```
from google.colab import files

uploaded = files.upload()
```

Choose Files No file chosen Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving most_subscribed_youtube_channels.csv to most_subscribed_youtube_channels.csv

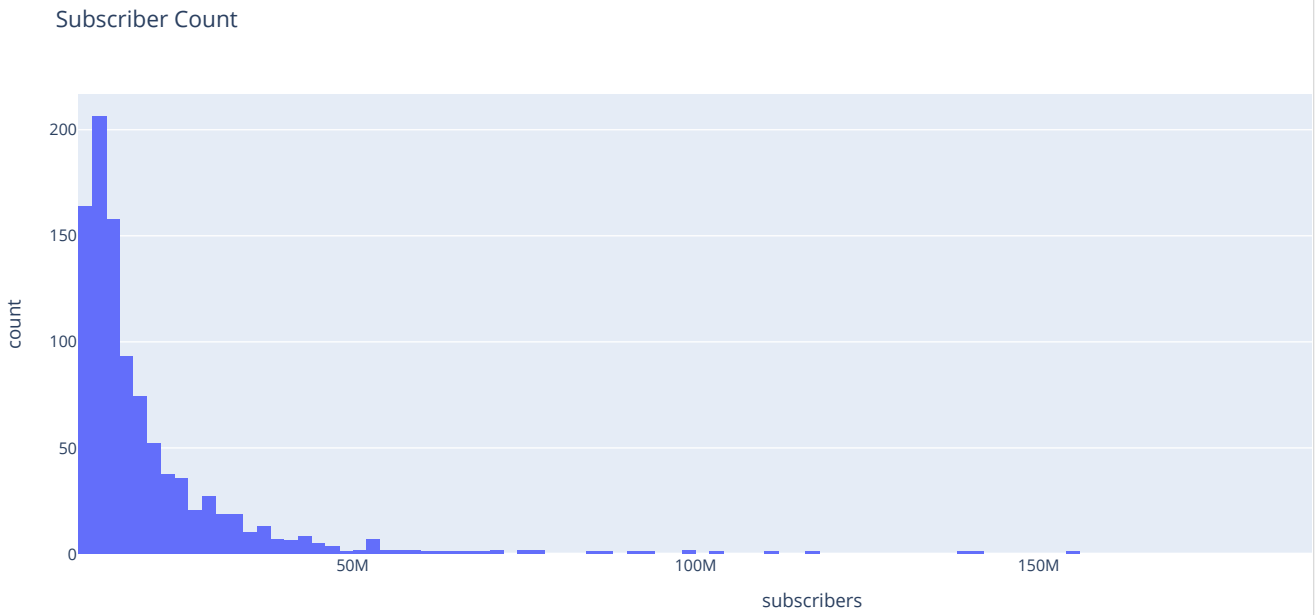
```
import pandas as pd
import io

# Read CSV and automatically handle commas in numbers
df = pd.read_csv(io.BytesIO(uploaded['most_subscribed_youtube_channels.csv']), thousands=',')

display(df)
```

	rank	Youtuber	subscribers	video views	video count	category	started
0	1	T-Series	222000000	198459090822	17317	Music	2006
1	2	YouTube Movies	154000000	0	0	Film & Animation	2015
2	3	Cocomelon - Nursery Rhymes	140000000	135481339848	786	Education	2006
3	4	SET India	139000000	125764252686	91271	Shows	2006
4	5	Music	116000000	0	0	NaN	2013
...	...	...	...	...	...	...	...
995	996	JP Plays	10900000	4609300218	3528	Gaming	2014
996	997	TrapMusicHDTV	10900000	4070521973	690	Music	2013
997	998	Games EduUu	10900000	3093784767	1006	Gaming	2011
998	999	Hueva	10900000	3040301750	831	Gaming	2012
999	1000	Dobre Brothers	10900000	2808411693	590	People & Blogs	2017

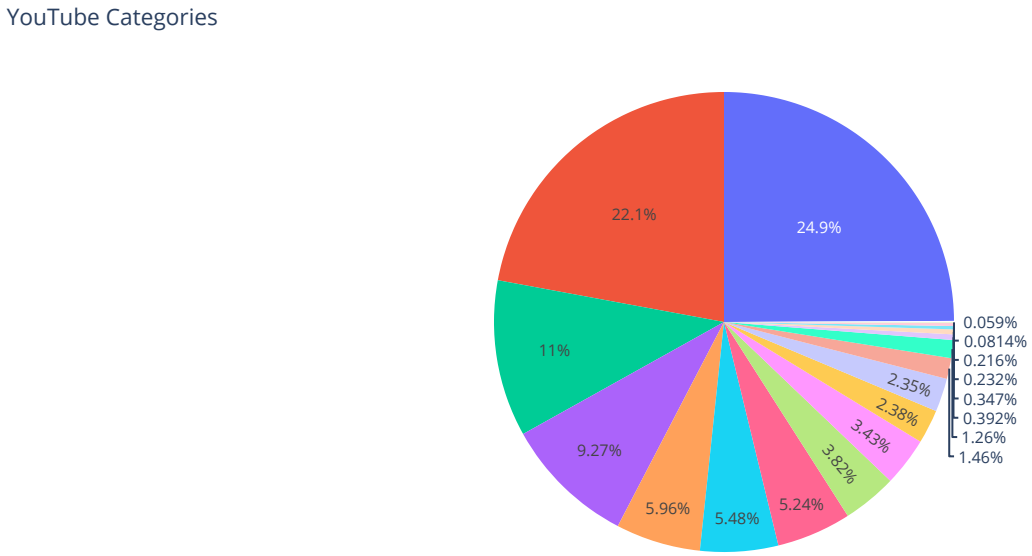
1000 rows × 7 columns

```
CURRENT_YEAR = 2022
df['Age'] = CURRENT_YEAR - df['started']

# Avoid division by zero by replacing 0 video counts with 1 before calculation
df['video count'] = df['video count'].replace(0, 1)
df['Avg_Views_Per_Video'] = df['video views'] / df['video count']
```

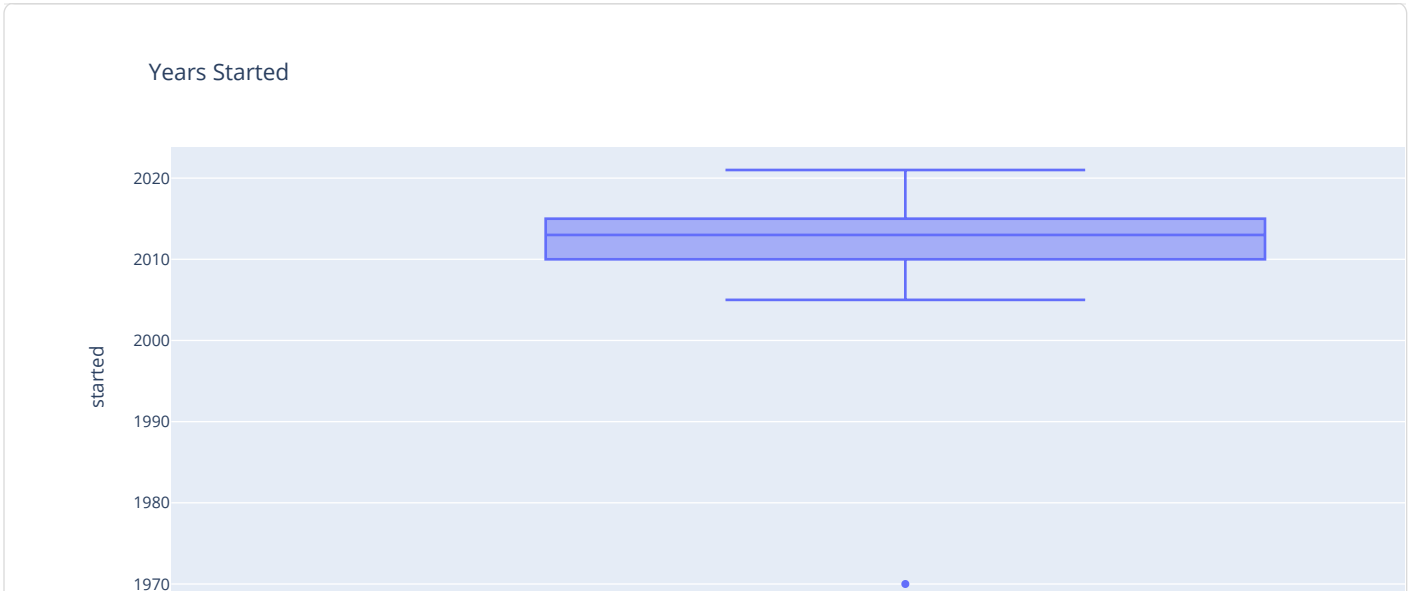
```
import plotly.express as px
fig = px.histogram(df, x='subscribers', title='Subscriber Count')
fig.show()
```




```
import plotly.express as px

fig = px.pie(df, values='subscribers', names='category', title='YouTube Categories')
fig.show()
```




```
import plotly.express as px

fig = px.box(df, y='started', title='Years Started')
fig.show()
```


Years Started
started
2020
2010
2000
1990
1980
1970