

Spring Data Project

For the Spring Data Project, you will give a presentation on your exploration of a data set. You will have the next week in class to work with one partner to try to find out as much interesting information as possible about your data. I am happy to help in any way I can with any questions you might have. You are expected to examine the data critically and ask questions that you can reasonably answer. You should prepare a rmd with all of the calculations, a PowerPoint slide show, and an 8-10 minute presentation. Your presentation should be a story that explains what you were trying to do with the data and what successes and failures you had. Do not just make every graph that you possibly can. I am looking to see that you understand what you are doing and why you are doing it.

Unless we decide otherwise your presentation must have: at least 1 multilinear regression and 1 ANOVA. You must have graphs of data (scatterplots, residual plots, histograms, boxplots, bar charts etc.) with explanations of what those graphs tell us. You are encouraged to analyze subsets of the population, create new statistics by combining variables, do a simulation if it seems appropriate, do hypothesis tests, create confidence intervals, or write a function to do something interesting.

I believe that there are lots of interesting things to find in the data sets provided. Do your best to ask interesting questions. The more creative you get with your questions the better your project will be. Do not worry if the data doesn't give you the exciting answers that you want. The questions are more important than the findings.

You will be graded on the quality of your presentation, the creativity of your questions, the accuracy of your work in R, and your willingness to take risks in analyzing the information.

The data sets are as follows:

[Electric Vehicle Population Data | Kaggle](#)

[American Housing Survey National Cleaned Version | Kaggle](#)

[Survey on Maternity\(NJ, USA\) | Kaggle](#)

[Fastfood Nutrition | Kaggle](#)

[MLB Hitting and Pitching Stats Through All Time | Kaggle](#)

[Smart Watch prices | Kaggle](#)

[Animal Crossing New Horizons Catalog | Kaggle](#)

[Life Expectancy \(WHO\) Fixed | Kaggle](#)

[!\[\]\(3dfb8d66e81160ad61421a3452093d1b_img.jpg\) Students' Perceptions of AI in Education | Kaggle](#)

[World Happiness Report 2021 | Kaggle](#)

[Women-in-Film - dataset by alice-c | data.world](#)

[2022 March Laptop data | Kaggle](#)

[Marvel Comics Characters Stats and Powers | Kaggle](#)

[Top rated English movies of this decade from IMDB. - dataset by saipranav | data.world](#)

[IMDB Movies Dataset | Kaggle](#)

Lots of baseball data sets: [Baseball Databank | Kaggle](#)

Car Fuel Consumption: [2022 Fuel Consumption Ratings | Kaggle](#)