**Python Capstone: Data Description**

The dataset being used includes the dataset that is provided by Coursera. The data has 37 attributes (columns) of numerous accidents and 194, 673 rows. A preliminary review of the data revelas that the severity of the accidents have been categorized in two categories: Severity of “1” and Severity of “2” where the level 2 accidents are the most severe. Accidents involving Severity of 1 include those that are “Property Damage Only Collisions” and accidents involving a Severity of 2 include those that are “Injury Collision.”

A preliminary review of the data also reveals that “Level 1, Property Damage Only Collisions” reveals that there are the following subcategories with the indicated number of accidents: blank: 3863, Angles: 21050, Cycles: 671, Head On: 1152, Left Turn: 8292, Other: 17591, Parked Car: 45325, Pedestrian: 672, Rear Ended: 19419, Right Turn: 2347, Sideswipe: 16103.

Level 2 (Injury Collision) accidents include the following features with the indicated number of accidents: (blank) 2082, Angles: 27248, Cycles: 9488, Head On: 1744, Left Turn: 10822, Other: 12224, Parked Car: 5324, Pedestrian: 11872, Rear Ended: 29342, Right Turn: 1218, and Sideswipe: 5012

The dataset can be imported into a data frame into a Jupyter notebook, pandas and numpy libraries can be imported into the notebook and the features of the dataset can be identified using Python’s “head” and “shape” commands. Below are screenshots that show features of the data:

A screenshot of a social media post

Description automatically generated

Above-mentioned features of Severity Level 1 and 2 accidents can be visualized as follows:

Severity Level 1: Property Damage Only Collisions:

Severity Level 2 Injury Collisions