

# BIOMEDICAL DATA SCIENCE HOMEWORK #1

## DUE DATE 10/2/2018

- Please turn your homework code through a GitHub Repository (see below).
- No late homework will be accepted (you will not be able to submit the repository link to Canvas after the due date). Do not change the repository after the due date; your homework will not be graded.
- 20 points total

Use the accompanied dataset for this homework. Read the dataset description carefully [here](#) and make sure you understand the dataset features and values.

- a) Load the dataset in an iPython notebook (3 points).
- b) Explore the dataset using a plot to show the relationship between each one of the categorical features and “comfort”. Interpret the results in a markdown cell (7 points).
- c) Perform any necessary preprocessing steps. Explain why each step is needed. Use the markdown cells of your notebook for explanations (8 points).
- d) Create a GitHub repository and share your code via GitHub with the instructor by submitting the link on Canvas (2 points).