

CS5044: Practical 1

Creating Interactive Visualisations Using Tableau Desktop

For this practical you will come up with a concept and design for an information visualisation of a given dataset and implement it in Tableau Desktop. As part of this practical, you will

- consider possible questions that could be explored – through an interactive visualisation – based on the given data,
- based on the above, follow the Five Design Sheet methodology to develop design ideas for an interactive visualisation through the process of sketching,
- design interactive features for the visualisation as part of your ideation process,
- implement the final design concept in Tableau Desktop, and, finally
- reflect on your design choices, in particular, the visual encodings of data attributes, and what types of insights your visualisation might provide.

WEIGHTING AND MARK DESCRIPTORS

This practical is weighted at **30% of the overall module grade**. The general mark descriptors apply and can be reviewed here:

https://info.cs.st-andrews.ac.uk/student-handbook/learning-teaching/feedback.html#Mark_Descriptors

The practical should be submitted to MMS no later than **February 22, 2023; 21:00** (UK-time).

RULES

You are expected to work on the practical on your own and to follow the Good Academic Practice policy: <https://www.st-andrews.ac.uk/students/rules/academicpractice/>. Lateness penalties follow the default School policy (Scheme B, 1 mark per 8-hour period or part thereof). You can review the policy on lateness penalties in the Coursework Penalties policy:

<https://www.st-andrews.ac.uk/media/teaching-and-learning/policies/penalties.pdf>.

The maximum 900 word limit for the report (see Report section below) is a compulsory word limit, and failure to adhere to it will be penalised according to penalty scheme A in the Coursework Penalties policy, i.e., “1 mark for work that is 10% over-length, then a further 1 mark per additional 10% over”.

The University views plagiarism very seriously. Ensure that any images you choose to use as material in your report are correctly attributed to their proper sources. Ensure that all text that you include in your report, that is not your own, is correctly cited as a quotation and carries an accompanying reference (including page numbers) to the original source. If your visualisation sketches are based on existing work, acknowledge their source and authors in the footnotes and/or references of your report. If you are using standard visualisation techniques such as bar charts or line graphs, it is enough to just name them properly in the report.

PRACTICAL DESCRIPTION

In this practical you will create a visualisation sketch and use Tableau Desktop to create an interactive visualisation of a given dataset that is available on Studres

CS5044/Coursework/Practical_1/data. Details on the dataset are provided further below.

The requirements for your visualisation are as follows:

1. PROMOTE INSIGHT THROUGH EXPLORATION

Your visualisation should facilitate the exploration of the data and, in this way, help answering/exploring at least **one open-ended question** that one may ask about the dataset. Your question(s) should target at least four distinct data attributes and be open-ended enough to require an interactive visualisation.

When building your visualisation, you may use any inspiration from the web, including existing visualisations or visualisation techniques or Tableau tutorials. Make sure to reference all external resources that you have used for this practical as part of your report. You should also correctly name any common visualisation techniques you apply.

Be sure to critically consider Tableau's built-in visualisation techniques (e.g., from the Show-me panel) and adjust them carefully for your particular data context and purpose.

2. CONSIDER DATA ATTRIBUTES TO SHOW, DEPENDING ON YOUR QUESTION(S)

Your visualisation should include the data attributes that are of interest given the question(s) you have decided to shed light on. Note that one aspect of the data can be included in several attributes. For example, the dataset may include several attributes describing a person's birthplace, at different levels: by latitude and longitude, by country, by continent, etc... It is up to you to decide what attributes to show in your visualisation and how. Depending on your question, it may make sense to show higher-level attributes (e.g., country or continent) or more fine-grained details (e.g., individual location data based on latitude and longitude).

3. PROVIDE MULTIPLE PERSPECTIVES ON THE DATA

It is difficult, if not impossible to show all aspects of a complex dataset in one single view. You will have to make decisions on the types of visual representations and "views" required to highlight particular data aspects in a meaningful way. As a rule of thumb, your visualisation should include **at least two different views**, that is, two different visual representations (e.g., a map and a timeline). Different views should be assembled into a **Tableau Dashboard** and interactively linked where appropriate.

4. INCLUDE INTERACTIVE ELEMENTS

Your visualisation should include basic interactive elements beyond the tooltips and legends that Tableau provides by default. Interactive elements may include, for example, filters or sliders to select certain attribute levels or interactive links between the different views to enable a coordinated exploration of the data from multiple perspectives.

REPORT (should not exceed 900 words)

Your report should start with the question(s) you have chosen to focus on, followed by three sections:

i) DESCRIPTION OF THE VISUALISATION

Describe your visualisation and all views it includes.

- Describe the visual mapping, that is, which visual variables you have used to encode the different data attributes. Provide a table¹ that lists all attributes represented in your visualisation, their attribute type, and their representation via visual variables (see example

¹ Please note that the table does not contribute towards the word count.

in Table 1). Your table should also indicate if your visual encodings are expressive and effective or not – a “maybe” may apply in some cases (see Munzner’s discussion of expressiveness [1]). **Clearly justify your choice of visual encoding** based on the expressiveness and effectiveness principles and other alternatives you have considered.

- Describe how one can interact with the visualisation. How do the interactive elements of your visualisation work and what effect do they have on the views themselves?
- Illustrate your descriptions with screenshots of the visualisation and/or your sketches where necessary. Make sure to add labels and descriptive captions to these figures and reference them in the text.

Attribute	Attribute Type	Visual Variable	Expressive (yes/no)
shirt size	ordinal	position	yes

Table 1

ii) INSIGHTS FROM THE VISUALISATION

Describe the insights that people can gather from your visualisation. Again, your visualisation should address at least one question that you clearly state at the beginning of your report. Make sure to not only discuss the insights that can be gathered, but also provide visual evidence of *how* one can arrive at these insights using the visualisation.

iii) CRITICAL DISCUSSION

Critically reflect on the strengths and limitations of your visualisation. Are there other questions that the visualisation (or your general work with the data) has brought up that you would like to explore in the future?

You are encouraged to include visualisation screenshots, sketches or other visual material to illustrate your argument.

Your report should not exceed 900 words, excluding references, tables and figure/table captions. Please state the word count at the start of the report.

VISUALISATION PROCESS

The dataset is large and complex, and, most likely, unfamiliar to you. Your first step should therefore be to familiarize yourself with the data [2]. This process can be guided by the following set of questions:

- What attributes are included in the data? How do these attributes relate to each other? Do some of them deal with the same aspect of the data?
- What are the types of attributes included in the dataset?
- What questions/ideas for stories does the data trigger in you? Write these ideas down.
- Does it make sense to sort or group certain data attributes? Could additional data be derived from existing data attributes to help tell an interesting story?

Follow the Five Design Sheet Method [3] to structure and document the sketch-based ideation and design process. To start exploring visual encodings of the data, pick a small sample of the data (<20 data points) and start your ideation process (Sheet 1). Try different types of visual representations dealing with the same idea/question(s)/data attributes. Use Sheets 2, 3 and 4 to flesh out three fundamentally different visualisation concepts covering the same ideas/question(s)/data attributes. Consider their advantages and limitations to cover insights about your selected question(s) before you decide on your final visualisation design. Use Sheet 5 to provide a detailed sketch of the final design.

Your visualisation sketches across all five sheets should be created manually (hand-drawn) using coloured pencils and paper. Alternatively, you may create digital sketches using a tablet. However, do not use a digital tool that constraints your visualisation design to particular visualisation techniques.

Digitize the five sheets (by scanning them or by taking a picture), and submit them as **a single multipage pdf file**. You can use PowerPoint to assemble them in different slides and save these slides as a pdf.

As discussed in the lectures and tutorials:

- It is important to critically think about your choice of visual encoding – no visual representation is perfect. You will probably have to make compromises. The Five Design Sheet Method will allow you to create many sketches to begin with without judging your ideas. Then you can start to narrow down your ideas and try to select the ones that show the most promise.
- Sketches do not have to look pretty, but they should be readable. They don't have to be decipherable to others. They are just for you to familiarize yourself with the data, to find interesting aspects in the dataset, and to help you to develop ideas on how to visualise them.

THE DATA

PANTHEON: A DATASET OF GLOBALLY FAMOUS BIOGRAPHIES [4]

Pantheon is a manually verified dataset of famous people that was created by Yu et al. [4]. It includes 11,341 biographies that have presence in more than 25 languages in Wikipedia by May 2013. The data, along with a more detailed description of the data, can be found on StudRes. Further information can be found in the paper by Yu et al [4].

DELIVERABLES

The deliverables of this practical include:

- Your visualisation as a Tableau Workbook file (.twbx). **Make sure to use the .twbx format as it groups your visualisation worksheets together with the underlying data.** This is particularly important, if you made any changes to the original data such as deriving additional attributes or combining it with additional external data sources.
- Your report as a pdf.
- Your 5 sheets of sketches as a multi-page pdf.

Your submission should be a zip file that includes these deliverables. Only zip files will be accepted.

POSSIBLE EXTENSIONS

- Linking additional external data resources to show additional interesting aspects of the provided dataset.
- Provide a video of your visualisation in-use that clearly explains how the visualisation works. Videos should not exceed 2 min. in length.
- ...

ASSESSMENT

The assessment of your practical will focus on the quality of your visualisation design and your report.

- Your visualisation should show multiple aspects of the data in an expressive and effective way and enable interactive exploration.
- Your report should include a concise yet clear description of your visualisation and its features as well as a good justification of your design decisions. Furthermore, it should include a critical discussion of the insights that your visualisation can provide as well as its limitations.
- You have to provide evidence of an ideation and design process following the Five Design Sheet Method, in the form of clear and readable paper and/or digital sketches.

An **adequate practical** will include a basic interactive visualisation in Tableau that includes at least two linked views that complement each other, some evidence of an ideation process, and a concise report that describes the visualisation, provides design justifications and insights and some critical reflection.

A **good practical** will include a well-designed visualisation consisting of at least two interactive views that are carefully designed and adequately customised (e.g., regarding titles describing the content of views, labels, tooltips, and filters, etc...). It will include a well-written and structured, concise report that describes the visualisation and its interactive features, justifies of design decisions, and insights that can be gathered (see above). The illustration of the ideation process will show that multiple design ideas have been critically considered.

An **excellent practical** will include all of the above plus one or multiple high-quality extensions that expand the basic solution by providing further insight and/or show evidence of further reading of InfoVis literature.

RESOURCES

- [1] Tamara Munzner. Visualisation Design and Analysis. 2015; Chapter 5: Marks and Channels.
- [2] M. Stefaner on the visualisation process. Useful techniques on how to approach visualizing unfamiliar datasets. <https://vimeo.com/28443920>.
- [3] Five Design Sheet Method. <http://fds.design/>
- [4] Yu, A. Z., et al. (2016). Pantheon 1.0, a manually verified dataset of globally famous biographies. *Scientific Data* 2:150075. doi: 10.1038/sdata.2015.75