



Interactive Visualization with Plotly - Plotly Express - 1

One should look for what is and not what he thinks should be. (Albert Einstein)

Plotly Express: Topic introduction

In this part of the course, we will cover the following concepts:

- Introduction to plotly express
- Organize and visualize data with plotly express

Warm up

- Check out this *interactive graphic* from The Guardian about dog breeds and how they are related to each other
- What surprised you?
- What did you already know?

Module completion checklist

Objective	Complete
Define Plotly Express	
Describe univariate plots in Plotly Express	

What is plotly?

- Plotly is a powerful graphing library that is used to create interactive, publication-quality graphs
- In `plotly`, you can create line plots, scatter plots, area charts, bar charts, error bars, box plots, histograms, heatmaps, subplots, multiple-axes, polar charts, and bubble charts
- [Click here](#) for more about `plotly`



What is different about plotly express?

- plotly express contains functions that can create entire figures at once
- It can be referred as plotly express or PX
- plotly express is a built-in part of the plotly library which is the recommended for creating most common figures
- Using the plotly express we can create a majority of the most commonly-used interactive visualizations in data science with single line of code
- With plotly express, you can make professional interactive visualizations easily and quickly

Module completion checklist

Objective	Complete
Define Plotly Express	✓
Describe univariate plots in Plotly Express	

Load the dataset and libraries

- In this course, we will be using the inbuilt tips dataset that's available in plotly

```
# Load the libraries and the dataset
import plotly.express as px

# Load the dataset
tips_dataset = px.data.tips()
# Top 5 entires of the dataset
tips_dataset.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

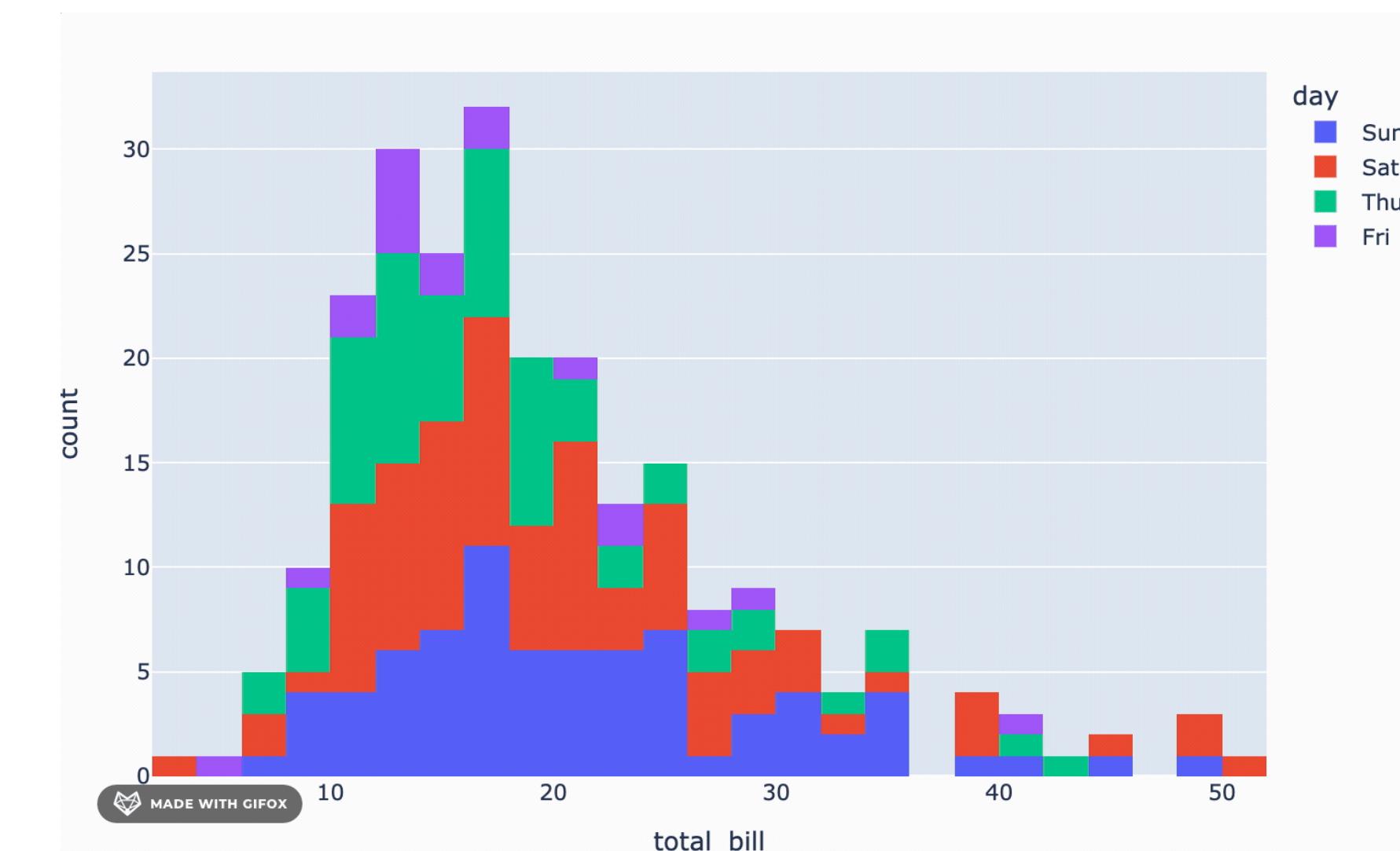
Describe univariate plots in plotly express

- A univariate plot shows the ins and outs of a single variable
- There are different type of univariate plots available in plotly express:
 - Histograms
 - Box plots
 - Strip plots
 - Violin plots
 - ECDF plots
- We will talk about a subset of these univariate plots

Histograms

- A histogram is a very easy way to visualize the distribution of a single variable. However, with interactive histograms we are not limited to a single variable
- **plotly express** allows us to create layered histograms, where layers can be toggled
- **plotly express** also has built-in datasets for use with plotting tools, which we will be using in next slides to create visualization

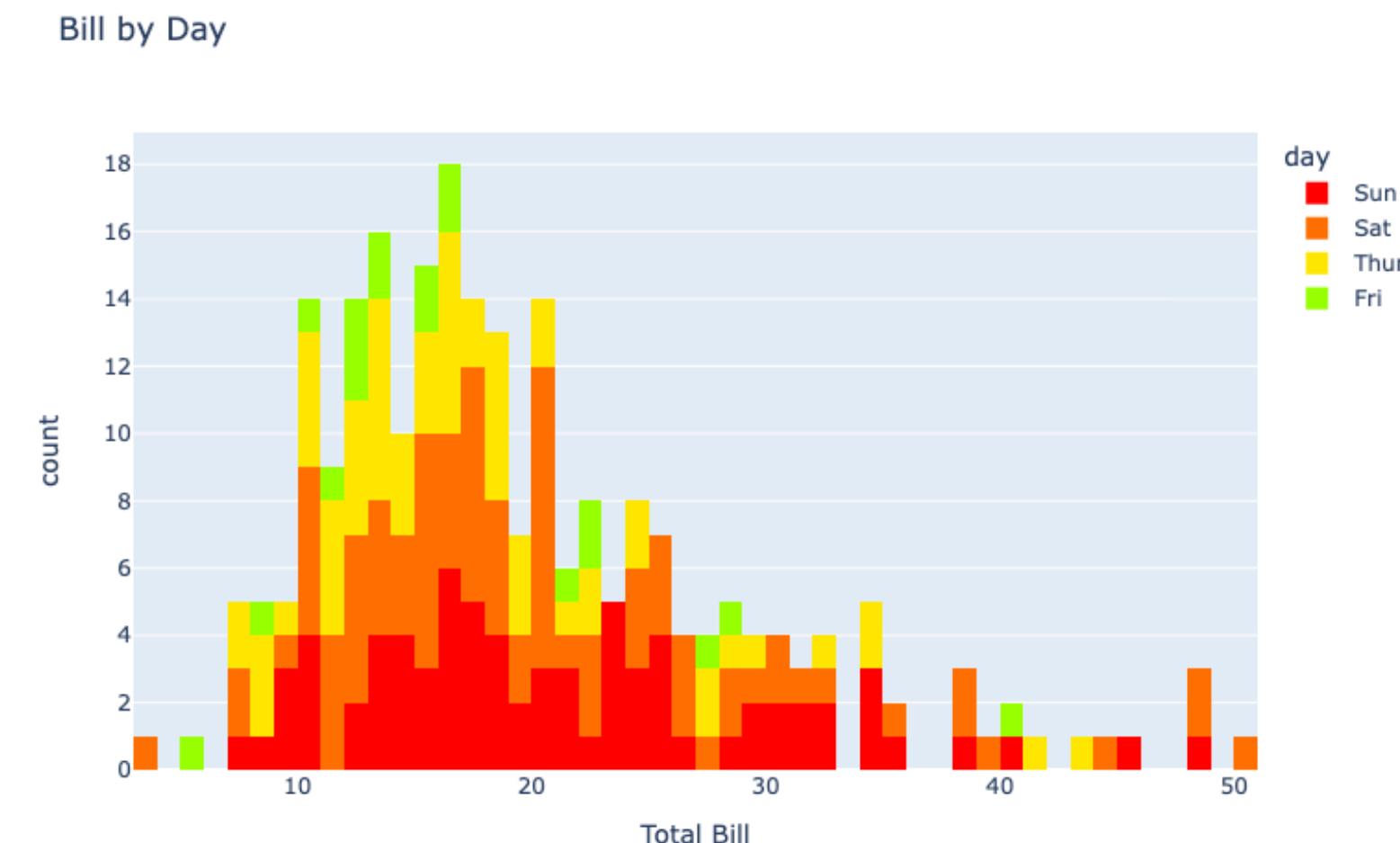
```
# Create a histogram using plotly function `histogram`  
fig = px.histogram(tips_dataset,  
                    dataset, #<- set  
                    x="total_bill", #<- set x  
                    variable, #<- set  
                    grouping, #<- add  
                    column_names)  
fig.show()
```



Histograms (cont'd)

- While that is a gorgeous plot, we might want to make some changes to make it look more professional
- Let's create the plot to measure bill on a daily basis and add the labels, the color scheme and the number of bins

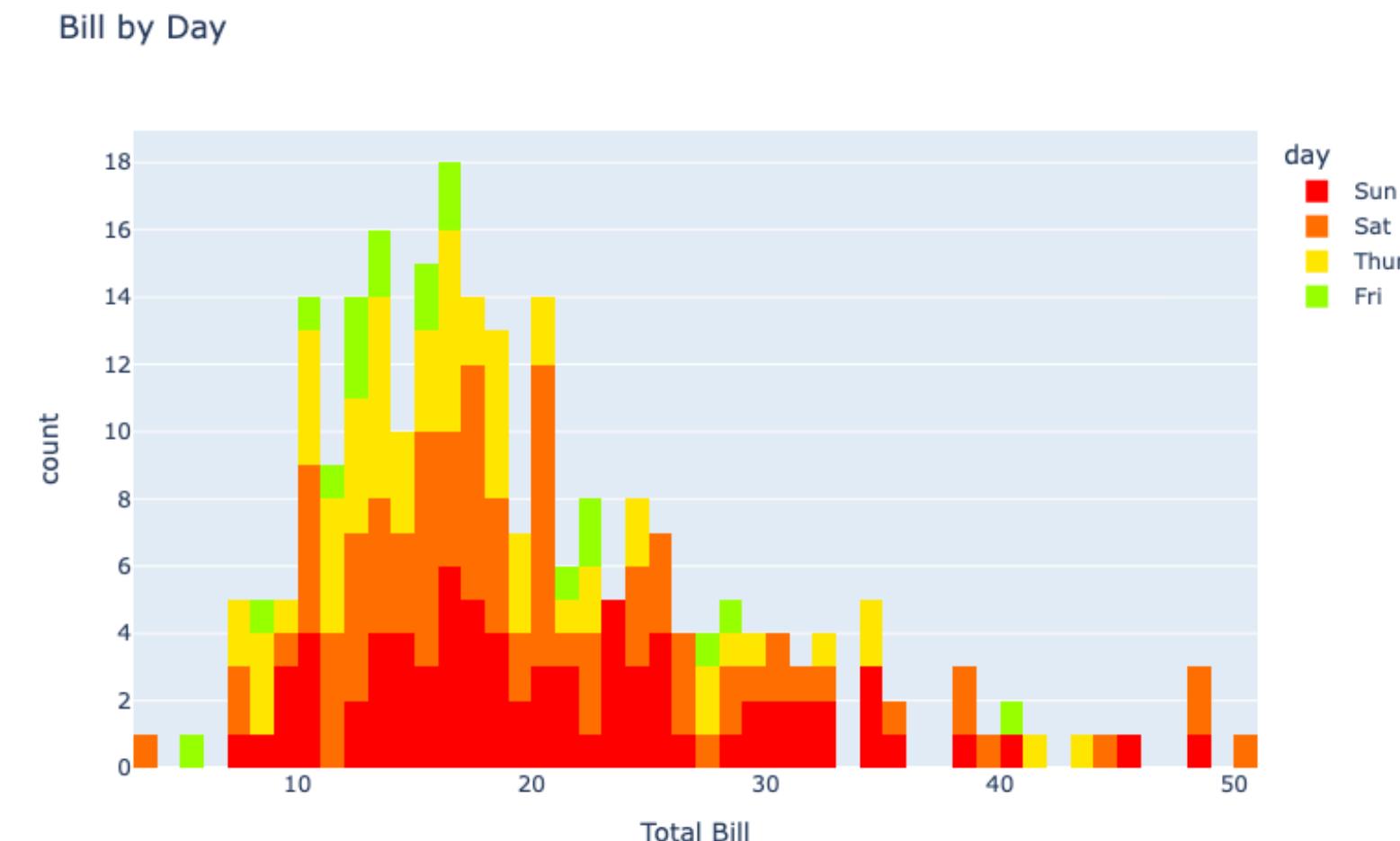
```
fig = px.histogram(tips_dataset,
                    x="total_bill",
                    color="day",
                    hover_data = tips_dataset.columns,
                    title='Bill by Day',
                    labels={'total_bill': 'Total Bill'},
                    nbins=50,
                    color_discrete_sequence =
px.colors.sequential.Rainbow_r)
fig.show()
```



Histograms (cont'd)

- Try the following:
 - toggling the different color boxes in the legend
 - hovering over a bin, drawing a box with your mouse on the plot
 - clicking the reset axes button

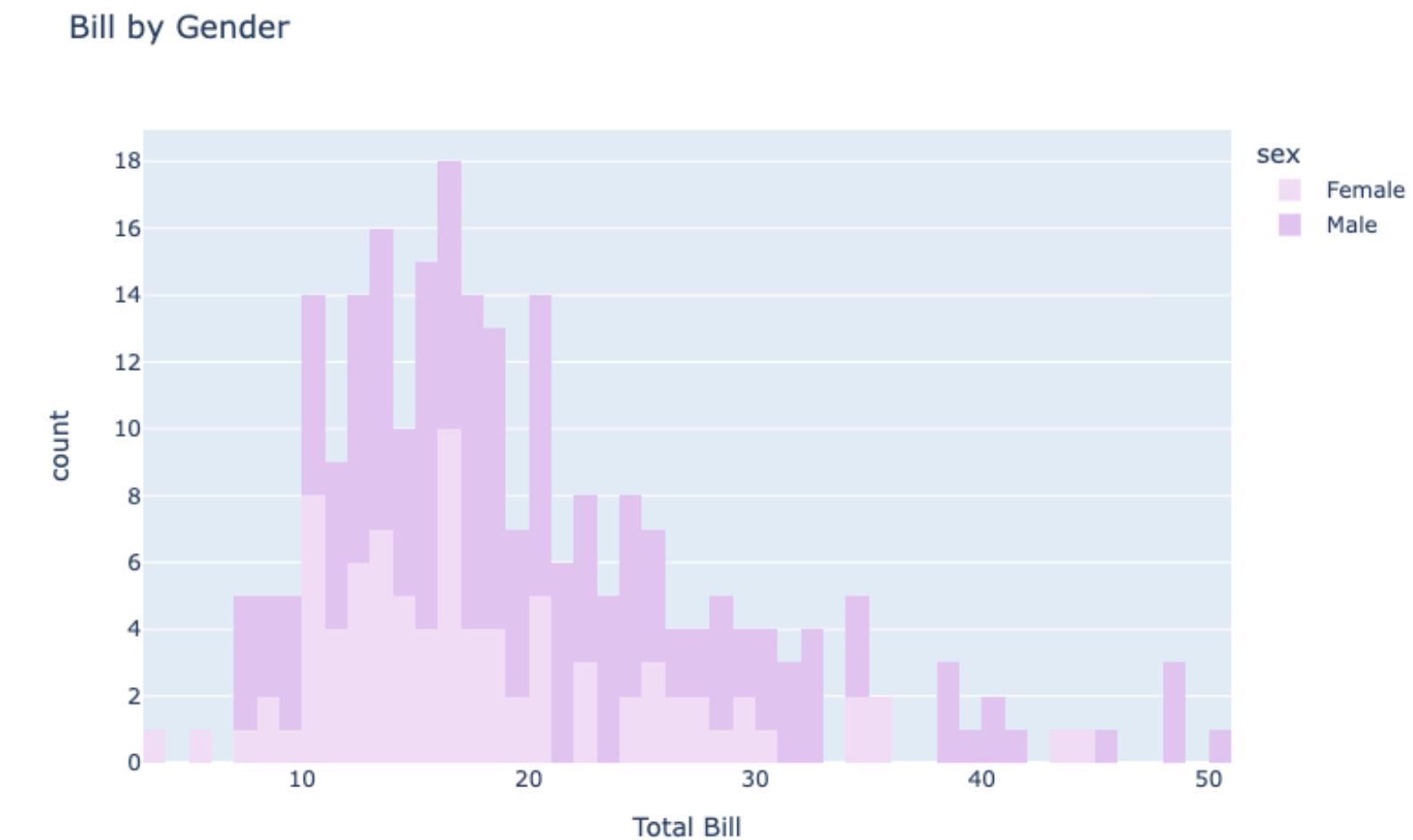
```
fig = px.histogram(tips_dataset,
                    x="total_bill",
                    color="day",
                    hover_data = tips_dataset.columns,
                    title='Bill by Day',
                    labels={'total_bill': 'Total Bill'},
                    nbins=50,
                    color_discrete_sequence =
px.colors.sequential.Rainbow_r)
fig.show()
```



Histograms (cont'd)

- Let's create histogram for gender wise total bill and try different `plotly` colors to build histograms
- For more info on colors in `plotly`, click [here](#)

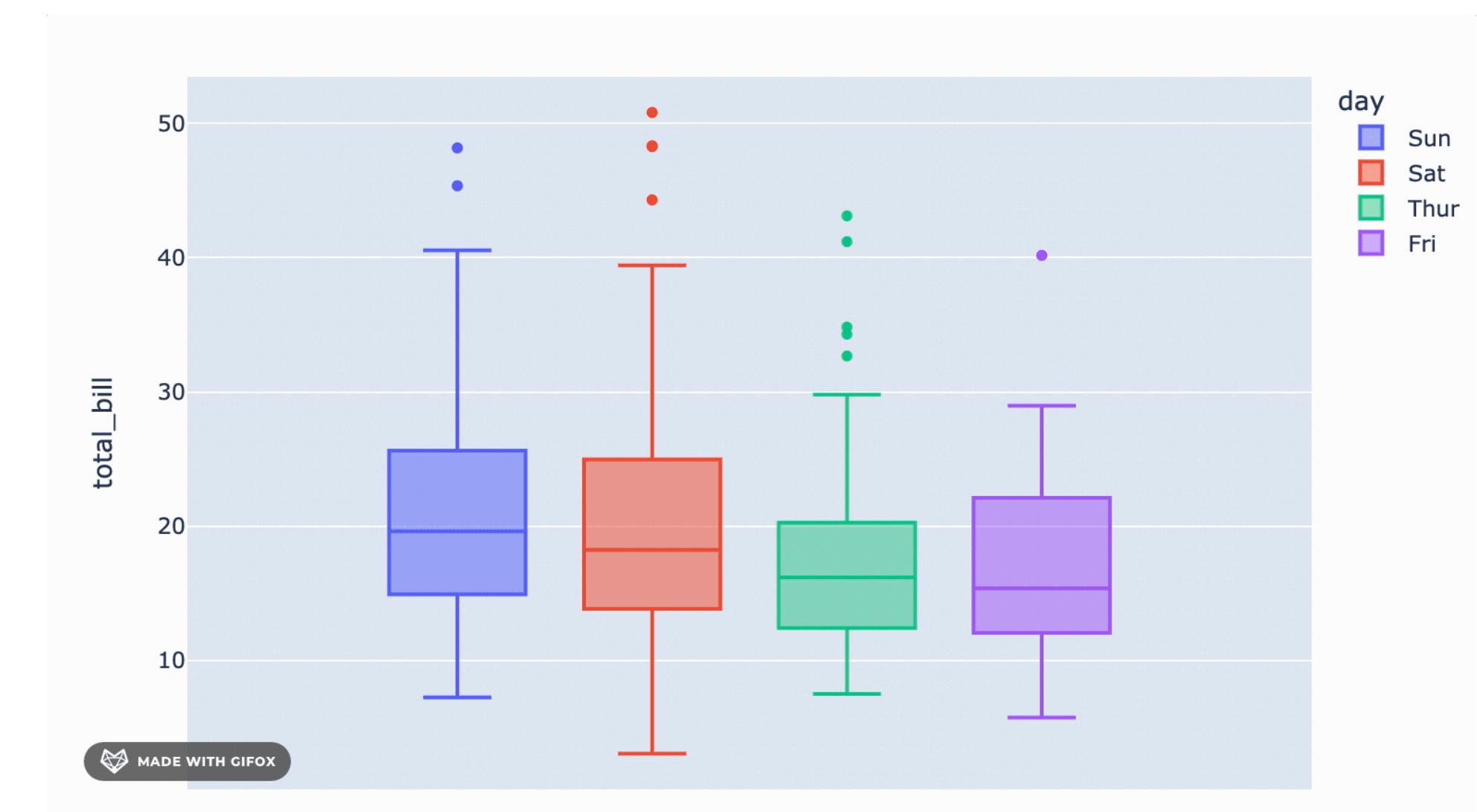
```
fig = px.histogram(tips_dataset,
                    x="total_bill",
                    color="sex",
                    hover_data = tips_dataset.columns,
                    title='Bill by Gender',
                    labels={'total_bill': 'Total Bill'},
                    nbins=50,
                    color_discrete_sequence =
px.colors.sequential.Purp)
fig.show()
```



Box plots

- With interactive box plots, each box plot can be toggled on or off, so all our variables can be visible at once
- We can choose to set the toggle using the `color` attribute
- In this case, we set the `color` attribute to `day` column so we can toggle to analyze the total bill across various days

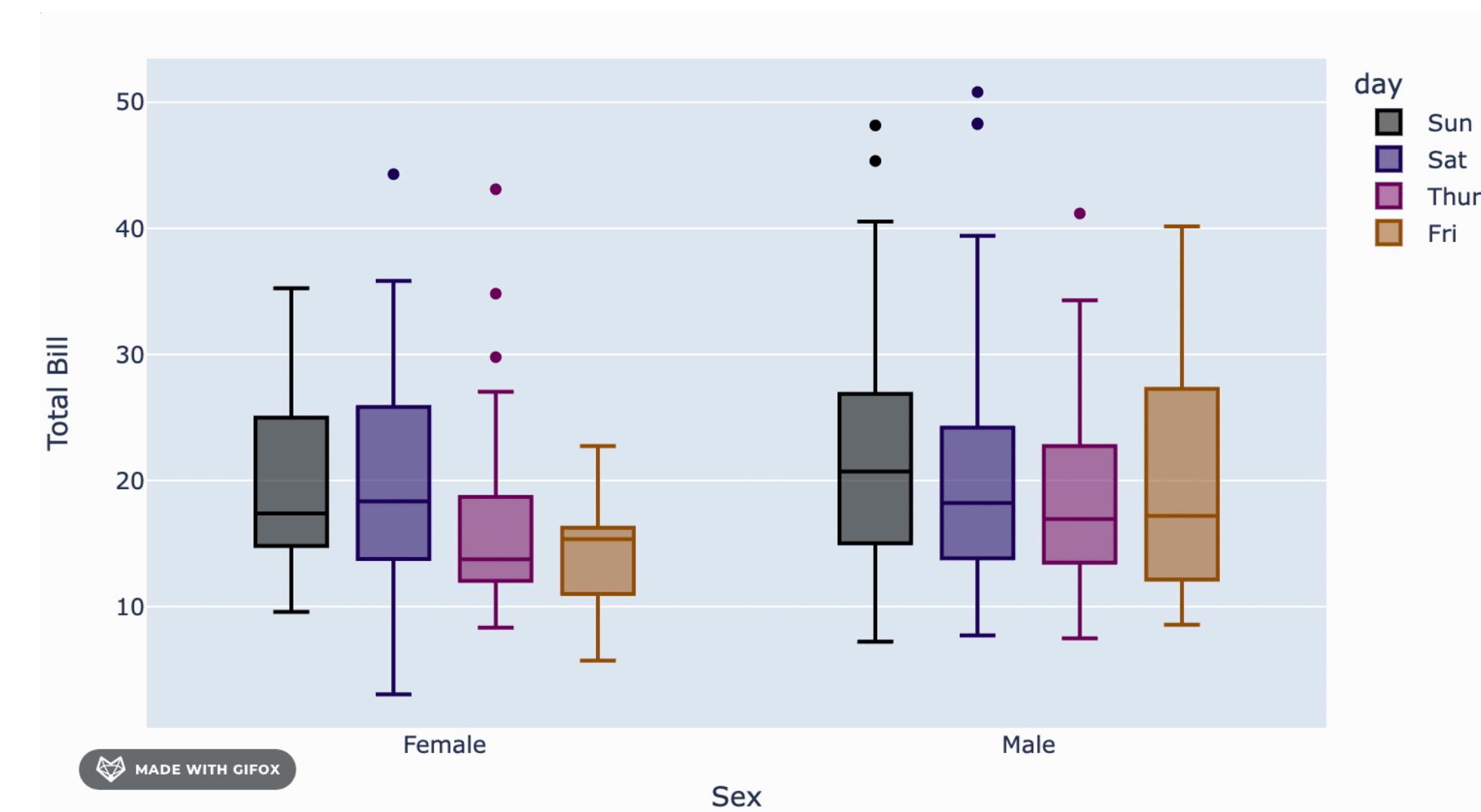
```
# Construct a box plot using plotly function `box`  
fig = px.box(tips_dataset,  
              y="total_bill",  
              color="day")  
fig.show()
```



Box plots (cont'd)

- This set of box plots is lovely, but let's increase the level of visual interest
- What if we changed the color sequence, the title, the labels, and added an extra grouping variable?
- While you can toggle the color variable, you can't toggle the x variable, so be sure which grouping you want to be able to toggle

```
fig = px.box(tips_dataset,
              x='sex',
              y='total_bill',
              color='day',
              labels={'total_bill':'Total Bill', 'sex':'Sex'},
              title='Total Bill grouped by Day and Sex',
              color_discrete_sequence =
px.colors.sequential.Electric)
fig.show()
```

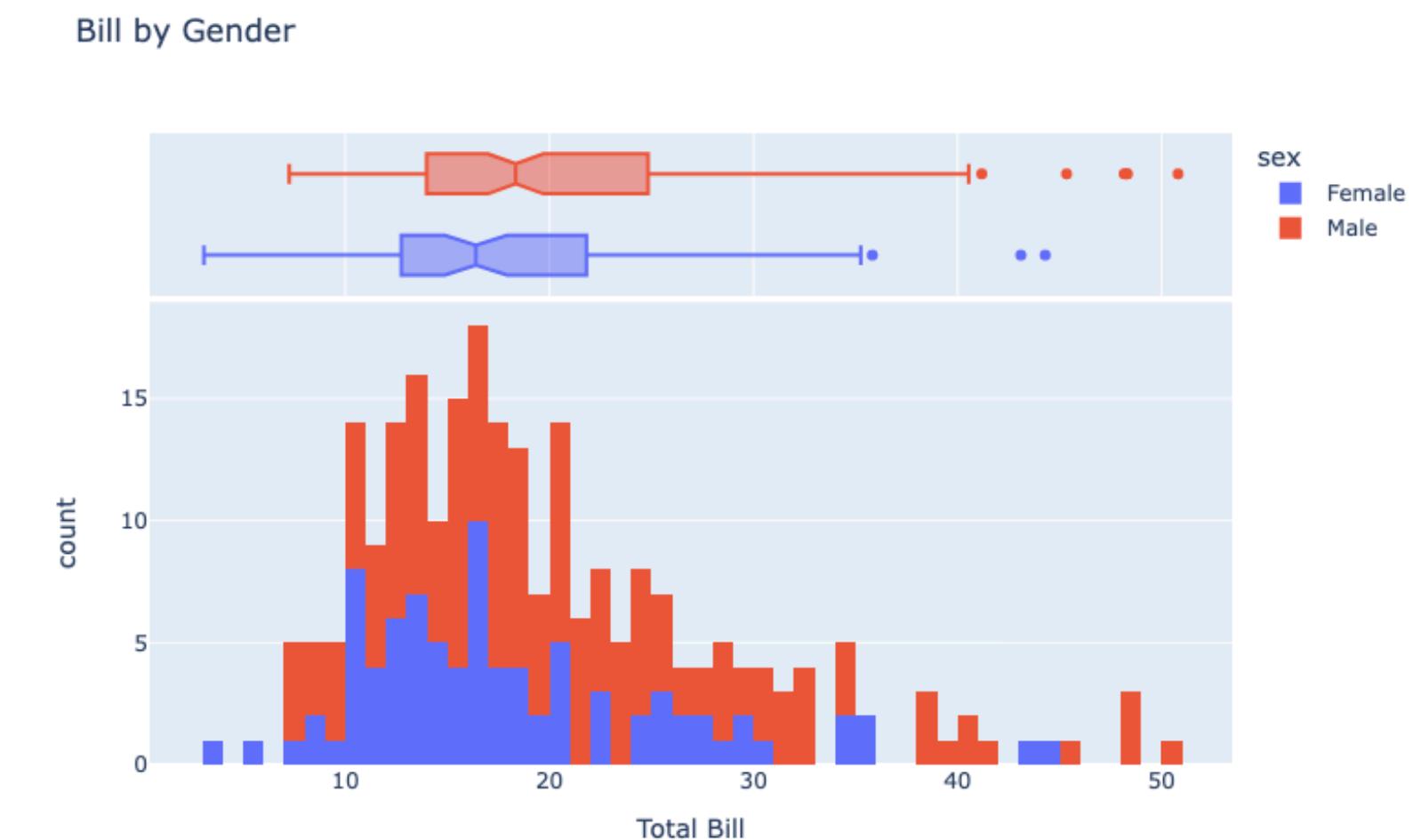


Marginal plots

- Marginal plots are small subplots above or to the right of a main plot, use the argument `marginal` to create small box plot above the main histogram

```
# Construct marginal plot by providing marginal argument
fig = px.histogram(tips_dataset, x="total_bill",
                    color="sex",
                    marginal="box",
                    hover_data = tips_dataset.columns,
                    title='Bill by Gender',
                    labels={'total_bill': 'Total Bill'},
                    nbins=50)

fig.show()
```



Knowledge check



Module completion checklist

Objective	Complete
Define Plotly Express	✓
Describe univariate plots in Plotly Express	✓

Congratulations on completing this module!

You are now ready to try Tasks 1-3 in the Exercise for this topic

