

## Homework #8

Please submit the solution in the form of R Markdown report, knitted into either of the available formats (HTML, pdf or Word). Provide all relevant code and output. Goal of this homework is to have you practice modeling categorical predictors (dummy variables); individual t-test and incremental F-test for significance of a categorical predictor; modeling interaction between predictors in explaining a response; your R coding.

### Problem #1

For *Credit* data set, proceed to

1. Figure out if there's statistical evidence of interaction between *Income* and *Rating*. Make sure to write down the full modeling equation along with hypotheses to be tested (all in parametric notation), make a conclusion. **If there's statistical evidence of interaction (otherwise - don't do that)**, interpret the effects of 1) *Income* on *Balance*; and 2) *Rating* on *Balance*
2. Figure out if there's statistical evidence of interaction between *Income* and *Age*. Make sure to write down the full modeling equation along with hypotheses to be tested (all in parametric notation), make a conclusion. **If there's statistical evidence of interaction (otherwise - don't do that)**, interpret the effects of *Income* and *Age* on *Balance*.
3. In case interaction term is not statistically significant, would it be beneficial to drop it? Or should we still retain it? Why (give a couple reasons)?

### Problem #2

For the *Wage* data set from *ISLR* package, in studying whether there's an interaction between

1. age and job class,
2. age and marital status

in explaining person's wage, proceed to do the following:

- a. Provide data visualization that will help determine if there's a strong interaction between respective variables in explaining the response. Comment on whether you observe an interaction. Explain. **NOTE(!)**: You shouldn't start fitting any models with interactions yet.
- b. Proceed to confirm your hunch from part 2(a) by actually writing out the appropriate model, fitting it, and conducting the respective statistical test for significance of **the overall** interaction. Make sure to write out all the steps of the respective hypothesis testing procedures, from stating the hypotheses, down to conclusion.
- c. If you found an overall significant interaction, proceed to interpret only the significant **individual** interaction terms. Otherwise, if the overall interaction is not significant - what should be your following steps (simply list them, no need to actually perform them)?