

**Topics:**

- Single Factor ANOVA and Multiple Comparisons Procedures
- Describing Visual Displays
- Lessons Covered: 31- 33
- Textbook Chapter (Optional): 10

**Grading:**

- Points are listed next to each question and should total 25 points overall.
- Grading will be based on the content of the data analysis as well as the overall appearance of the document.
- Late assignments will not be graded.

**Deadlines:**

- Final Submission: **Monday, February 18<sup>th</sup>**. All submissions must be PDF files.

**Instructions:**

- Clearly label and **type answers** to the questions on the proceeding pages, **without** question prompts, in Word, Google Docs, or other word processing software.
- Insert **diagrams or plots as a picture** in an appropriate location.
- Math Formulas need to be typed with Math Type, LaTeX, or clearly using key board symbols such as +, -, \*, /, sqrt() and ^
- Submit assignment to the Canvas link as a PDF. Verify the correct document has been uploaded. If not, resubmit. You can submit up to three times.

**Allowances:**

- You may use any resources listed or posted on the Canvas page for the course.
- You are encouraged to discuss the problems with other students, the instructor and TAs, however, all work must be your own words. Duplicate wording will be considered plagiarism.
- Outside resources need to be cited. Websites such as Chegg, CourseHero, Koofers, etc. are discouraged, but if used need to be cited and used within the boundaries of academic honesty.

### Part 1. (25 points)

Single Factor ANOVA is a method we use when we want to compare a quantitative variable among more than two categories. It evaluates whether the means of different treatment groups, or populations, are equivalent. When we only have two populations then we can perform a two-sample t procedure, but when we have more populations we need to examine the data with Single Factor ANOVA.

In the R script `DA6_Single_Factor_ANOVA.R`, follow along with the analysis that compares average number of roommates between majors for the ST314 online students. You will need to upload the student information dataset `ST314SISW19.csv`.

Once you have reviewed the example analysis, conduct your own by choosing one of the following three options:

**Option 1:** Is there evidence average number of terms a student has been studying at OSU differs between majors for ST314 students?

**Option 2:** Is there evidence average weekly gaming hours per week differs between majors for ST314 students?

**Option 3:** Is there evidence average anticipated salary differs between majors for ST314 students?

For the option you selected, answer the questions below. Use a significance level of 0.10.

- a. **(3 point)** Create side-by-side boxplot of the data and add color and a title to your plot. Paste your plot.
- b. **(2 point)** From the side-by-side box plot, does there look to be a difference between the averages? Explain your reasoning.
- c. **(2 point)** State the appropriate null and alternative hypothesis for the Single Factor ANOVA F test.
- d. **(3 points)** State the conditions for the Single Factor ANOVA F Test. Is it reasonable to seem that these conditions are satisfied? *If not, still proceed.*
- e. Perform the Single Factor ANOVA F test in R.
  1. **(2 point)** Paste the ANOVA table.
  2. **(2 points)** From the ANOVA table, what is the average between group variability and the average within group variability, respectively the  $MSTr$  and  $MSE$ ?
- f. Use the F statistic and p-value from the ANOVA table to state whether there is a significant difference between at least two of the group means.
  1. **(2 points)** State whether to reject the null. State the test statistic and p-value.
  2. **(2 points)** Include a statement in terms of the strength of evidence in terms of the alternative. Include context.
- g. Using the Tukey's Multiple Comparison procedure output. Are there any individual comparisons that are significant at the 0.10 significance level?
  1. **(2 point)** Paste R output for the multiple comparisons procedure.
  2. **(2 point)** List all comparisons that are significant (or state those that are not).
  3. **(3 points)** Interpret the 90% F-W confidence interval for the difference with the smallest p-value (even if it is not significant).