Assignment 28: Plotly - Utkarsh Gaikwad

Assignment pdf link

Question 1

Question 1: Load the "titanic" dataset using the load_dataset function of seaborn. Use Plotly express to plot a scatter plot for age and fare columns in the titanic dataset.

Answer :

Loading titanic dataset

In [1]:
```
import seaborn as sns
titanic = sns.load_dataset('titanic')
titanic.head()
```

Out[1]:

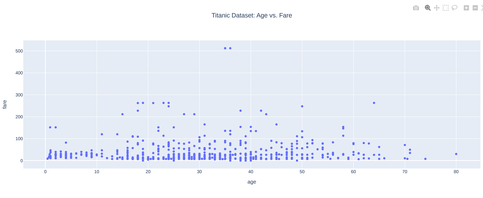
	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True

Plotting the scatter plot

In [2]:
```
import plotly.express as px
# Plot a scatter plot for "age" and "fare" columns using Plotly express
fig = px.scatter(titanic,
                 x="age",
                 y="fare",
                 title="Titanic Dataset: Age vs. Fare")

# update the layout to center the title above the plot
fig.update_layout(
    title={
        'text': "Titanic Dataset: Age vs. Fare",
        'y':0.95,
        'x':0.5,
        'xanchor': 'center',
        'yanchor': 'top'})

fig.show()
```



Question 2

Question 2: Using the tips dataset in the Plotly library, plot a box plot using Plotly express

Answer :

Importing tips dataset from plotly express

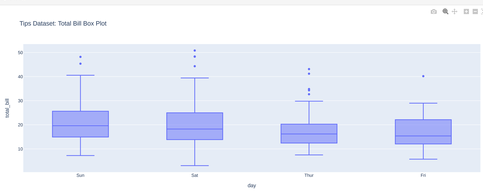
In [3]:
```
import plotly.express as px
tips = px.data.tips()
tips.head()
```

Out[3]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Plotting Boxplot

In [4]:
```
# Plot a box plot for the "total_bill" column using Plotly express
fig = px.box(tips,x='day',y='total_bill', title="Tips Dataset: Total Bill Box Plot")
fig.show()
```



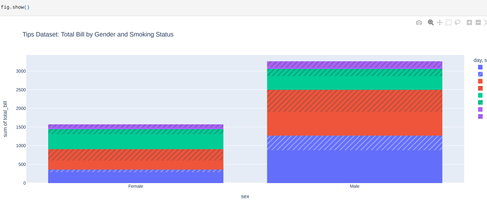
Question 3

Question 3: Using the tips dataset in the Plotly library, Plot a histogram for x= "sex" and y="total_bill" column in the tips dataset. Also, use the "smoker" column with the pattern_shape parameter and the "day"column with the color parameter.

Answer:

In [6]:
```
# Plot a histogram for "sex" and "total_bill" columns using Plotly express
fig = px.histogram(tips, x="sex", y="total_bill",
                   color="day", pattern_shape="smoker",
                   title="Tips Dataset: Total Bill by Gender and Smoking Status")

fig.show()
```



Question 4

Question 4: Using the iris dataset in the Plotly library, Plot a scatter matrix plot, using the "species" column for the color parameter.

Note : Use "sepal_length", "sepal_width", "petal_length", "petal_width" columns only with the dimensions parameter.

Answer :

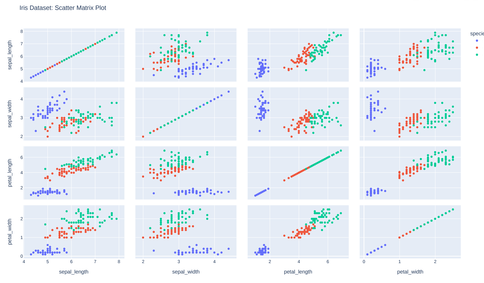
In [8]:
```
import plotly.express as px

# Load the "iris" dataset from Plotly
iris = px.data.iris()

# Plot a scatter matrix plot using Plotly express
fig = px.scatter_matrix(iris,
                        dimensions=["sepal_length", "sepal_width", "petal_length", "petal_width"],
                        color="species", title="Iris Dataset: Scatter Matrix Plot")

# Updating the height of figure
fig.update_layout(height=600)

fig.show()
```



Question 5

Question 5: What is Distplot? Using Plotly express, plot a distplot

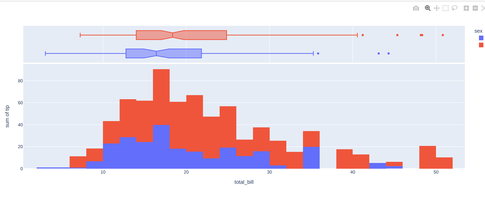
Answer :

A Distplot or distribution plot, depicts the variation in the data distribution. Seaborn Distplot represents the overall distribution of continuous data variables.

The Seaborn module along with the Matplotlib module is used to depict the distplot with different variations in it. The Distplot depicts the data by a histogram and a line in combination to it.

Plotly express has px.histogram as the function

In [7]:
```
import plotly.express as px
tips = px.data.tips()
fig = px.histogram(tips, x="total_bill", y="tip", color="sex",
                   marginal="box", # or violin, rug
                   hover_data=tips.columns)
fig.show()
```



Can also plot distplot like below

In [8]:
```
import plotly.figure_factory as ff
import numpy as np
np.random.seed(2)

x = np.random.randn(1000)
hist_data = [x]
group_labels = ['distplot'] # name of the dataset

fig = ff.create_distplot(hist_data, group_labels)
fig.show()
```