



Project Details

This project is divided into two major parts. In the first part, you will conduct an **exploratory** data analysis on a dataset of your choosing. You will use Python data science and data visualization libraries to explore the dataset's variables and understand the data's structure, oddities, patterns and relationships. The analysis in this part should be structured, going from simple univariate relationships up through multivariate relationships, but it does not need to be clean or perfect. There is no one single answer that needs to come out of a given dataset. This part of the project is your opportunity to ask questions of the data and make your own discoveries. It's important to keep in mind that sometimes exploration can lead to dead ends, and that it can take multiple steps to dig down to what you're truly looking for. Be patient with your steps, document your work carefully, and be thorough in the perspective that you choose to take with your dataset.

In the second part, you will take your main findings from your exploration and convey them to others through an **explanatory** analysis. To this end, you will create a slide deck that leverages polished, explanatory visualizations to communicate your results. This part of the project should make heavy use of the first part of the project. Select one or two major paths in your exploration, choose relevant visualizations along that path, and then polish them to construct a story for your readers to understand what you found.

Step 1.1: Choose your Dataset

First, you will choose a dataset from the Dataset Options.

Download the Dataset Options file for full details & descriptions from the Resources Tab.

1. Click on Resources in the leftmost panel of your classroom
2. Click the File Name to start download

Quick List Below:

[Dataset Options](#)

[Ford GoBike System Data](#)

[Flights](#)

[Loan Data from Prosper](#) with [Prosper Data Dictionary to Explain Dataset's Variables](#)

