

Regression Project Milestone #1

PART 1: Nice to meet you!

- Exchange contact information and write up a brief plan for collaboration (1 page max). Include times that you'll be available to work with each other, what tools you'll use for collaboration (Google docs is great!), and how work will be equally distributed.

Example of submission:

Names: Hello Kitty and Minnie Mouse

Emails: hkitty@willamette.edu , mmouse@willamette.edu

Collaboration plan: We plan to meet after class on Tuesdays to work on our project. Since we are remote students we will be meeting on Zoom. We are going to make a google drive folder to share. We will also make a Google doc so that we can both edit it simultaneously.

Equitability: We plan to program together using screen sharing. We are also dividing sections of the paper.

PART 2: Look for a dataset!

Discuss topics

What are you passionate about? Find a topic that excites you!

Data Sources:

- a) SWD list of free public datasets: [LINK HERE](#)
- b) UC Irvine Machine Learning Repository <http://archive.ics.uci.edu/ml/index.php>
- c) Kaggle Datasets <https://www.kaggle.com/datasets>

Find **three possible datasets** to use for your regression project (i.e. needs a numeric outcome)

- Describe the variables used in the data set and what types of variables there are
- There must be at least 5 features
 - 3 categorical variables
 - 2 numeric variables

Please provide the following information in your group submission:

- Description of the dataset
- Link to where the data was found
- What is the outcome variable?
- What are the features and what types of variables are they?