Visualization in Python

Line Plots for Video Games Sales

Presented by Shreshtha Barthwal

Line Chart and its significance

What is a line plot?

It is a plot that displays information as a series of data points connected by straight lines.

When to use a line plot?

- Line plots are useful for visualizing trends and changes over time, making them a popular choice for time-series data, such as changes in stock prices, website traffic, or temperature fluctuations.
- You can also use a line plot to show relationships between two variables.
- They can also be used to compare multiple data series on one chart.
- Line plots can also effectively highlight sudden changes or anomalies in data

There are various Python libraries which can be used for Line Plots like Pandas, Matplotlib, Seaborn, Plotly ,Bokeh ,Panel etc

Glance through the data

- Data presents the sales figures of video games in multiple region.
- These numbers are pivoted by three categories –
  - Platforms
  - Publishers
  - Genres
- This data has following attributes –

Rank - Ranking of overall sales	Name - The games name	Platform - Platform of the games release	Year - Year of the game's release
Genre - Genre of the game	Publisher - Publisher of the game	NA_Sales - Sales in North America	EU_Sales - Sales in Europe
JP_Sales - Sales in Japan	Other_Sales - Sales in the rest of the world	Global_Sales - Total worldwide sales	

Analysis Goals

1. What are the global sales looking like over time?.
2. What platforms(consoles/pc) are responsible for the bulk of the games being played?.
3. Who are the top publishers who have hold on different market.
4. And finally, what categories(Genre's) of games are people playing?
5. Exploring different libraries in Python to see how we can show the same data trends in different ways.

Sample Data

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37

Data Cleaning

```
df.isna().sum()
```

```
Rank            0
Name            0
Platform        0
Year          271
Genre           0
Publisher      58
NA_Sales        0
EU_Sales        0
JP_Sales        0
Other_Sales     0
Global_Sales    0
dtype: int64
```

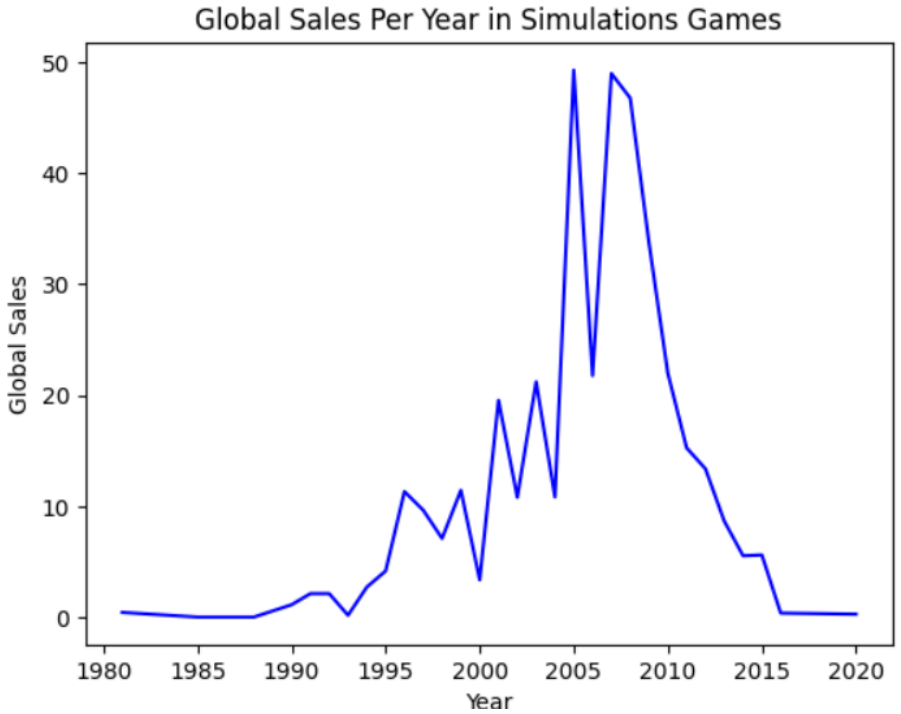
```
df1 = df.dropna(axis=0, subset=['Publisher'])
```

```
df2 = df1.dropna(axis=0, subset=['Year'])
```

```
# Converting float year type to int
df_cleansed['Year'] = df_cleansed['Year'].astype(int)
df_cleansed['Year'].dtype
```

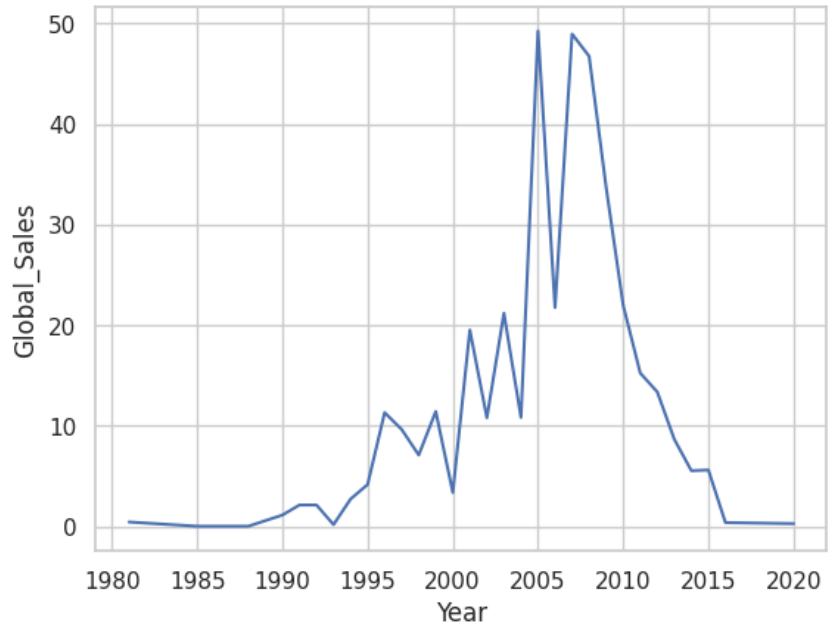
Global Sales Per Year in Simulations Games

```
plt.plot(df_SimulationSales['Year'], df_SimulationSales['Global_Sales'], color='b')
plt.title('Global Sales Per Year in Simulations Games')
plt.xlabel('Year')
plt.ylabel('Global Sales')
plt.show()
```



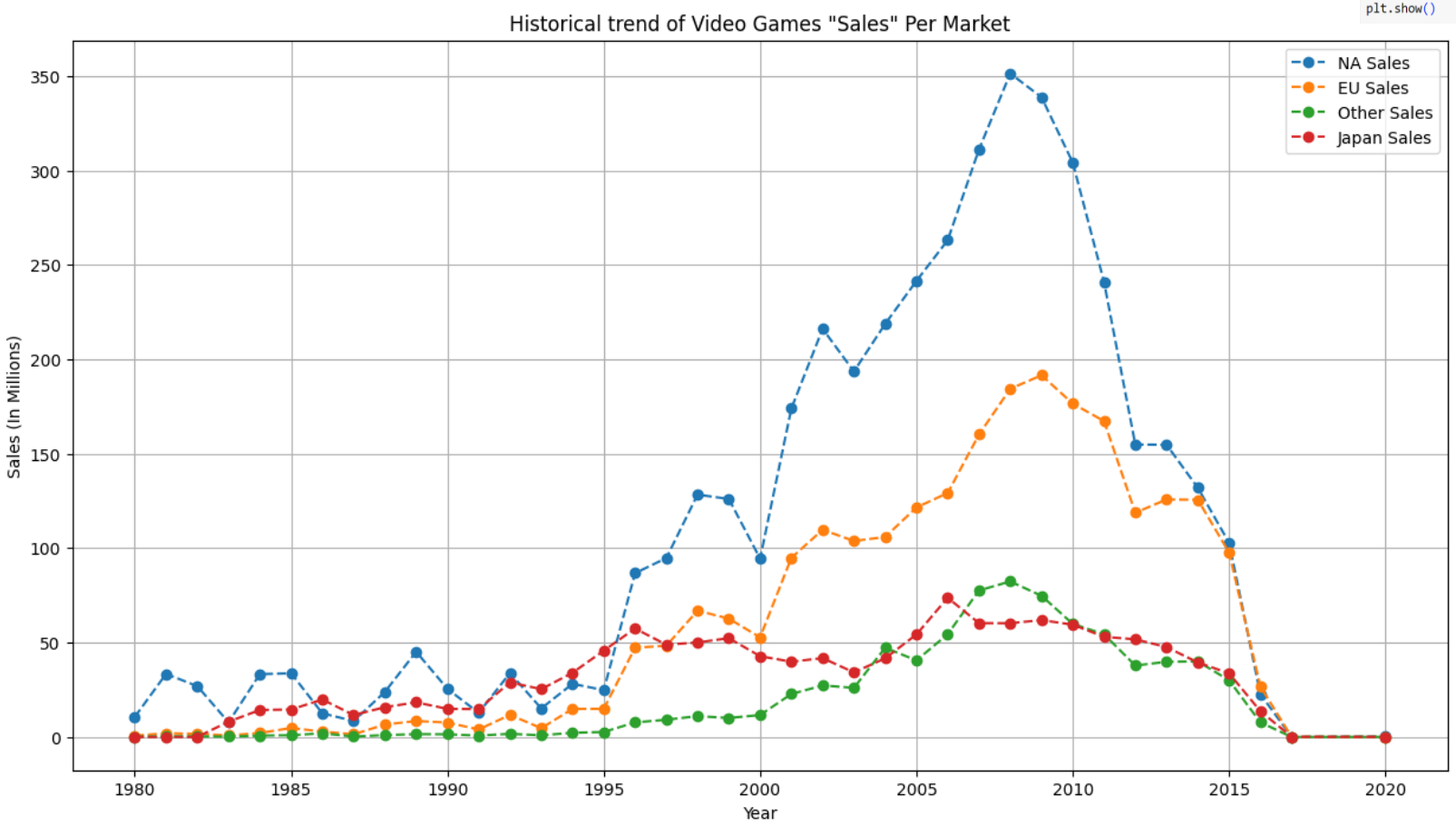

```
import seaborn as sns

sns.lineplot(x='Year', y='Global_Sales', data=df_SimulationSales)
```



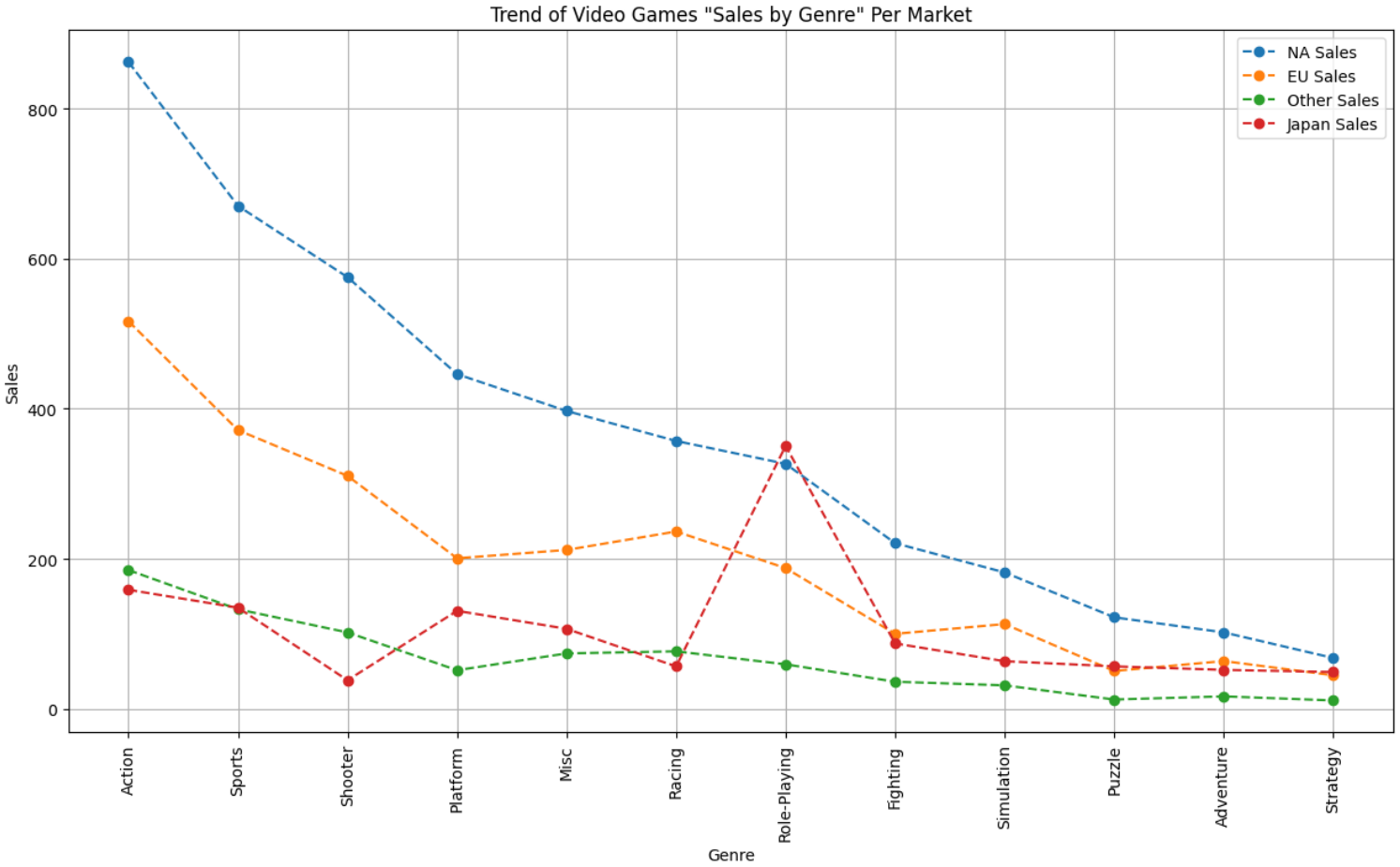
Video Games Sales Per Market

```
plt.figure(figsize=(15, 8))
plt.plot(df_group_year['Year'], df_group_year['NA_Sales'], label='NA Sales', marker='o', linestyle='dashed')
plt.plot(df_group_year['Year'], df_group_year['EU_Sales'], label = 'EU Sales', marker='o', linestyle='dashed')
plt.plot(df_group_year['Year'], df_group_year['Other_Sales'], label = 'Other Sales', marker='o', linestyle='dashed')
plt.plot(df_group_year['Year'], df_group_year['JP_Sales'], label = 'Japan Sales', marker='o', linestyle='dashed')
plt.title('Historical trend of Video Games "Sales" Per Market')
plt.legend()
plt.xlabel('Year')
plt.ylabel('Sales')
plt.grid()
plt.show()
```



The graph shows the trend of video games sales in different market with North America being a big market of video games.
The video game sales have boomed over b/w 1995 - 2007 after which there has been decrease till 2016. We do not have enough data to analyze after 2017-2020.

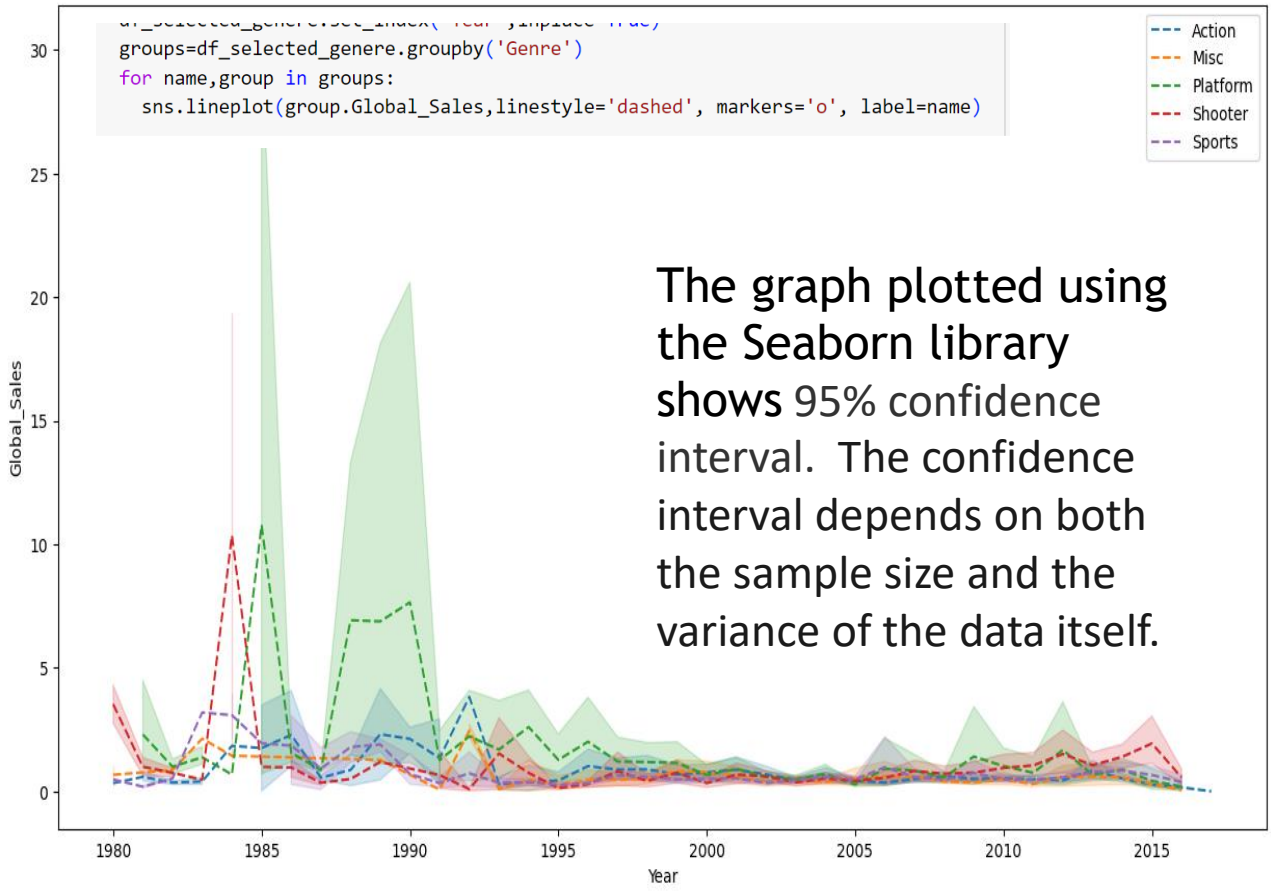
Video Games Sales By Genre Per Market



The graph shows the top 10 categories of video games with Action and sports being the most popular types.
An interesting fact to note is popularity of role playing games in the Japan market.

Trend in Global Sales of Top 5 Products using Seaborn & Matplotlib libraries

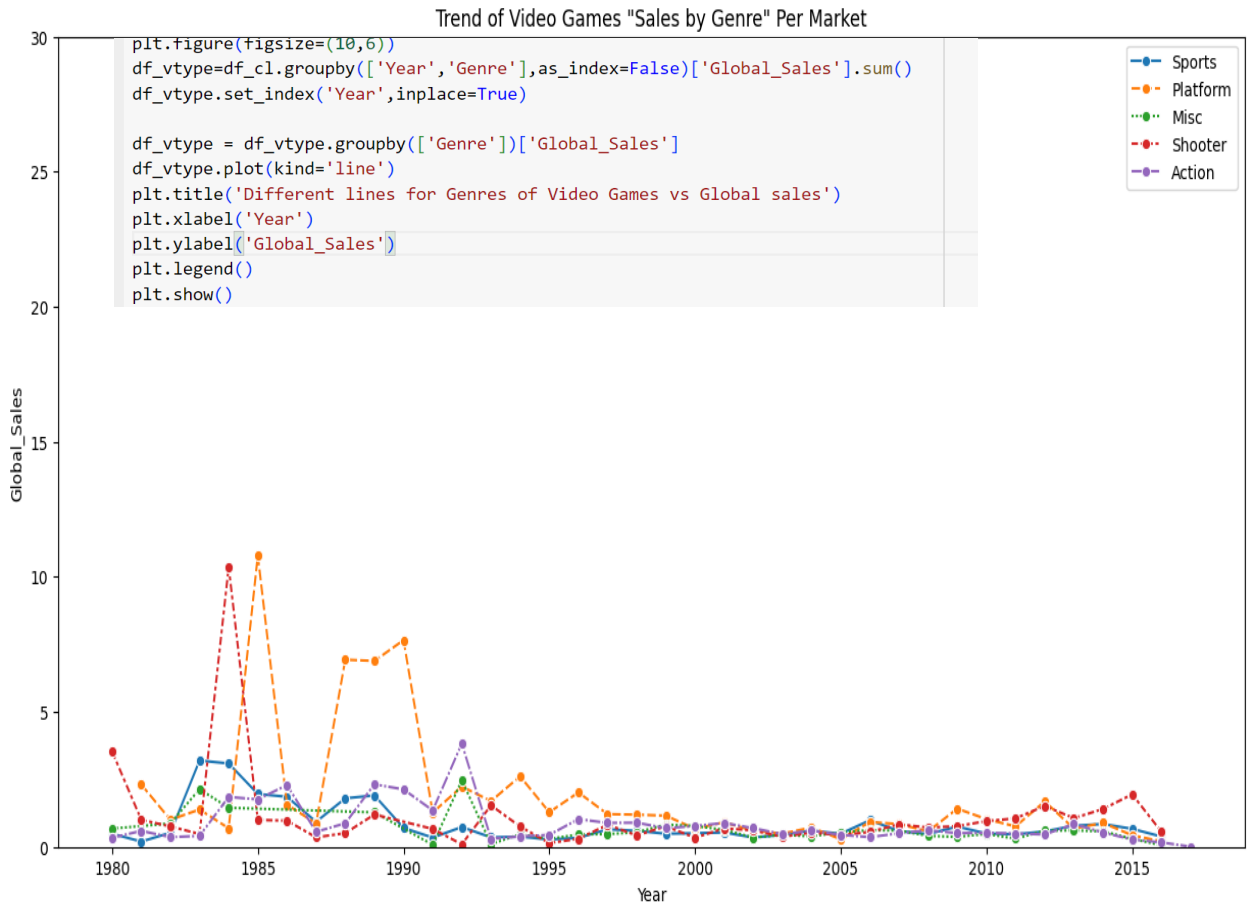
```
groups=df_selected_genere.groupby('Genre')
for name,group in groups:
  sns.lineplot(group.Global_Sales,linestyle='dashed', markers='o', label=name)
```



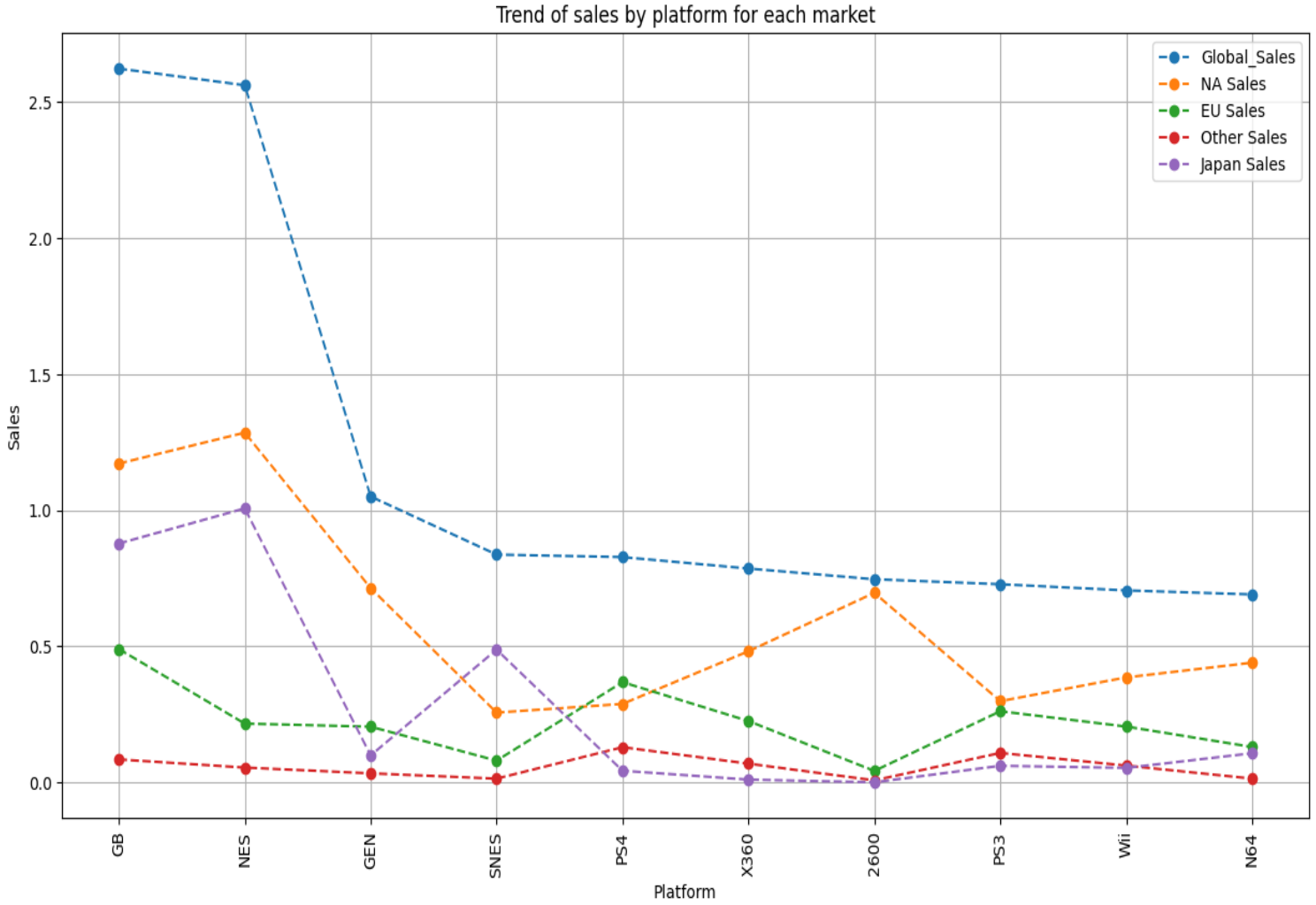
The graph plotted using the Seaborn library shows 95% confidence interval. The confidence interval depends on both the sample size and the variance of the data itself.

```
plt.figure(figsize=(10,6))
df_vtype=df_cl.groupby(['Year','Genre'],as_index=False)['Global_Sales'].sum()
df_vtype.set_index('Year',inplace=True)

df_vtype = df_vtype.groupby(['Genre'])['Global_Sales']
df_vtype.plot(kind='line')
plt.title('Different lines for Genres of Video Games vs Global sales')
plt.xlabel('Year')
plt.ylabel('Global_Sales')
plt.legend()
plt.show()
```

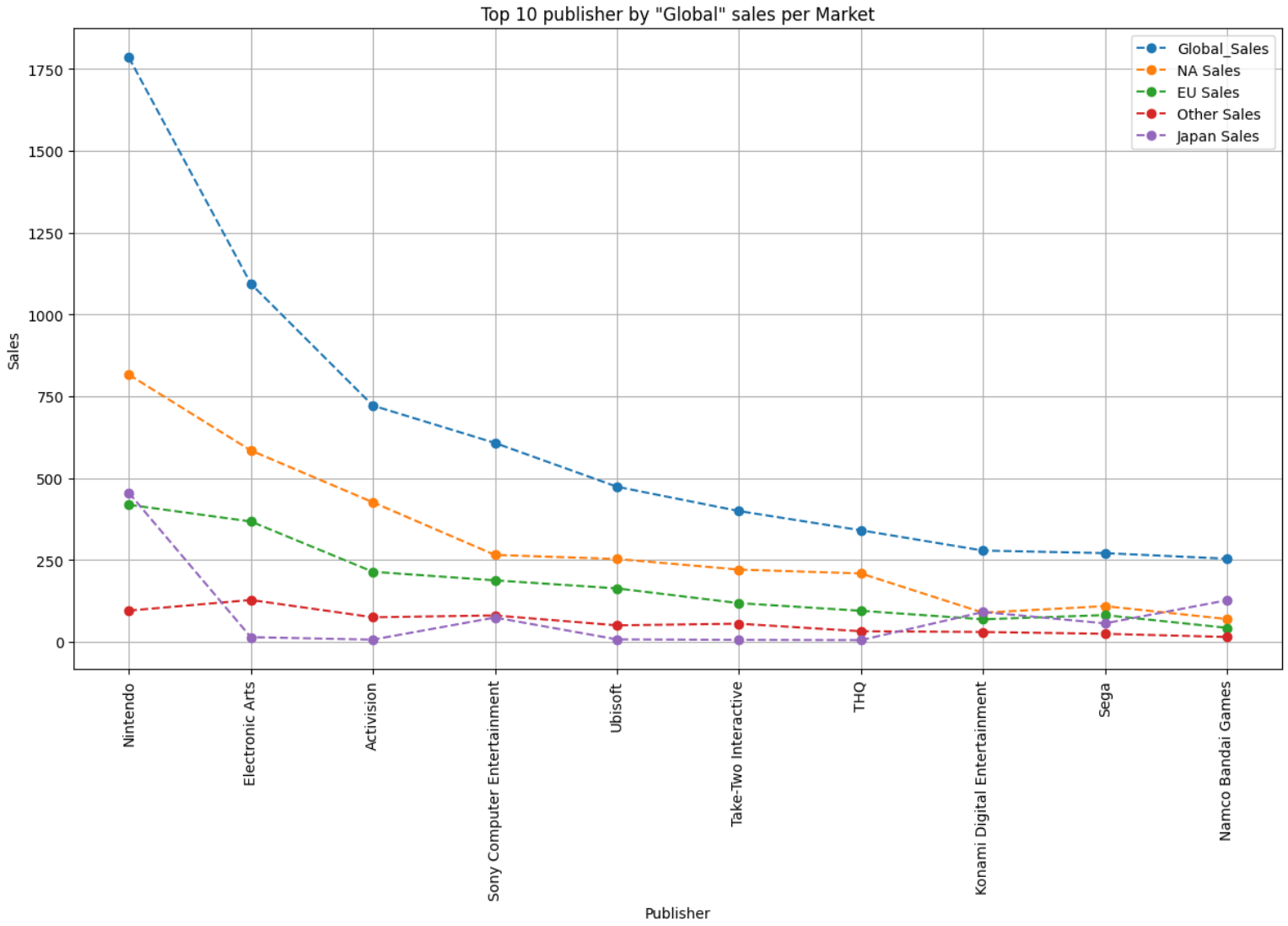


Trend in Sales of Top 10 Platforms



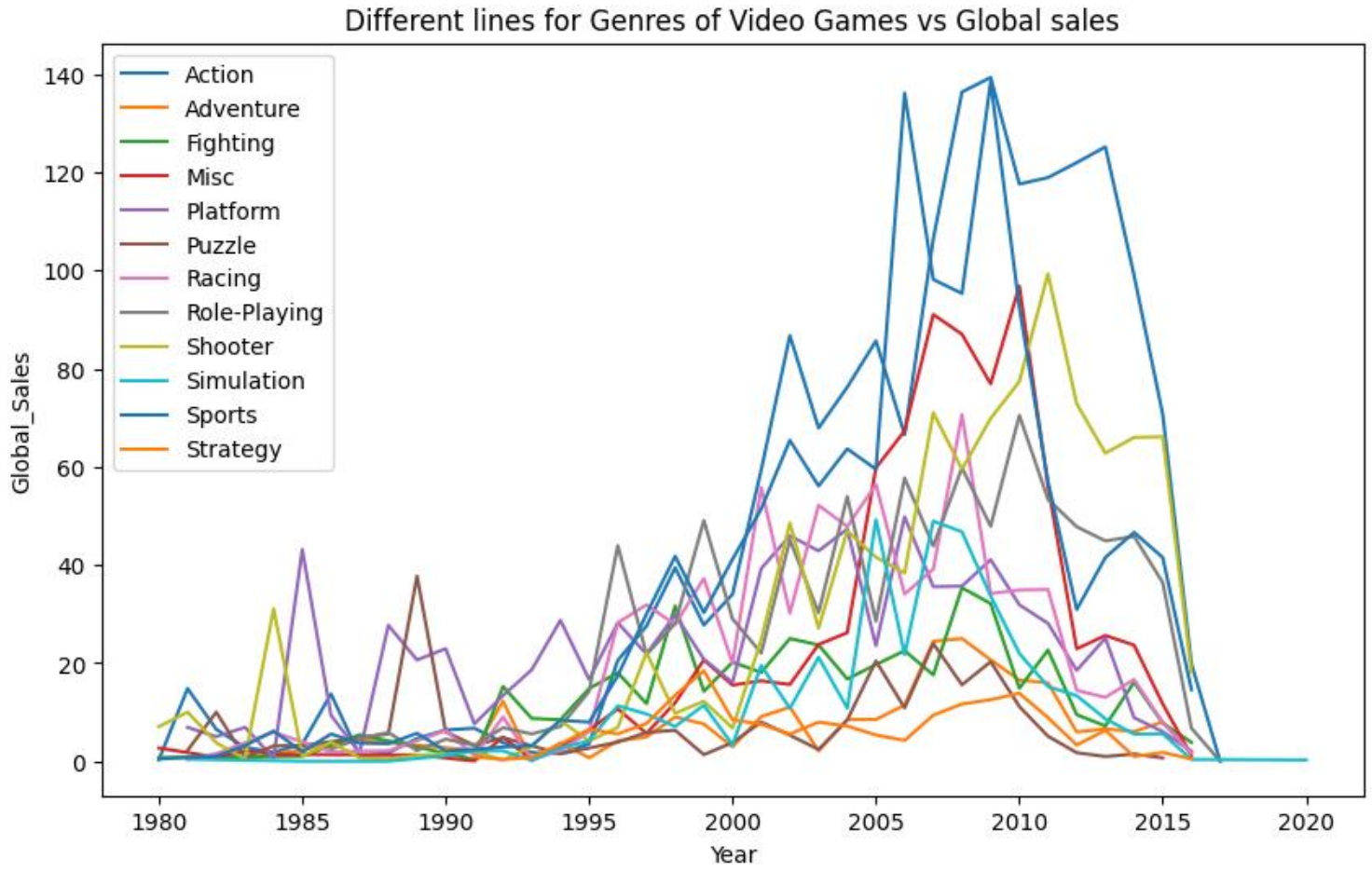
The market leader Platforms are GB and NES especially in NA and Japan.
PS4 seems to be doing well in EU market while 2600 holds strong in NA then compared to other markets.

Trend in Sales of Top 10 Publishers



Nintendo is the leading publisher of video games in all the markets.

Grouping variable to produce different colours



There is sharp increase in sales of action, shooter and misc games post 2005.

Plotly: An overview

- Interactive, open-source plotting library
- Supports over 40 unique chart types
- Includes various types of charts
- Visualizations can be:
  - Displayed in Jupyter notebook
  - Saved to HTML files
  - Used in developing Python-built web applications

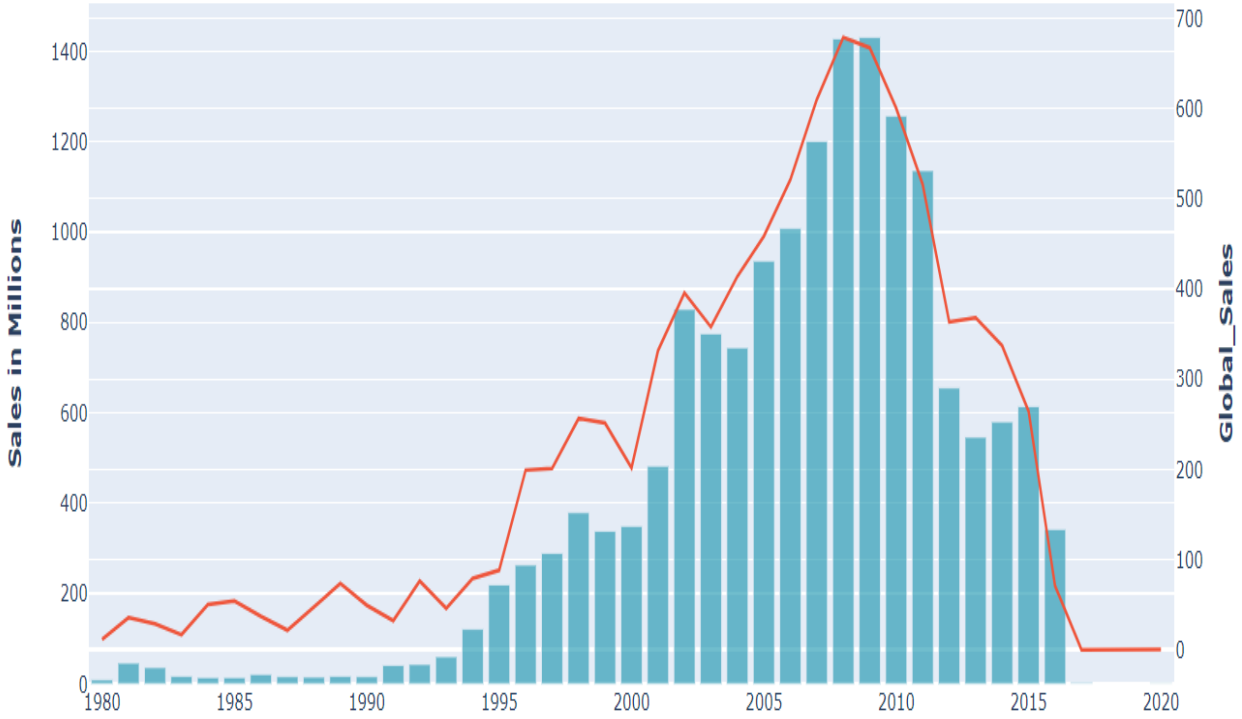
Plotly Sub-modules

- Plotly Graph Objects: Low-level interface to figures, traces, and layout

```
plotly.graph_objects.Figure
```

- Plotly Express: High-level wrapper

Using the go function in Plotly library




```
import plotly.graph_objects as go
```

```
df_yearcount =df_cleansed.groupby(df_cleansed['Year'])[['Rank']].count().rename(columns={'Rank':'counts'})
df_yearsales =df_cleansed.groupby(df_cleansed['Year'])[['Global_Sales']].sum()

fig = make_subplots(specs=[[{"secondary_y": True}]])
fig.add_trace(
    go.Bar(x=df_yearcount.index, y=df_yearcount['counts'], marker=dict(color='rgba(17, 145, 171, 0.6)'), name = 'counts'),
    secondary_y=False,
)
fig.add_trace(
    go.Scatter(x=df_yearsales.index, y=df_yearsales['Global_Sales'], name='Global_Sales'),
    secondary_y=True,
)
fig.update_xaxes(title_text="Year")
fig.update_yaxes(title_text="<b>Sales in Millions</b>", secondary_y=False)
fig.update_yaxes(title_text="<b>Global_Sales</b>", secondary_y=True)
fig.show()
```

Dashboarding tools

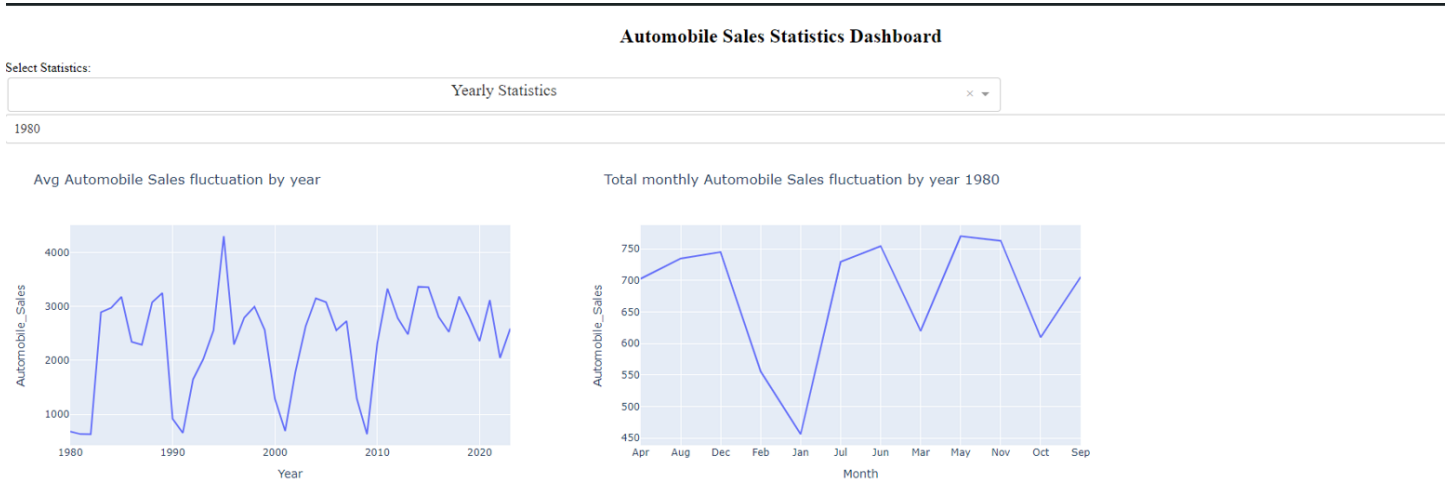
Just about any Python library can be used to create a “static” PNG, SVG, HTML, or other output that can be pasted into a presentation, sent in an email, published as a figure in a paper, and so on. Many people also want or need to create “live” Python-backed applications or dashboards that a user can interact with to explore or analyze some data. Python offers several libraries for this purpose. The four main tools designed specifically for web-based dashboarding in Python are:

- Dash (from Plotly)
- Bokeh is a plotting library, a widget and app library, and a server for both plots and dashboards.
- Panel is built on Bokeh, providing a higher-level toolkit specifically focused on app and dashboard creation and supporting multiple plotting libraries (not just Bokeh).

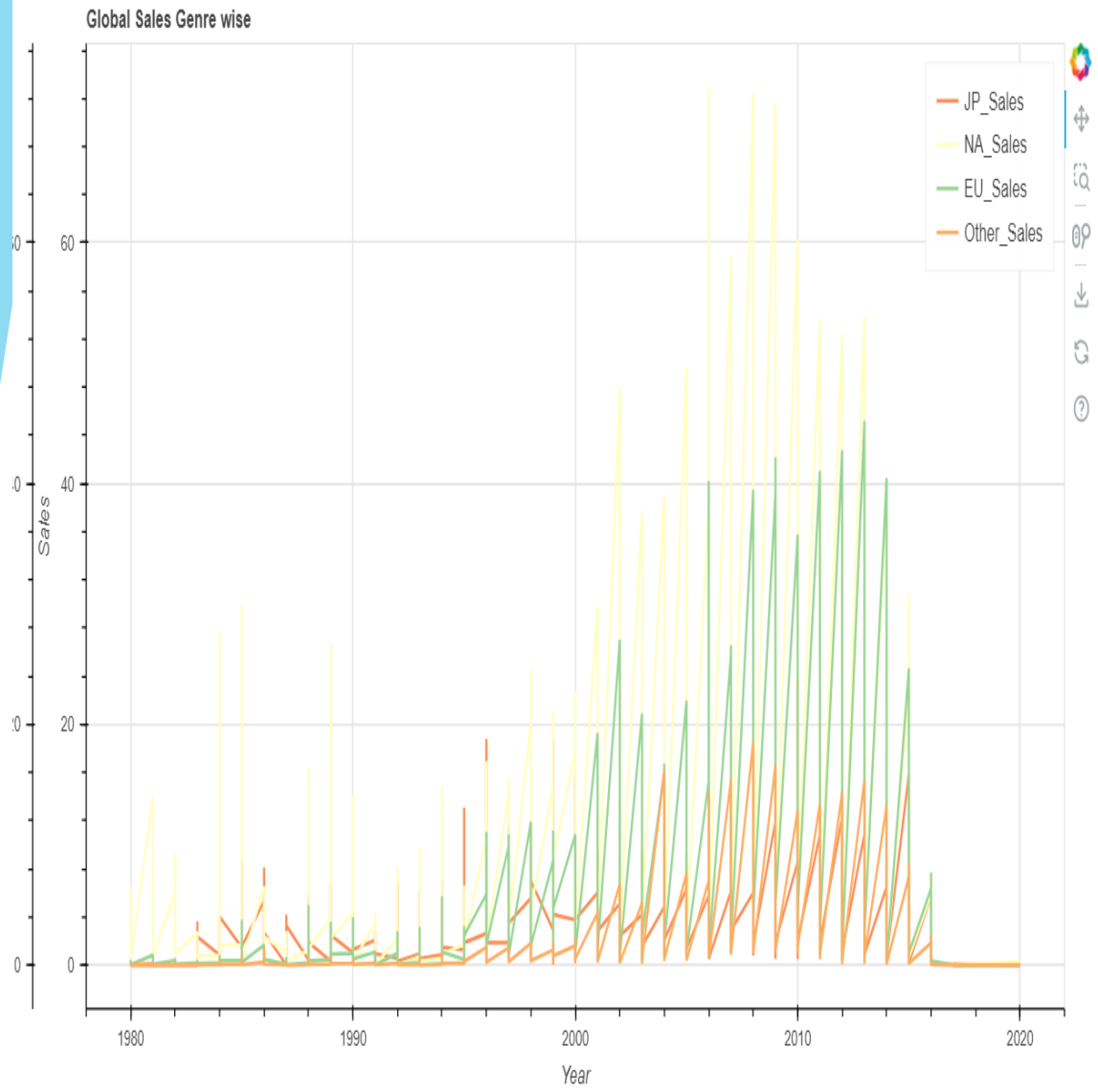
```
import dash
from dash import dcc
from dash import html
from dash.dependencies import Input, Output
import pandas as pd
import plotly.graph_objs as go
import plotly.express as px
```

```
#plot 1 Yearly Automobile sales using line chart for the whole period.
        yas= data.groupby('Year')['Automobile_Sales'].mean().reset_index()
        Y_chart1 = dcc.Graph(figure=px.line(yas,'Year','Automobile_Sales',title="Avg Automobile Sales fluctuation by year"))

# Plot 2 Total Monthly Automobile sales using line chart.
        mas= yearly_data.groupby('Month')['Automobile_Sales'].sum().reset_index()
        Y_chart2 = dcc.Graph(figure=px.line(mas,'Month','Automobile_Sales',title="Total monthly Automobile Sales fluctuation by year {}".format(input_year))
```



Using the Bokeh library




```
from bokeh.plotting import figure,show
from bokeh.palettes import Spectral3,Spectral4
from bokeh.resources import INLINE
from bokeh.models import ColumnDataSource, CDSView, GroupFilter, HoverTool,Legend,LegendItem,LinearAxis,Grid,BoxAnnotation
import bokeh.io
```

```
 df_cleansed.groupby(['Year','Genre'])['JP_Sales','NA_Sales','Other_Sales','EU_Sales'].sum().reset_index()
df_g)
['JP_Sales','NA_Sales','EU_Sales','Other_Sales']
t1[0])
gure(width=1000)
(x=df_g['Year'], y=df_g['JP_Sales'], line_width=2,color=Spectral3[2])
(x=df_g['Year'], y=df_g['NA_Sales'], line_width=2,color=Spectral3[1])
(x=df_g['Year'], y=df_g['EU_Sales'], line_width=2,color=Spectral3[0])
(x=df_g['Year'], y=df_g['Other_Sales'], line_width=2,color=Spectral4[2])
e.text = 'Global Sales Genre wise'
s.axis_label = 'Year'
s.axis_label ='Sales'
= LinearAxis()
layout(xaxis, 'below')
= LinearAxis()
layout(yaxis, 'left')
layout(Grid(dimension=0, ticker=xaxis.ticker))
layout(Grid(dimension=1, ticker=yaxis.ticker))

LegendItem(label='JP_Sales', renderers=[p.renderers[0]])
LegendItem(label='NA_Sales', renderers=[p.renderers[1]])
LegendItem(label='EU_Sales', renderers=[p.renderers[2]])
LegendItem(label='Other_Sales', renderers=[p.renderers[3]])

1 = Legend(items=[li1, li2, li3,li4], location='top_right')
layout(legend1)

)
```

Conclusion

- Through our visualization journey ,we can conclude that video game sales grew from 1995 onwards and 2000-2007 was the golden period and Nintendo was and is a market leader. The action games being the most popular.
- Today, video games make up a $100 billion global industry, and nearly two-thirds of American homes have household members who play video games regularly.
- In 1983, the North American video game industry experienced a major “crash” due to a number of factors, including an oversaturated game console market, competition from computer gaming, and a surplus of over-hyped, low-quality games,.
- The video game home industry began to recover in 1985 when the Nintendo Entertainment System (NES), called Famicom in Japan, came to the United States. The NES had improved 8-bit graphics, colors, sound and gameplay over previous consoles.
- Nintendo, a Japanese company that began as a playing card manufacturer in 1889, released a number of important video game franchises still around today, such as *Super Mario Bros.*, *The Legend of Zelda*, and *Metroid*.

Reference :https://www.history.com/topics/inventions/history-of-video-games

Thank You