Course Name	**: Introduction to Data Science**
Exam Type	**:** Midterm I
Submission Deadline	**:** 05.04.2016, 13:00
Files to be submitted	**:** nameSurname.doc, nameSurname.ipynb and your data files

Student Name Surname :   Score:

Student No   :

DATA MANAGEMENT & EXPLORATORY DATA ANALYSIS MIDTERM I PROJECT

This study acquaints you with the data and encourages you to develop a research question. You are free to choose the data to work with; however you are required to clearly describe your data and the association that you would like the study. You are also required to support your hypothesis by making statistical analysis on the dataset of your choice. Below are the steps to fulfill this:

Step1 Defining the dataset:

- Choose the dataset that you would like to work with and then give the details about it.
- What is it about, how many variables are there in your dataset, what are the types of each of those variables, are there any missing values, etc.

Step2 Managing the data: (include both detailed explanation and your python codes)

- Does your dataset comes from separate files? If so, how did you merge them?
- How did you handle missing values?
- Which variables (and which values) were included in your analysis, and how did you choose them?
- Did you have to create groups of some quantitative variables?
- Did you recode some variables?

Step3 Providing basic descriptive statistics: (include both detailed explanation and your python codes)

- Frequency distributions of counts and percentages (for numerical / categorical variables)
- Draw univariate bar graph of each variable
- Draw a bivariate plot that shows the relationship between two variables (for numerical variables)
- Draw a bivariate plot that shows the relationship between two variables (for categorical variables)
- Do the graphical results support your hypothesis?