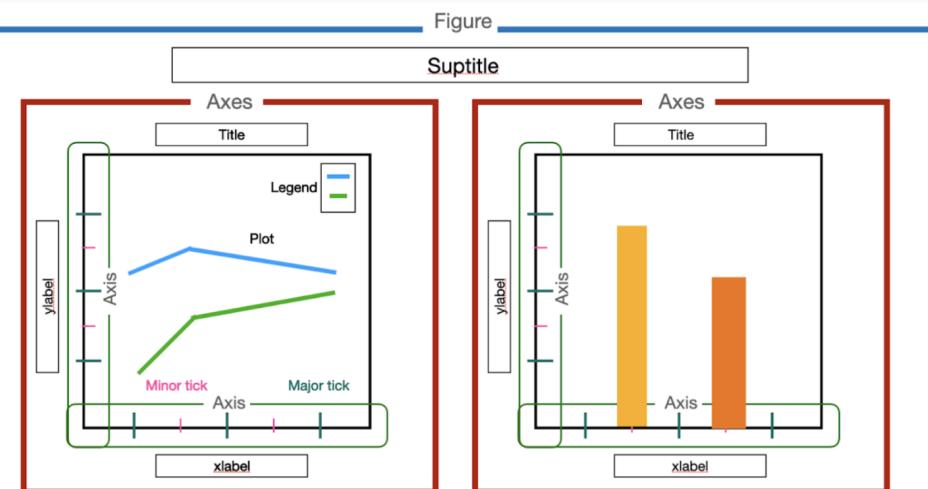


DATA VISUALIZATION CHEAT SHEET

SCALER 6

01. IMPORT CONVENTION



- FIGURE:** The overall window or page that everything is drawn on.
- AXIS:** Simply the `x-axis` and `y-axis`
- _AXES:** It is the area on which the data is plotted
- X-LABEL:** Name of x-axis
- Y-LABEL:** Name of y-axis
- MAJOR TICKS:** subdivides the axis into major units.
- MINOR TICKS:** subdivides major tick units.
- TITLE:** Title of each plot (Axes)
- LEGEND:** describes the elements in the plot, blue and green curves in this case
- SUPTITLE:** The common title of all the plots

Figure `plt.figure(figsize=(x,y))`
E.g. `plt.figure(figsize=(15,10))`

Labels `plt.xlabel('x label name')`
`plt.ylabel('y label name')`

Title `plt.title('Title of the plot')`

Suptitle `fig.suptitle('Title of the whole figure')`
Used in case of subplots

Ticks `plt.xticks(rotation=90)`
For x ticks
`plt.yticks(rotation=90)`
For y ticks

Legend `plt.legend()`

01. IMPORT CONVENTION

NUMERICAL DATA:

Discrete : Can only take finite values
(E.g. Year, 2010, 2011, etc.)

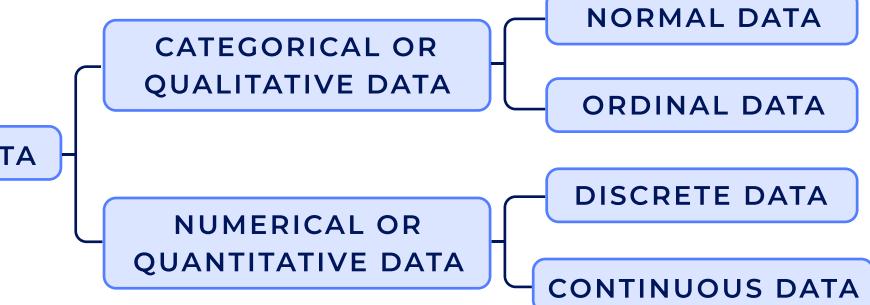
Continuous : Can take any numerical value
(E.g. temperature, pressure, etc.)

CATEGORICAL DATA:

Ordinal: Categorical Data with an order
(E.g. low, medium, high)

Non-ordinal/nominal: Categorical Data without any order (example gender as Male/Female)

TYPES OF DATA



03. INSTALLING AND IMPORTING MATPLOTLIB AND SEABORN

Installing

Pip Install Matplotlib
Pip Install Seaborn

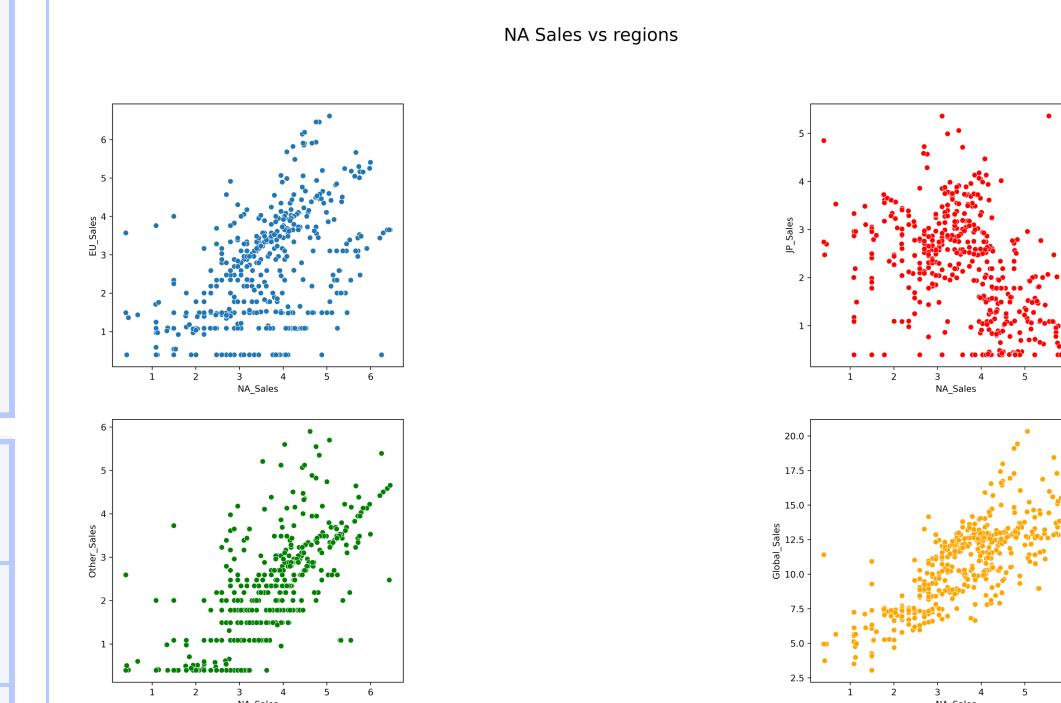
Importing Convention

Import Matplotlib.Pyplot As Plt
Import Seaborn As Sns

MISC PLOTS

SUB PLOTS

DIRECT ACCESSING THE FUNCTIONS



We can plot multiple plots in a single figure

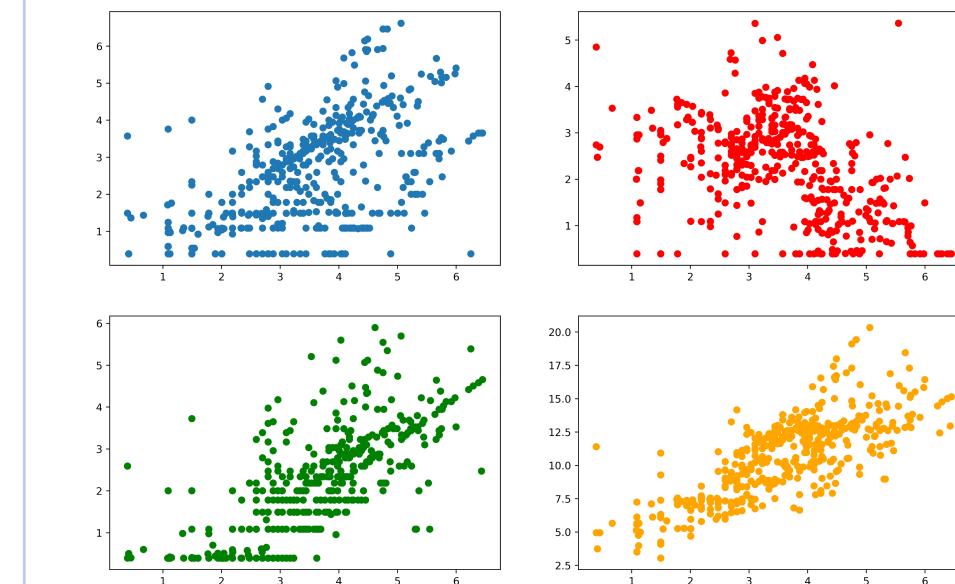
```
plt.figure()
plt.subplot(row,col,1)
# plot
...
E.g.
plt.figure(figsize=(20,12)).suptitle("NA Sales vs regions", fontsize=20)
# Using a 2x3 subplot
plt.subplot(2, 3, 1)
sns.scatterplot(x='NA_Sales', y='EU_Sales',
data=top3_data)

plt.subplot(2, 3, 3)
sns.scatterplot(x='NA_Sales', y='JP_Sales',
data=top3_data, color='red')

plt.subplot(2, 3, 4)
sns.scatterplot(x='NA_Sales', y='Other_Sales',
data=top3_data, color='green')
```

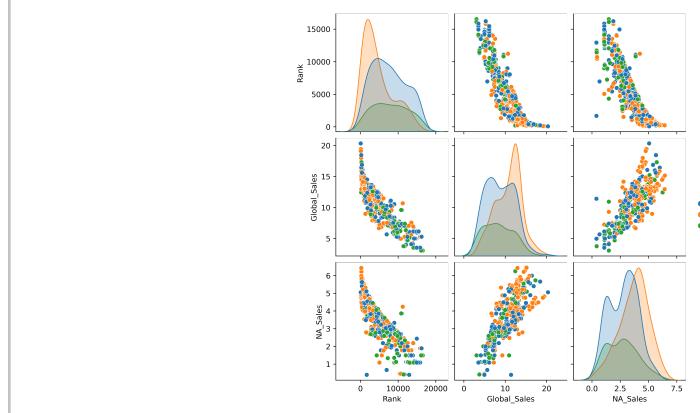
```
plt.subplot(2, 3, 6)
sns.scatterplot(x='NA_Sales', y='Global_Sales',
data=top3_data, color='orange')
```

USING OBJECT



PAIRPLOT

Displays a scatterplot for each pair of attributes, can provide a hue for each category too



`sns.pairplot(data=df, hue='hueCol')`

E.g.

`sns.pairplot(data=df, hue='Genre')`

CORRELATION AND HEATMAPS

Plot a heatmap of correlation between various variables

`sns.heatmap(df.corr(), cmap='colour_map')`



E.g.

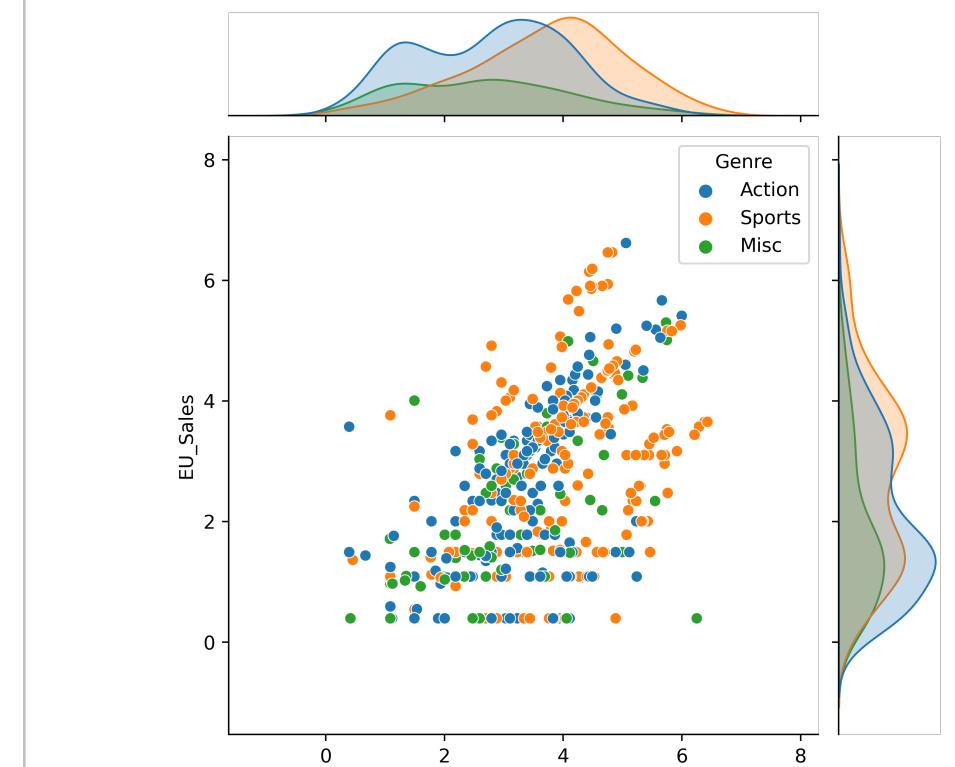
`sns.heatmap(top3_data.corr(), cmap= "Blues", annot=True)`

plots a heatmap of the data with the correlation coefficient values annotated

JOINT PLOTS

Draws multiple types of plot of two variables in the same plot

`sns.jointplot(x='xcol', y='ycol', data=data, hue='hueCol')`

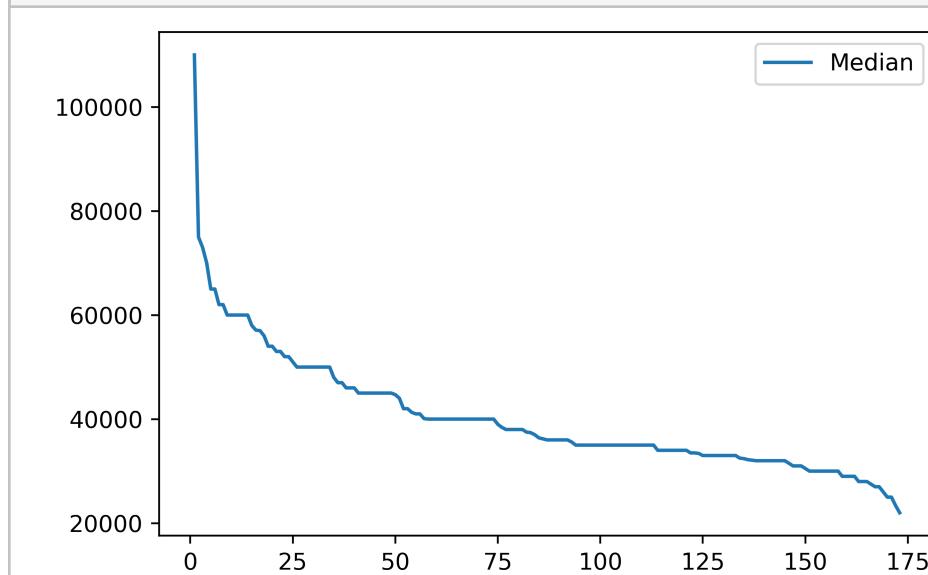


E.g.
`sns.jointplot(x='NA_Sales', y='EU_Sales', data=top3_data, hue='Genre')`

PLOTTING WITH PANDAS

LINEPLOT

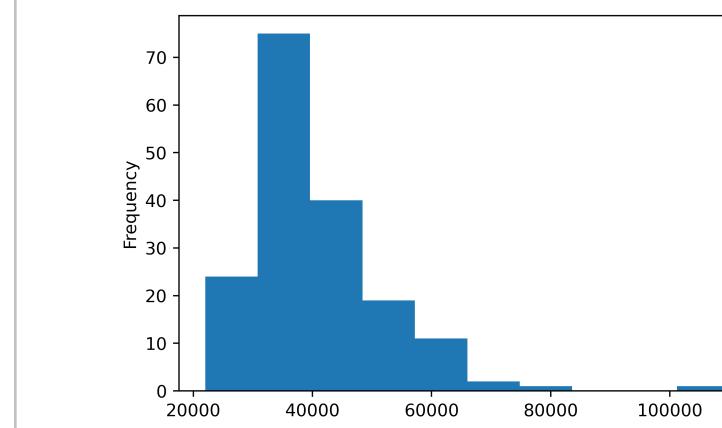
`df.plot(x='col1', y='col2')`



E.g.
`df.plot(x="Median", y = "Unemployment_rate", kind="scatter")`

HISTOGRAM

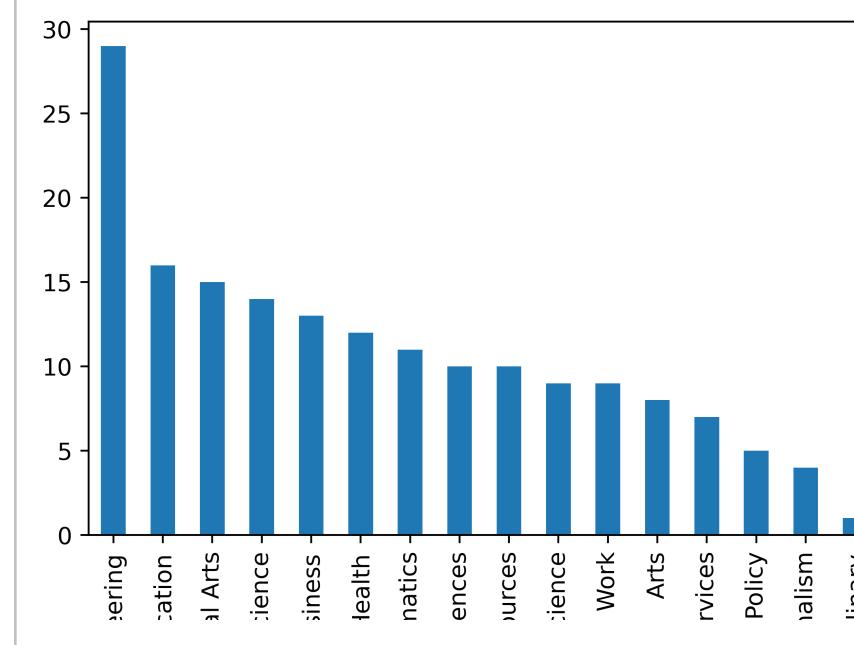
`df['col'].plot(kind='hist')`



E.g.
`df['Median'].plot(kind="hist")`

BARPLOT

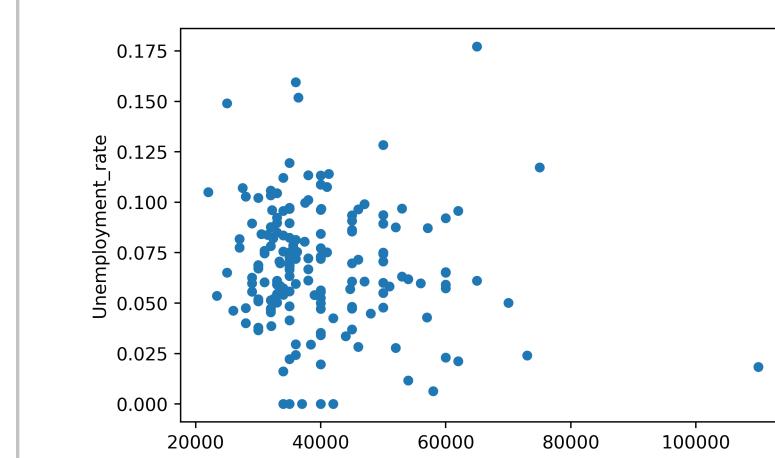
`df['col'].value_counts().plot(kind='bar')`



E.g.
`df['Major_category'].value_counts().plot(kind = 'bar')`

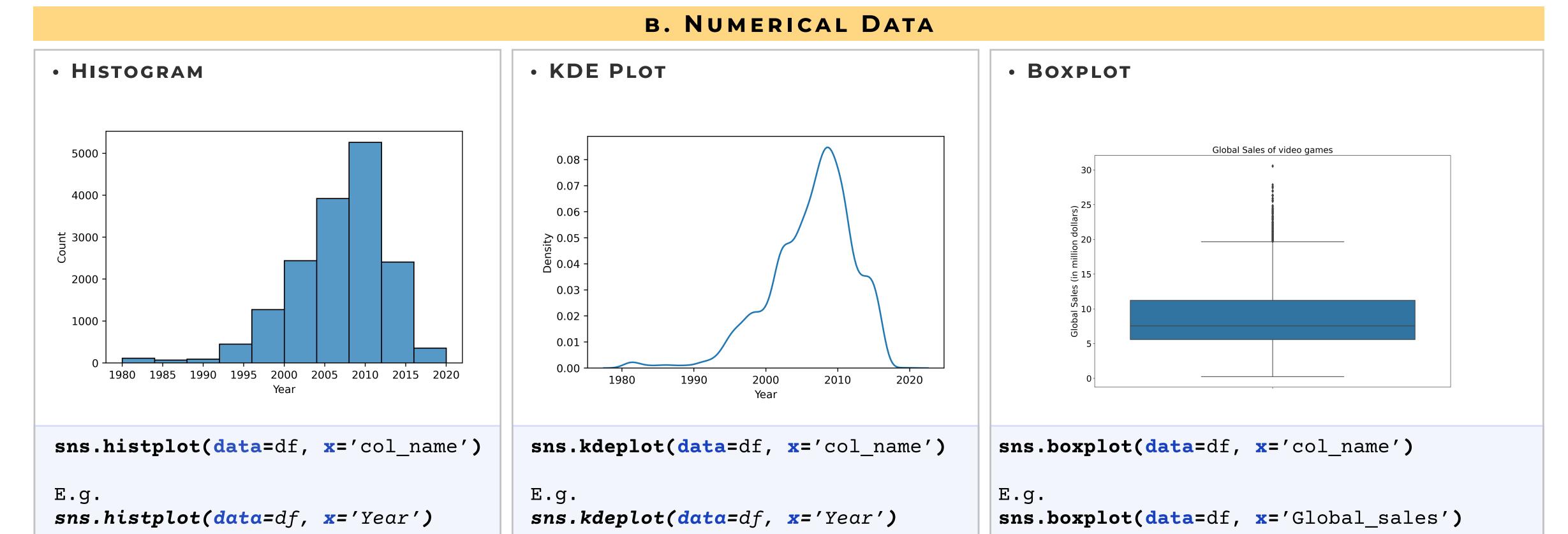
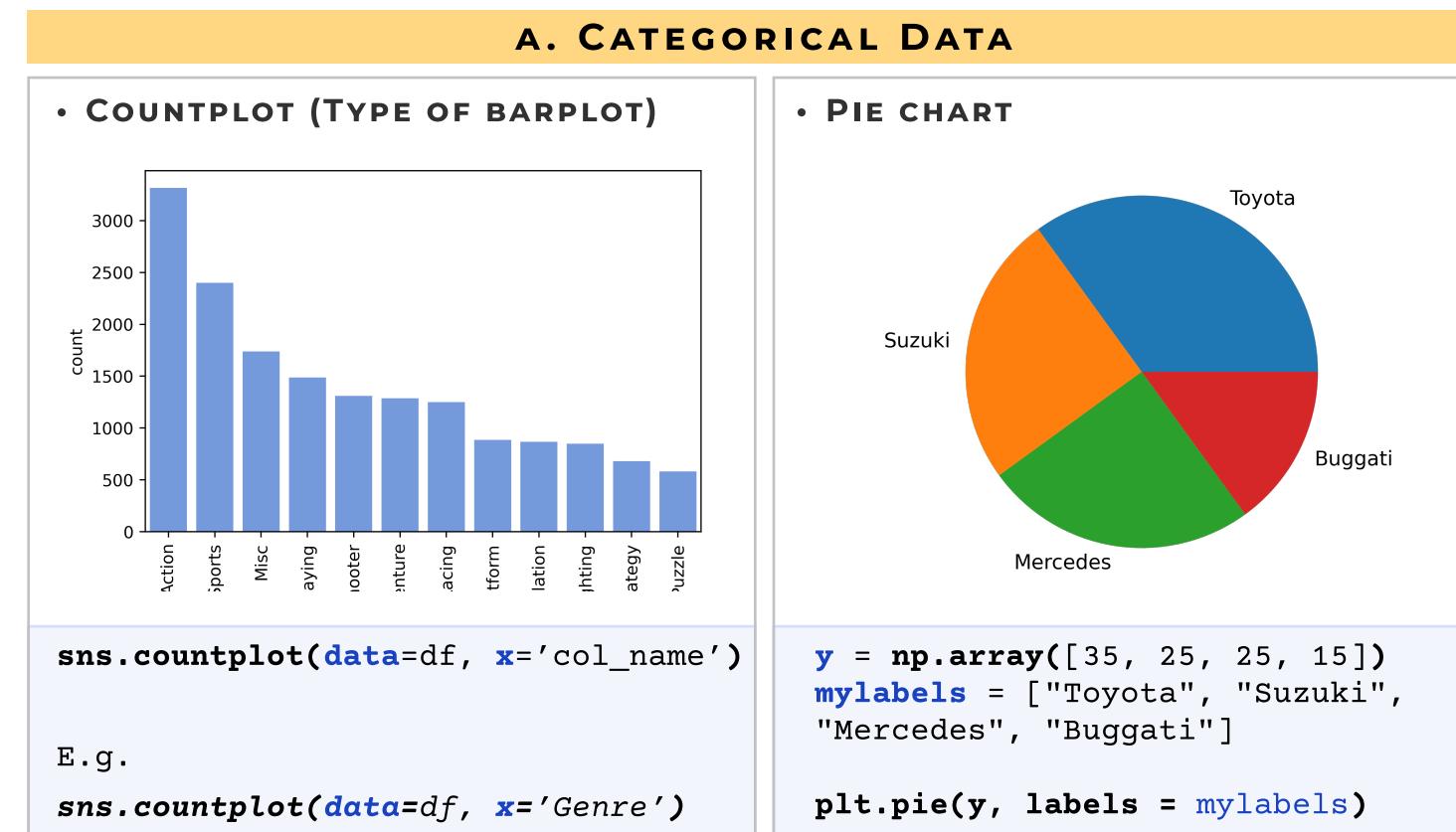
SCATTERPLOT

`df.plot(x='col1', y='col2', kind='scatter')`

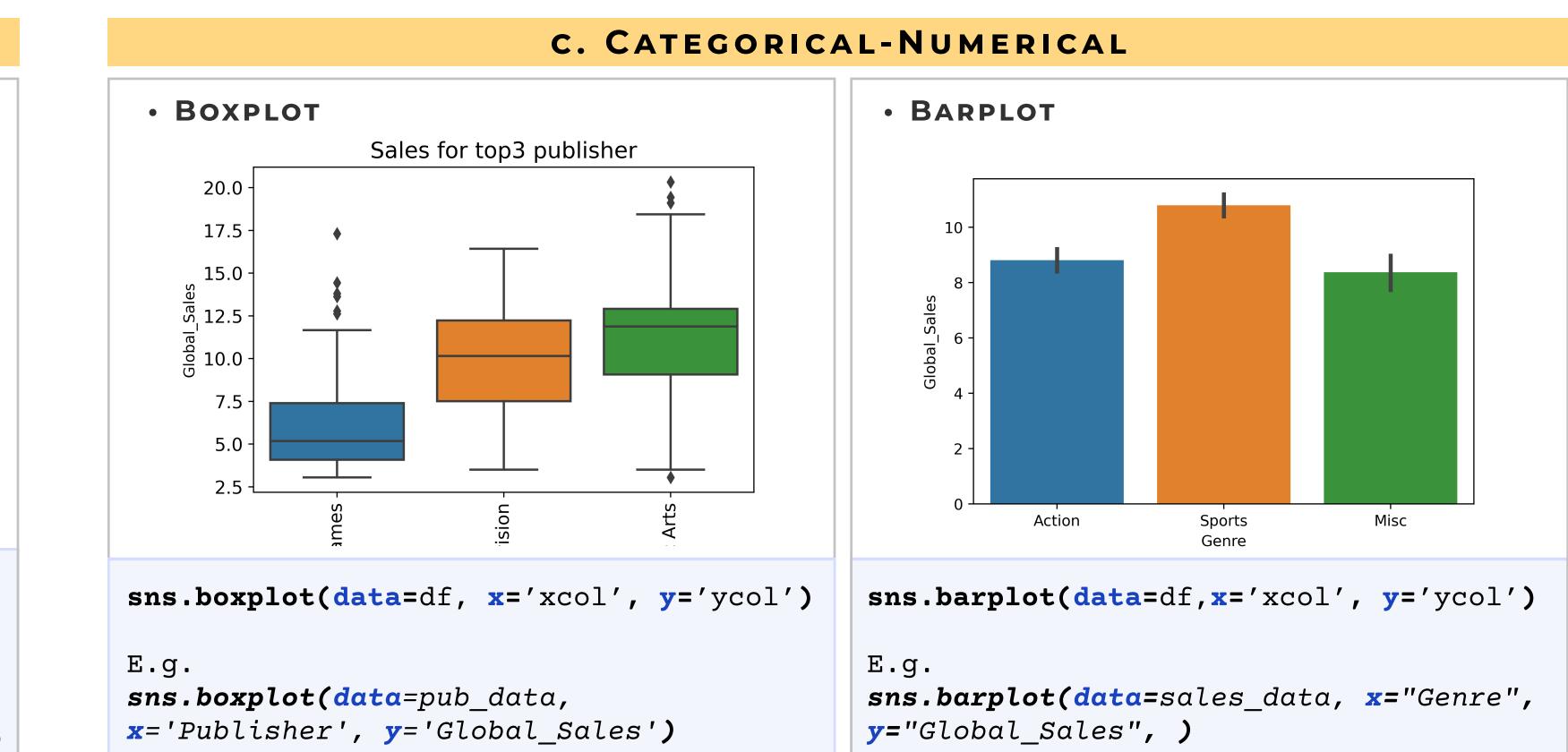
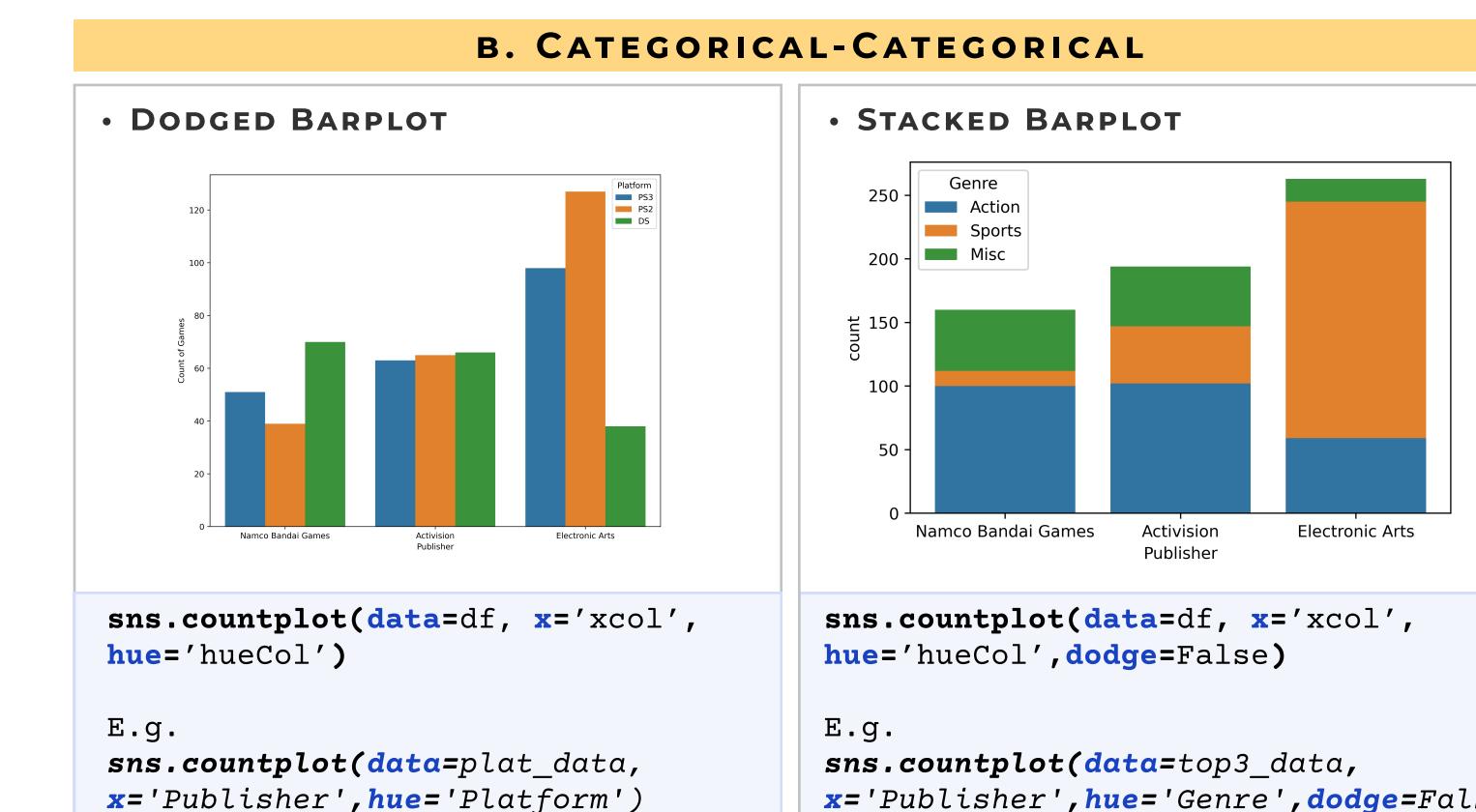
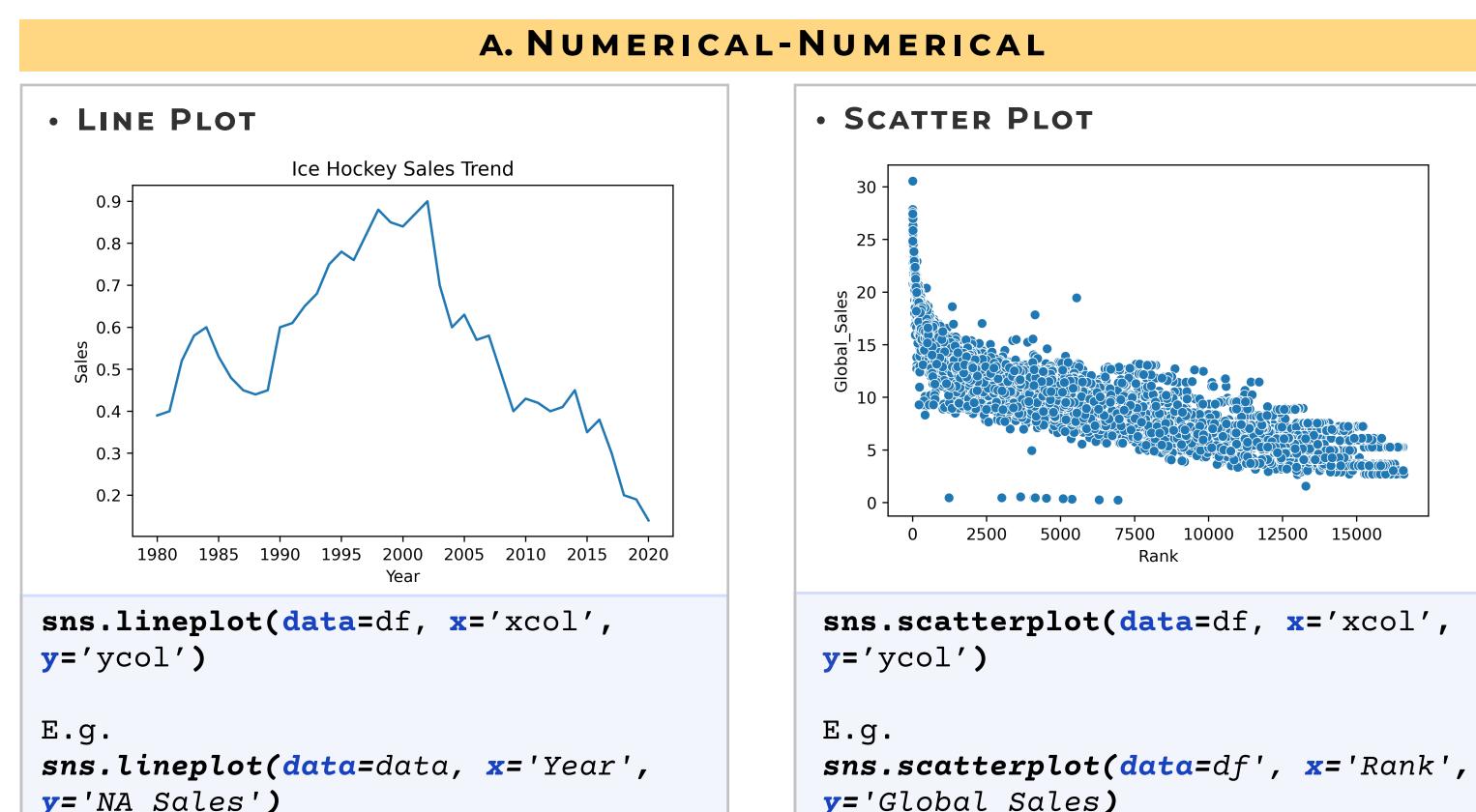


E.g.
`df.plot(x="Median", y = "Unemployment_rate", kind="scatter")`

04. UNIVARIATE DATA ANALYSIS:



05. BIVARIATE DATA ANALYSIS:



06. MULTIVARIATE DATA ANALYSIS:

