˅ Assignment #9

In this assignment, you will practice using the plotly express library. https://plotly.com/python/plotly-express/

Your goal is to recreate the following graphics below using plotly express. You should attempt to recreate them as close as possible.

You may work individually or with a group

```
import plotly.express as px
df = px.data.tips()
```

```
df.head()
```

|	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

Next steps: [Generate code with `df`]   [⬤ View recommended plots]
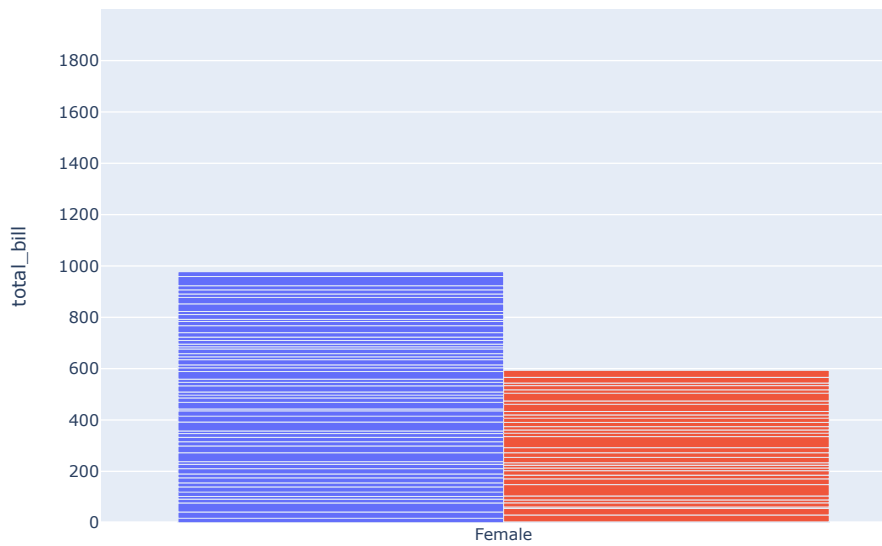
˅ Bar Plot

A barplot shows the relationship between a numeric and a categoric variable. Each entity of the categoric variable is represented as a bar. The size of the bar represents its numeric value.

```
df = px.data.tips()
fig=px.bar(df,
          x="sex",
          y="total_bill",
          color="smoker",
          barmode="group")
```
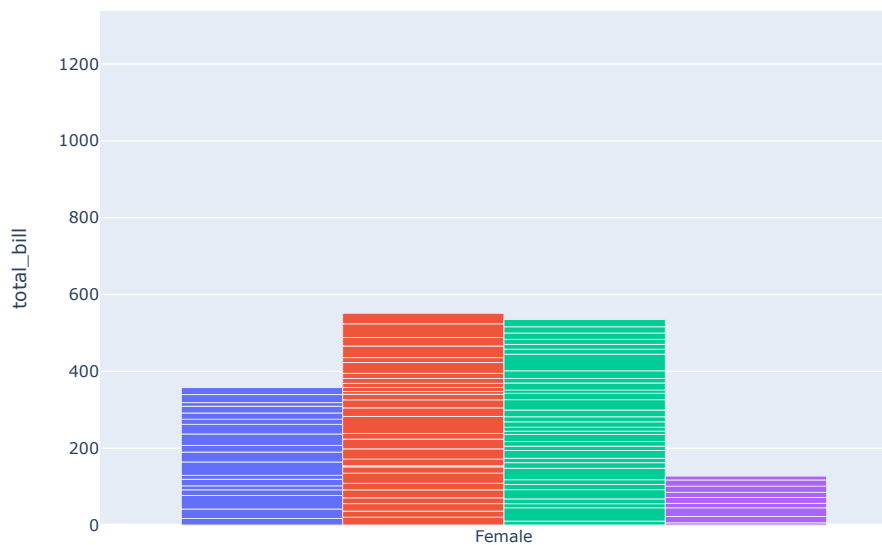
```
fig.show()
```
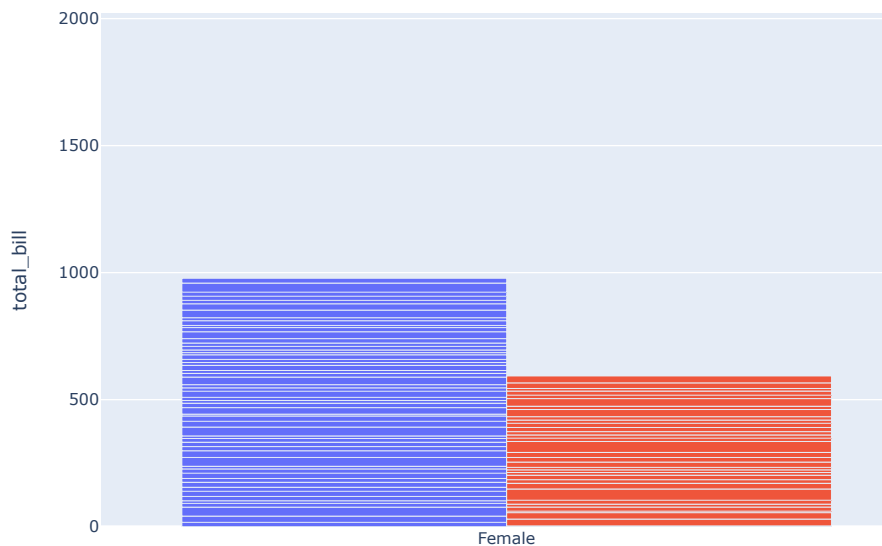
Start coding or generate with AI.


```
df = px.data.tips()
fig=px.bar(df,
          x="sex",
          y="total_bill",
          color="day",
          barmode="group")

fig.show()
```


Start coding or generate with AI.

⌄ Scatter Plot

Scatter plots are used to check the relationship between the variables and the distribution of the data. A scatterplot displays the relationship between 2 numeric variables. For each data point, the value of its first variable is represented on the X axis, the second on the Y axis

```
fig = px.scatter(df,  'total_bill','tip' , color = 'smoker', facet_col='sex')
fig.update_layout(plot_bgcolor='rgb(255,255,255)')

fig.show()
```
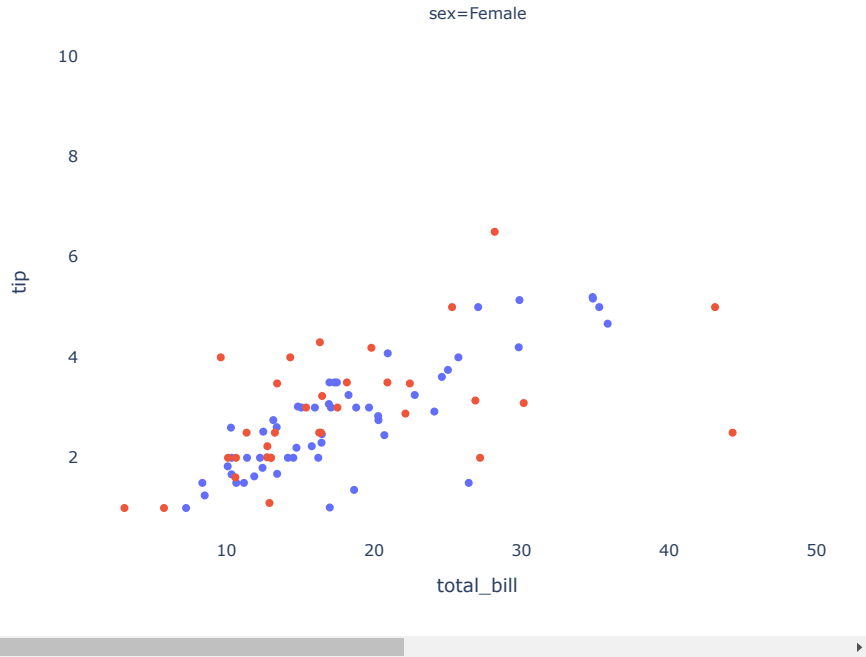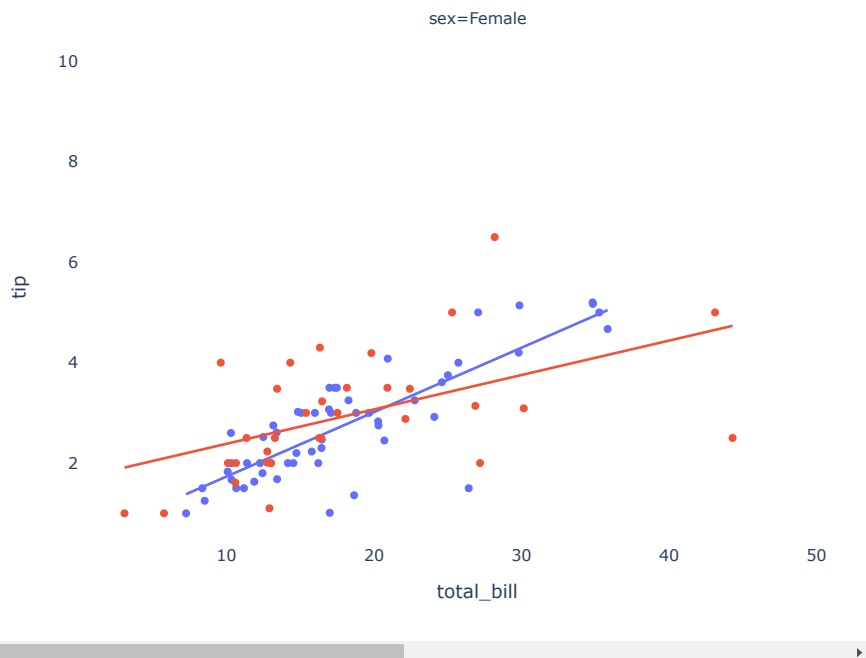


```
#adding trend lines

fig = px.scatter(df,  'total_bill','tip' , color = 'smoker', facet_col='sex', trendline='ols')
fig.update_layout(plot_bgcolor='rgb(255,255,255)')

fig.show()
```
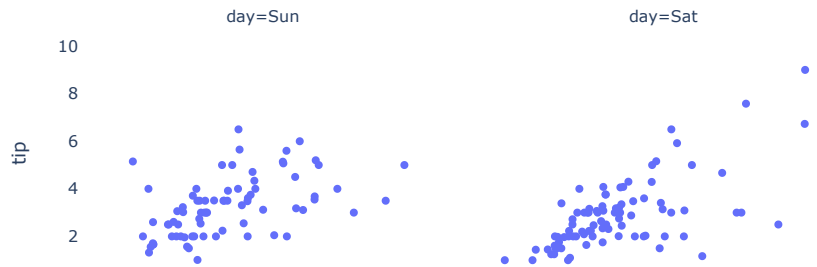
```
#exploring vs. day

fig = px.scatter(df,  'total_bill','tip', 'time', facet_row = 'time',facet_col='day')
fig.update_layout(plot_bgcolor='rgb(255,255,255)')

fig.show()
```


Start coding or generate with AI.

⌄ Histogram

A histogram takes as input a numeric variable only. The variable is cut into several bins, and the number of observation per bin is represented by the height of the bar. It is possible to represent the distribution of several variable on the same axis using this technique.

```
#Exploring distrubution of tip
fig = px.histogram(df, 'tip', nbins=10, marginal='rug')
fig.update_layout(plot_bgcolor='rgb(255,255,255)')
fig.show()
```
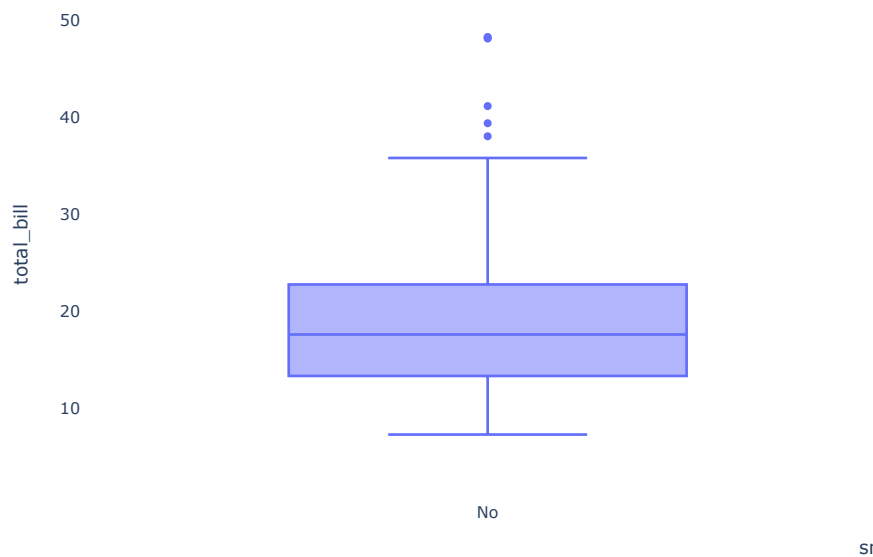
Start coding or generate with AI.

Boxplot

A boxplot gives a nice summary of one or several numeric variables. The line that divides the box into 2 parts represents the median of the data

```
fig = px.box(df,'smoker','total_bill', color='smoker')
fig.update_layout(plot_bgcolor='rgb(255,255,255)')
fig.show()
```


Start coding or generate with AI.