CS510 Data Science Exploration, Project Plan
Name: Qiushi Teng

## Project Objective (1-2 sentences)

In this project, I plan to compare how traffic has changed in the last four years using data from Portal. The system I plan to use is Apache Spark.

## Project Approach (1-2 paragraphs)

For this project, I will use the data form Portal to compare how traffic has changed in the last four years. More specifically, I may compare the speed and volume columns from the data of highway US26 or OR217. I will do a general comparison between these four years. Then depending on the size of the data set, I may use data from a specific month of all four years instead of an entire year to make some further comparisons, such as if there is a significant difference between the peak time traffic.

The system I plan to use is Spark, which is a popular system used for big data workloads. Other systems should also work fine for this project. Since I have already worked on projects using PostgresQL, and MongoDB in previous courses, I think it will be interesting to try and learn some other new systems.

## Team Structure (3-4 sentences)

I will work on my own.

## Project Milestones (~5 milestones)

- Get the Spark system set up and ready for this project
- Learn how to perform different kinds of queries using Spark
- Gather data needed for this project
- Perform different queries for further analysis
- Analyze and compare how traffic has changed