Problem 3.1. (14 points) **Pizza & ice cream.**

Recently you were sent a link to a survey regarding pizza and ice-cream preferences. There were altogether 94 people who responded to the survey (there was hope for a higher response from the student population). The results, as Google reports them, are in the spreadsheet you received as an attachment. To see if there is any evidence of association between pizza and ice-cream preferences, please do the following:

(i) (5 points) “Clean up” the spreadsheet so that you have more manageable entries in the cells. You can do this in R or using some Excel-like software. Then, create and display a two-way table summarizing the results of our survey.

(ii) (4 points) Graph the data from the two-way table you obtained above. Creative data presentation will earn bonus points. Do not be afraid to download additional R libraries.

(iii) (5 points) Perform that $\chi^2$ test to see if there is an association between your subjects’ preferences.

Problem 3.2. (41 points) **Walkability vs. rent.**

Go to [https://www.walkscore.com/TX/Austin](https://www.walkscore.com/TX/Austin). Then, either look at Austin as a whole, or at the neighborhood in which you currently live, or at a neighborhood in which you would like to live but do not (yet!). For the latter two options you will need to scroll a bit down the webpage. Concentrate on one bedroom apartments. This is essential! Do not mix-and-match different properties!!!

(i) (2 points) Pick out 10 different apartments in the above category however you want. For instance, you can pick them to your liking. A better scenario would be to outsource this job to a friend/colleague/room-mate instructing them to choose as randomly as possible. Do not overthink this phase, please.

For each of the chosen apartments, record the following in a table: the monthly rent, the Walk Score, and the Transit Score. Make your table neat and display it.

(ii) (8 points) Make a scatterplot of the Walk Score vs. monthly rent. Perform a simple linear regression analysis and report your findings. The best way to do this is to output the summary from R and accompany it with your own conclusions/comments about the obtained result.

(iii) (8 points) How much in rent do you predict to have to pay if your required Walk Score is 70? Provide both the confidence interval for the mean response and the prediction interval, please!

(iv) (2 points) Imagine that your friend wants an apartment with a WalkScore of 90. What would you report to him/her based on your statistical analysis?

(v) (8 points)

Make a scatterplot of the Transit Score vs. monthly rent. Perform a simple linear regression analysis and report your findings. Again, the best way to do this is to output the summary from R and accompany it with your own conclusions/comments about the obtained result.

(vi) (8 points) How much in rent do you predict to have to pay if your required Transit Score is 60? Provide both the confidence interval for the mean response and the prediction interval, please!

(vii) (5 points) Think of an improvement to the statistical analysis you used above. I am not referring to data presentation, but to the statistical tools you used. Describe your proposed improvement, and explain why you think it would be an improvement.

---

**Instructor:** Milica Ćudina