

4th Mar Assignment

March 14, 2023

1 Assignment 28

Q1. 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.

Ans.

```
[1]: import seaborn as sns
```

```
[2]: t=sns.load_dataset('titanic')
```

```
[3]: t.head()
```

```
[3]:   survived  pclass    sex  age  sibsp  parch    fare embarked  class \
0         0      3   male  22.0     1     0   7.2500         S  Third
1         1      1 female  38.0     1     0  71.2833         C  First
2         1      3 female  26.0     0     0   7.9250         S  Third
3         1      1 female  35.0     1     0  53.1000         S  First
4         0      3   male  35.0     0     0   8.0500         S  Third
```

```
   who  adult_male  deck  embark_town  alive  alone
0  man         True  NaN  Southampton    no  False
1 woman        False   C   Cherbourg   yes  False
2 woman        False  NaN  Southampton   yes   True
3 woman        False   C   Southampton   yes  False
4  man         True  NaN  Southampton    no   True
```

```
[5]: pip install plotly
```

Collecting plotly

Downloading plotly-5.13.1-py2.py3-none-any.whl (15.2 MB)

15.2/15.2 MB

61.4 MB/s eta 0:00:0000:0100:01

Collecting tenacity>=6.2.0

Downloading tenacity-8.2.2-py3-none-any.whl (24 kB)

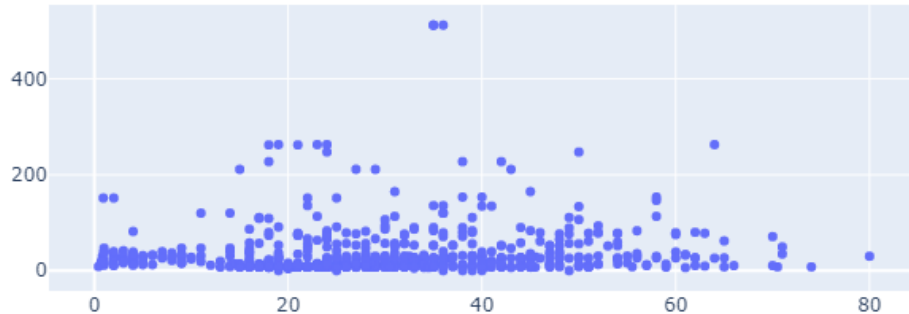
Installing collected packages: tenacity, plotly

Successfully installed plotly-5.13.1 tenacity-8.2.2

Note: you may need to restart the kernel to use updated packages.

```
[5]: import plotly.graph_objects as go
```

```
[11]: fig=go.Figure()  
fig.add_trace(go.Scatter(x=t.age,y=t.fare,mode='markers'))
```



Q2. Using the tips dataset in the Plotly library, plot a box plot using Plotly express.

Ans.

```
[14]: tips=sns.load_dataset('tips')
```

```
[15]: tips.head()
```

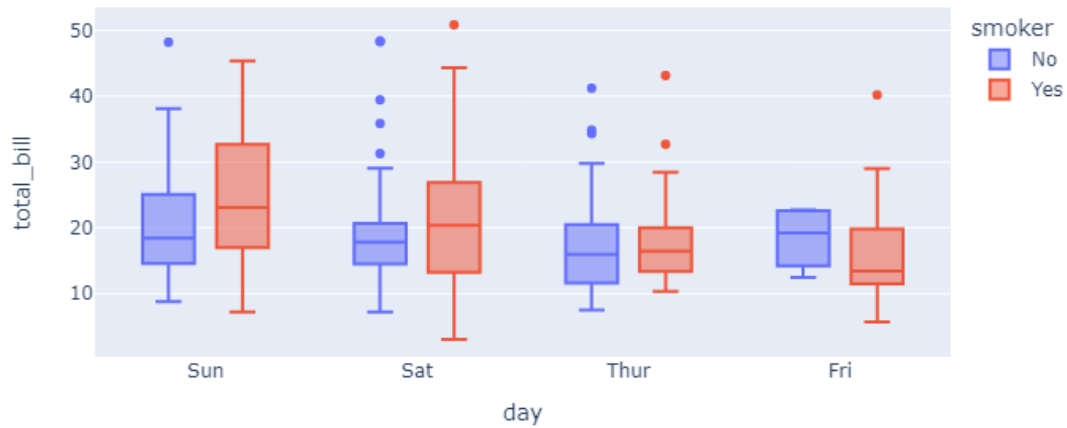
```
[15]:
```

	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

```
[16]: import plotly.express as px
```

```
[26]: fig=px.box(data_frame=tips,x='day',y='total_bill',color='smoker')
```

```
[27]: fig.show()
```

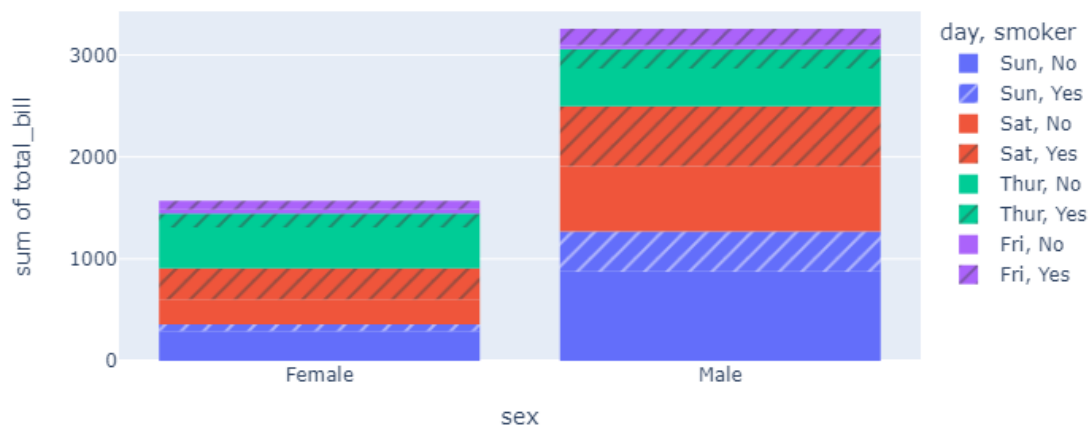


Q3. Using the tips dataset in the Plotly library, Plot a histogram for x= “sex” and y=“total_bill” column in the tips dataset. Also, use the “smoker” column with the pattern_shape parameter and the “day” column with the color parameter.

Ans.

```
[28]: fig=px.histogram(data_frame=tips,x='sex',y='total_bill', color="day",
    ↪pattern_shape="smoker")
```

```
[29]: fig.show()
```



Q4. 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.

Ans.

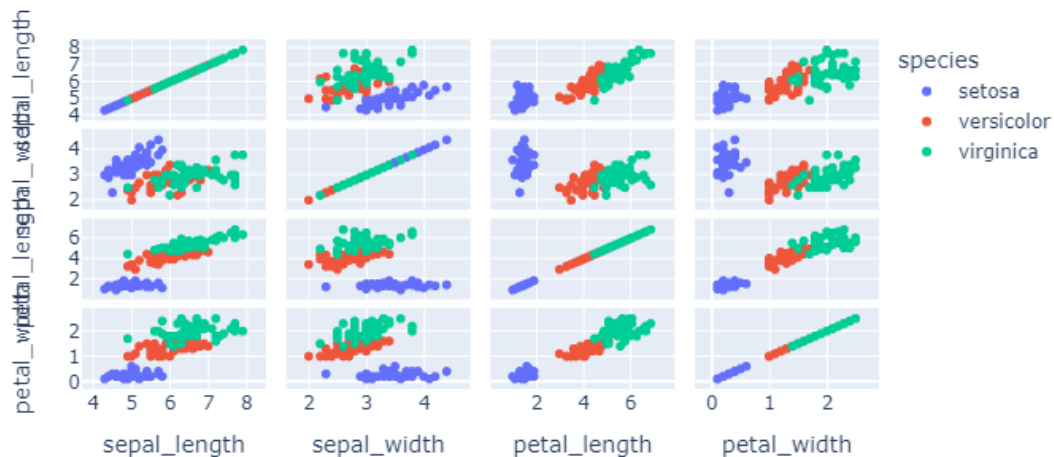
```
[30]: iris=sns.load_dataset('iris')
```

```
[31]: iris.head()
```

```
[31]:   sepal_length  sepal_width  petal_length  petal_width  species
0         5.1         3.5         1.4         0.2   setosa
1         4.9         3.0         1.4         0.2   setosa
2         4.7         3.2         1.3         0.2   setosa
3         4.6         3.1         1.5         0.2   setosa
4         5.0         3.6         1.4         0.2   setosa
```

```
[33]: fig=px.scatter_matrix(data_frame=iris,dimensions=['sepal_length','sepal_width',
↪ 'petal_length', 'petal_width'],color='species')
```

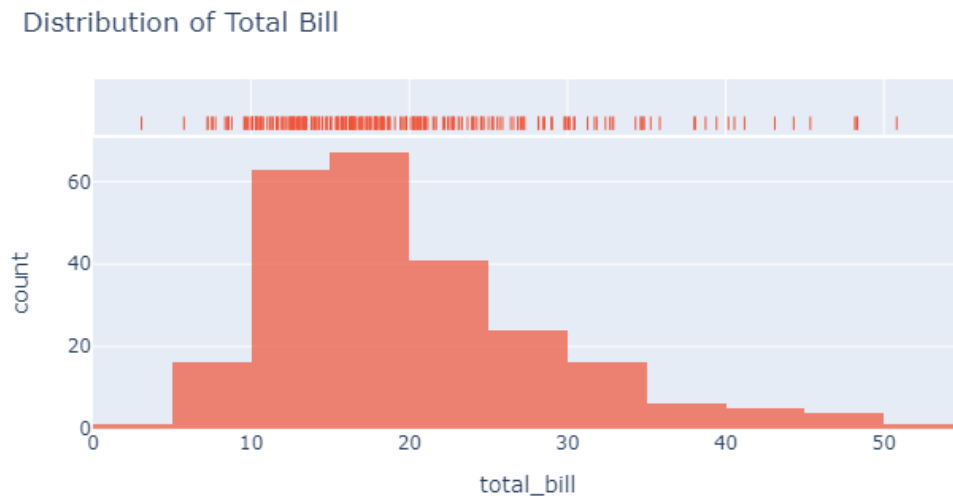
```
[35]: fig.show()
```



Q5. What is Distplot? Using Plotly express, plot a distplot.

Ans.distplot is a type of visualization in seaborn library that combines a histogram with a kernel density estimate (KDE) plot. It provides a way to visualize the distribution of a dataset.

```
[38]: fig = px.histogram(tips, x="total_bill", nbins=20, marginal="rug", opacity=0.7,
    ↪color_discrete_sequence=["#EF553B"])
fig.update_layout(title_text="Distribution of Total Bill")
fig.show()
```



```
[ ]:
```