

Linear regression

For this exercise, we will learn how to create a line of best fit. We will use the csv file auto-mpg.csv. And you will need to import numpy, pandas and matplotlib.

1.

Read in the auto-mpg.csv and save it to a variable called auto. Check that auto is a pandas dataframe.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2}$$

where \bar{x} and \bar{y} indicate the average value of x and y respectively.