Module 4: Introduction to Numpy, Pandas & Matplotlib
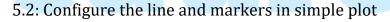
Case Study

edureka!

edureka!

Case Study

1. You are given a dataset, which is present in the LMS, containing the number of hurricanes occurring in the United States along the coast of the Atlantic. Load the data from the dataset into your program and plot a Bar Graph of the data, taking the Year as the x-axis and the number of hurricanes occurring as the Y-axis.

2. The dataset given, records data of city temperatures over the years' 2014 and 2015. Plot the histogram of the temperatures over this period for the cities of San Francisco and Moscow.

3. Create csv file from the data file available in LMS which goes by the name *'M4_assign_dataset'* and read this file into a pandas data frame

4. Let the x axis data points and y axis data points are

    X = [1,2,3,4]
    y = [20, 21, 20.5, 20.8]

    5.1: Draw a Simple plot

    5.2: Configure the line and markers in simple plot

    5.3: configure the axes

    5.4: Give title of Graph & labels of x axis and y axis

    5.5: Give error bar if y_error = [0.12, 0.13, 0.2, 0.1]

    5.6: define width, height as figsize=(4,5) DPI and adjust plot dpi=100

    5.7: Give a font size of 14

    5.8: Draw a scatter graph of any 50 random values of x and y axis

    5.9: Create a dataframe from following data

    'first_name': ['Jason', 'Molly', 'Tina', 'Jake', 'Amy'],
    'last_name': ['Miller', 'Jacobson', 'Ali', 'Milner', 'Cooze'],
    'female': [0, 1, 1, 0, 1],
    'age': [42, 52, 36, 24, 73],
    'preTestScore': [4, 24, 31, 2, 3],
    'postTestScore': [25, 94, 57, 62, 70]

Draw a Scatterplot of preTestScore and postTestScore, with the size of each point determined by age

5.10: Draw a Scatterplot from the data in question 9 of preTestScore and postTestScore with the size = 300 and the color determined by sex