Chapter 1

Information visualization and interaction

The focus in this assignment lies on *visual encoding* for information visualization, *interaction*, *multiple views*, and the *exploratory analysis process*. You should also bring in elements from the lectures on perception and design principles to evaluate your work.

The assignment comes in two variants, one that involves programming (corresponds to the type of work a visualization expert would do in practice) and one that does not (corresponds to the type of work a data analyst would do in practice). Choose one of the variants. For variant 1, expect to spend most time on building an interactive prototype (the dataset may be relatively simple). For variant 2, expect to spend most time on data acquisition, augmenting datasets with additional data from other resources, cleaning data, and aggregating data (the dataset needs to be sufficiently complex).

Note that, in addition to a report, you have to make a **screencast** and include a **breakdown of how the work was split among the group members**. See also the assignment descriptions below.

1.1 Assignment variant 1

- Choose data as described in Section 1.3.1.
- Derive and document aspects of the data that could be of interest to an analyst. Formulate a set of tasks that an analyst might want to perform with the data, and some specific questions. Make sure that at least some of the tasks and questions *require* interaction and/or multiple linked views in order to be performed or solved.
- Consider various interactive visualization techniques and combinations thereof (as linked views) that support these tasks, and that are suitable to analyze this data. Discuss pros and cons.
- Build an interactive application in any language that you prefer to visualize the data with the techniques chosen (using D3 is a viable option, see Section 1.3.3). You should have at least three different techniques in your approach that are linked interactively and are shown simultaneously.
- Use the application to make interesting observations about the data. Document how you came to these observations and how the application design was beneficial (or not) to your discoveries. Emphasize on interaction aspects, and also explain why you think your approach is good from a perceptual point of view and how it takes design principles

into account. Provide some support by citing relevant literature. Since describing an interactive process is not so easy on paper, you should also make a screencast of 3–5 minutes (not longer!) that shows how interaction helps to do some of the tasks. In the report, you should include a breakdown of how the work was split among the group members.

1.2 Assignment variant 2

- Choose data as described in Section 1.3.1. Look for additional sources of data/information to enrich the dataset (use at least one, more is better). The dataset you choose should be suffiently complex to allow meaningful aggregation/clustering. Preprocessing is unavoidable, so be prepared to do data cleansing, curation, filtering, etc.
- Derive and document aspects of the data that could be of interest to an analyst. Formulate a set of tasks that an analyst might want to perform with the data, and some specific questions. Make sure that at least some of the tasks and questions *require* interaction and/or multiple linked views in order to be performed or solved. Apply also at least one form of data aggregation (use whatever tool you have available).
- Consider various interactive visualization techniques and combinations thereof (as linked views) that support these tasks, and that are suitable to analyze this data. Discuss pros and cons.
- Choose an existing information visualization tool, for example the free GGobi (http://www.ggobi.org/), a scientific prototype like iVisDesigner (https://donghaoren.org/ivisdesigner/) or request a demo license from a commercial package such as Tableau (http://www.tableausoftware.com). Your choice should be based on the considerations that you made in the previous step.
- Use the application to make interesting observations about the data. Document how you came to these observations and how the application design was beneficial (or not) to your discoveries. Emphasize on interaction aspects, and also explain why you think your approach (making use of facilities offered by the chosen tool) is good from a perceptual point of view and how it takes design principles into account; or write a critique about the lack of these aspects in the tool. Provide some support by citing relevant literature. Since describing an interactive process is not so easy on paper, you should also make a screencast of 3–5 minutes (not longer!) that shows how interaction helps to do some of the tasks. In the report, you should include a breakdown of how the work was split among the group members.

1.3 Resources

1.3.1 Datasets

You can use a dataset that you may already have from different courses or have obtained otherwise (it should not be the one from assignment 1). Make sure that there is sufficient complexity in the dataset, so there is actually something to explore.

There are also many online data resources that provide open data (note that some are in Dutch only):

- http://opendatanederland.org

- https://data.overheid.nl
- 30 places to find Open Data on the web, http://blog.visual.ly/data-sources/
- http://datacatalog.worldbank.org

BMT students may be more interested in working with medical data. There are not so many rich resources as for other types of data, but you may have access to some in your own department. Please contact Michel Westenberg if you want to do a medical visualization project.

1.3.2 Screencast tools

There are a number of tools available to capture either your whole screen or a part of it.

- Linux: Simple Screen Recorder, http://www.maartenbaert.be/simplescreenrecorder/. It does unfortunately not provide video editing facilities.
- Windows/OS X: Camtasia, http://www.techsmith.com/camtasia.html. A fully functional trial version can be used for 30 days.

1.3.3 D3

The JavaScript library *D3.js* (http://d3js.org) provides a wealth of visualization techniques that can be run directly in your web browser. The website has many examples and code snippets, all of which you are free to use in this assignment. There are also many other good resources on D3, so it should not be hard to get started.

Useful resources:

- The D3 website itself has lots and lots of resources: http://d3js.org.
- A huge number of examples is available here too: http://christopheviau.com/d3list/.
- Scott Murray's book on Interactive Data Visualization for the Web, can be read online for free at http://chimera.labs.oreilly.com/books/1230000000345/index.html. The book also has an introduction in JavaScript, CSS, and HTML.
- For a quick intro, browse the D3 workshops slides: http://bost.ocks.org/mike/d3/workshop/.