Python – Read Excel using the Pandas Library

Install Pandas Library: -

A picture containing text

Description automatically generated

Text

Description automatically generated

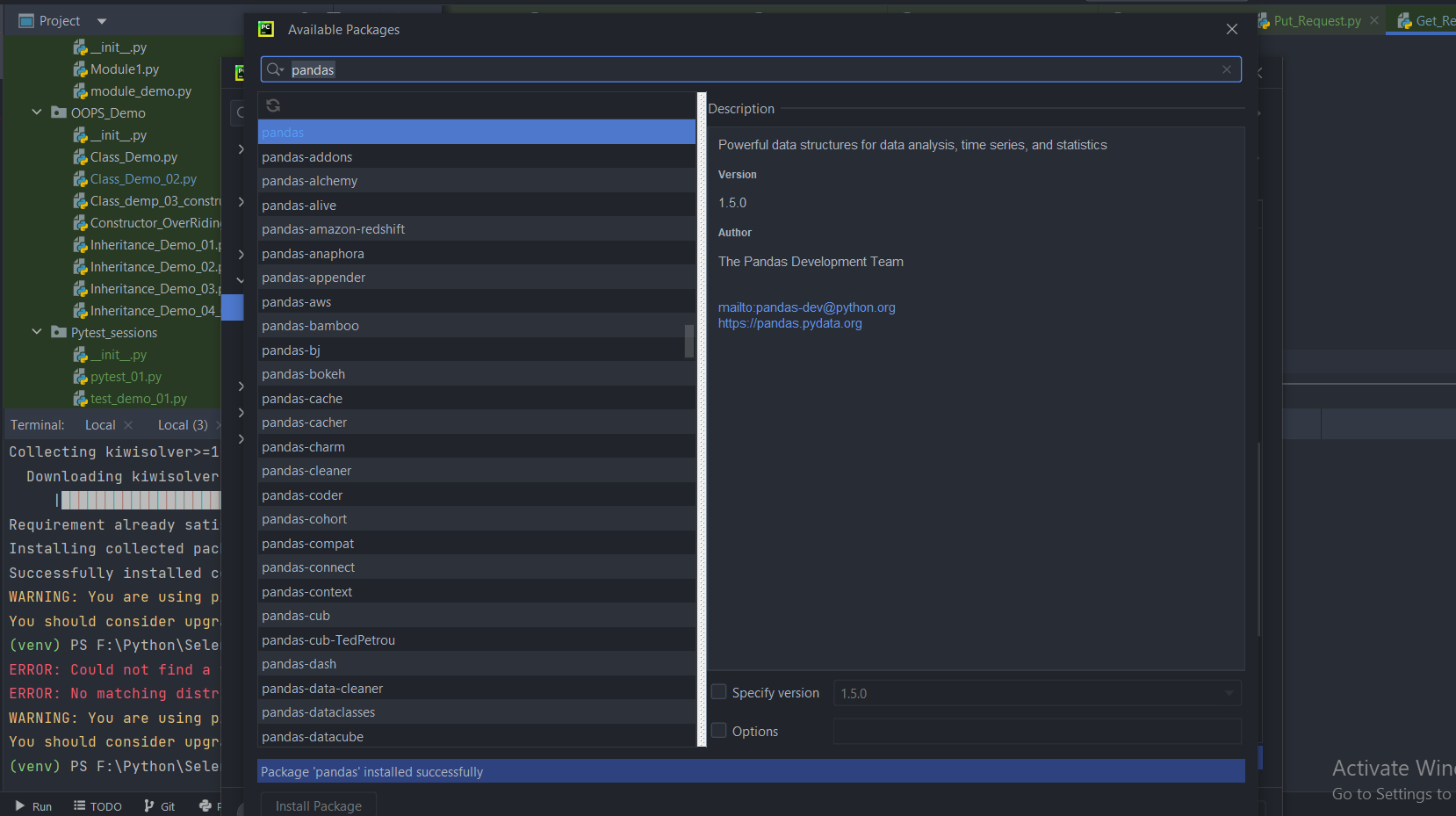
Install numpy Library: -

Text

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence



Numpy Installed successfully

Graphical user interface, application

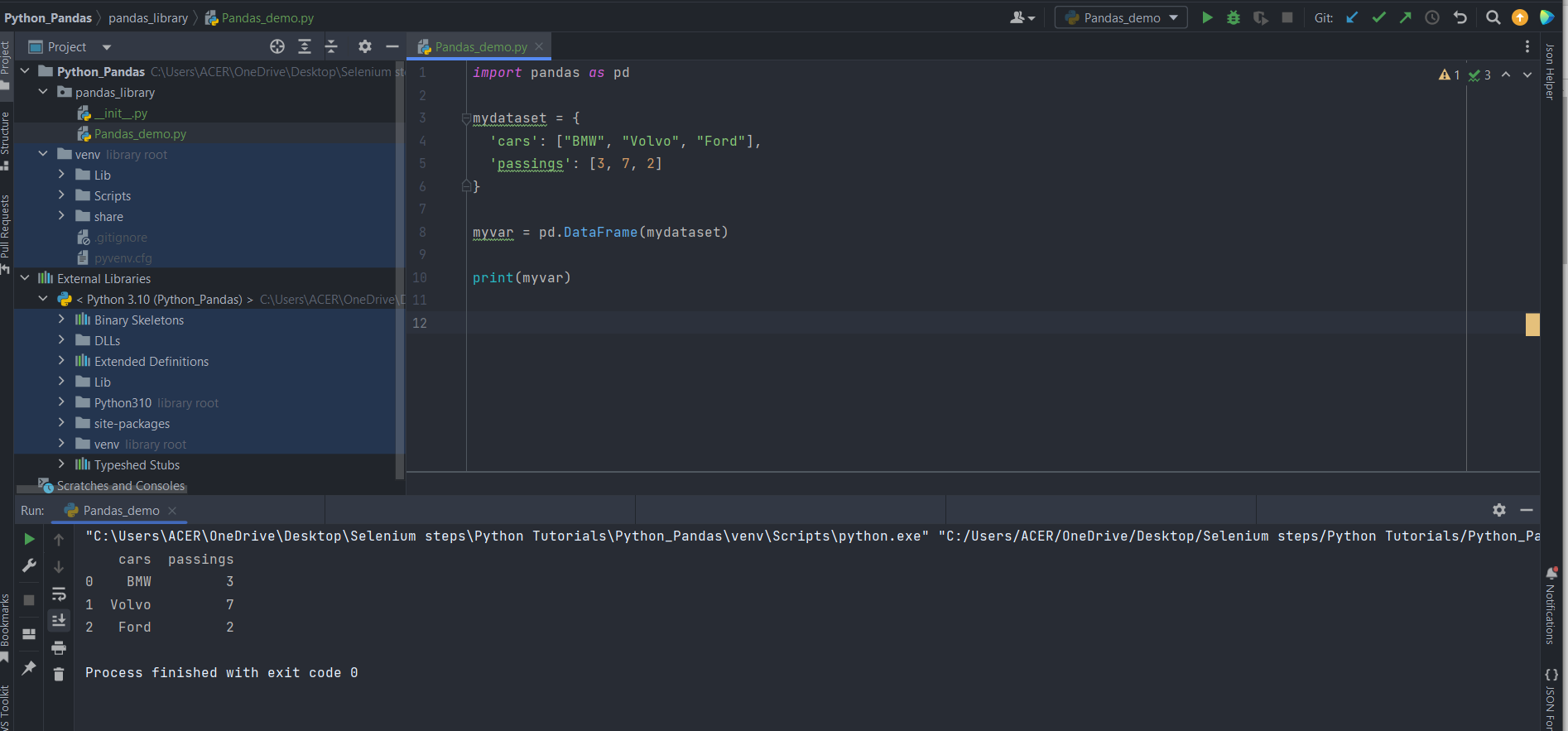
Description automatically generated

Install the matplotlib library package

A screenshot of a computer

Description automatically generated with medium confidence

Try to use **pandas – library – dataset function**



To get the pandas version

Graphical user interface, application, website

Description automatically generated

Text

Description automatically generated

Series - A Pandas Series is like a column in a table.

It is a one-dimensional array holding data of any type.

Text

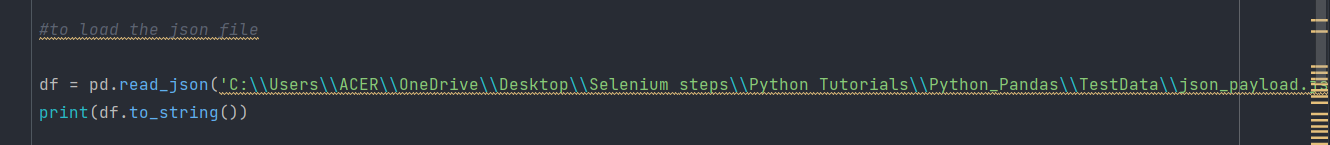
Description automatically generated

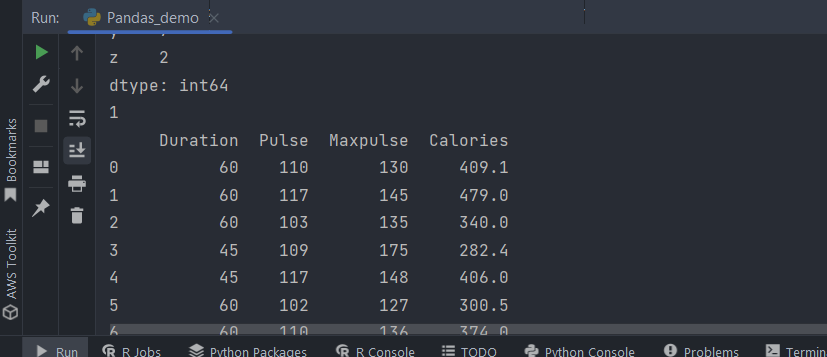
Output:-

A screenshot of a computer

Description automatically generated with medium confidence

To read the json and print it as data frame





It will print the json into the Data Frame

**Let’s load the JSON as Dictionary Object**

A picture containing graphical user interface

Description automatically generated

Shape

Description automatically generated with medium confidence

Output: -

Graphical user interface

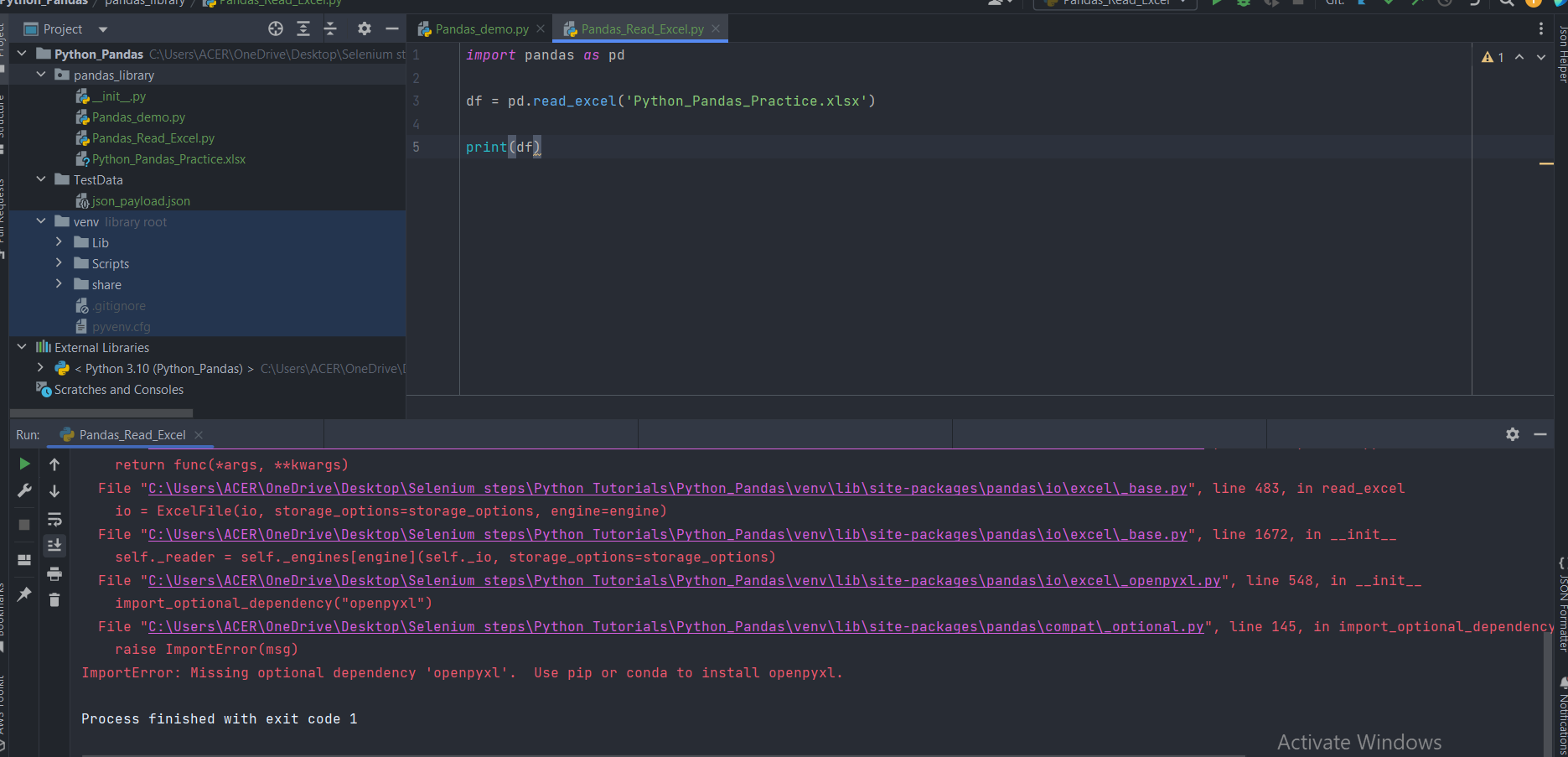
Description automatically generated

**Python – Pandas Basics** for reading the excel: -

**Link for more details**

<https://pythonbasics.org/read-excel/>

If we try to read the excel without installing the xlrd package we will be getting error:-



Graphical user interface, text, application, email

Description automatically generated

**For reading the excel with pandas** – **xlrd library should be installed**

Text

Description automatically generated

Text

Description automatically generated

Xlrd package installed successfully

A screenshot of a computer screen

Description automatically generated with medium confidence

After the xlrd package, install **the openpyxl library too for reading the excel**

Text

Description automatically generated

Now with the help of **read excel method – we can read the excel file easily**

Let’s give the sheet name – and read the specific sheet details

Text

Description automatically generated

Output – It will print **the sheet as data frame**

Graphical user interface, text

Description automatically generated with medium confidence

Now lets get the value of the city column

Text

Description automatically generated

Output:-

Text

Description automatically generated

Let’s try to get the value of the city – Barbados – we just need to pass the index value as 01

Text

Description automatically generated

Output: -

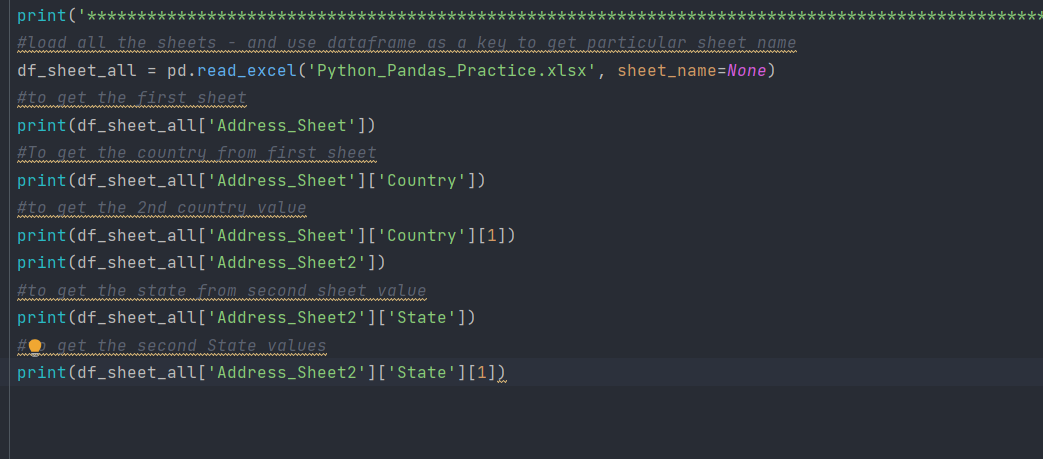
A screenshot of a computer screen

Description automatically generated with medium confidence

Load All the Sheets – in one go

Graphical user interface, text, application, email

Description automatically generated

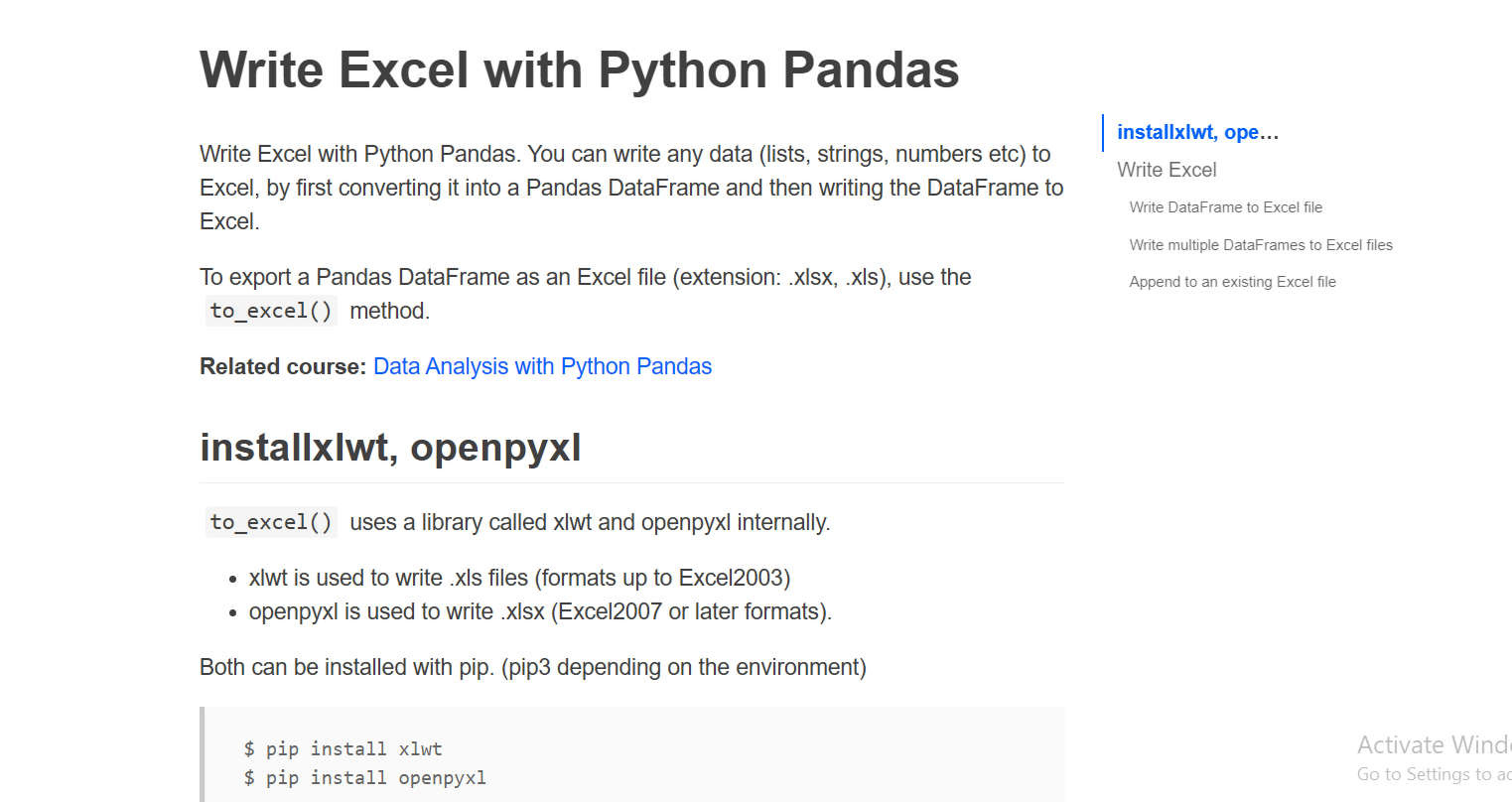


Get all the values – Output

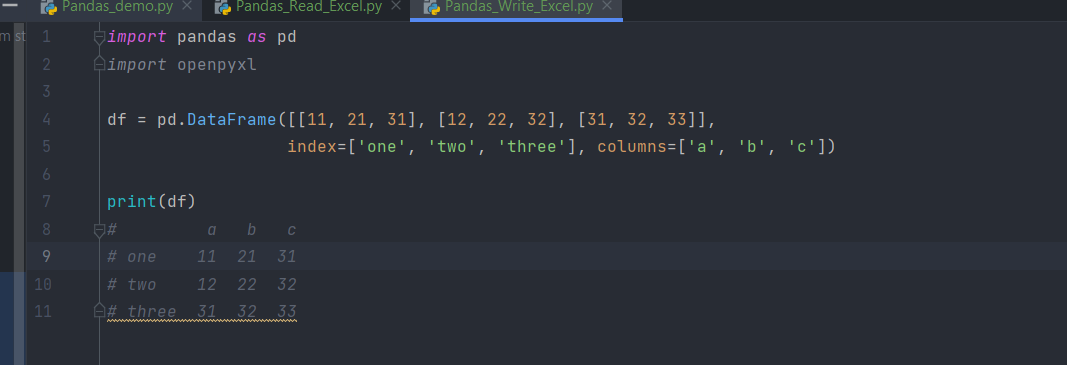
Graphical user interface, text

Description automatically generated

**Write Excel Data – with the help of Pandas Data Frames**

****

Pass the data which we want to **write as Data Frame and specify the column names which we want to write**



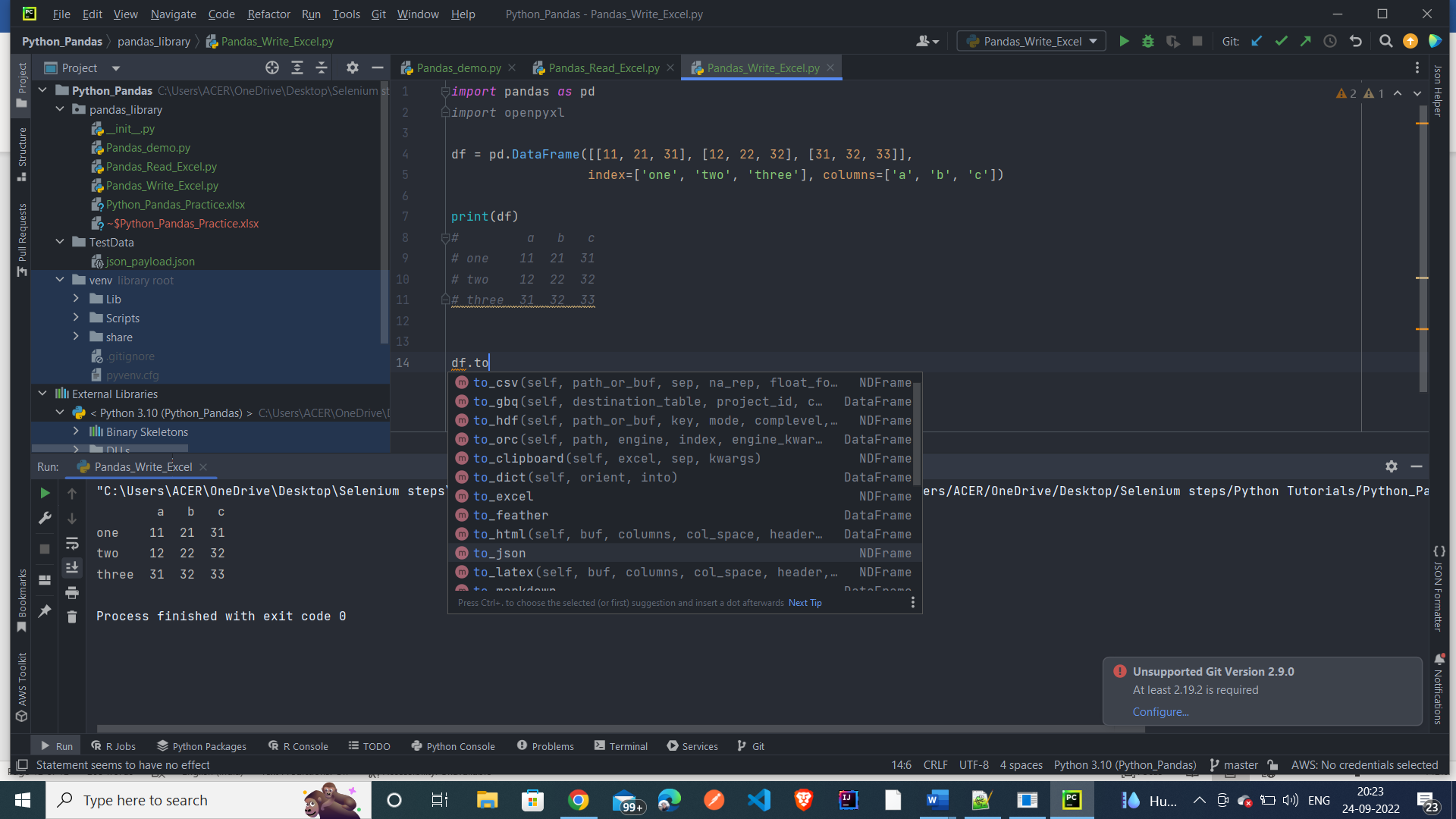
Output :-

Text

Description automatically generated

Let’s write the particular data Frame inside the Excel using to\_excel method

Df -has so many methods to play with – to convert to excel , json ,csv



Text

Description automatically generated

This new excel will be saved inside the package name

Text

Description automatically generated

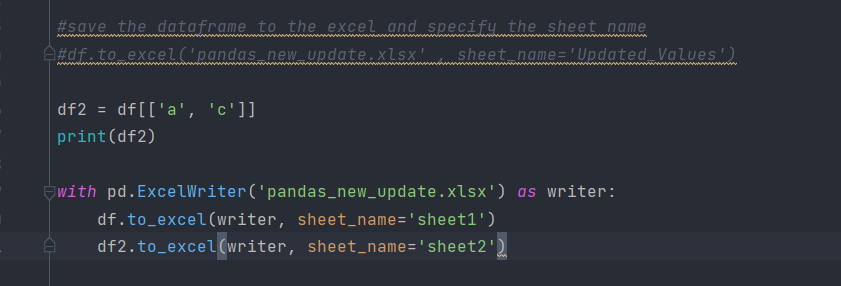
New Excel file got created

Graphical user interface, application, table, Excel

Description automatically generated

We can see the new excel is having the Sheet name updated – having the new values updated

Lets create another df -data frame and write it into another sheet



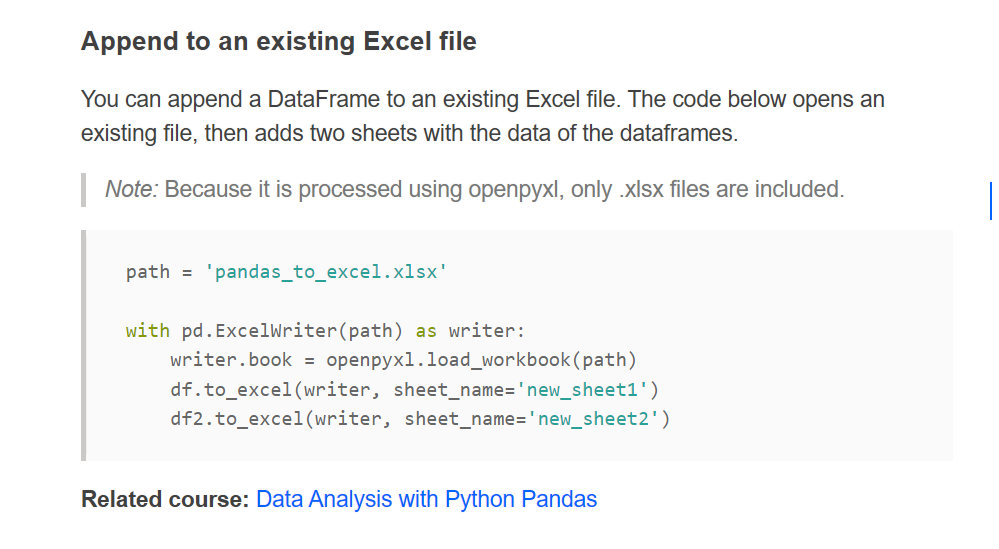
With block can also be used to write the DataFrames into multiples Sheets – it will write into the different sheets

Graphical user interface, application, table, Excel

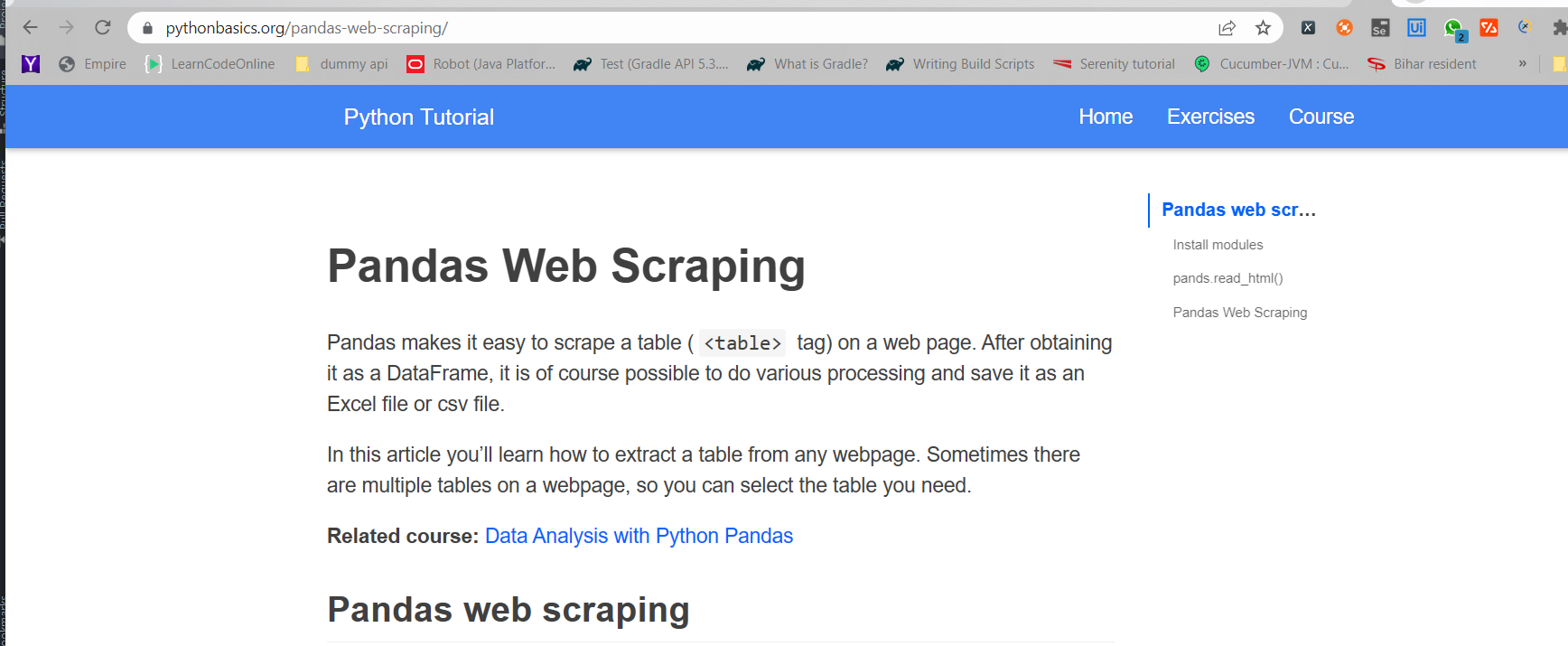
Description automatically generated

We can see two different sheet is created with Sheet 1 and Sheet 2 having different values

Append to an existing file



Pandas – Web Scrapping



Install the following library

pip install lxml html5lib beautifulsoup4

Graphical user interface

Description automatically generated

Graphical user interface, table

Description automatically generated

Text

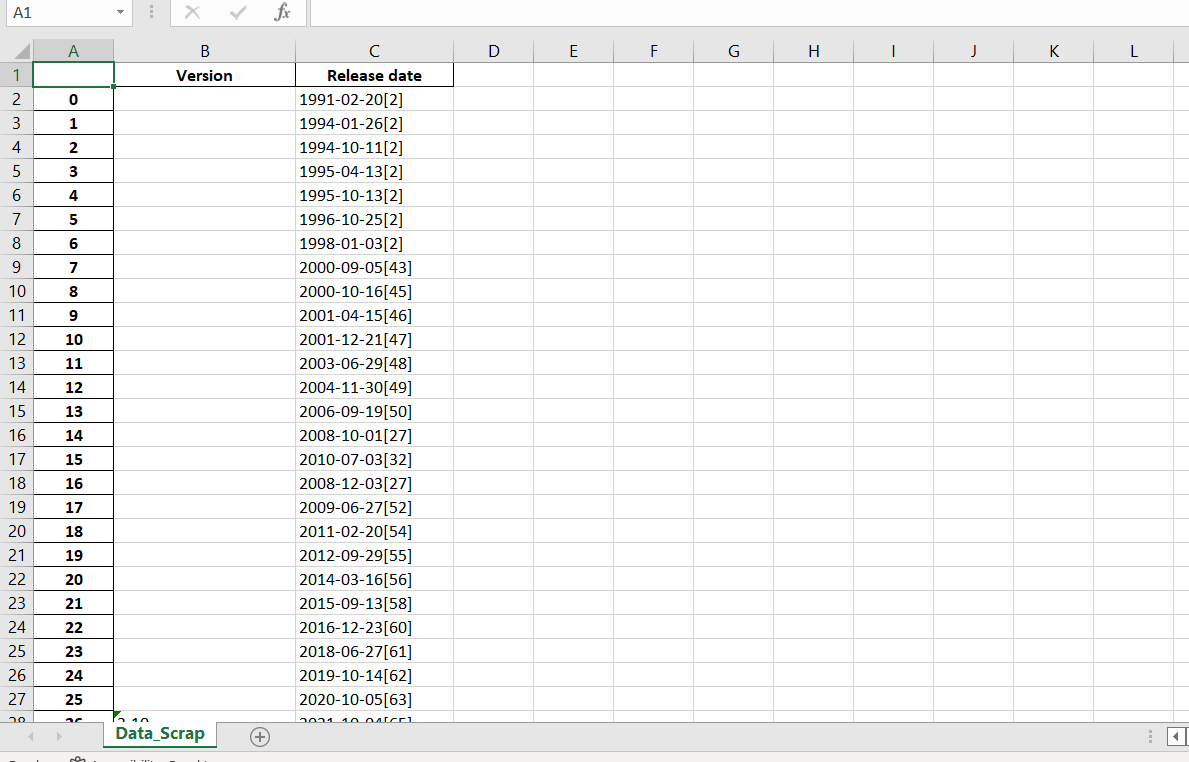
Description automatically generated

It will get the webtable extract from the url

We can save it inside the excel

Text

Description automatically generated



It saved inside the Excel

Lets load the json from the url and save it inside the excel

Text

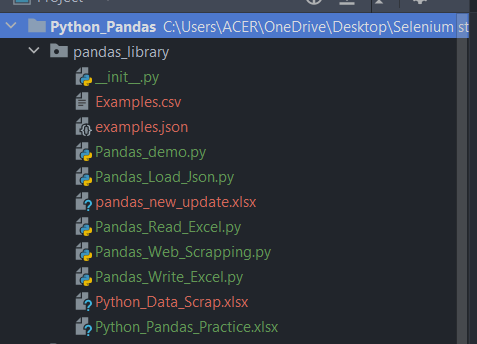
Description automatically generated

The Response will be saved inside the json file

Graphical user interface

Description automatically generated

After reading the json – save it inside the csv file



Csv file is created

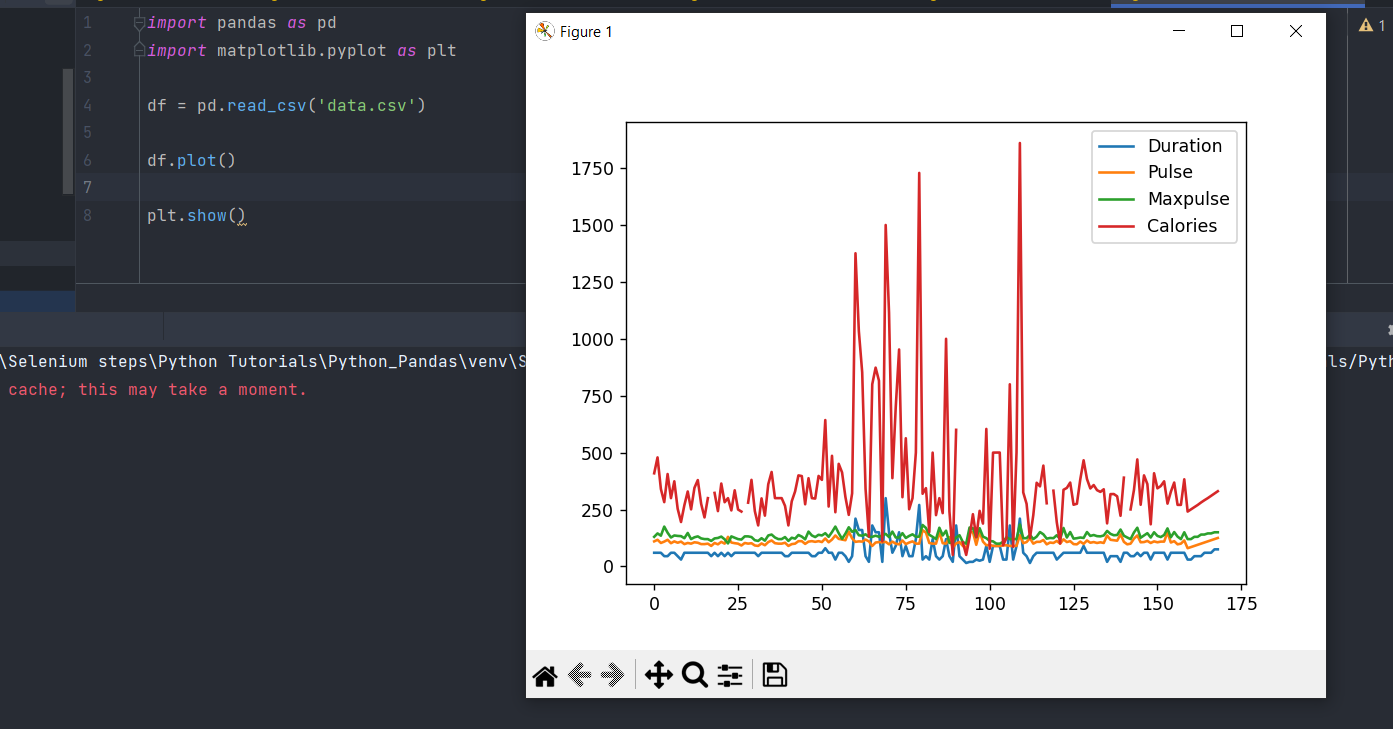
Text

Description automatically generated

**Pandas – Plot for creating the graphs**

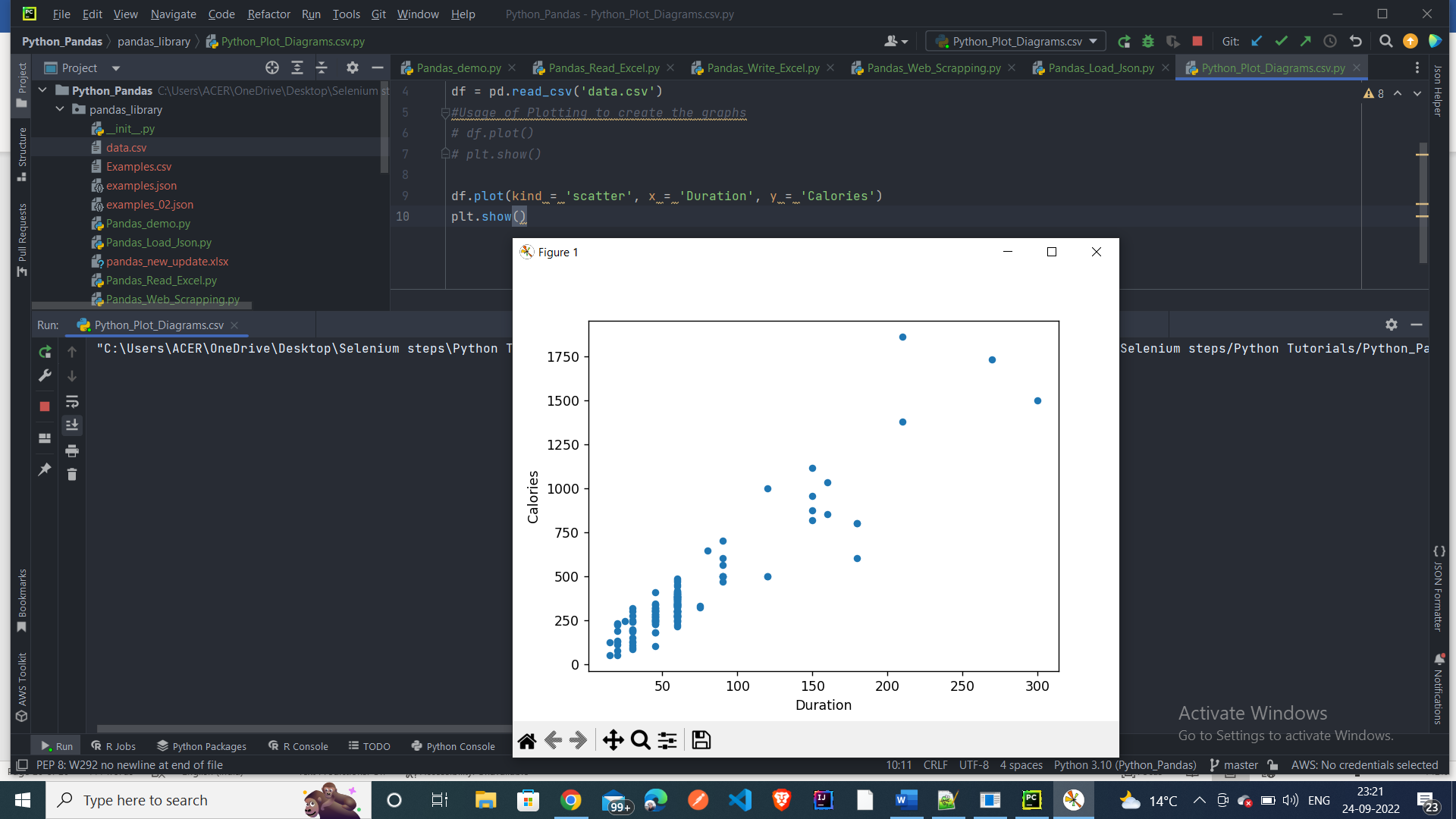
Text

Description automatically generated with medium confidence

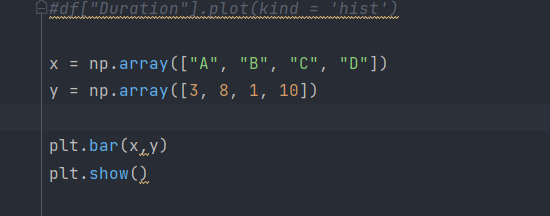


Graphical user interface, text, application, email

Description automatically generated



Create the bar graph



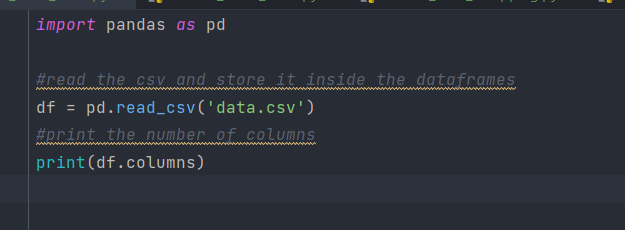
Output :-

Graphical user interface

Description automatically generated with low confidence

**Pandas – Data Analysis**: -

Let’s import the csv using the pandas library: -

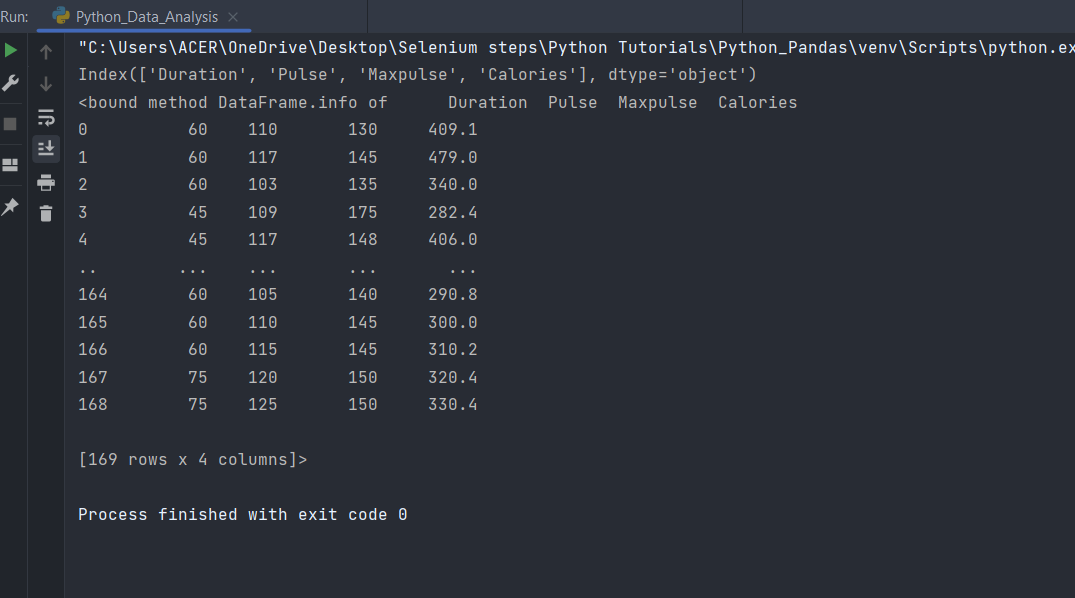


Print the columns – **column names**

Text

Description automatically generated

To have the data info



**Describe method** will talk about the max, mean, count of the data

Graphical user interface, text

Description automatically generated

Graphical user interface, application

Description automatically generated

Output: -

Text

Description automatically generated

Group by – **get the max value of the column**

Graphical user interface, application

Description automatically generated

Output

Text

Description automatically generated

**CSV File Output**: -

Max Calories with respect to Pulse is printed as output

Graphical user interface, application, table, Excel

Description automatically generated

To get the occurrence of particular value from the column

Text

Description automatically generated

Output: -

A picture containing text, scoreboard

Description automatically generated

Inside this Output if we want to get the Calories value – 390

Graphical user interface, application

Description automatically generated

Output

