Quantitative exercise

Preparation: Read through the instructions for this exercise. Load R or another tool that you have chosen for your data analysis into your computer and import (or insert) some of your data (or generate some data similar to what you expect to collect).

Tasks

Step 1: Each group should:

1. Do the pre-processing necessary to load the data into the analysis tool that is to be used for their project. While using R is recommended, the actual choice of tool is up to each of the groups. (Note that this pre-processing should be done *before* the start of the lab.)
2. Load your data into the analysis tool.
3. Identify which are your independent variables and which are your dependent variables (in the data which you have collected or generated). Write a description about the *expected* properties of each of these variables.
4. Perform exploratory data analysis. Does your data have the expected properties? If not, can you identify why it does not?
5. Identify what statistical tools you will apply to analyze your data. If you have a model for the relationship between the independent and the dependent variables fit your data to this model. If you do not have a model, what could you to do identify a model, for example using principle component analysis.
6. Generate some visual aids (such as tables or graphs) to present your data to others.

Step 2: Each group should exchange the output of the above with another group, then comment on the other group's output. The goal is to help the other group improve their analysis and presentation.

Note that an important goal of this exercise is to help groups determine what data they need to collect and how much and what type of data needs to be collected.

Step 3: Submit a summary of your planned data analysis and perhaps some illustrative images (plots, graphs, images) as appropriate.

Hints

Manuals about R can be found at:

http://cran.r-project.org/doc/manuals/R-intro.html

Sample Data

Sample data files - with .PCAP files from Wireshark*
(https://kth.instructure.com/courses/189/pages/sample-data-files-with-pcap-files-from-wireshark-2?module_item_id=2142)
The wireshark page: http://wiki.wireshark.org/SampleCaptures

R Packages

gplots Various R programming tools for plotting data
princomp {stats} Principal components analysis on a matrix

PCA analysis:

- Thiago G. Martins, 'Computing and visualizing PCA in R', R-bloggers. 28-Nov-2013 [Online]. Available: http://www.r-bloggers.com/computing-and-visualizing-pca-in-r/. [Accessed: 13-Sep-2015]
- Robert I. Kabacoff, 'Principal Components and Factor Analysis', Quick-R: Factor Analysis, 03-Oct-2014. [Online]. Available: http://www.statmethods.net/advstats/factor.html. [Accessed: 13-Sep-2015]