Categorical Data Analysis
STAT 653, Fall 2017

Data Analysis Project

The data analysis project is to be completed by all students in STAT653. The project will result in a **project report** and **presentation** given at the end of semester (see the timeline below).

Project Goals:

- Learn theoretical methods and develop practical skills of statistical data analysis;
- Practice in formulating realistic research questions for a given data set;
- Get experience with exploring, analyzing, and interpreting real-life data sets;
- Learn how to prepare a scientific data analysis report;
- Practice oral presentation skills.

Methods to use:

The project should combine theoretical and numerical methods studied in the class, which can be divided into the three broad categories:

- Exploratory data analysis and descriptive statistics
- Contingency tables
- Generalized Linear Models

A project would necessarily include exploratory data analysis and data description, as well as some methods from at least one other category. It might happen that you need (want) to apply methods outside of the scope of this class; this is highly encouraged and should be discussed with instructor.

Project report:

A project report should be written according to scientific standards; the project format and requirements will be discussed in class.

Project presentation:

The project results will be presented before the class in December.
## Data Analysis Project Timeline

<table>
<thead>
<tr>
<th>To do item</th>
<th>Pts</th>
<th>Deadline</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Think what data sets to use. In this class we mainly study methods to detect association between random variables. Hence, stick to  
• Several sets that might be related (whether they are actually related might be the question of your research)  
• Categorical data  
• Large number of observations (>1000)  
• Something you are interested in! |     |          |          |
| Meet w/instructor to discuss the data, possible research questions, and the methods to apply. | 10  | Sep 21   |          |
| Finalize the data set for analysis, Formulate research questions (you might expand/refine/update them later), Start working on data analysis. |     | Oct 5    |          |
| Data analysis is mostly completed, You know how to interpret the data and answer all the research questions, You might still need to do some analysis to improve results/figures, etc., Meet w/instructor to discuss the progress | 5   | Nov 16   |          |
| Project report draft submitted                                           | 5   | Nov 30   |          |
| Final project report submitted, Presentation ready                        | 10  | Dec 7    |          |
| Project presentation                                                     | 30  | Dec 12   |          |
| Project report evaluated                                                 | 40  |          |          |
| **TOTAL points**                                                         | 100 |          |          |