PAI

IU9 Python Pandas – Lab

Overview

In this lab, you will learn how to create, run and visualise using Python.

What You'll Need

To complete this lab, you will need the following:

- Azure Data Studio
    - Notebook using Python Kernel
    - The files for this lab
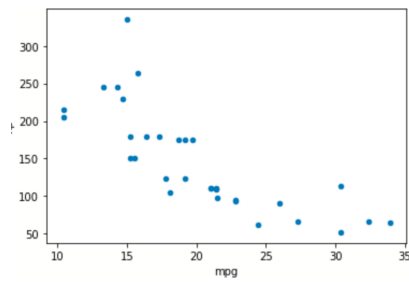
Performing a Python visualization

In this exercise, you will perform upload of a csv file and visualise using Python.

1. From the folder where you extracted the lab files for this module (for example, C:\PAI\IU9), open the **mtcars.csv** file, using either a spreadsheet application such as Microsoft Excel, or a text editor such as Microsoft Windows Notepad.
2. View the contents of the **mtcars.csv** file, noting that it contains data on 16 or 32 cars, documenting mpg(miles per galaon) and cyl(cylinder). Then close the text file without saving any changes.
3. From Azure Data studio , browse to file->add file to workspace, select the folder where mtcars.csv file located in your laptop
4. Import pandas package as pd
5. Use function pd.read_csv to upload mtcars.csv
    1. Use_cols (all the columns except car_names)
   2. Index_col = car_names (expected output as below)

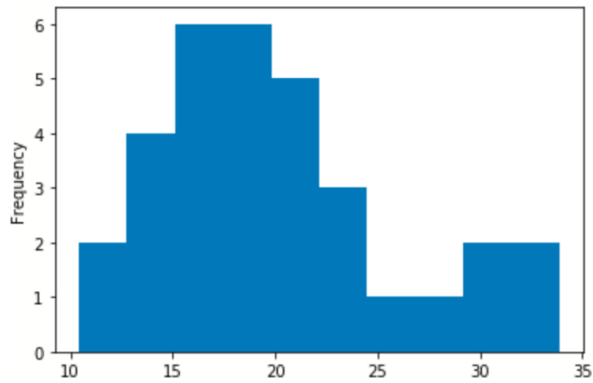| car_names	mpg	cyl	hp	am
Mazda RX4	21.0	6	110	1
Mazda RX4 Wag	21.0	6	110	1
Datsun 710	22.8	4	93	1
Hornet 4 Drive	21.4	6	110	0
Hornet Sportabout	18.7	8	175	0
Valiant	18.1	6	105	0
Duster 360	14.3	8	245	0
Merc 240D	24.4	4	62	0
Merc 230	22.8	4	95	0
Merc 280	19.2	6	123	0
Merc 280C	17.8	6	123	0
Merc 450SE	16.4	8	180	0

   3. Expected output
6. Import matplotlib package as plt

7. Create a scatter plot with mpg and hp
8. Provide xlabel and ylabel (expected output as below)

1.
9. Create crosstab for cyl(cylinder) and columns as count (expected output as below)

1.

10. Create crosstab for cyl(cylinder) and am and create a bar plot (expected output as below)

1.
11. Create groupby , bar plot (expected output as below)

1.

12. Create a pie chart for cyl count (expected output as below)
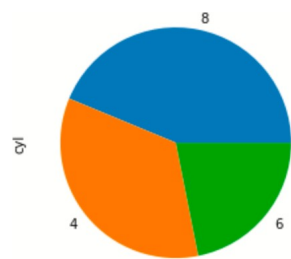
1.

13. Create a Histogram for MPG (expected output as below)

1.