain and will be interested in the proportion rather than absolute values.

It's easy to see which category has the most members. The colours used are consistent to avoid confusion. For finer grained detailed, absolute values can be viewed in the tooltips.

The birth year chart (F) is a line graph showing the number of books for each birth year. There is one colour-coded line per domain. This chart can be overwhelming if the number of authors is too high, but becomes effective once filtered down to a smaller number of books. This chart enables the user to view frequency and domain trends over time.

The views bar chart shows the total number of views per selected book, breaking this down into English and non-English views. Again, this is less effective when a large number of authors are selected, due to the limited expressiveness of a bar chart with too many items, however, once a more limited author selection is made, the chart is useful for comparing views across authors, and language distribution of views.

The final view (G) is a list of all the selected books. The list is coloured coded by domain, and also contains gender symbols for quick reference. These encodings are not particularly effective in visual terms; however, they do enable the other visuals to be further filtered, and detailed information is displayed when the entries are hovered over. If an author is clicked, the tooltip contains a hyperlink that can be used to navigates to the relevant Wikipedia page, as shown in figure 3.
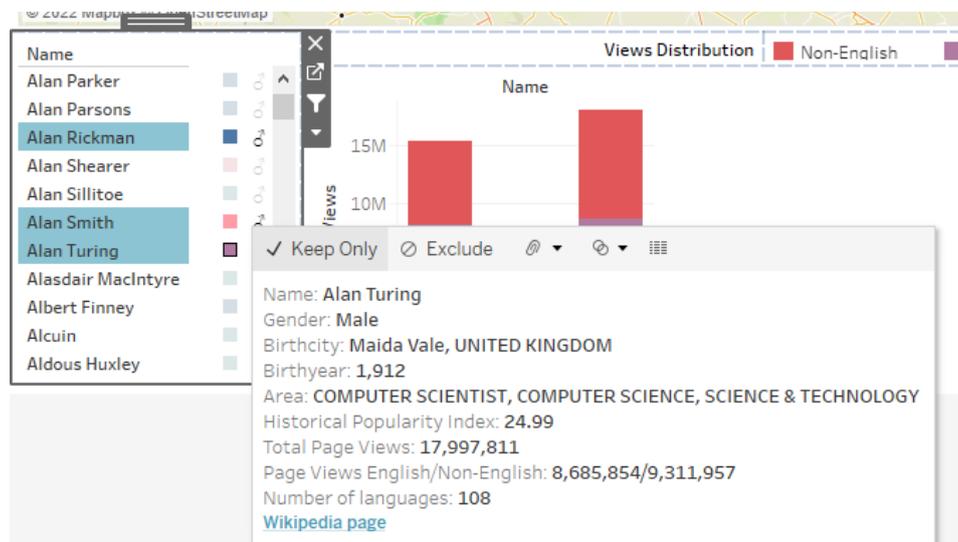


*Figure 3 Author List (G)*

# Insights

## Question 1
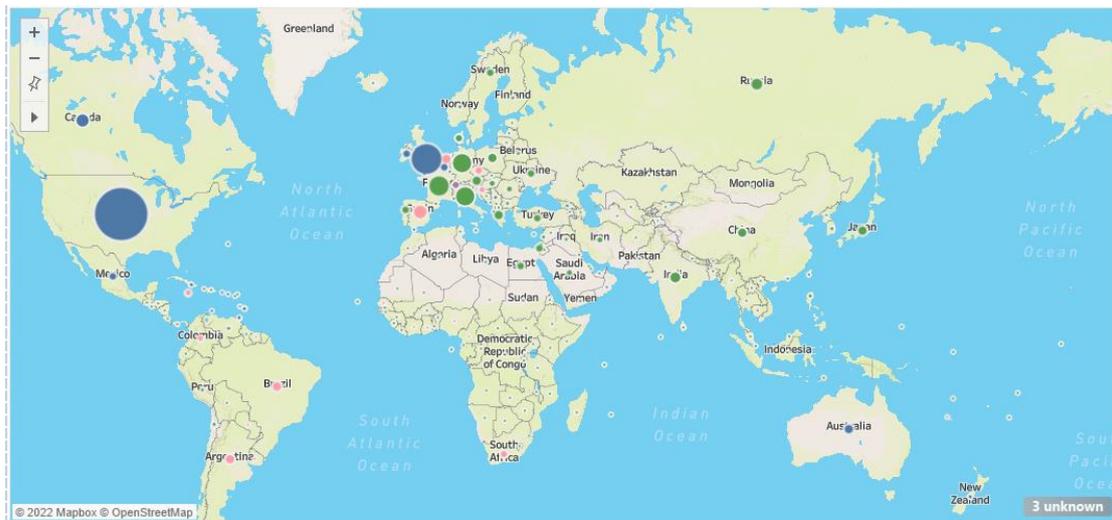


*Figure 4 Global distribution*

We can see visually from figure 4 that the largest contribution is made by the US, followed by Britain, France, Italy, and Germany. The only other nations with more then 200 contributions are Russia and Turkey.



*Figure 5 USA*



*Figure 6 UK*



*Figure 7 France*



*Figure 8 Turkey*



*Figure 9 Germany*



*Figure 10 Italy*
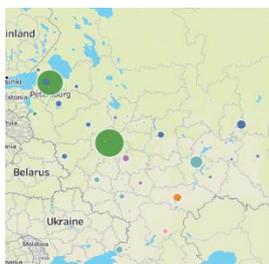


*Figure 11 Russia*

Figures 5 – 11 show the city-level distribution. In nearly every case, the city with the highest contribution is the capital. The only exception to this is the USA. Usually, the larger contributions are limited to one or two cities, however in the case of Italy and Germany, a much wider variety of cities make large contributions. This is likely because these countries were until recently, a collection of smaller countries.

## Question 2



*Figure 12 Global*

*Figure 13 USA*

*Figure 14 UK*

*Figure 15 France*

*Figure 16 Turkey*

*Figure 17 Germany*

*Figure 18 Italy*

*Figure 19 Russia*

The pie charts in figures 12 to 19 compare the global distribution of domains with the top 7 countries. The largest domain globally is institutions (green) with arts closely following this (blue). Art is the most popular in the UK and USA, with institutions most popular in all the other countries. As shown in figure 4, institutions are the most prevalent domain in Europe, Africa, Oceania and Asia; art is most prevalent in North America and sport is most prevalent in South America.

## Question 3



*Figure 20 3500 BC to 2005 AD*

*Figure 21 3500 BC to 0 AD*

*Figure 22 0 AD to 500 AD*

*Figure 23 500 AD to 1000 AD*

*Figure 24 1000 AD to 1500 AD*

*Figure 25 1500 AD to 2000 AD*

*Figure 26 1805 AD to 2005 AD*

*Figure 27 1905 AD to 2005 AD*

*Figure 28 1975 AD to 2005 AD*

Figures 20 to 28 show the distribution of gender in the dataset. Figure 20 shows the complete time period, whilst the others show a narrower slice. We can see that from 3500 BC through to 1000 AD the share of female authors reduces. After 1000 AD the female share gradually increases, eventually reaching around 25% from 1975 to 2005.

## Evaluation

The visualisation provides the user with an ability to visually explore most of the attributes present in the dataset in an open-ended way. It has powerful filtering abilities and many interactive features. The visualisation is most effective in answering question 1 as it is able to present location data visually on a map. The visualisation is less effective in answering questions 2 and 3 as it requires the user to visually compare pie charts. It could be more effective at answering these questions if it enabled the user to plot all the data points simultaneously on a different type of chart. It would also be interesting to compare the English vs non-English views distribution across a broader dataset then the bar chart enables.