University of Melbourne
Melbourne School of Engineering
GEOM90007 Spatial Visualization
Assignment 1: R graphics

Overview

Your task is to develop one, visually appealing and communicative, data graphic using R based on data from the Guardian DataBlog, http://www.guardian.co.uk/news/datablog.

Assessment

This exercise is to be completed individually. The assessment is worth 10% of your final subject mark. You must submit the following through the course web site:

1. a pdf version of your data graphic (one page);
2. a zipped file containing any data sets you used. The total size of all zipped data sets must be less that 5MB; and
3. a commented and working R script that generates the data graphic, including clear acknowledgment of any code used or adapted from other sources.

The submission deadline will be clearly stated on the course web site. **No late submissions will be accepted**.

Assessment Criteria

The key assessment criteria are (out of 100%):

- **Submission** (10%):
  - The submission is in the correct formats,
  - the R code works when run on the marker's machine, and
  - contains all the necessary data and files, and conforms with all the submission criteria above
- **Basic design** (40%):
  - The data graphic conforms to basic visual design principles covered in lectures/practicals, and
  - is well-presented with evidence of care and attention to detail.
- **Technical challenge** (25%): Thinking beyond the tutorial, i.e.,
  - the R code relies on new or different programming techniques not encountered in class;
  - has particularly compact, well-structured, or well-presented code;
  - uses data that required new techniques to use or additional effort to resource; and/or
  - demonstrates clear evidence of advanced and independent work.
- **Design innovation** (25%):
  - The data graphic involves design elements that are innovative (i.e., beyond standard elements available in a spreadsheet, innovative design not already on the Guardian datablog);
  - reveals interesting or meaningful patterns;
  - is notably aesthetically pleasing or striking;
  - demonstrates independent background research into the research literature;
  - and/or existing demonstrates clear evidence of original thinking and advanced understanding of graphical design principles.

Generalized marking criteria:

- <50%: *Inadequate work* that in one or more respects fails to meet basic technical standards or apply basic design principles.
- 50-60%: *Satisfactory work* that is a correctly submitted basic data graphic of simple data set.
- 60-70%: *Good work* that involves marginal additional challenge or marginal design innovation, and moderate levels of design quality.
- 70-80%: *Excellent work* that involves clear additional technical challenge and additional design innovation and high levels of design quality.

- >80%: *Outstanding work* that demonstrates substantial additional technical challenge, substantial design innovation, flawless design, and involves work that clearly goes beyond that normally expected in class.

Hints

- You are free to design any type of data graphic (you do not need to design a graphic or map that contains spatial data, although you are most welcome to do so if you wish).
- You should aim to design your own data graphic, not simply duplicate a data graphic that is already on the Guardian DataBlog (copying will be penalized under all categories and is a form of plagiarism.)
- You are encouraged to research widely, to find interesting and engaging ways of constructing your data graphic (this might be where most of your time is spent).
- Your R code should contain the commands necessary to install and use any packages necessary for your R code to run. You will be penalized if the marker has to manually install any necessary packages, and heavily penalized if the marker is unable to get your code working with reasonable effort.
- Spelling and grammar are part of this assessment. Your code commenting should exhibit attention to detail, and should be free of errors.

Data

Datasets can be found in articles linking to the Datablog or on the main Guardian website. The focus in Assignment 1 is to create a data graphic, NOT analyze a big dataset - that's why we're using pre-packaged data which has already been formatted ready for visualization. Example datasets from The Guardian include (but are not limited to):

- **University rankings** (UK, 2016): https://docs.google.com/ spreadsheets/d/1qy- n4Hftekk3weuwRwBe8zcLgpBOPzABD ugrjiXIysQ/edit?usp=sharing
- **Deaths from police** (USA, 2015): http://www.theguardian.com/us- news/ng-interactive/2015/jun/ 01/about-the-counted
- **Heights of world leaders**: https://docs.google.com/ spreadsheets/d/1U-MIoTV- PrKoJCc4QzV4bvWQ6qB8ZH2jaSKTz3 3b0wQ/edit?hl=en_GB#gid=0
- ...a huge data dump (UK centric, 2011): https://www.theguardian.com/ news/datablog/2011/jan/27/ data-store-office-for- national-statistics

Plagiarism

In short: you must clearly acknowledge any material you have used in your assessment.

Plagiarism is copying and use of another's work without proper acknowledgment (can be both known and unknown). The university has a clear policy prohibiting any form of plagiarism. Further information can be found at: http://www.services.unimelb.edu.au/plagiarism/.

Note that it is acceptable to reuse ideas and code you have found on the web *as long as the source is clearly acknowledged, and that use is permitted by any license restrictions*. If properly acknowledged, using other people's code and ideas can count as independent background research (see grade related criteria above). If not properly acknowledged, using other people's code and ideas is plagiarism and will result in a mark of zero for this assessment. In serious cases plagiarism may also result in failure of the entire subject and further University disciplinary action.