

Group Project 2014

Group Project

- To visualize and analyze a data set of your own choice
- Work in groups of 2 or 3
- Graded individually
- Choose your own data set
- Visualize and analyze

Proposal

- Stage 1 (Due: Oct 3)
- **Send your group list to the TAs by Sep 24**
- Project Proposal (20%) – max 2 pages
- Project objectives
- What is to be visualized?
- The target visualization
- The question you would ask
- Division of labour
- Choice of tools

Final Report

- Stage 2 (Due: Nov 7)
- Final report (60%) – max. 20 pages
- Difficulties that you have encountered
- Different methods that you have tried
- Justification of your choice
- Are your questions answered by your visualization?
- Any new insights?
- Anything you wanted to do but haven't?

Presentation

- State 3 (Nov 12, 19 & 26 during the lecture)
- Project Presentation (about 2-3 mins for each student)
- Q&A for each group presentation
- Each project is assessed by the Instructor, the TAs, as well as your peers
- Details to be given later

Assessment Criteria

- The assessment will be based on the followings:
 - * how you process and compile your data for the purpose of a meaningful visualization
 - * how you apply proper visualization design rationales and techniques
 - * how you make use of your visualization as a tool to explore your data set and obtain new insights from it
- Hence, you should provide sufficient explanation and justification of your decisions, and demonstrate the efficacy of your visualization with appropriate examples.

Reference Data Sets

- Data Sets from Interactive Data Visualization Book

<http://www.idvbook.com/teaching-aid/data-sets/>

- Census and Statistics Department, HKSAR

<http://www.censtatd.gov.hk/home>

- Hong Kong Environmental Statistics

[http://www.epd.gov.hk/epd/english/resources pub/
envir info/envir info.html](http://www.epd.gov.hk/epd/english/resources_pub/envir_info/envir_info.html)

Reference Data Sets

- The movie database IMDb
<http://www.imdb.com/interfaces#plain>
- Network data (**SNAP**)
<http://snap.stanford.edu/data/>
- The World DataBank
<http://databank.worldbank.org/data/home.aspx>
- Wikipedia dumps
http://en.wikipedia.org/wiki/Wikipedia:Database_download

Reference Data Sets

- Google's Ngram Viewer
words and phrases (i.e., ngrams) and their
usage frequency over time of over 8 million
books
[http://storage.googleapis.com/books/ngrams/
books/datasetv2.html](http://storage.googleapis.com/books/ngrams/books/datasetv2.html)
- The Datawrangling blog: A data set collection
[http://www.datawrangling.com/some-datasets-
available-on-the-web](http://www.datawrangling.com/some-datasets-available-on-the-web)

Some Sample Visualizations

- Gap Minder World
<http://www.gapminder.org/world>
- How Different Groups Spend Their Day
<http://www.nytimes.com/interactive/2009/07/31/business/20080801-metrics-graphic.html>
- Examples in Many Eyes, Tableau