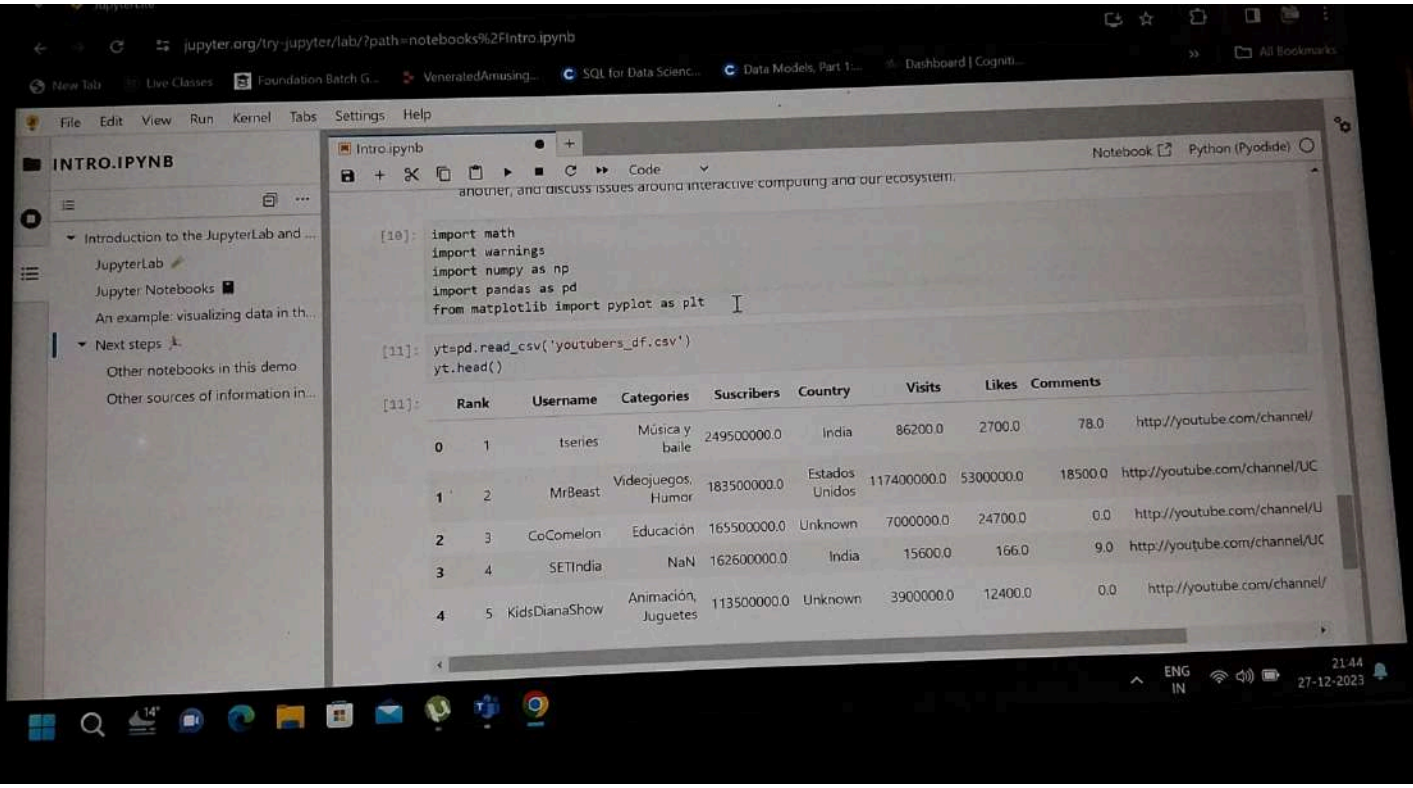

jupyter.org/try-jupyter/lab/?path=notebooks%2FIntro.ipynb
File Edit View Run Kernel Tabs Settings Help
INTRO.IPYNB
Intro.ipynb
Code
Notebook
Python (Pyodide)
Introduction to the JupyterLab and ...
JupyterLab
Jupyter Notebooks
An example: visualizing data in th...
Next steps
Other notebooks in this demo
Other sources of information in...
another, and discuss issues around interactive computing and our ecosystem.
[10]:
import math
import warnings
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
[11]:
yt=pd.read_csv('youtubers_df.csv')
yt.head()
[11]:
Rank Username Categories Suscribers Country Visits Likes Comments
0 1 tseries Música y baile 249500000.0 India 86200.0 2700.0 78.0 http://youtube.com/channel/
1 2 MrBeast Videojuegos, Humor 183500000.0 Estados Unidos 117400000.0 5300000.0 18500.0 http://youtube.com/channel/UC
2 3 CoComelon Educación 165500000.0 Unknown 7000000.0 24700.0 0.0 http://youtube.com/channel/U
3 4 SETIndia NaN 162600000.0 India 15600.0 166.0 9.0 http://youtube.com/channel/UC
4 5 KidsDianaShow Animación, Juguetes 113500000.0 Unknown 3900000.0 12400.0 0.0 http://youtube.com/channel/
ENG IN
21:44
27-12-2023

er/lab/?path=notebooks%2FIntro.ipynb

Batch G... VeneratedAmusing... SQL for Data Scienc... Data Models, Part 1:... Dashboard | Cogniti...

Settings Help

Intro.ipynb

Code

Notebook Python (Pyodide)

[10]:
```
import math
import warnings
import numpy as np
import pandas as pd
from matplotlib import pyplot as plt
```

[11]:
```
yt=pd.read_csv('youtubers_df.csv')
yt.head()
```

[11]:

	Rank	Username	Categories	Suscribers	Country	Visits	Likes	Comments	
1	1	tseries	Música y baile	249500000.0	India	86200.0	2700.0	78.0	http://youtube.com/channel/UCq
2	2	MrBeast	Videojuegos, Humor	183500000.0	Estados Unidos	117400000.0	5300000.0	18500.0	http://youtube.com/channel/UCX6O
3	3	CoComelon	Educación	165500000.0	Unknown	7000000.0	24700.0	0.0	http://youtube.com/channel/UCbC
4	4	SETIndia	NaN	162600000.0	India	15600.0	166.0	9.0	http://youtube.com/channel/UCpEh
5	5	KidsDianaShow	Animación, Juguetes	113500000.0	Unknown	3900000.0	12400.0	0.0	http://youtube.com/channel/UCk

ENG IN 22:23 27-12-2023

Intro.ipynb

Code

Notebook

Python (Pyodide)

[12]:
```
yt.columns
```

[12]:
```
Index(['Rank', 'Username', 'Categories', 'Suscribers', 'Country', 'Visits',
       'Likes', 'Comments', 'Links'],
      dtype='object')
```

[14]:
```
yt.isnull().sum()
```

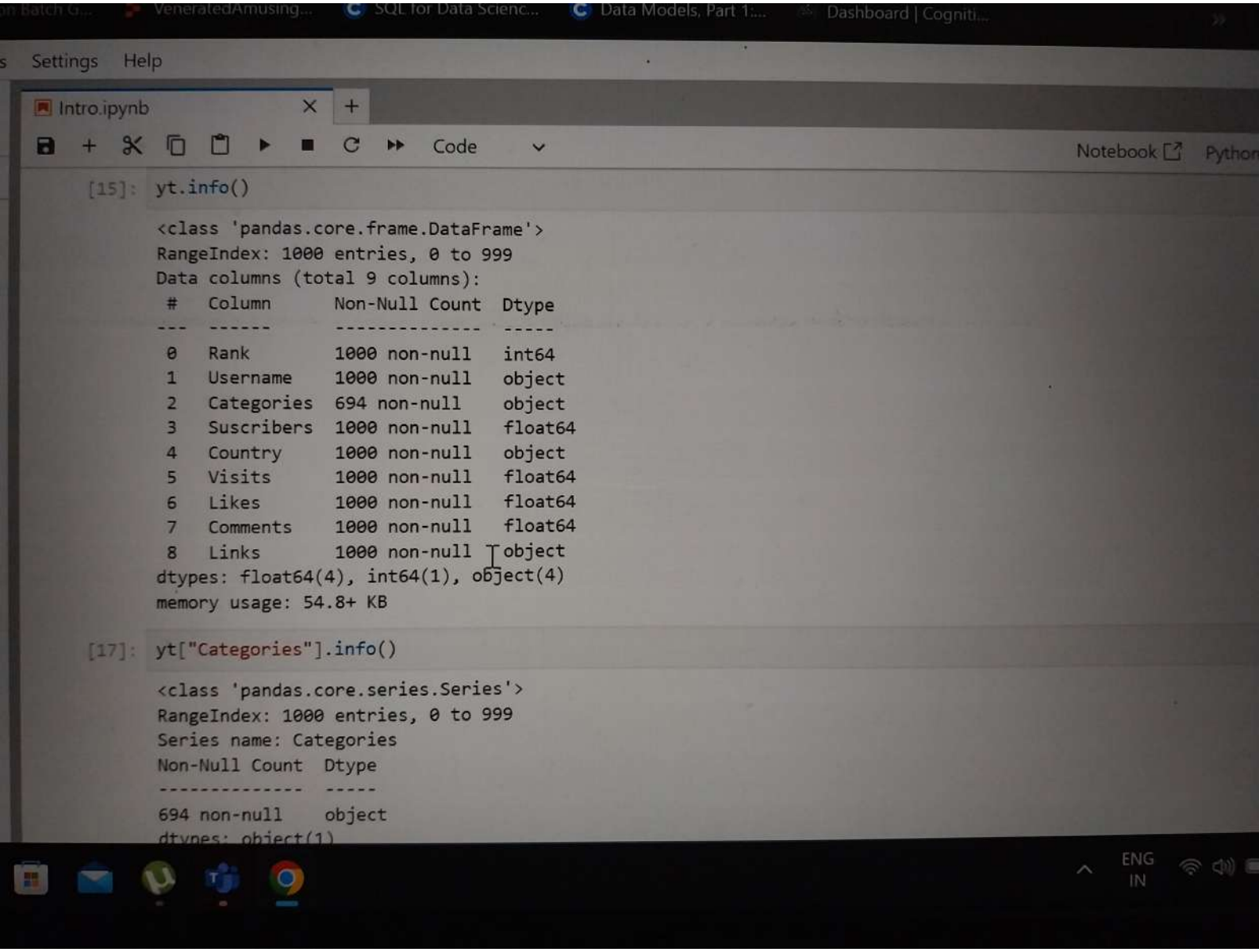
[14]:
```
Rank            0
Username        0
Categories    306
Suscribers      0
Country         0
Visits          0
Likes           0
Comments        0
Links           0
dtype: int64
```

[15]:
```
yt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Rank        1000 non-null   int64
```

ENG
IN

27-12-


Settings
Help
Intro.ipynb
Code
Notebook
[15]: yt.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 9 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 Rank 1000 non-null int64
1 Username 1000 non-null object
2 Categories 694 non-null object
3 Suscribers 1000 non-null float64
4 Country 1000 non-null object
5 Visits 1000 non-null float64
6 Likes 1000 non-null float64
7 Comments 1000 non-null float64
8 Links 1000 non-null object
dtypes: float64(4), int64(1), object(4)
memory usage: 54.8+ KB
[17]: yt["Categories"].info()
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
Series name: Categories
Non-Null Count Dtype
-------------- -----
694 non-null object
dtypes: object(1)
ENG
IN

Intro.ipynb

Code

Notebook

[17]:
```
yt["Categories"].info()
```

```
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
Series name: Categories
Non-Null Count  Dtype
--------------  -----
694 non-null    object
dtypes: object(1)
memory usage: 4.0+ KB
```

[20]:
```
yt["Country"].info()
```

```
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
Series name: Country
Non-Null Count  Dtype
--------------  -----
1000 non-null   object
dtypes: object(1)
memory usage: 4.0+ KB
```

[21]:
```
yt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
```

ENG
IN

Intro.ipynb

Code

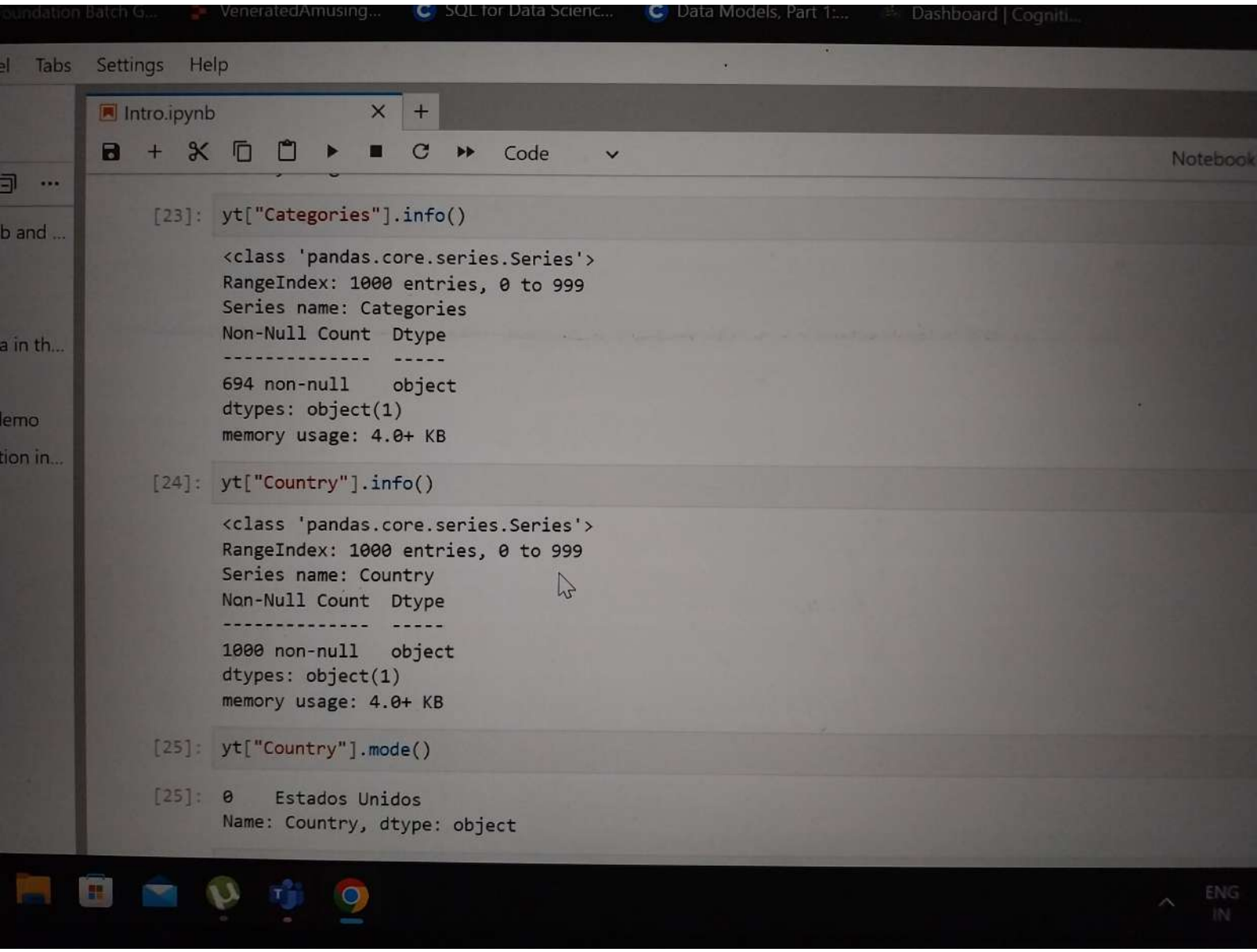
```
--------------  -----
1000 non-null   object
dtypes: object(1)
memory usage: 4.0+ KB
```

[21]:
```
yt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Rank        1000 non-null   int64
 1   Username    1000 non-null   object
 2   Categories  694 non-null    object
 3   Suscribers  1000 non-null   float64
 4   Country     1000 non-null   object
 5   Visits      1000 non-null   float64
 6   Likes       1000 non-null   float64
 7   Comments    1000 non-null   float64
 8   Links       1000 non-null   object
dtypes: float64(4), int64(1), object(4)
memory usage: 54.8+ KB
```

[23]:
```
yt["Categories"].info()
```

```
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
```


Tabs
Settings
Help
Intro.ipynb
Code
Notebook
[23]: yt["Categories"].info()
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
Series name: Categories
Non-Null Count Dtype
694 non-null object
dtypes: object(1)
memory usage: 4.0+ KB
[24]: yt["Country"].info()
<class 'pandas.core.series.Series'>
RangeIndex: 1000 entries, 0 to 999
Series name: Country
Non-Null Count Dtype
1000 non-null object
dtypes: object(1)
memory usage: 4.0+ KB
[25]: yt["Country"].mode()
[25]: 0 Estados Unidos
Name: Country, dtype: object

rnel Tabs Settings Help

Intro.ipynb

Code

[25]:
```
yt["Country"].mode()
```

[25]:
```
0    Estados Unidos
Name: Country, dtype: object
```

[26]:
```
country_youtubers_count=yt["Country"].value_counts().head(10)
country_youtubers_count
```

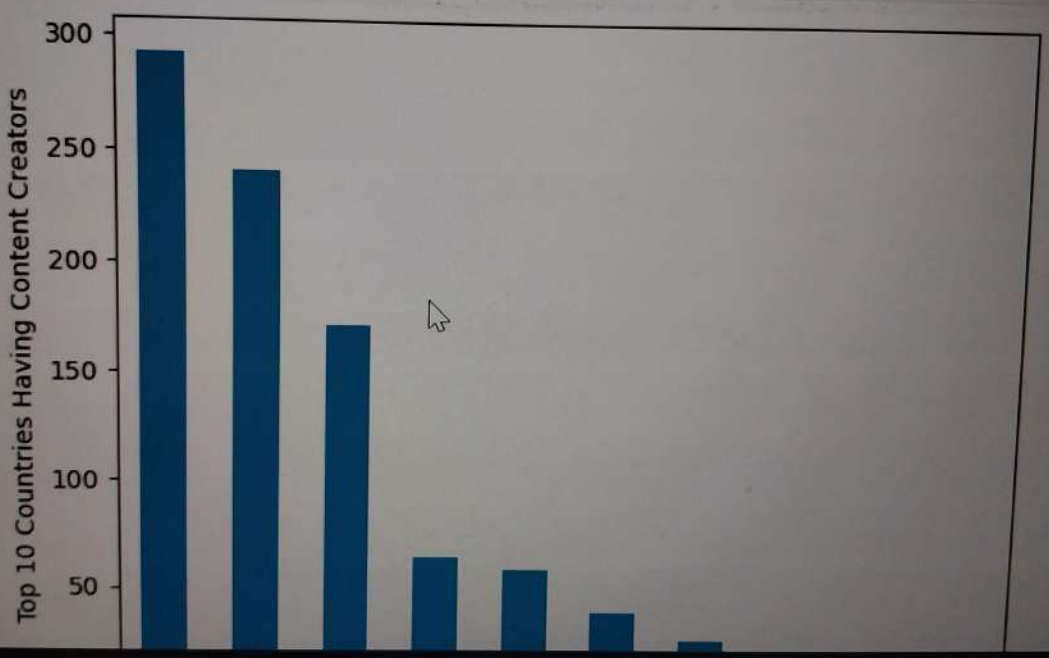
[26]:
```
Estados Unidos    293
India             241
Unknown           171
Brasil             64
México             58
Indonesia          38
Rusia              25
Tailandia          18
Colombia           16
Filipinas          13
Name: Country, dtype: int64
```

[28]:
```
import matplotlib.pyplot as plt
country_youtubers_count.plot(kind="bar")
plt.xlabel("Country")
plt.ylabel("Top 10 Countries Having Content Creators")
plt.show()
```

Intro.ipynb

Code

[28]:
```
import matplotlib.pyplot as plt
country_youtubers_count.plot(kind="bar")
plt.xlabel("Country")
plt.ylabel("Top 10 Countries Having Content Creators")
plt.show()
```

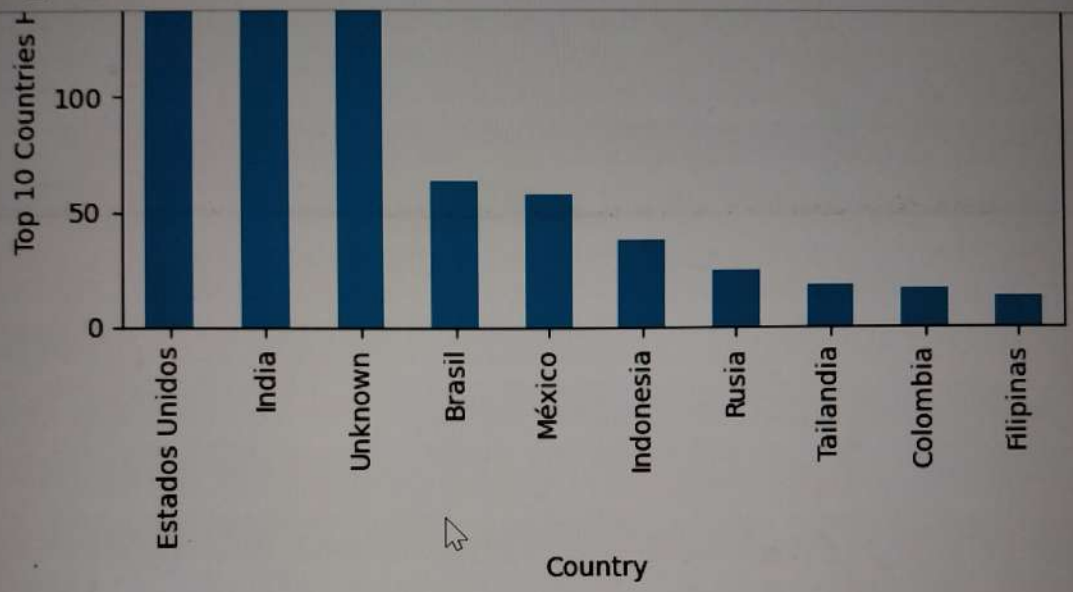


Tabs Settings Help

Intro.ipynb

Code

Notebook



[ ]:

[29]:
```
total_categories=yt["Categories"].value_counts()
total_categories.head(15)
```

[29]:
```
Música y baile                160
Películas, Animación           61
```

ENG
IN

Tabs Settings Help

Intro.ipynb

Code Notebook

[29]:
```
total_categories=yt["Categories"].value_counts()
total_categories.head(15)
```

[29]:
```
Música y baile                160
Películas, Animación           61
Música y baile, Películas      41
Vlogs diarios                  37
Noticias y Política            36
Películas, Humor               34
Animación, Videojuegos         34
Animación, Juguetes            29
Animación, Humor               27
Películas                      24
Educación                      24
Animación                      22
Videojuegos                    19
Videojuegos, Humor             17
Música y baile, Animación      16
Name: Categories, dtype: int64
```

[31]:
```
yt.dropna(subset='Categories',inplace=True)
yt.describe()
```

[31]:

	Rank	Suscribers	Visits	Likes	Comments
count	694.000000	6.940000e+02	6.940000e+02	6.940000e+02	694.000000

ENG IN

Foundation Batch G... VeneratedAmusing... SQL for Data Scienc... Data Models, Part 1:... Dashboard | Cogniti...

ernel Tabs Settings Help

Intro.ipynb

Code Notebook

[31]:
```
yt.dropna(subset='Categories',inplace=True)
yt.describe()
```

[31]:

	Rank	Suscribers	Visits	Likes	Comments
count	694.000000	6.940000e+02	6.940000e+02	6.940000e+02	694.000000
mean	495.298271	2.241556e+07	1.210730e+06	5.347360e+04	1558.793948
std	289.222212	1.824123e+07	6.038274e+06	2.979711e+05	7967.470234
min	1.000000	1.170000e+07	0.000000e+00	0.000000e+00	0.000000
25%	244.250000	1.380000e+07	3.692500e+04	5.685000e+02	2.000000
50%	492.500000	1.680000e+07	1.587000e+05	3.550000e+03	78.000000
75%	746.750000	2.390000e+07	1.339000e+05	2.377500e+04	499.750000
max	1000.000000	2.495000e+08	1.174000e+08	5.300000e+06	154000.000000

[32]:
```
yt.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 694 entries, 0 to 999
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
```

ENG IN

Foundation Batch G... VeneratedAmusing... SQL for Data Scienc... Data Models, Part 1:... Dashboard | Cogniti...

Kernel Tabs Settings Help

Intro.ipynb

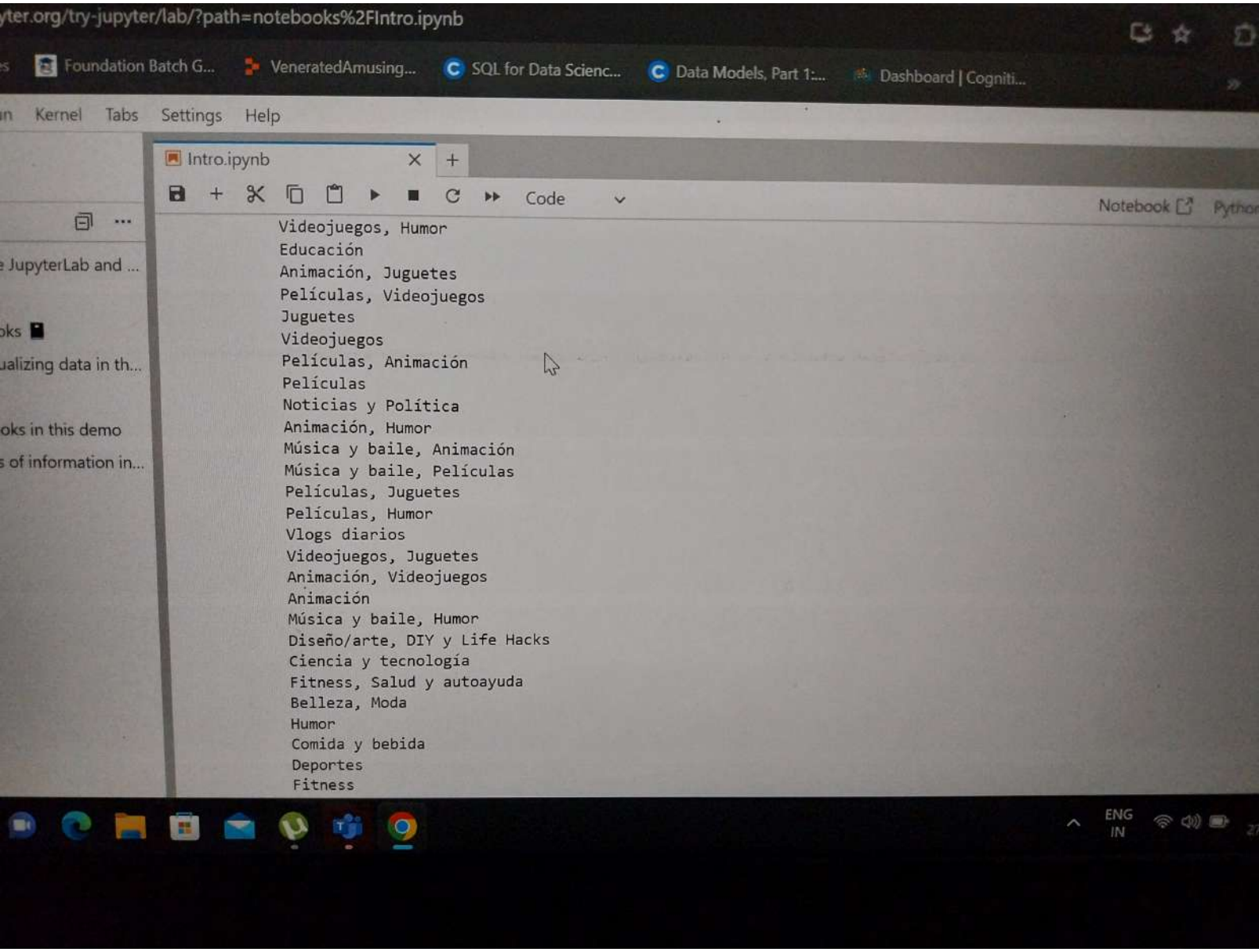
Code Notebook

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 694 entries, 0 to 999
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   Rank        694 non-null    int64
 1   Username    694 non-null    object
 2   Categories  694 non-null    object
 3   Suscribers  694 non-null    float64
 4   Country     694 non-null    object
 5   Visits      694 non-null    float64
 6   Likes       694 non-null    float64
 7   Comments    694 non-null    float64
 8   Links       694 non-null    object
dtypes: float64(4), int64(1), object(4)
memory usage: 43.4+ KB
```

[35]:
```
categories=yt['Categories'].unique()
print("Categories in the dataset:")
for categorie in categories:
    print(categorie)
```

```
Categories in the dataset:
Música y baile
Videojuegos, Humor
Educación
Animación, Juguetes
```

ENG IN


yter.org/try-jupyter/lab/?path=notebooks%2FIntro.ipynb
Foundation Batch G...
VeneratedAmusing...
SQL for Data Scienc...
Data Models, Part 1:...
Dashboard | Cogniti...
Kernel
Tabs
Settings
Help
Intro.ipynb
Code
Notebook
Videojuegos, Humor
Educación
Animación, Juguetes
Películas, Videojuegos
Juguetes
Videojuegos
Películas, Animación
Películas
Noticias y Política
Animación, Humor
Música y baile, Animación
Música y baile, Películas
Películas, Juguetes
Películas, Humor
Vlogs diarios
Videojuegos, Juguetes
Animación, Videojuegos
Animación
Música y baile, Humor
Diseño/arte, DIY y Life Hacks
Ciencia y tecnología
Fitness, Salud y autoayuda
Belleza, Moda
Humor
Comida y bebida
Deportes
Fitness
ENG
IN

jupyter.org/try-jupyter/lab/?path=notebooks%2FIntro.ipynb

Foundation Batch G... | VeneratedAmusing... | SQL for Data Scienc... | Data Models, Part 1:... | Dashboard | Cogniti...

New Run Kernel Tabs Settings Help

Intro.ipynb

Code

Notebook

[41]:
```
suscribers=yt.head(8)
plt.figure(figsize=(10,6))
plt.bar(suscribers['Username'],suscribers['Suscribers'])
plt.xlabel="suscribers"
plt.ylabel="Youtuber"
plt.title="Top 8 Youtube Suscribers"
plt.show()
```

Foundation Batch G... VeneratedAmusing... SQL for Data Scienc... Data Models, Part 1:... Dashboard | Cogniti...

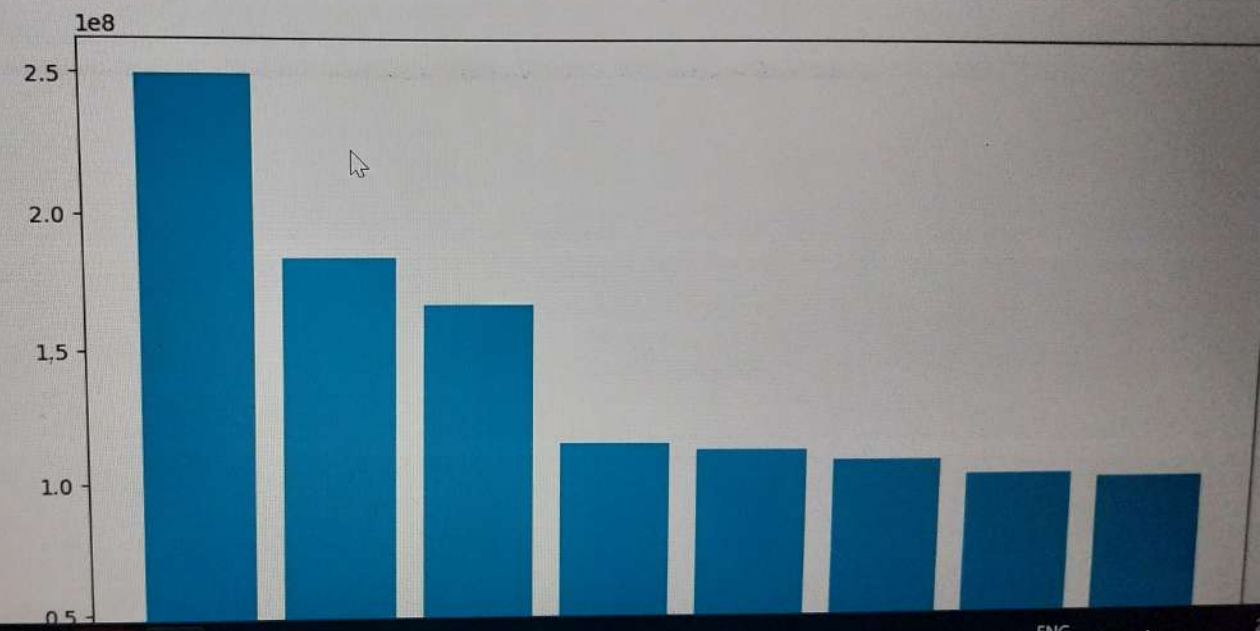
Kernel Tabs Settings Help

Intro.ipynb

Code

Notebook Python (Py

```
plt.title="Top 8 Youtube Suscribers"
plt.show()
```



ENG IN

youtube-dataset · GitHub Topic

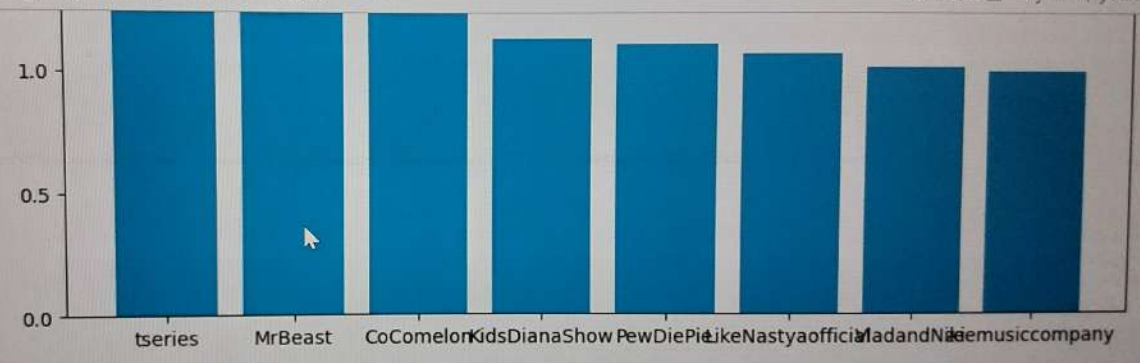
jupyter.org/try-jupyter/lab/?path=notebooks%2FIntro.ipynb

Classes | Foundation Batch G... | VeneratedAmusing... | SQL for Data Scienc... | Data Models, Part 1:... | Dashboard | Cogniti... | All Bookmarks

Run Kernel Tabs Settings Help

Intro.ipynb

Code | Notebook | Python (Pyodide)



[ ]:

ENG IN 22:26 27-12-2023