# DS 3000 – Dataset

In this deliverable, you will identify your dataset (if you haven’t already) and retrieve your data, which may involve you downloading a data file or (even better) scraping your own dataset from an online source. Follow the instructions below.

**What to submit**

* Complete the sections on the following pages.
* Before submitting your completed document, delete this first page containing the instructions.
* Submit your completed document on Blackboard.
* Submit your dataset as a CSV file.
* Submit your Jupyter Notebook file if you’ve scraped your data.

## Project Topic Idea: <I*nsert title here>*

Of the project ideas you submitted in the previous deliverable, select the one you want to work on throughout this project. In making your decision, please refer to the feedback you have on the previous deliverable. If all ideas were found plausible, feel free to select the one that you liked the best **as a team**.

Complete this section based on the previous deliverable. After deciding on your topic idea, simply copy and paste same information from the Topic Proposals document here.

### Problem Statement

* Describe the problem you would like to tackle.
* What is the topic of your project?
* What do you want to learn about it?

### Significance of the Problem

* Why is it important to tackle this problem in your project?
* In what ways could the insights from this project be useful?

### Dataset(s)

* Describe where you obtained your data. Provide a link to the original source.
* This should be the dataset(s) you are using and should correspond to the attached dataset.

## Dataset File

Download or scrape your data from the source you identified above. Save your dataset as a CSV file. The first row of the file should contain variable names.

Describe your variables below (add more rows if necessary):

|  |  |
| --- | --- |
| **Variable name in file** | **Description** |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

Submit a CSV file, or multiple files, containing your data.

If you have scraped your data, you should also submit a Jupyter Notebook containing your Python code used to scrape the data. Please be reasonable and comment your code out whenever it makes sense to do so.