**Data visualization:** Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from.

The main goal of data visualization is to make it easier to identify patterns, trends, and outliers in large data sets. The term is often used interchangeably with others, including information graphics, information visualization, and statistical graphics.

Data visualization is important because Data visualization is

* The ability to absorb information quickly, improve insights and make faster decisions;
* An increased understanding of the next steps that must be taken to improve the organization;
* An improved ability to maintain the audience's interest with the information they can understand;
* An easy distribution of information that increases the opportunity to share insights with everyone involved;
* Eliminate the need for data scientists since data is more accessible and understandable; and
* An increased ability to act on findings quickly and, therefore, achieve success with greater speed and fewer mistakes.

There are also a lot of benefits of data visualization. Data visualization positively affects an organization’s decision-making process with interactive visual representations of data. Businesses can now recognize patterns more quickly because they can interpret data in graphical or pictorial forms.

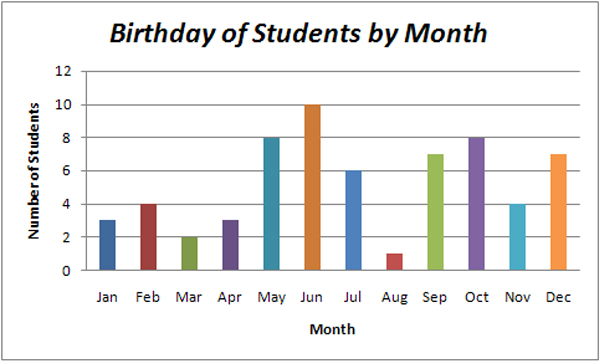
Here are some more specific ways that data visualization can benefit an organization:

* Correlations in Relationships: Without data visualization, it is challenging to identify the correlations between the relationship of independent variables. By making sense of those independent variables, we can make better business decisions.
* Trends Over Time: While this seems like an obvious use of data visualization, it is also one of the most valuable applications. It’s impossible to make predictions without having the necessary information from the past and present. Trends over time tell us where we were and where we can potentially go.
* Frequency: Closely related to trends over time is frequency. By examining the rate, or how often, customers purchase and when they buy gives us a better feel for how potential new customers might act and react to different marketing and customer acquisition strategies.
* Examining the Market: Data visualization takes the information from different markets to give you insights into which audiences to focus your attention on and which ones to stay away from. We get a clearer picture of the opportunities within those markets by displaying this data on various charts and graphs.
* Risk and Reward: Looking at value and risk metrics requires expertise because, without data visualization, we must interpret complicated spreadsheets and numbers. Once information is visualized, we can then pinpoint areas that may or may not require action.
* Reacting to the Market: The ability to obtain information quickly and easily with data displayed clearly on a functional dashboard allows businesses to act and respond to findings swiftly and helps to avoid making mistakes.

There are many types of graphs we can use to represent data. They are as follows,

* **Bar graph**

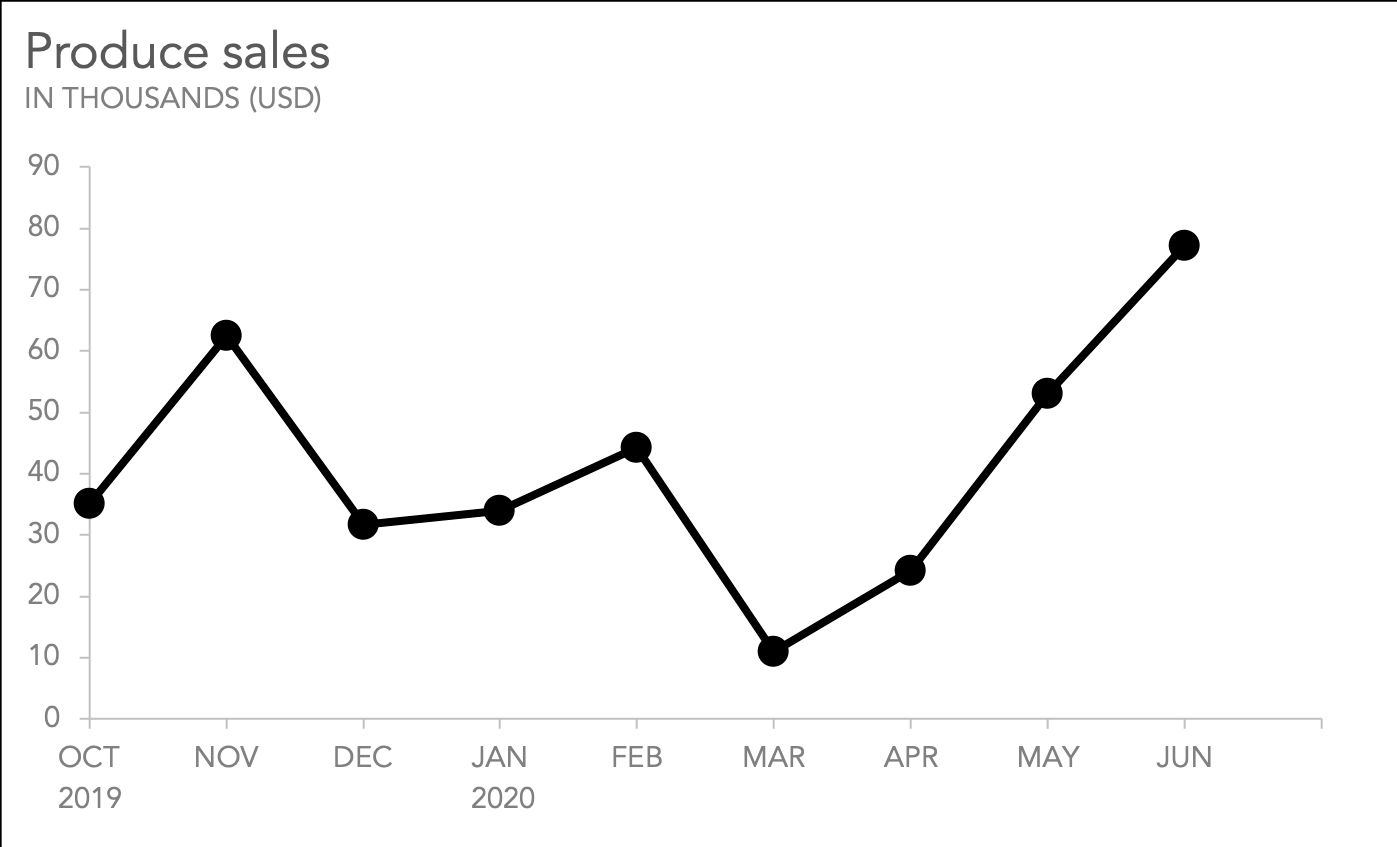
A bar graph or chart is a way to represent data by rectangular column or bar. The heights or length of the bar is proportional to the values. *A bar graph represents categorical variables.*



* **Line graph**

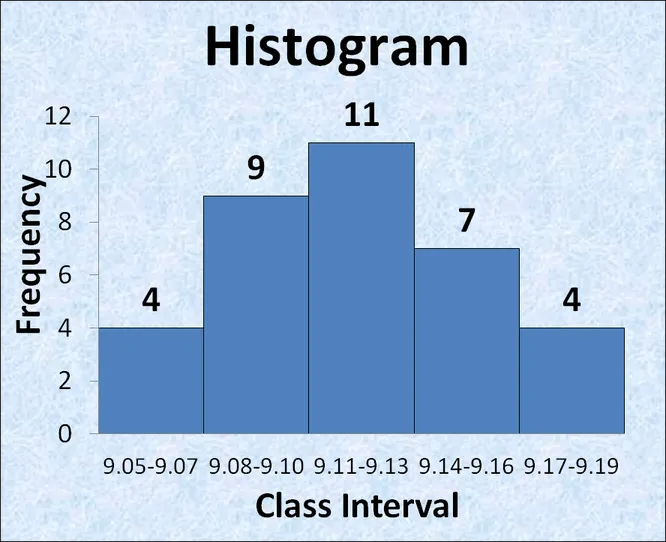
A line graph is a type of graph where the information or data is plotted as some dots which are known as markers and then they are added to each other by a straight line.

*The line graph is normally used to represent the data that changes over time.*

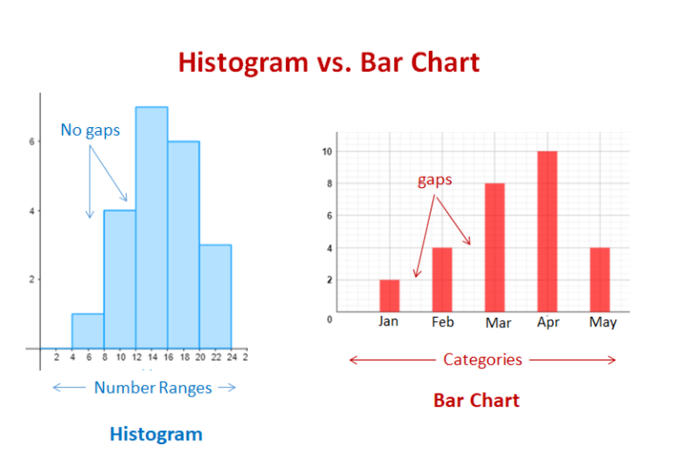
****

* **Histogram**

A histogram graph is a graph where the information is represented along with the height of the rectangular bar. Though it does look like a bar graph, there is a fundamental difference between them.

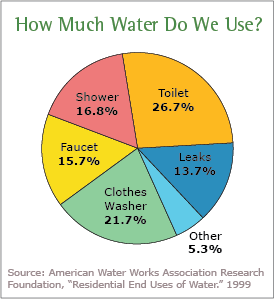


*With the histogram, each column represents a range of quantitative data when a bar graph represents categorical variables.*



* **Pie chart**

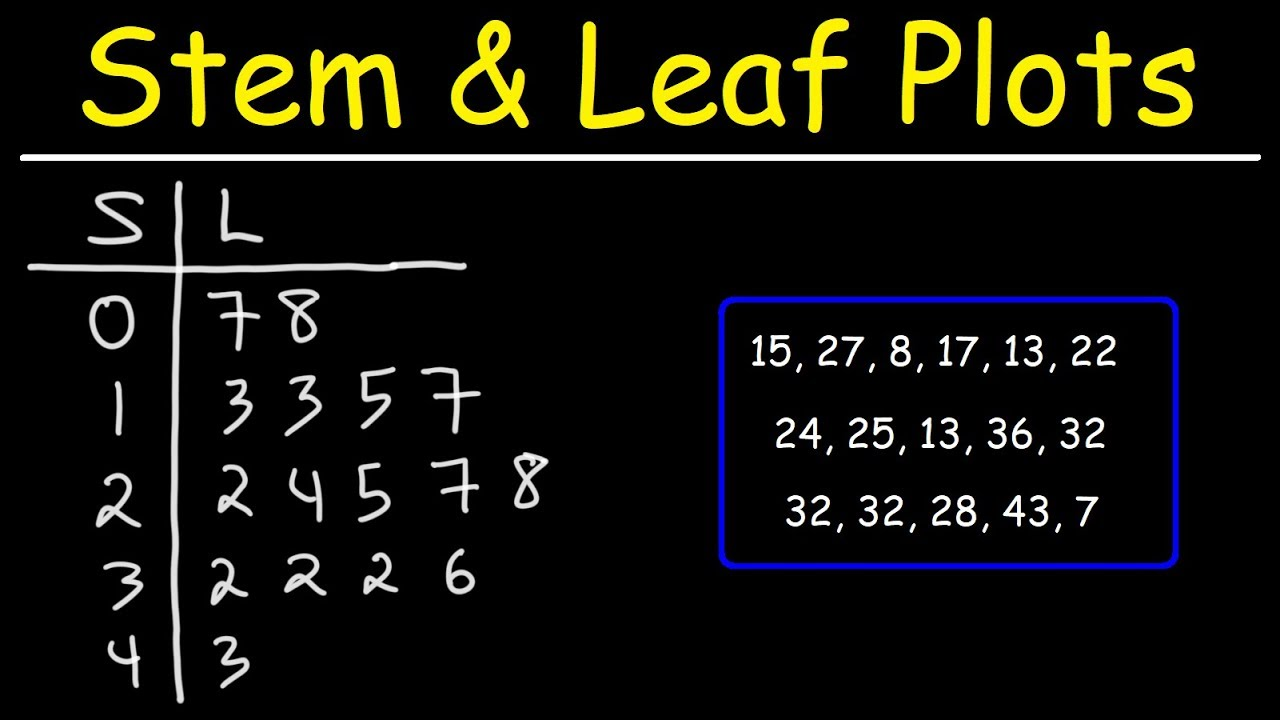
The other name of the pie chart is a circle graph. It is a circular chart where numerical information represents as slices or in fractional form or percentage where the whole circle is 100%.



* **Stem and leaf plot**

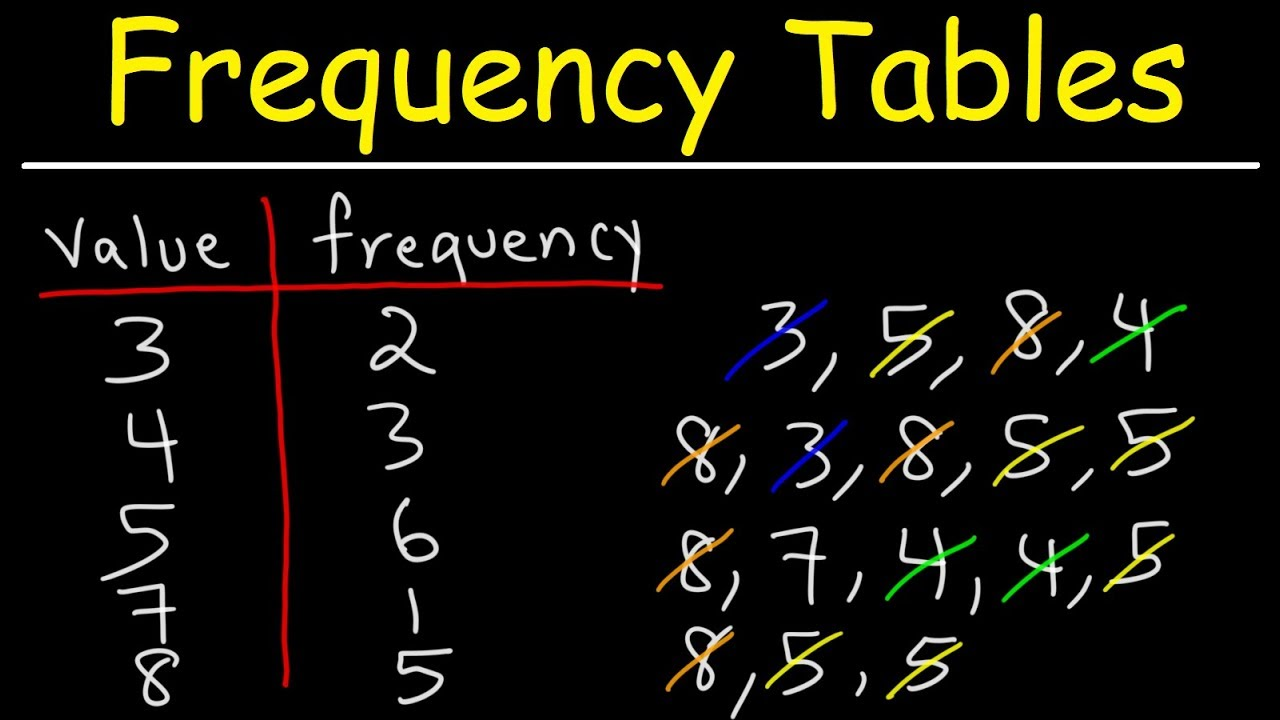
The stem and leaf plot is a way to represents quantitative data according to frequency ranges or frequency distribution.

In the stem and leaf plot, each data is split into stem and leaf, which is 32 will be split into 3 stems and 2 leaves.



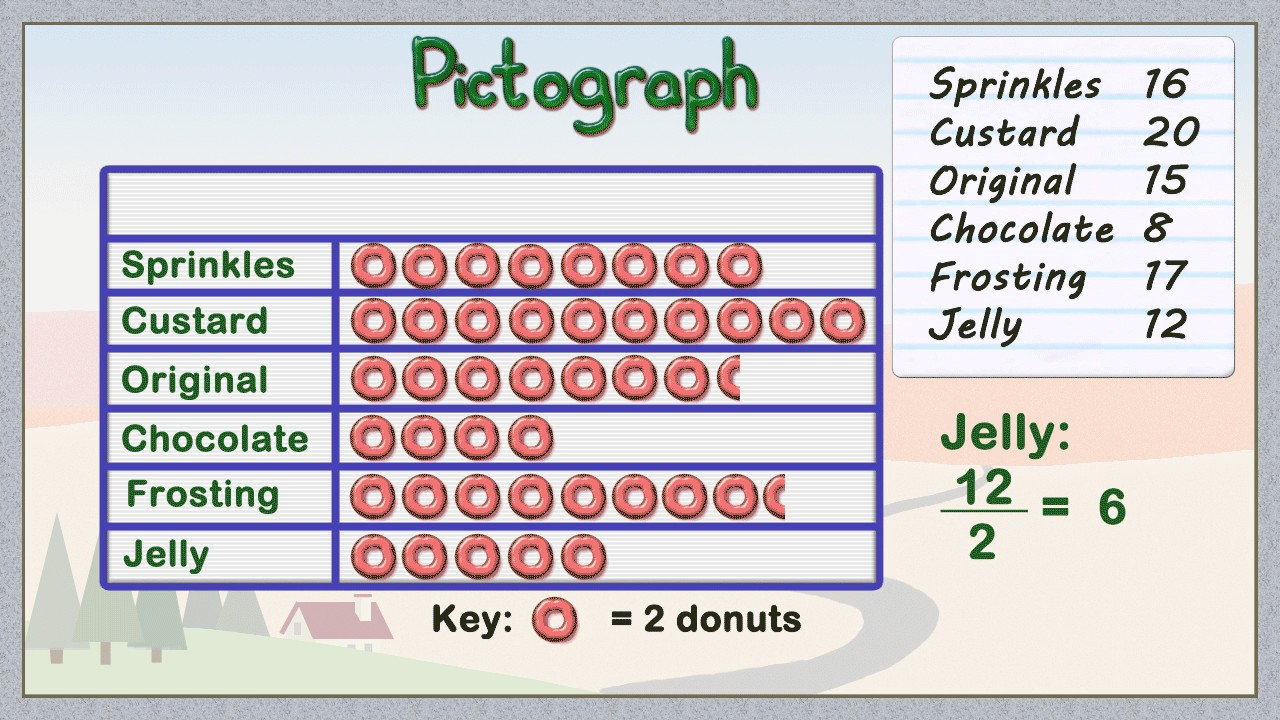
* **Frequency table**

Frequency means the number of occurrences of an event. A frequency distribution table is a graph or chart which shows the frequency of events. It is denoted as **‘f’**.



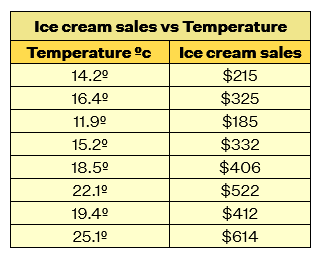
* **Pictograph**

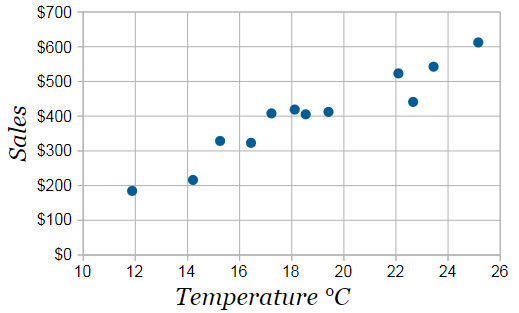
Pictograph or Pictogram is the earliest way to represents data in a pictorial form or by using symbols or images. And each image represents a particular number of things.



* **Scatter diagrams**

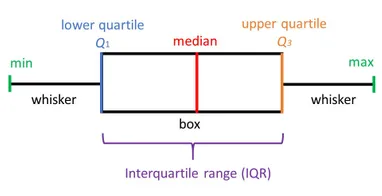
Scatter diagram or scatter plot is a way of graphical representation by using cartesian coordinates of two variables. The plot shows the relationship between two variables. Below there is a data table as well as a Scattergram as per the given data.

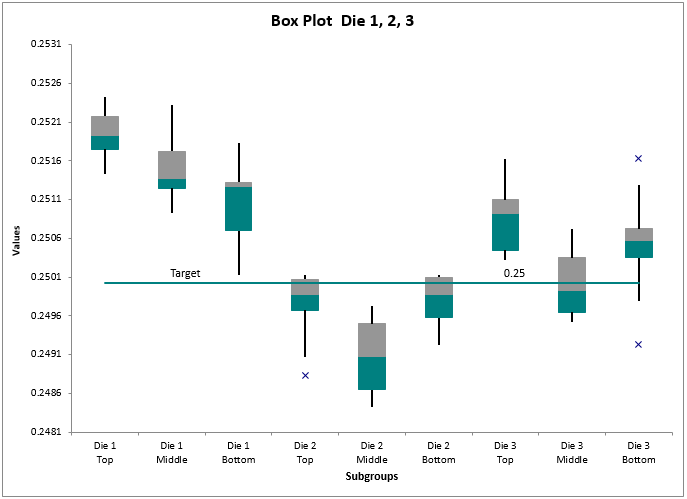




* **Box plots**

In descriptive statistics, a box plot or boxplot is a method for graphically depicting groups of numerical data through their quartiles. Box plots may also have lines extending from the boxes indicating variability outside the upper and lower quartiles, hence the terms box-and-whisker plot and box-and-whisker diagram.





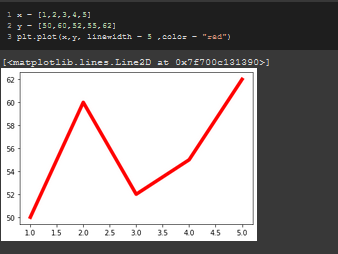
**Line plots**

* **.plot()**

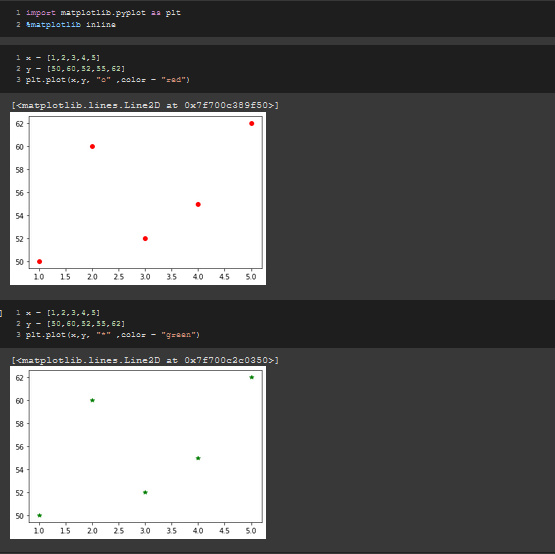
**Syntax:**



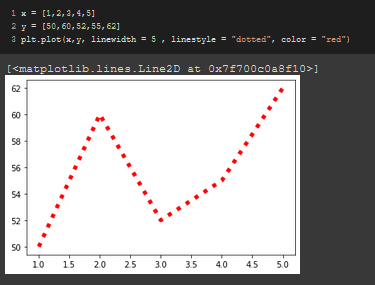
If nothing else is mentioned, It’s usually a line plot. We can also control the width of the line.



We can also use different markers.

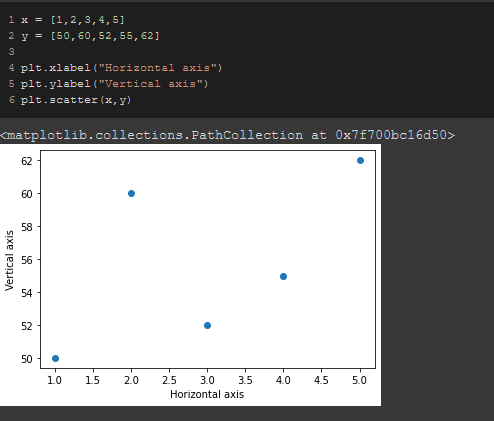


We can also make the plot look unique by using different styles in the ***“linestyle”*** arguement.

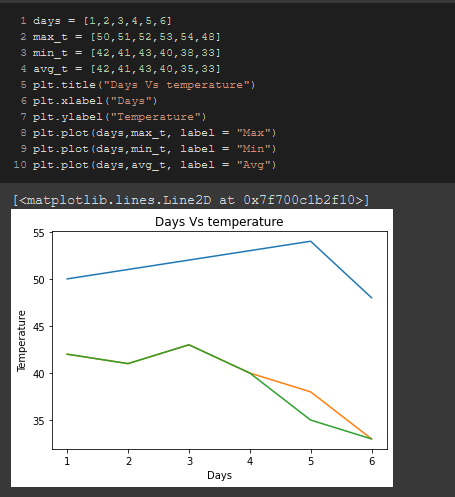


* ***.scatter()***

We can also make a scatter plot using the “***.scatter()***” method.



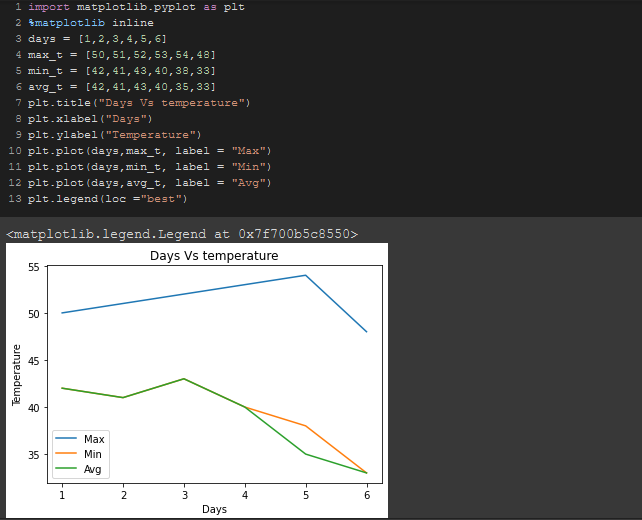
In the figure below, We can’t tell which one’s which.



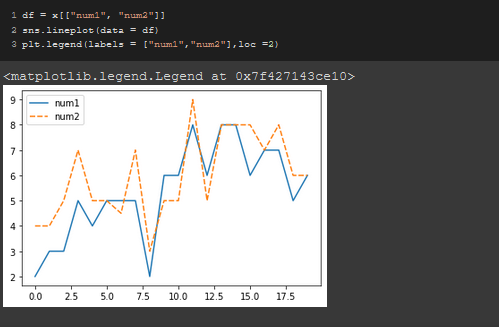
To distinguish one from the other, We use the ***“.legend()”*** method.

But in order to do that, We need to use the **“*label*”** method and label a tag to see that using the ***“.legend()”*** method.

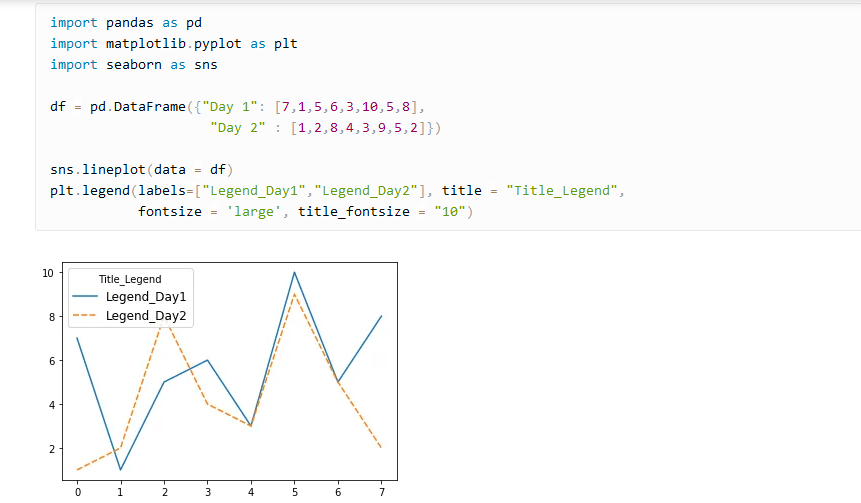
We use the **“*loc = “best”* ”** argument to set the legend snippet in the best position.



We can also set **“loc"** to be 1,2,3,4 and it will set the legend block at anti-clockwise positions. Like Here 2 means leftmost up!



We can also set the labels in a list in the legend method as well as the title, font size and all.

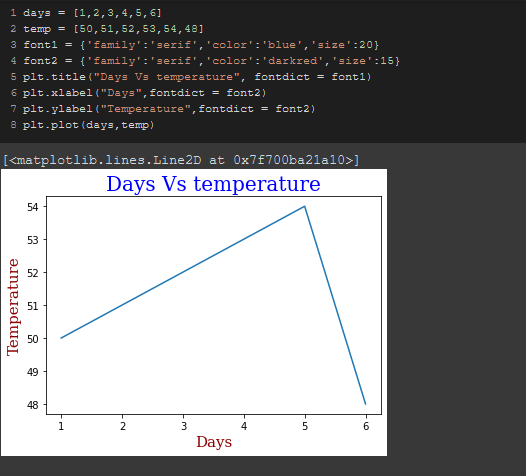


**Labeling a plot:**

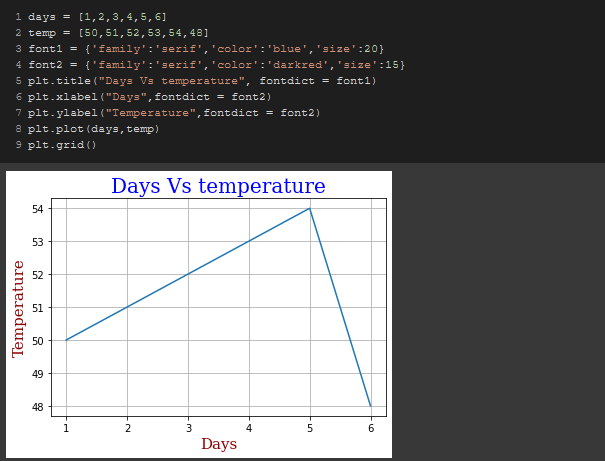
We can use the **“*.xlabel()*”** & **“*.ylabel()*”** to label the “X” & “Y” axises respectively.

We can use the **“*.title()*”** method to set the title of the figure.

We can also use customized fonts using the ***fontdict*** argument but to do that, We need to use define the font first in details.

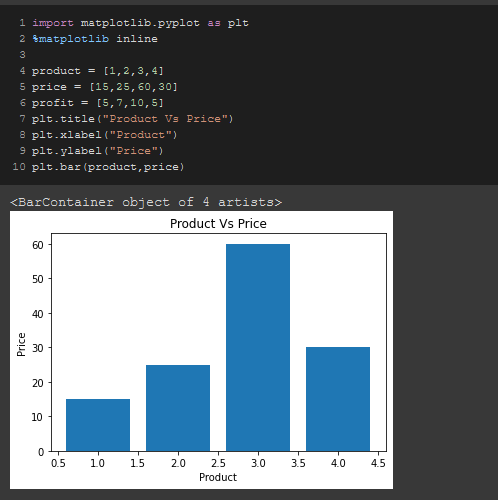


To make it gridded, We use the **“*.grid()*”** method.

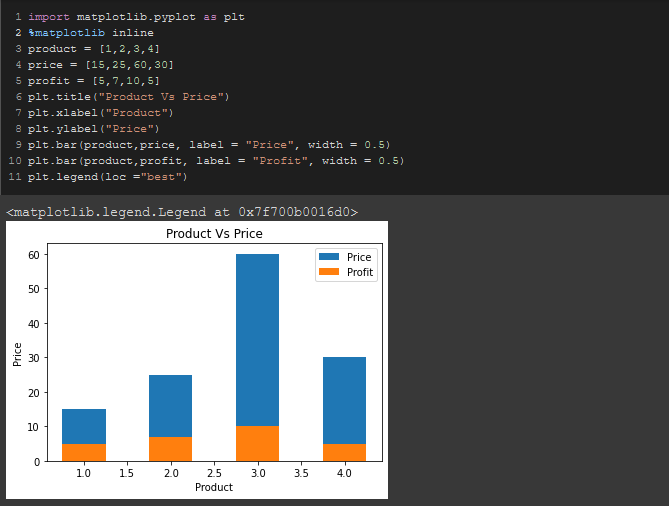


**Bar plots**

We can simply use **“*.bar*”** method to draw a bar chart.



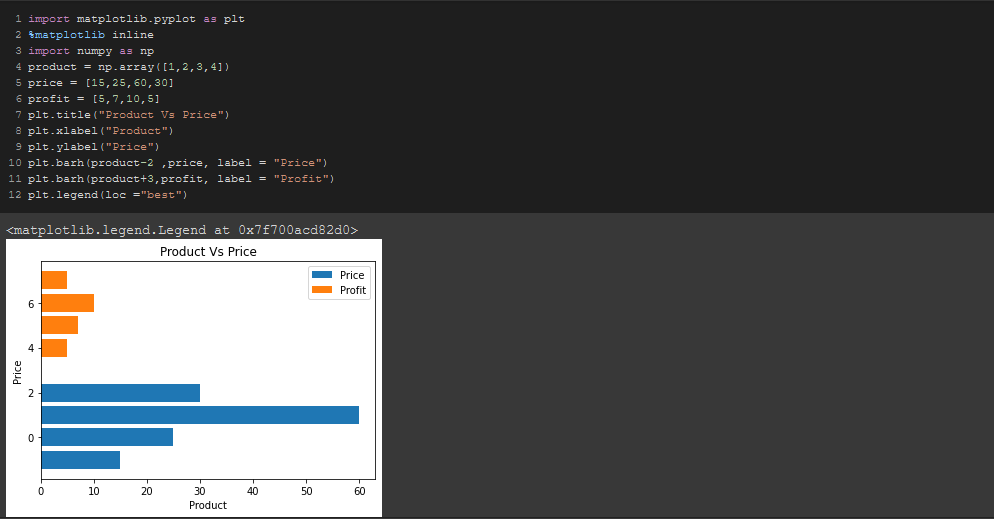
We can plot multiple bar charts in one place and later label them to distinguish them from one another.



We can separate the 2 graphs using **“+”** & **“-”** but in order to do that, we need to import the ***numpy***library first and convert one of the lists to an array.



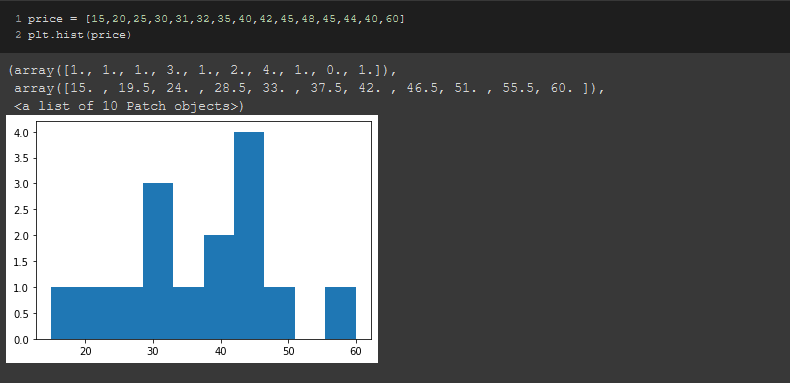
To make these bars horizontal, We need to use the **“*.barh*”** method.



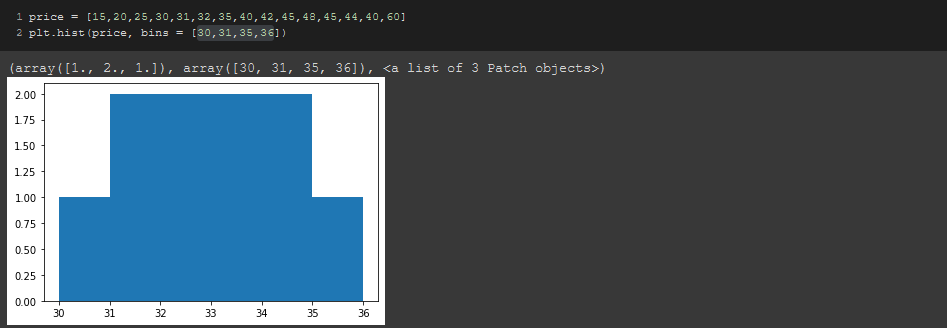
**Histograms**

We can ue the ***“.hist()”*** method to plot a histogram which shows the frequency of an element in a list/array.

For example, This is showing the frequency of the values in the **“*Price*”** list.



In the histogram, If we want a portion to be shown equally distributed, We use the ***“bins”*** method.



**Note:**

If *bins* is a sequence, it defines the bin edges, including the left edge of the first bin and the right edge of the last bin; in this case, bins may be unequally spaced. All but the last (righthand-most) bin is half-open. In other words, if *bins* is:

[1, 2, 3, 4]

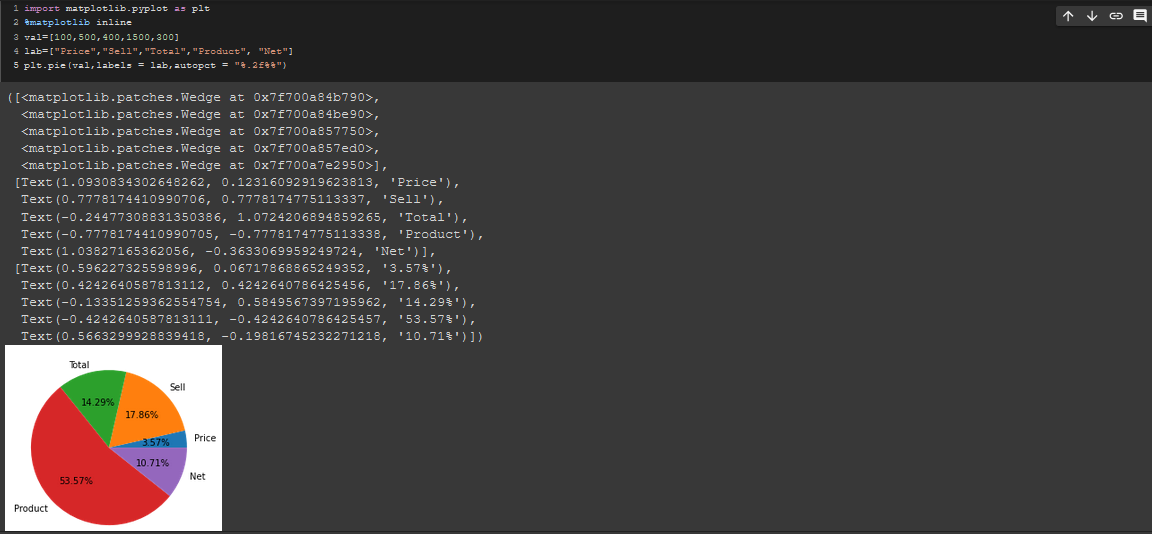
then the first bin is [1, 2) (including 1, but excluding 2) and the second [2, 3). The last bin, however, is [3, 4], which *includes* 4.

**Pie charts**

We can use the ***“.pie()”*** method to plot a pie chart that shows the amount of the elements.

To label the chart, We need to assign the tags in a list and assign that list to the ***“label”*** argument in the ***“.pie()”*** method.

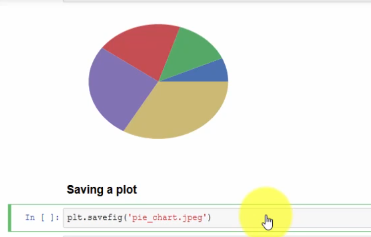
We use the arguement ***“autopct”*** to limit the decimal point



**Saving a figure**

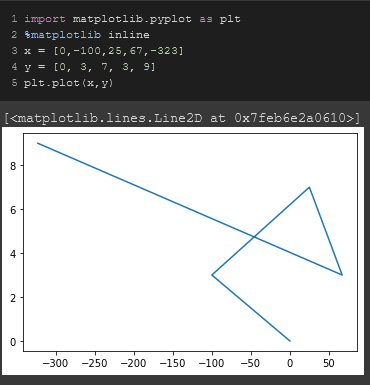
To save a figure, We use the syntax below.

***.savefig(“file\_name.format”)***

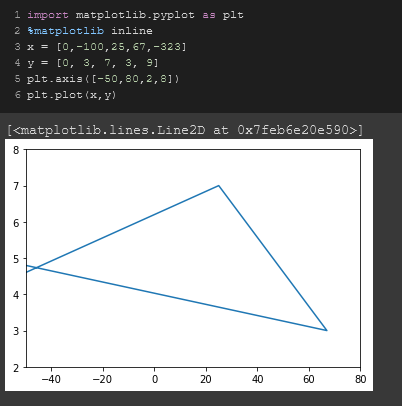


**Zooming in a plot**

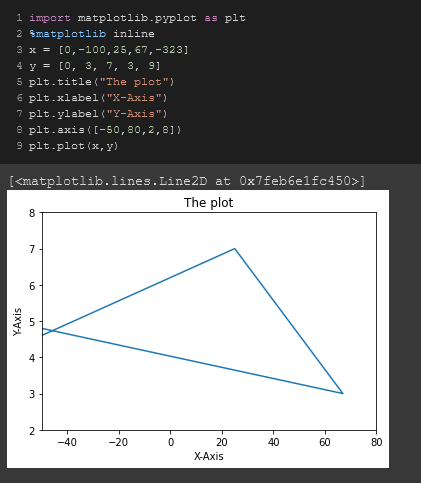
Let’s say, we plot a graph like this.



Now, I want to zoom into the triangle part. For that, I need to use **plt.axis(x-min,x-max,y-min,y-max)**. So, after putting the corresponding values, we get



To label the two axes and to give a reading to the graph, we use.



**Code:**

import matplotlib.pyplot as plt

%matplotlib inline

x = [0,-100,25,67,-323]

y = [0, 3, 7, 3, 9]

plt.title("The plot")

plt.xlabel("X-Axis")

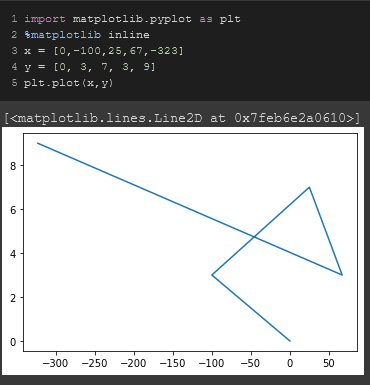
plt.ylabel("Y-Axis")

plt.axis([-50,80,2,8])

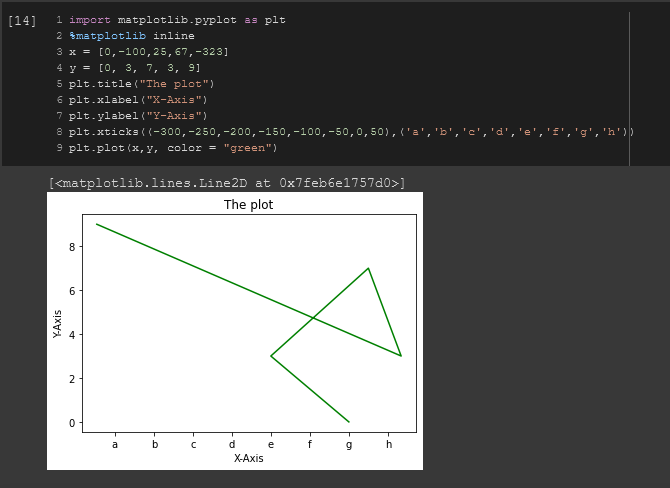
plt.plot(x,y)

**Changing the ticks**

Let’s say, we want to change the ticks of this figure below.



To change the ticks of an axis with customized ticks, we use **plt.xticks** method.



**Code:**

import matplotlib.pyplot as plt

%matplotlib inline

x = [0,-100,25,67,-323]

y = [0, 3, 7, 3, 9]

plt.title("The plot")

plt.xlabel("X-Axis")

plt.ylabel("Y-Axis")

plt.xticks((-300,-250,-200,-150,-100,-50,0,50),('a','b','c','d','e','f','g','h'))

plt.plot(x,y, color = "green")