Seminar no. 1

Problem 1

In the data set *Movies.RData*, there are 1322 respondent answers to the question: „How do you assess the impact of nowadays movies on the youth?“ The answers were on the 5 value scale: The impact I see as:
Very positive(1) - Positive(2) - Neutral(3) - Negative(4) - Very negative(5).
Using suitable descriptive statistics characterize the data and the data categorized with regard to gender.
(General Social Survey, USA, 1990)

Problem 2

In the data set *Household_marriage.RData*, there are 1346 respondent answers to two questions:
1. answer: „How is it important that the youth do not live with their parents together at home?“
2. answer: „How is it important to get married?“
The answers were on the 5 value scale: Very important(1) - Quite important(2) - To some extent important(3) - Not very important(4) - Absolutely not important(5).
Using suitable descriptive statistics characterize both variables and compare the distribution of the two variables.
(General Social Survey, USA, 2002)

Problem 3

Decide whether the opinion about the impact of the movies on the youth is different as far as gender is concerned in the problem 1.1. Compare and interpret p values of the all possible tests. In terms of two sample t test, add confidence interval of the difference between the populations' means.

Problem 4

Imagine the following question regarding the problem 1.2 (X variable). Let classify the respondents into groups of education (nominal variable Y) of following values: lower than high school - high school or junior college - bachelor - graduate.
The data are in the *Household_education.RData* set, use the data in the *fdegree4* column for the Y categorical variable.
Detect if the distribution of the opinions about own household is different regarding particular groups of education. Compare and interpret p values of all possible tests. Concerning ANOVA keep figuring out with the Tukey multi-comparison method (the Scheffé test is also possible).