PLOTLY

By:

Team Orange

Venkat Sairam Veeranki

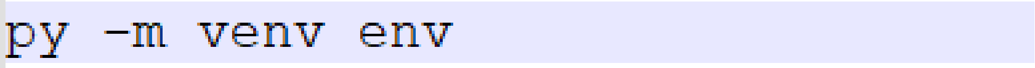
Anandram

Rahul Teja Bolloju

CONTEXT

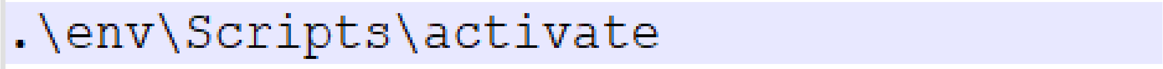
1. Creating Python virtual environment:

- In windows, the machine creates a new folder.
- Open the command prompt and cd~ to the folder path.
- The next step would be to create a python virtual environment in the path, while working with any python projects it's always recommended to create a virtual environment.



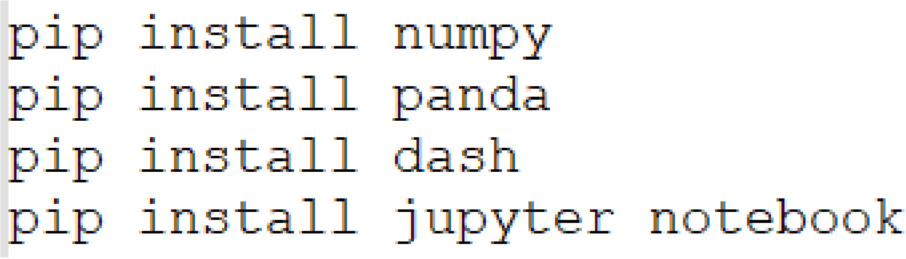
```
py -m venv env
```

- We need to activate the virtual environment,



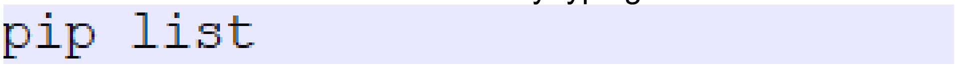
```
.\env\Scripts\activate
```

- Now as a final step we need to install a couple of libraries in the environment if we are not mentioning the versions in pip install it directly takes the latest versions from pypi,



```
pip install numpy
pip install panda
pip install dash
pip install jupyter notebook
```

- We can check the installed libraries by typing the command

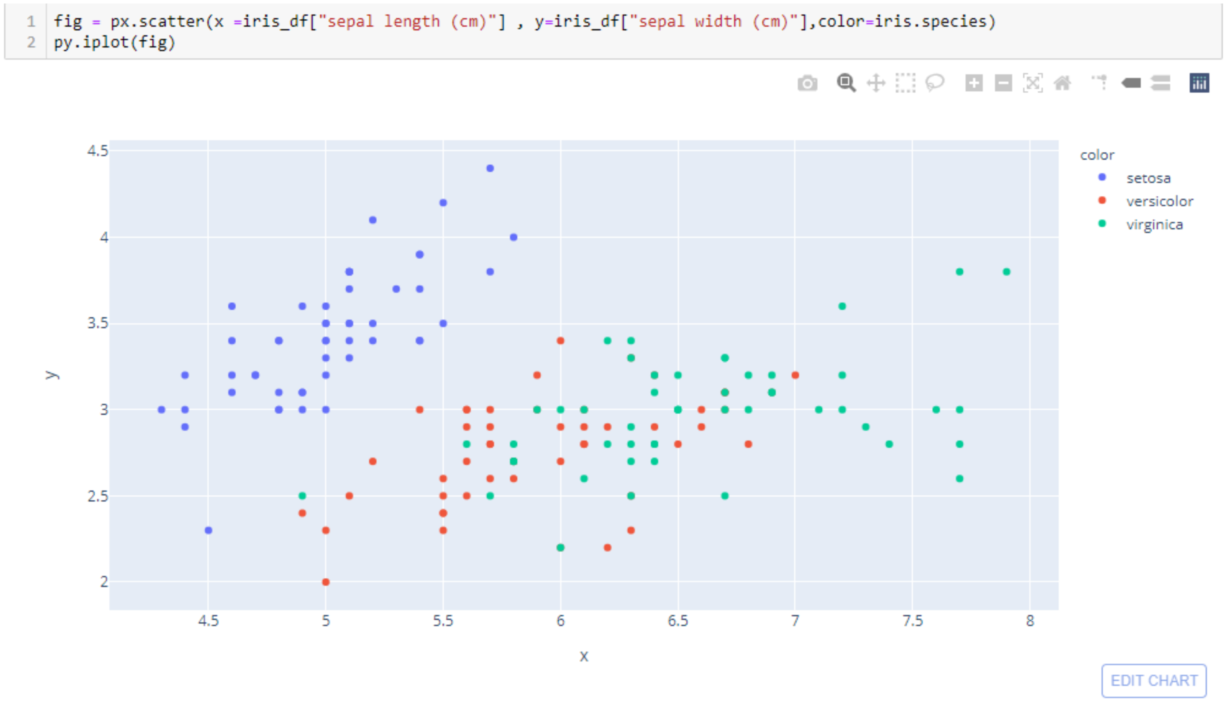


```
pip list
```

- As a final step, we can open Jupiter notebook from the environment

2. Online plotting feature

- Create an account in the link plot signup.
- Once we log in from the profile setting we can get the username and api_key values.
- Using these credentials we can connect to a chart studio and project the locally plotted chart into an online chart studio.
- Once we move the plot to the chart studio we can make different types of edits to the plot.
- We can also import the data to the chart studio and create different types of plots.



```
fig = px.scatter(x =iris_df["sepal length (cm)"] , y=iris_df["sepal width (cm)"],color=iris.species)
py.iplot(fig)
```

3. TOOLS AND DATASET

The tools that we have used are Plotly Express tools and Plotly Graph objects.

Plotly Express Tools:

Plotly express tool is a built-in module in the plotly library. These modules help to plot the entire figure at once. There are more than 30 different functionalities that can plot different types of plots. Basic plots such as the dist plot and histogram. 2-dimensional distribution plots such as density heatmap and density contour. Scatter 3D plots and 3D surface plots are the 3-dimensional distribution plots.

Plotly Graph Objects:

The figures or plots that are created by the plotly python library are like tree-like data structures, which are serialized in JSON. there are 3 main features in tree data structures. They are

- Figures
- Data
- Layouts

In figures, we specify the plots that we want to plot, in data we import the dataset on which we want to plot the figures and in layouts, we specify the height, width, and anime for the plots.

Plotly graph objects consider precise data visualization. For example, if we provide a wrong name or key value to the object it rises an error message with a useful description of the problem. This will not happen in the case if the plots are plotted by using normal python libraries.

4. DATASET

As we have already worked on the iris dataset before, we are using the same dataset to plot the graphs. In the iris dataset, there are 160 rows and 6 columns. In columns, there are Id, Sepal Length, Sepal Width, Petal Length, Petal Width, and Species.

We are also using the “Temperature_change_Data.csv" dataset for some particular plots. There are 16756 rows and 4 columns. In columns, they are Country Code, Country Name, year, and tem_change.

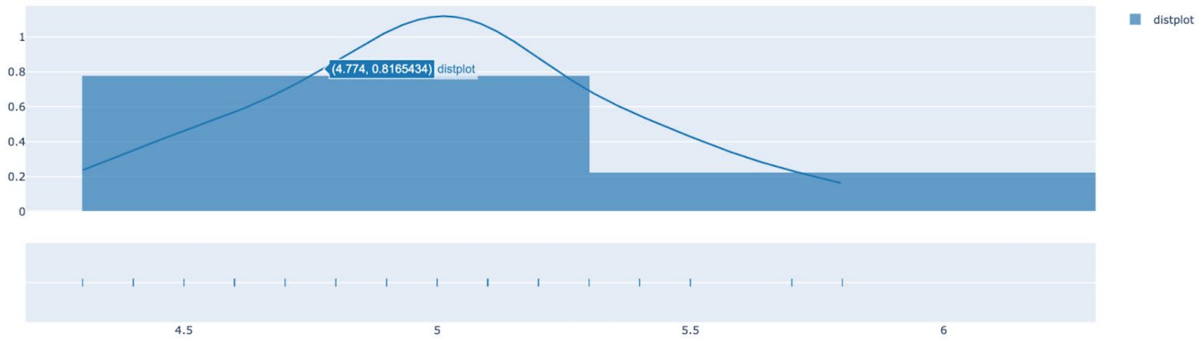
For a better understanding of certain plots, we are creating instance data for plotting those certain plots.

5. PLOTS

5.1. 1Dimensional Distribution

Dist plot:

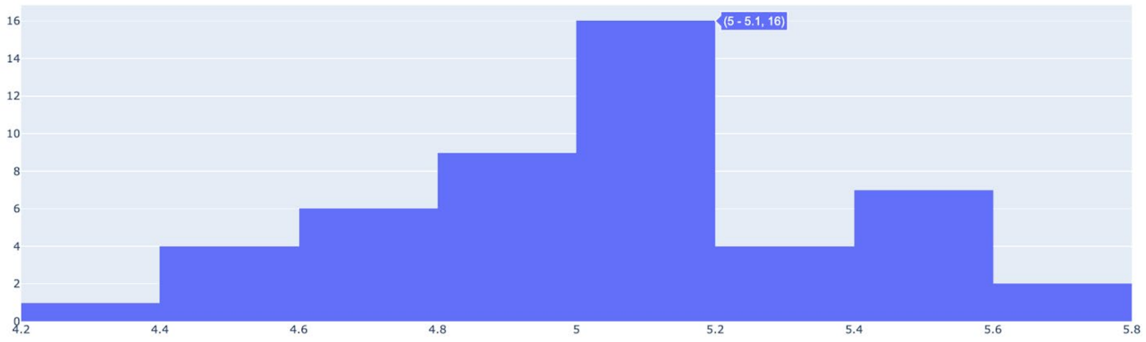
A distribution plot, often known as a distplot, shows how the distribution of the data varies. Comparing distribution and the range of numerical data using the distribution plot is useful.



Histogram:

The histogram displays quantitative or numerical data.

It is utilized to encapsulate discrete or continuous data that are measured using an interval scale.

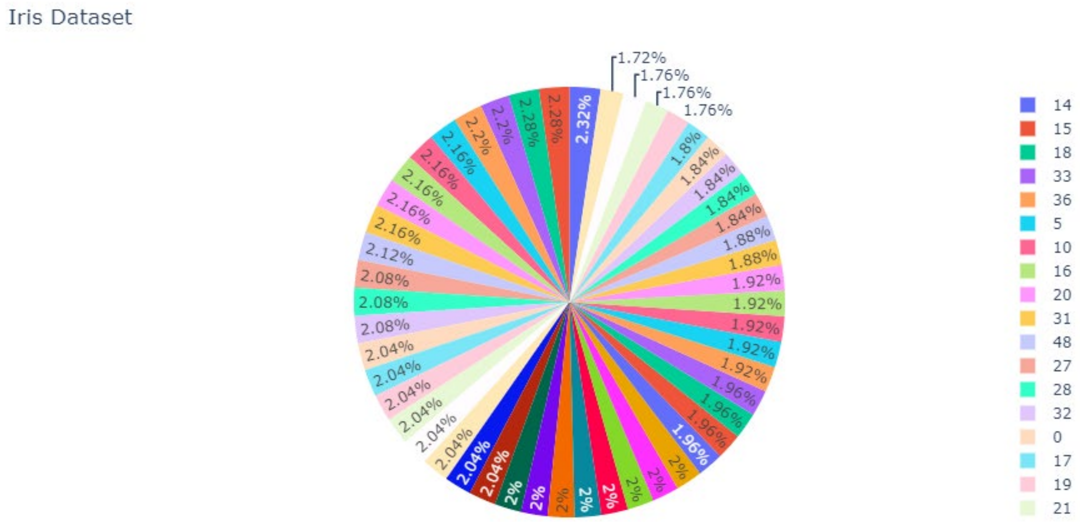


5.2. 2Dimensional Distribution

Pie Chart:

Represents a single series of data. The graph objects have the function go.Pie().

Within this function 2 variables are considered namely names and values. Here, I have taken names as the types of the species and labels to be one of the parameters of measurement being Sepal length.

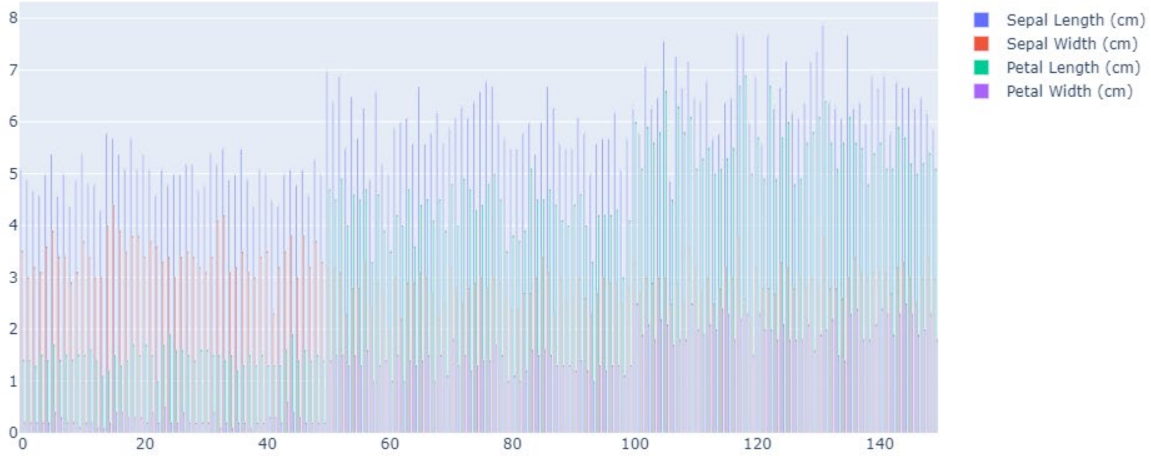


Bar Chart:

Representation of categorical data with the rectangular bars either being heights or widths that is proportional to the data.

Shows a comparison between discrete values.

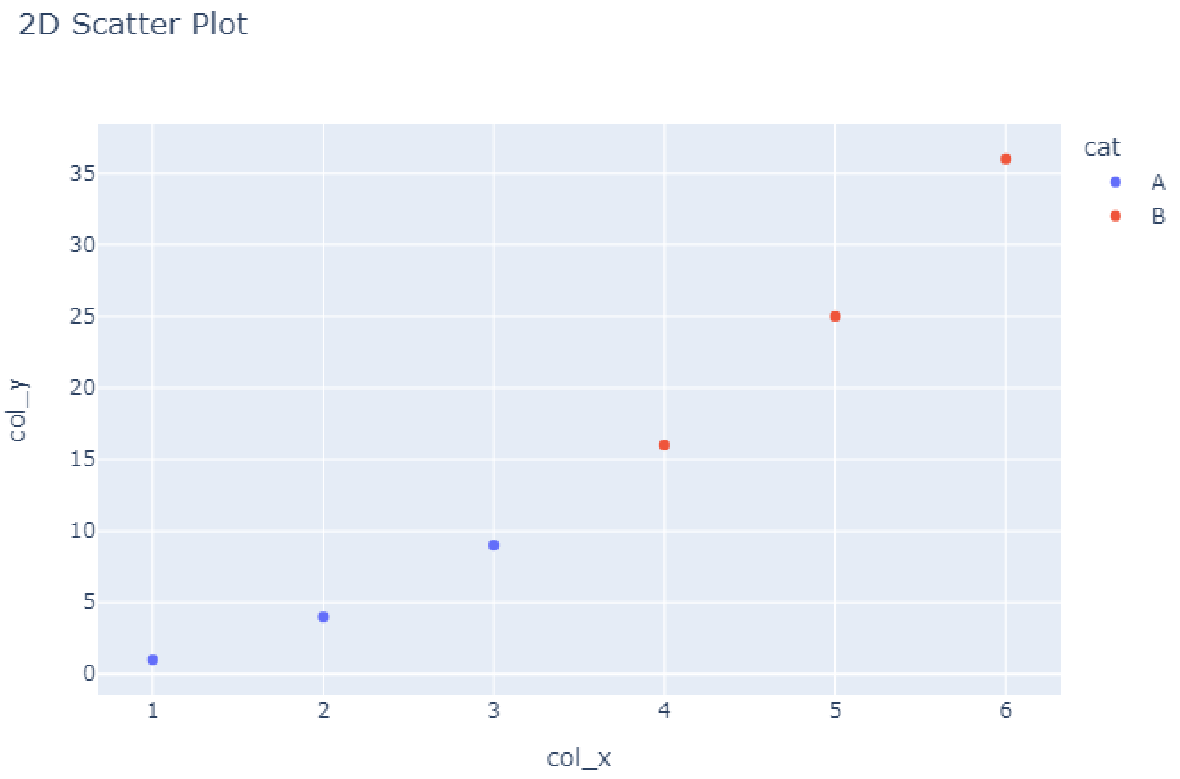
One of the axes represents the values to be compared and the other represents measured value. Here we have utilized the Sepal length as the parameter for analysis along with the other values like petal length, petal width and sepal width.



2D Scatter plot:

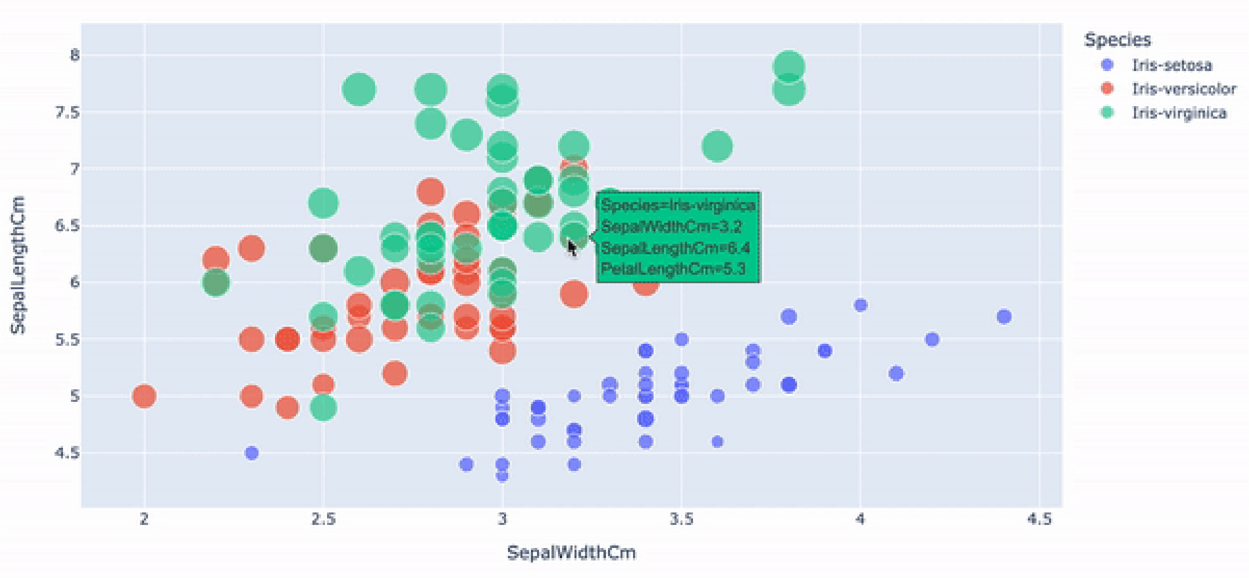
Scatter plots are utilized to know how one parameter affects another. Here I will be taking Species and the sepal and petal parameters to analyze how one is affecting another.

I will be utilizing Sepal length here to understand the correlations of each of them by making use of the go.Scatter.

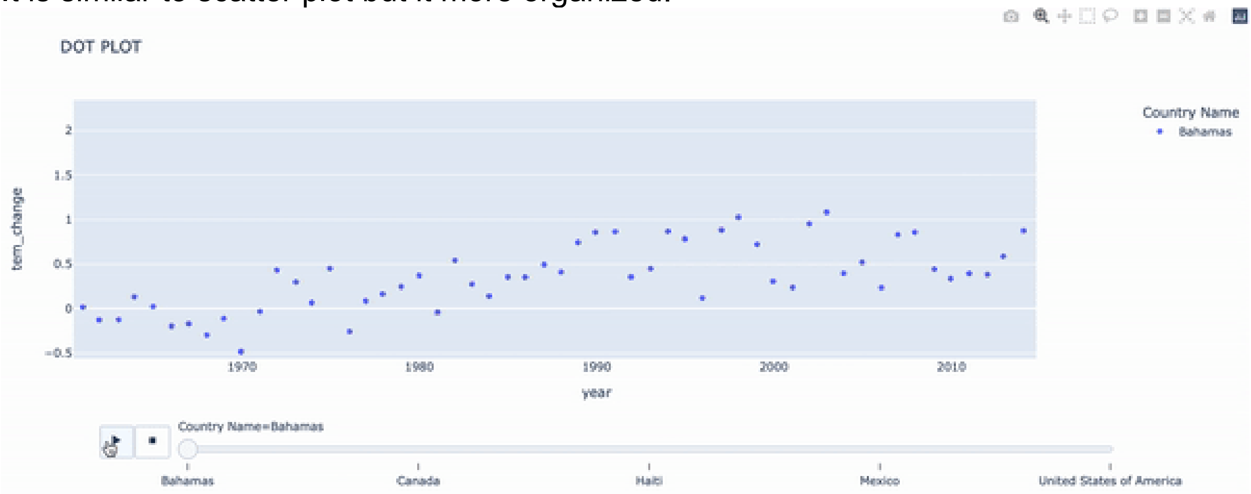


Bubble plot
An extension of the scatter plot used to examine correlations between three numerical variables is a bubble chart, often known as a bubble plot.

If your data consists of three data series, each of which contains a set of values, you may use a bubble chart instead of a scatter chart. The numbers in the third data series are used to calculate the sizes of the bubbles. Financial data is frequently represented using bubble charts.



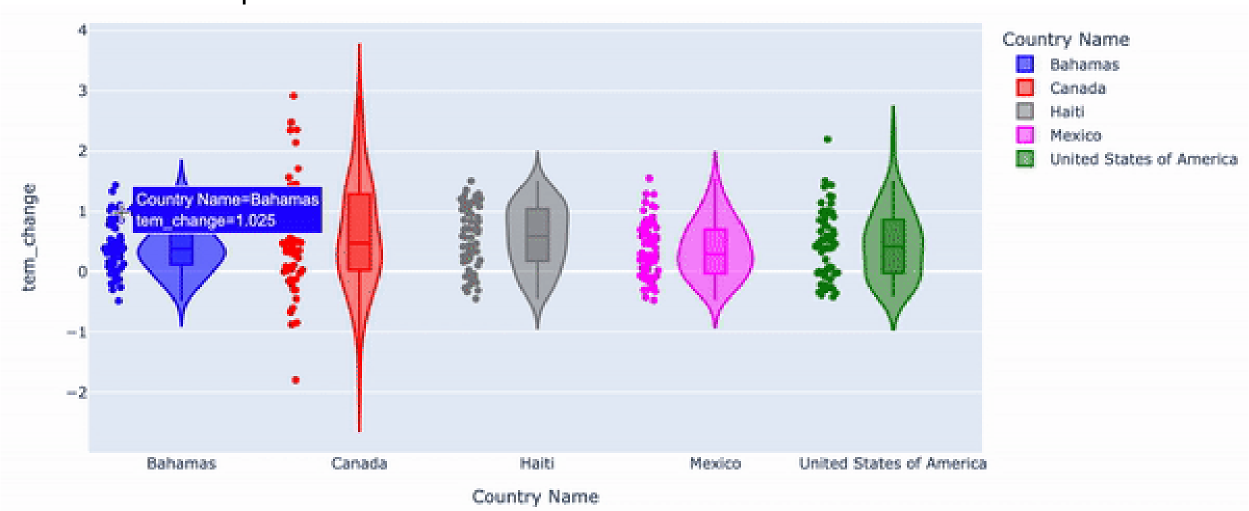
Strip plot
It is also known as strip plot
It is similar to scatter plot but it more organized.



- Visualize relations between any number of attributes.
- Can observe range of data values measured
- Can observe shape of all data points
- Gives more sense when compared with scatter plot
- Helps to observe data more accuracy
- enables to analyze data more clearly, which is a significant and important work as a data scientist

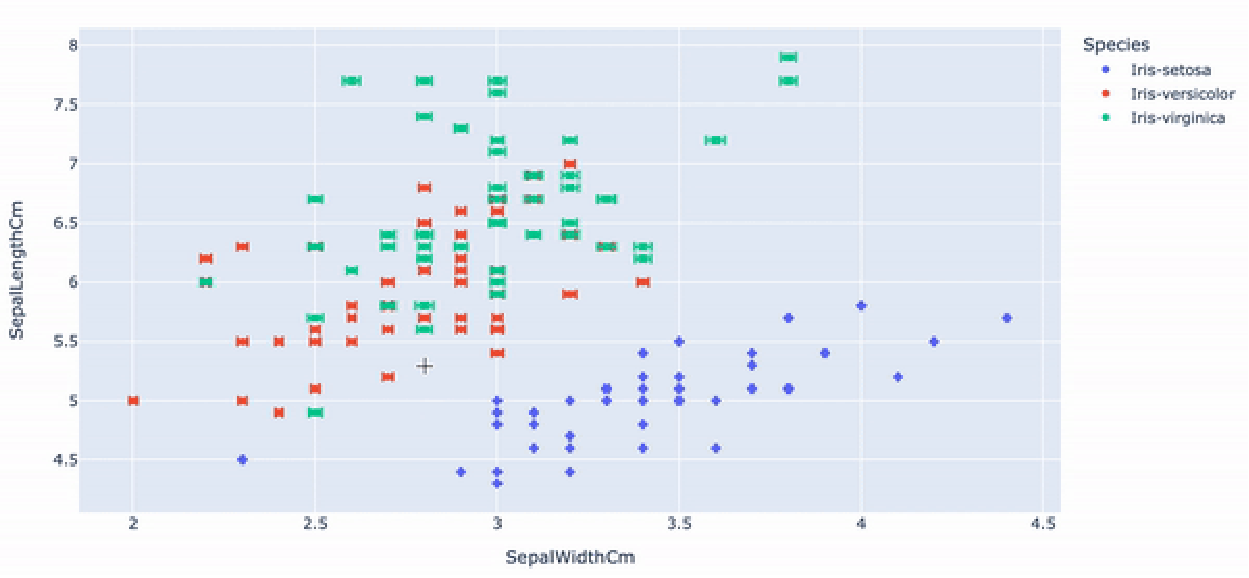
Box and Violin
A violin plot uses density curves to represent the distributions of numerical data for one or more groups.

When analyzing numerical data from different groups, box plots are a popular approach to show how the patterns of the data are distributed.



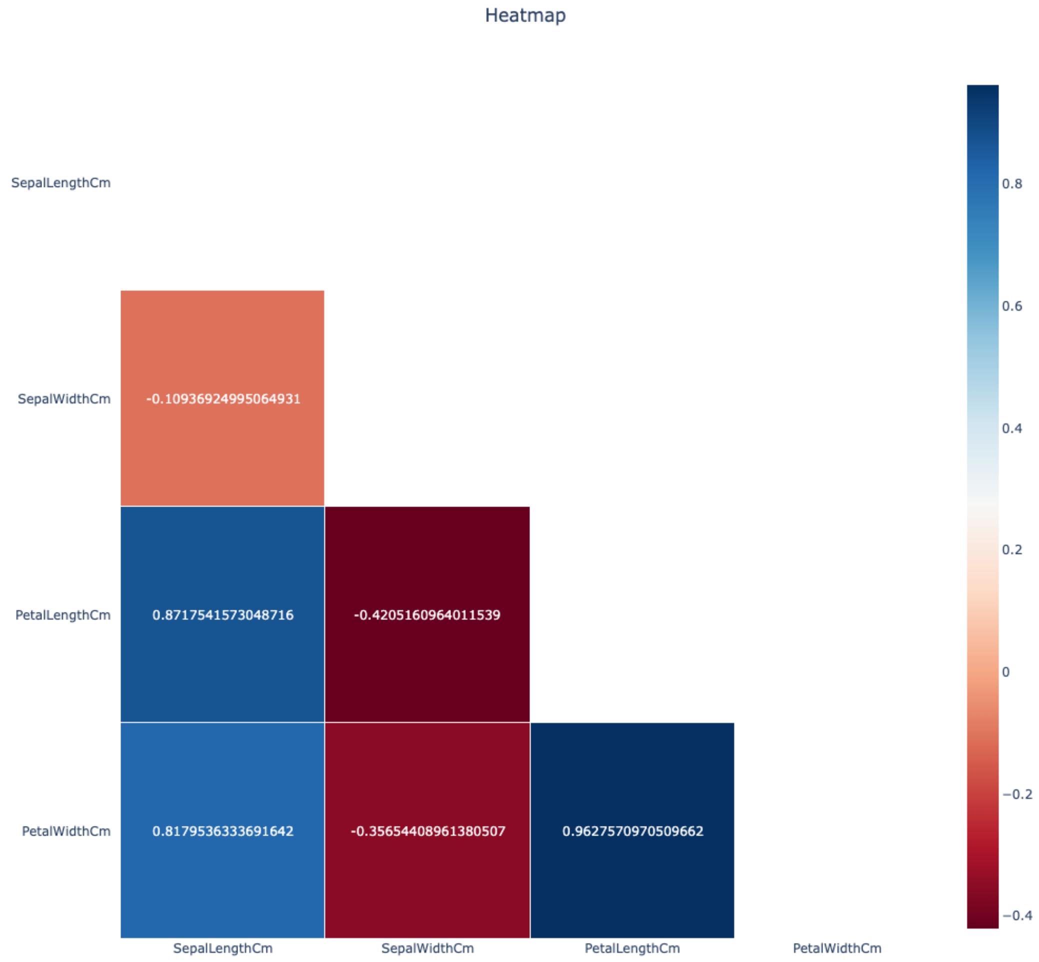
Error plot

Since the data are concentrated and the displayed average value is more likely, a narrow error bar indicates that. A large error bar, on the other hand, suggests that the data are less reliable and more distributed.



Heatmap

Heat Maps are visual displays of data that make use of color coding. Heat Maps' main goal is to make it easier to see how many events are there in a dataset and to help us to focus on the parts of data visualizations that are most important.

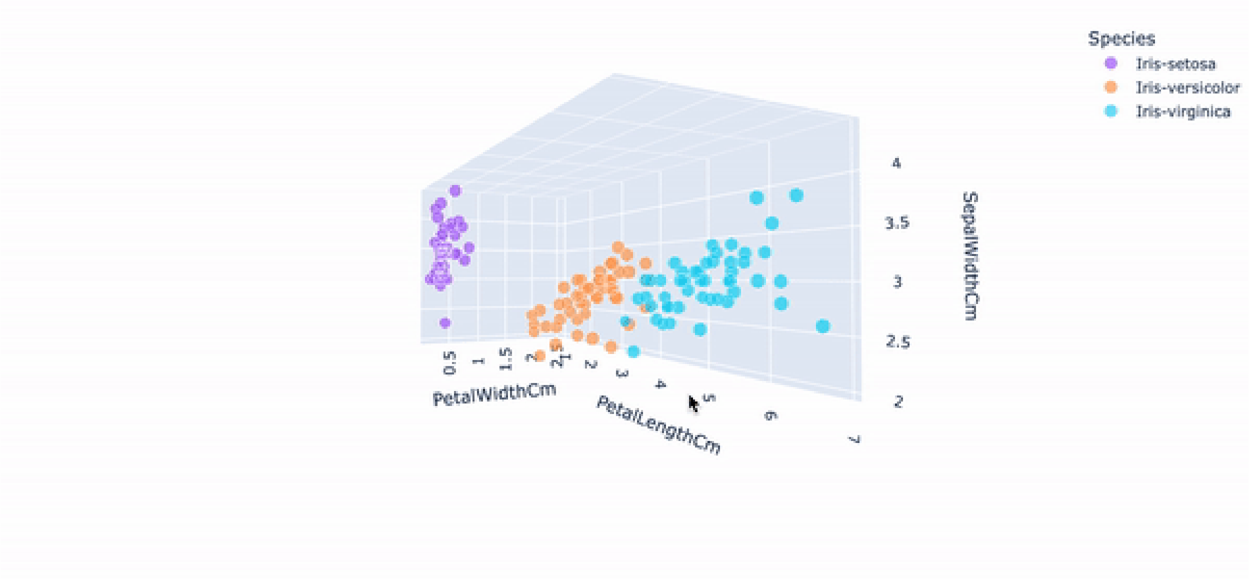


5.3. 3Dimensional Distribution

3D Scatter plot

A 3D scatter plot adds depth to your scatter plot by allowing you to examine both on x and y axis, as well as their relationship in a third dimension (z-axis).

One of the most popular types of charts used in business intelligence reports is the scatter plot chart.

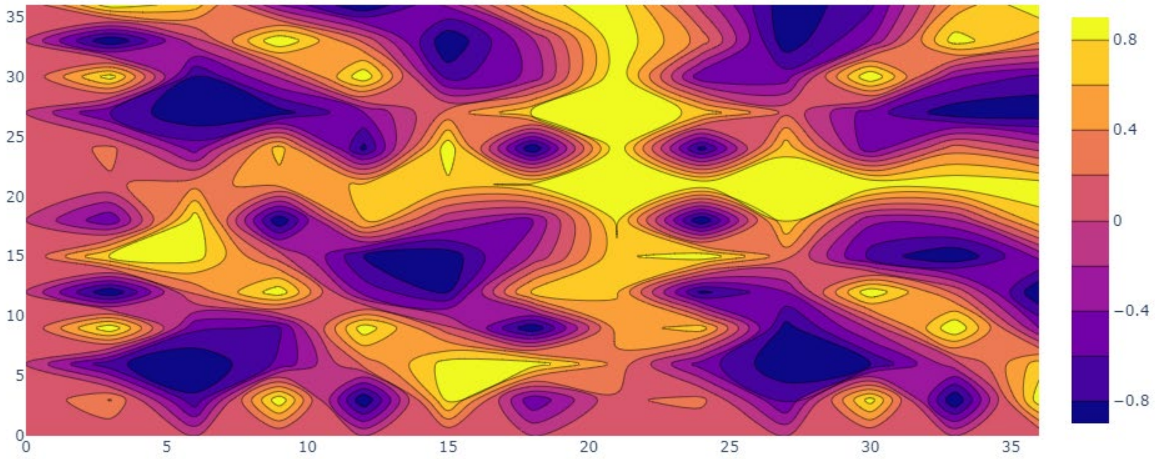


5.4. Contour

The contour plot represents the 2D contour lines of the 2-dimensional numerical array Z. More precisely it can be understood as slicing of Z values while visualizing the 3D format of contours.

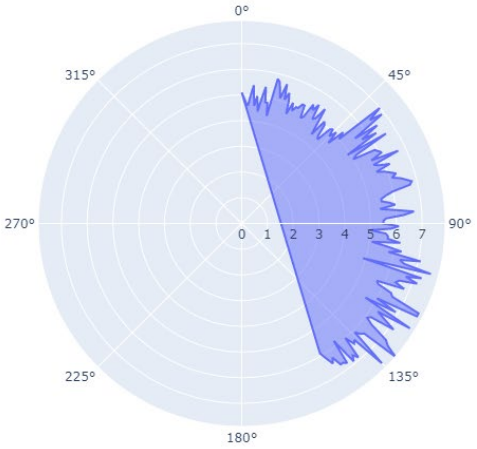
Assuming a function Z = f (X ,Y) (where X and Y represents the horizontal axes)are represented in a meshgrid and then given to a function with the go.Contour command.

In the tutorial we have created a sine function along with the generated data.



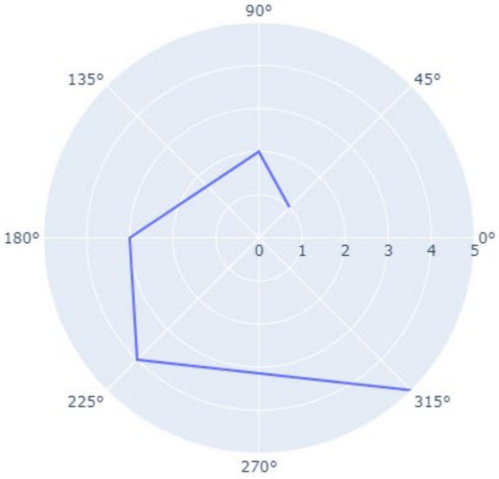
5.5. Polar

This type of charts represents a closed curve which is connected by points in a coordinate system.Data is represented on the radial and the angular axes and therefore, r and theta are the arguments for go.Scatterpolar(). Here, I have assumed the values of r to be the sepal and petal values while theta being the types of the species.



5.6. Radar

It is in similar lines with polar charts, where the data is represented quantitively which has the variables originating from the center. I am using plotly express here with the same parameters but visually better.


90°
45°
0°
315°
270°
225°
180°
135°
0
1
2
3
4
5