Introduction:

In this project, you will put that knowledge to use!

Put yourself in the shoes of a data scientist being given a data set and asked to draw conclusions from it. Your job will be to understand what the data is showing you, design the analyses you need, justify those choices, draw conclusions from running the analyses, and explain why they do (or do not) make sense.

We are deliberately not giving you detailed directions on how to solve these problems, but feel free to come to office hours and recitation hours to brainstorm.

Objectives

There are two possible paths through this project:

1. You may use data set #1, which captures information about bike usage in New York City. See below for the analysis questions we want you to answer.

Problem:

Path 1: Bike traffic

The NYC_Bicycle_Counts_2016_Corrected. csv gives information on bike traffic across a number of bridges in New York City. In this path, the analysis questions we would like you to answer are as follows:

1. You want to install sensors on the bridges to estimate overall traffic across all the bridges. But you only have enough budget to install sensors on three of the four bridges. Which bridges should you install the sensors on to get the best prediction of overall traffic?

Expected Solution:

report. pdf : A project report, which should consist of:

- A section with the names of the team members (maximum of two), your Purdue username(s), and the path (1 or 2) you have taken.
- A section describing the dataset you are working with.
- A section describing the analyses you chose to use for each analysis question (with a paragraph or two justifying why you chose that analysis and what you expect the analysis to tell you).
- A section (or more) describing the results of each analysis, and what your answers to the questions are based on your results. Visual aids are helpful here, if necessary to back up your conclusions. Note that it is OK if you do not get "positive" answers from your analysis, but you must explain why that might be.

• All Python . py code files you wrote to complete the analysis steps.

Note: You may encounter instances in the data where you would need to make a decision on including extra information, changing non number values in files to appropriate numbers. How you choose to handle this is up to you. You may visit office hours to seek guidance.