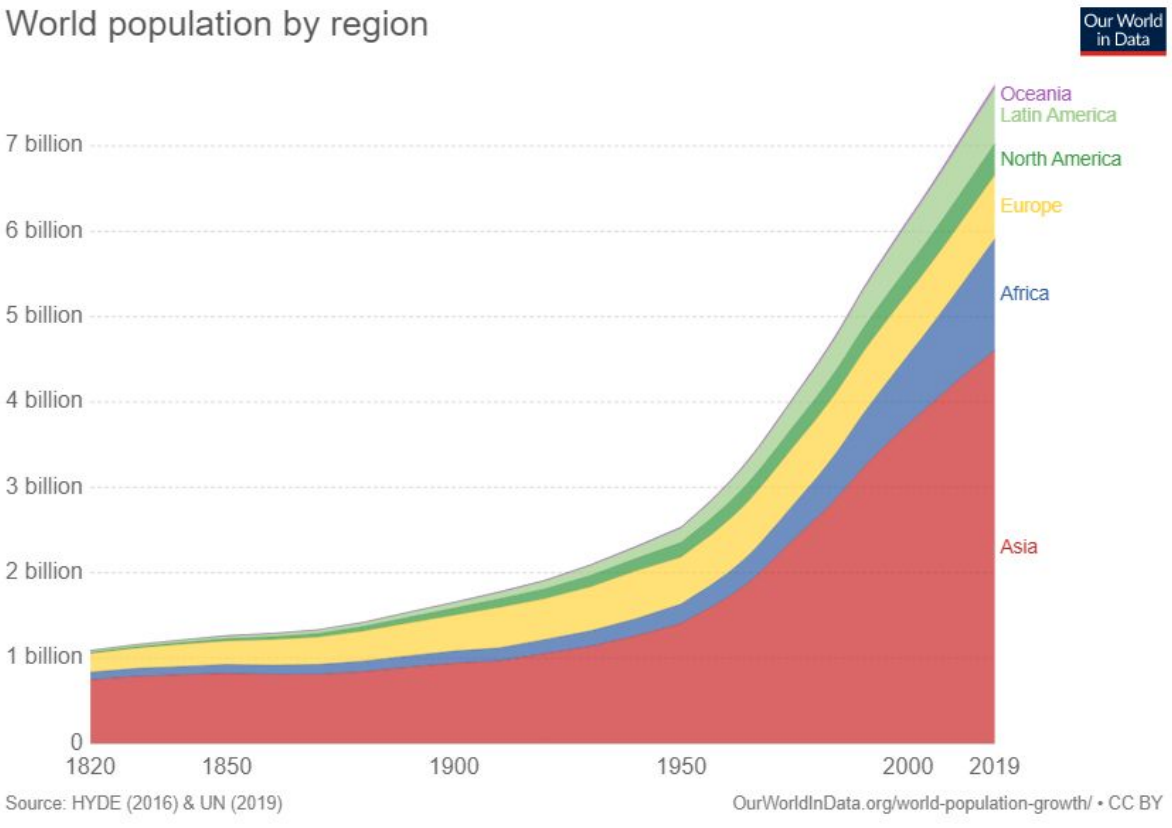
World’s population by region data in tabular format

Country	Start
Afghanistan	3.32 million in 1820
Africa	89.18 million in 1820
Albania	438,671.00 in 1820
Algeria	2.71 million in 1820
American Samoa	19,000.00 in 1950
Andorra	2,700.00 in 1820
Angola	1.60 million in 1820
Anguilla	5,000.00 in 1950
Antigua and Barbuda	37,000.00 in 1820
Argentina	570,719.00 in 1820
Armenia	423,527.00 in 1820
Aruba	38,000.00 in 1950
Asia	745.26 million in 1820
Australia	334,002.00 in 1820
Austria	3.39 million in 1820
Azerbaijan	901,772.00 in 1820
Bahamas	27,404.00 in 1820
Bahrain	64,474.00 in 1820
Bangladesh	20.12 million in 1820
Barbados	85,456.00 in 1820
Belarus	2.41 million in 1820
Belgium	3.48 million in 1820
Belize	25,556.00 in 1820
Benin	653,348.00 in 1820
Bermuda	37,000.00 in 1820
Bhutan	91,144.00 in 1820
Bolivia	1.12 million in 1820
Bonaire Sint Eustatius and Saba	7,000.00 in 1950
Bosnia and Herzegovina	871,970.00 in 1820
Botswana	121,000.00 in 1820
Brazil	4.58 million in 1820
British Virgin Islands	7,000.00 in 1950
Brunei	2,251.00 in 1820
Bulgaria	2.19 million in 1820
Burkina Faso	1.70 million in 1820
Burundi	917,804.00 in 1820
Cambodia	2.10 million in 1820
Cameroon	1.87 million in 1820
Canada	879,432.00 in 1820
Cape Verde	57,062.00 in 1820
Cayman Islands	6,000.00 in 1950
Central African Republic	485,538.00 in 1820
Chad	1.43 million in 1820
Channel Islands	102,000.00 in 1950
Chile	820,889.00 in 1820
China	380.06 million in 1820
Colombia	1.27 million in 1820
Comoros	57,537.00 in 1820
Congo	320,596.00 in 1820
Cook Islands	15,000.00 in 1950
Costa Rica	63,375.00 in 1820

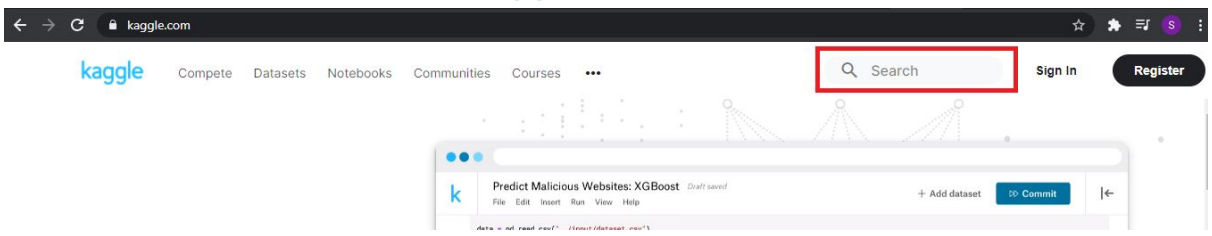
Data visualization region wise as per the population of the above data



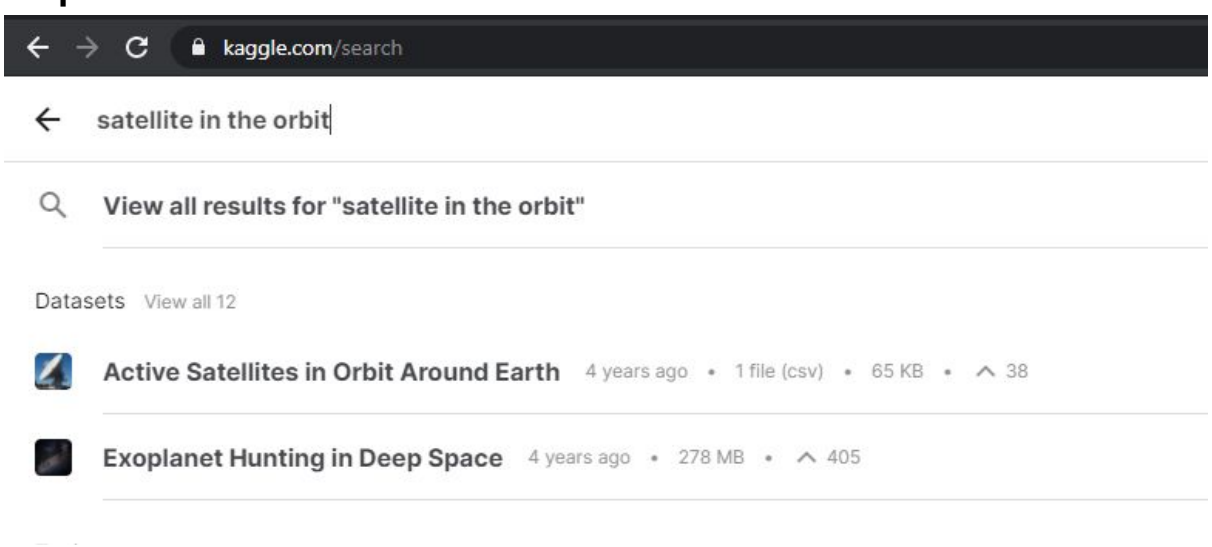
How to find a dataset to work on?

You can find sample datasets on https://www.kaggle.com/ , Kaggle is the world's largest data science community with powerful tools and resources to help you achieve your data science goals.

Step 1: Go to https://www.kaggle.com/

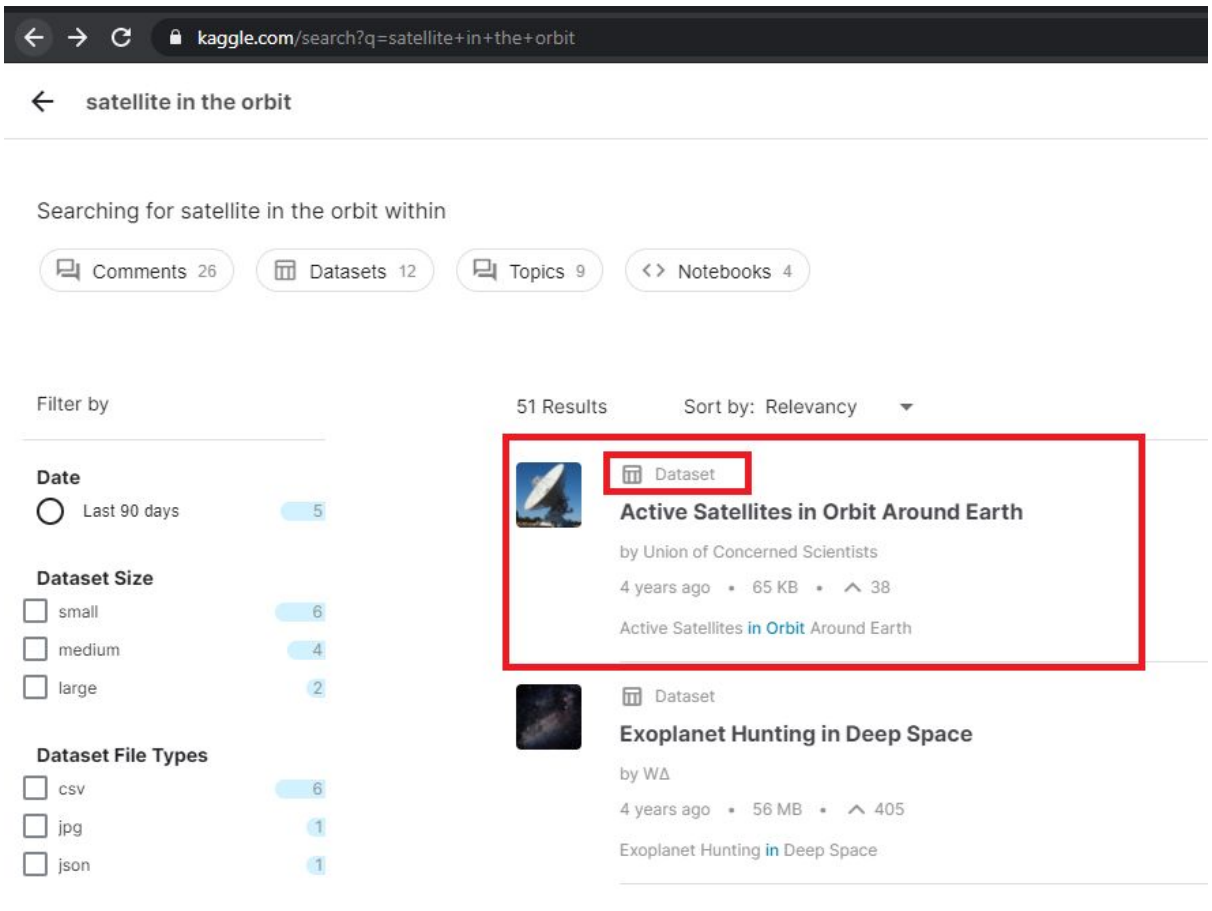


Step 2: Click on the search bar and enter - “satellite in the orbit” and press enter



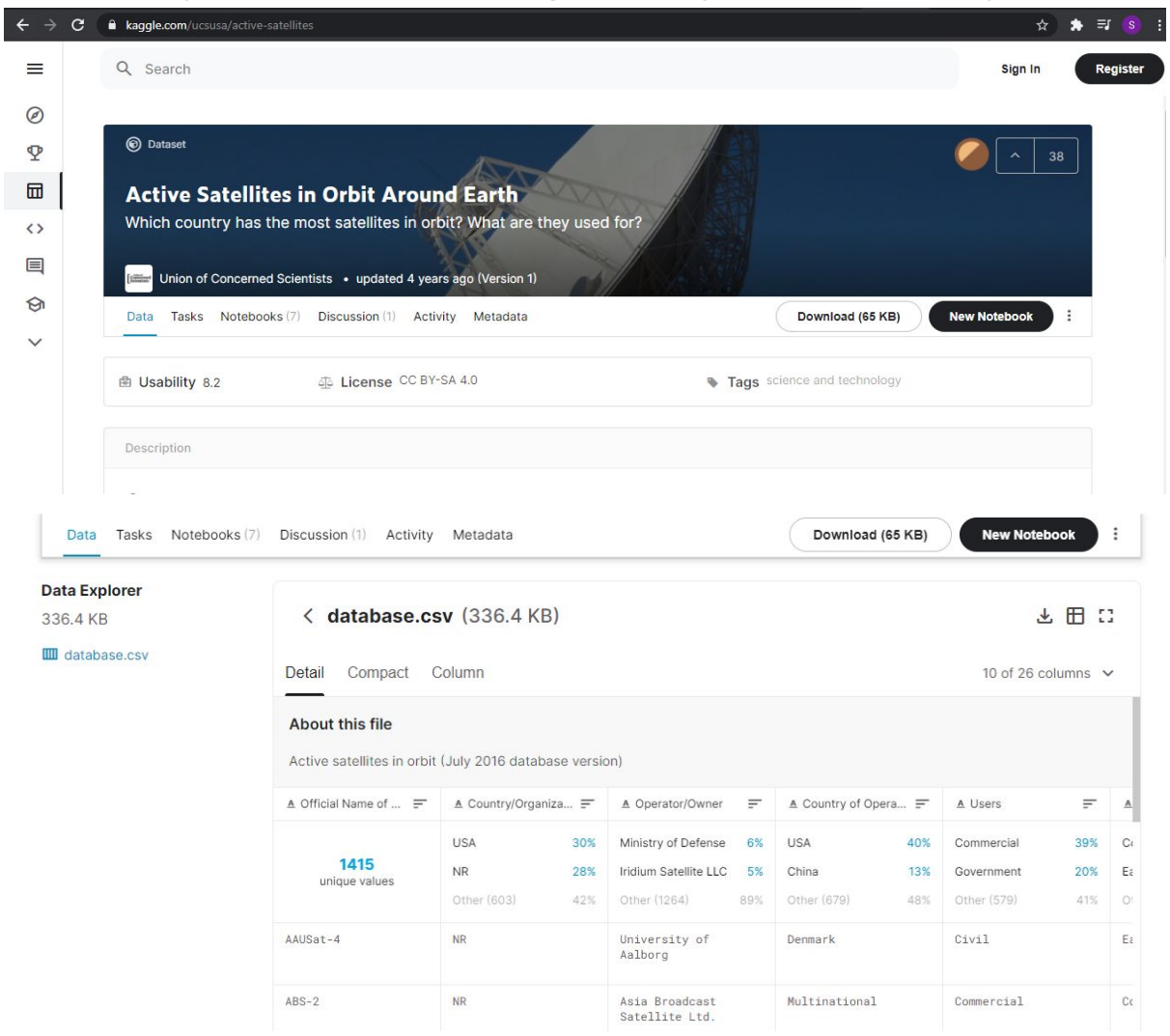
You will get a list of dataset, topics and notebooks but we want a dataset

Step 3: select a dataset



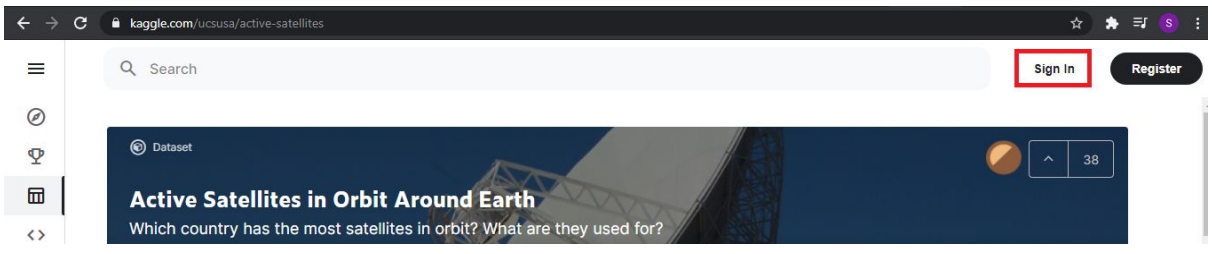
Click on this dataset to select it

It will take you to the dataset page and if you scroll down you can find the data in tabular format

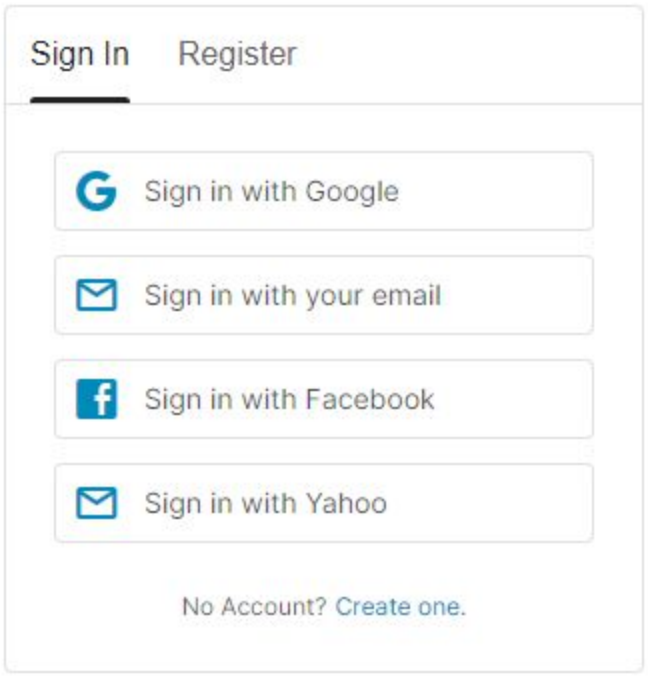


We check if this is the data that we required and download it

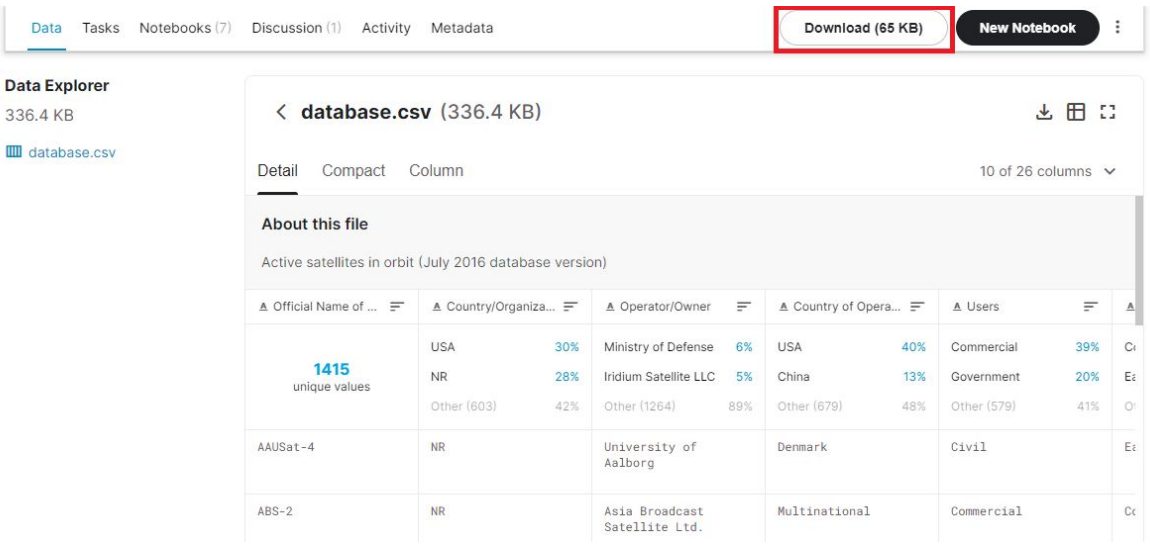
Step 4: Sign in to download the dataset



Select the sign in option

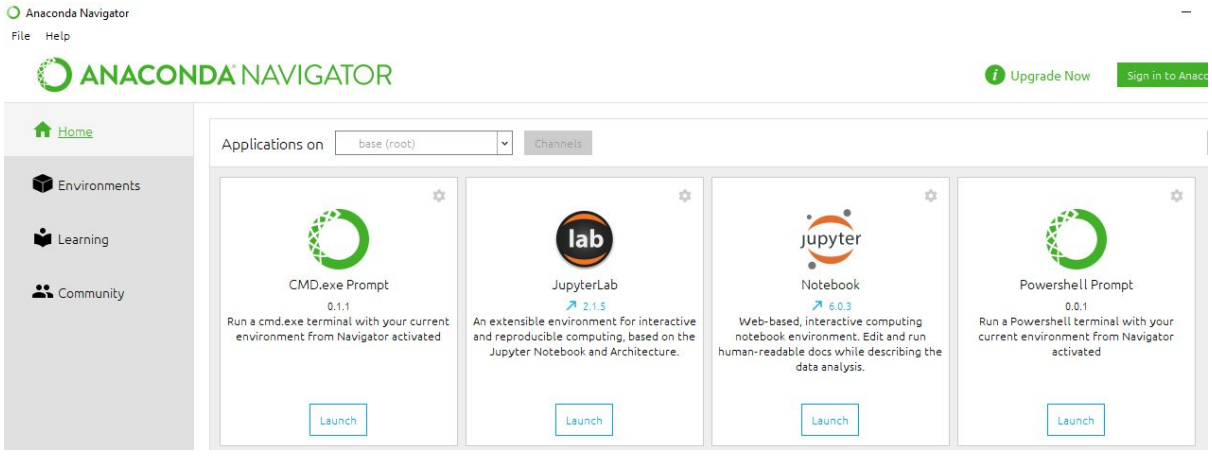


Step 5 : Once signed in you can download the dataset



How to open a Jupyter notebook.

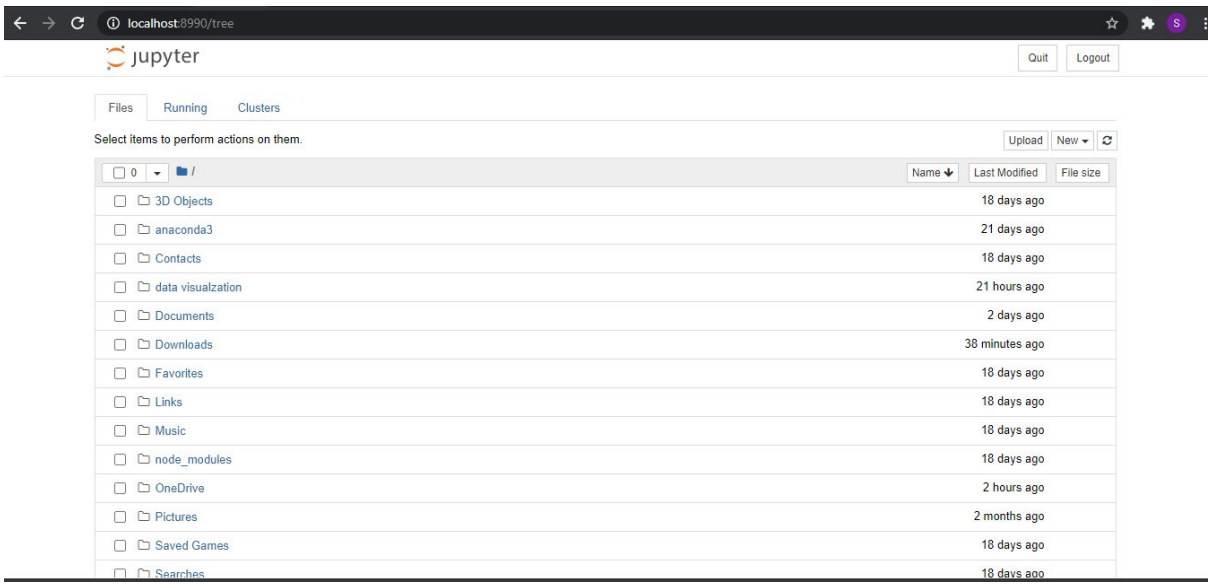
Open Anaconda Navigator



You can see we have a jupyter notebook there below it and we have a launch button. Click on the launch

button and it will open the jupyter notebook in your browser.

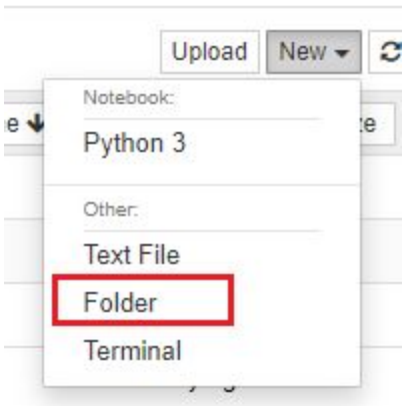
Opens a jupyter notebook like this -



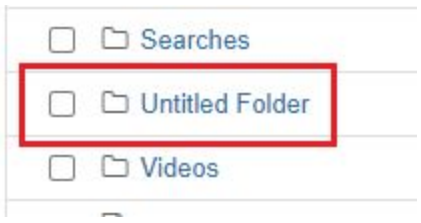
Creating a new folder in jupyter notebook here

Click on new dropdown, it will open the dropdown menu,

The dropdown menu contains folder and files , first we will create a folder



When clicked on folder it will create a folder named “Untitled Folder”



you can rename it by clicking on the checkbox beside it  **and click on rename button**

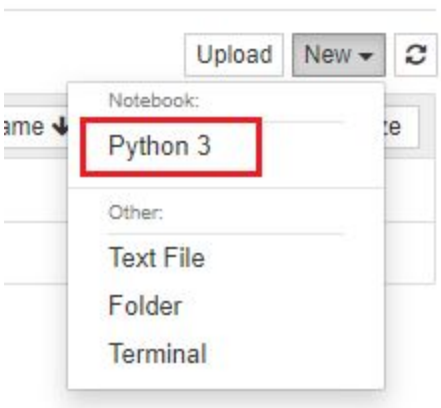
 **to rename it**

Once the folder is created double click on the folder to open it and create a python file

Creating a new file in jupyter notebook

Click on new dropdown, it will open the dropdown menu,

The dropdown menu contains folders and files , we will now create a file.



This will open a python notebook in the new tab of your browser. This is were you can start with your code.

Class Activities

1) **Activity -1 - Plot a bar graph for top 5 Country/Organization who has the most number of Satellites In Space**

Code:

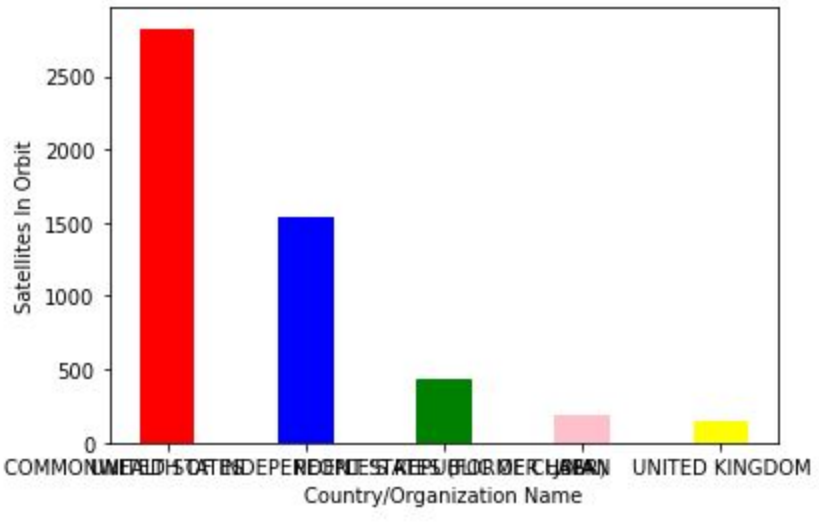
```
#Q - Plot a bar graph for top 5 Country/Organization who has the most number of Satellites In Space

import  numpy as np
import pandas as pd
from matplotlib import pyplot as plt


dataframe = pd.read_csv('country_satellites.csv')

top_5 = dataframe.head(5)
print(top_5)
name = top_5['Country/Organization Name']
number = top_5['Satellites In Orbit']


plt.xlabel("Country/Organization Name")
plt.xticks(rotation='vertical')
plt.ylabel("Satellites In Orbit")


label = name
value = number
plt.bar(label, value,width=0.4, color=('red','blue','green','pink','yellow')) #bar-grap
```

Output:

This is the output of print(top_5) - as the top_5 variable holds the 1st 5 rows of the dataset(table)

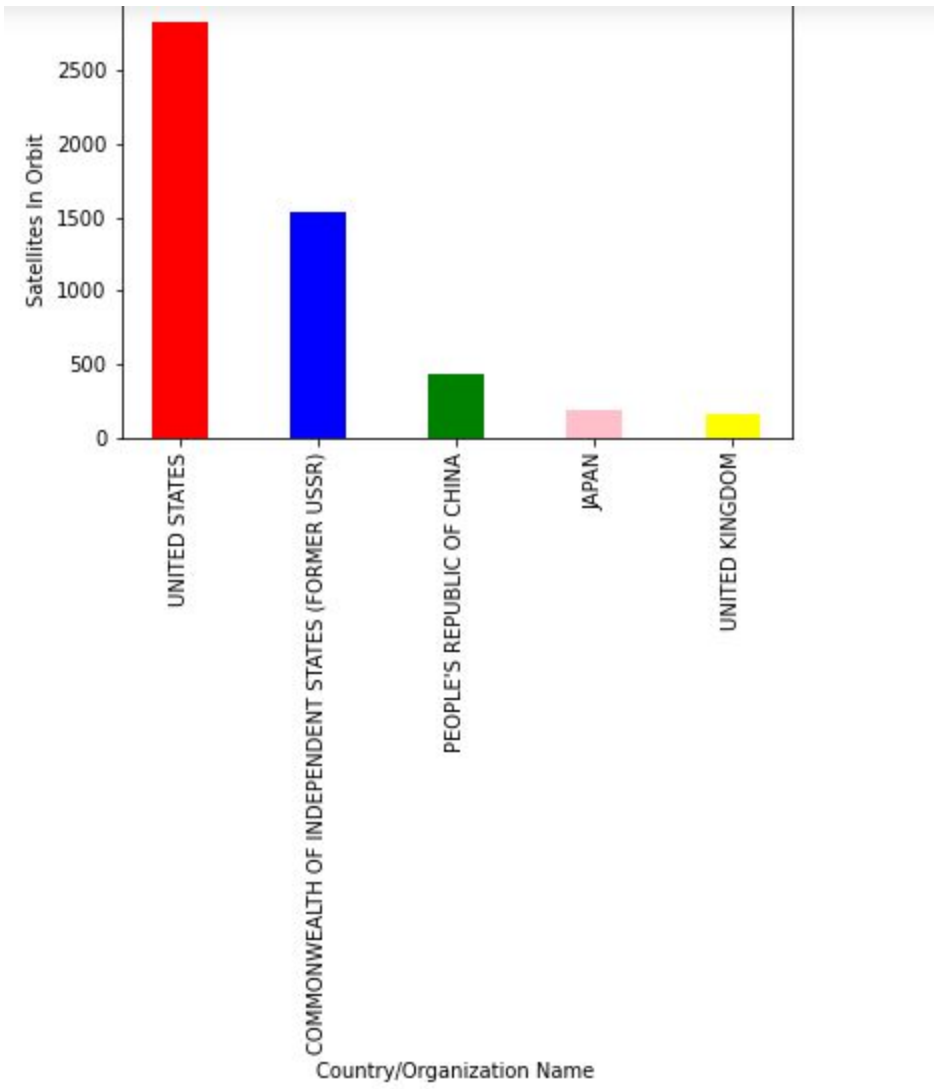
	Country/Organization Name	Satellites In Orbit
0	UNITED STATES	2825.0
1	COMMONWEALTH OF INDEPENDENT STATES (FORMER USSR)	1535.0
2	PEOPLE'S REPUBLIC OF CHINA	436.0
3	JAPAN	188.0
4	UNITED KINGDOM	154.0

Before using xticks(rotation = “vertical”)



And the plot displaying top 5 Country/Organization who has the most number of Satellites In Space

After using xticks(rotation = “vertical”) -



2) Activity -2 - Plot a bar graph for 20 Country/Organization who has the least number of Satellites In Space

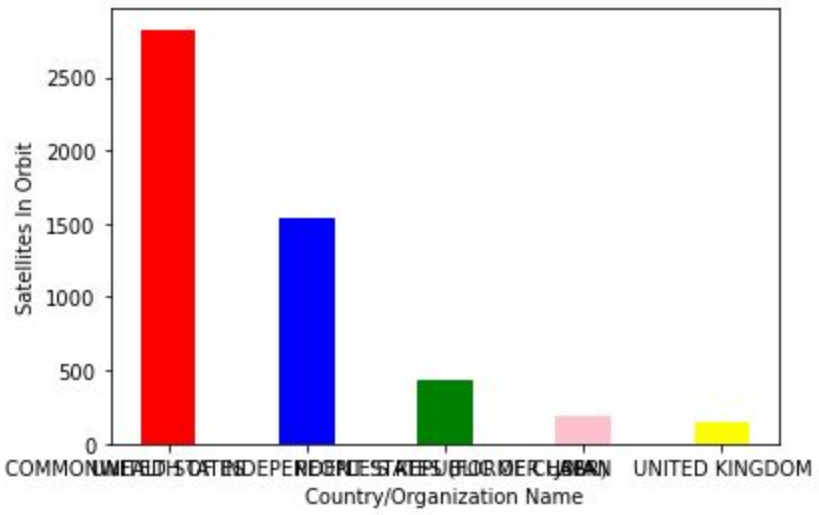
```
#Q - Plot a bar graph for 20 Country/Organization who has the least number of Satellites In Space
import  numpy as np
import pandas as pd
from matplotlib import pyplot as plt

dataframe = pd.read_csv('country_satellites.csv')
df = dataframe.dropna()

least_20 = df.tail(20)
print(least_20)

name = least_20['Country/Organization Name']
number = least_20['Satellites In Orbit']

plt.xlabel("Country/Organization Name")
plt.xticks(rotation='vertical')
plt.ylabel("Satellites In Orbit")


label = name
value = number
plt.bar(label, value,width=0.4, color=('red','blue','green','pink','yellow')) #bar-grap
```

Output:

As we had printed the least_20 variable which holds the last 20 rows of the dataset(table)

	Country/Organization Name	Satellites In Orbit
71	REPUBLIC OF SLOVENIA	2.0
72	SINGAPORE/TAIWAN	2.0
73	BANGLADESH	1.0
74	BOLIVIA	1.0
75	BULGARIA	1.0
76	DEMOCRATIC SOCIALIST REPUBLIC OF SRI LANKA	1.0
77	FEDERAL DEMOCRATIC REPUBLIC OF NEPAL	1.0
78	IRAQ	1.0
79	LAOS	1.0
80	LATVIA	1.0
81	NEW ICO	1.0
82	NEW ZEALAND	1.0
83	QATAR	1.0
84	REPUBLIC OF RWANDA	1.0
85	SEA LAUNCH	1.0
86	SLOVAKIA	1.0
87	TURKMENISTAN/MONACO	1.0
88	UNITED STATES/BRAZIL	1.0
89	URUGUAY	1.0

Before using xticks(rotation = “vertical”) -



And the plot displaying 20 countries/organization who has the least number of Satellites In Space
After using xticks(rotation = “vertical”)

```
Out[5]: <BarContainer object of 19 artists>
```