1. Statistics
   1. Let’s explore a dataset
      1. Import the sklearn datasets module
      2. From the module, load the Boston dataset (load\_boston)
      3. Print the keys from the Boston dataset
      4. Print the feature names from the dataset
      5. Print the data from the dataset
      6. Convert the data to a Pandas Dataframe
      7. Add the column headers to the Pandas Dataframe (Hint: they’re in the feature\_names property of the sklearn Bunch)
      8. Add the PRICE column to the Pandas Dataframe (Hint: it’s in the target property of the sklearn Bunch)
      9. Print the summary statistics from the Pandas Dataframe
      10. Plot the PRICE feature
   2. Let’s predict the house price as a function of the other available features
      1. Split the data into two: the output variable and the input variables
      2. Train a linear regression model using your training data
      3. Score your model (make a prediction) using your test data
      4. Draw a scatter plot to show predicted versus actual prices
      5. What comment could you make about the linearity of the plot?
      6. Show how good (or otherwise) a fit our model is by calculating the Mean Square Error
      7. Show how much of the observed variance in our data is explained by our model by calculating the coefficient of determination (r2\_score)
      8. Think about what the MSE are r2\_score is telling us and we’ll discuss.