Lab 09

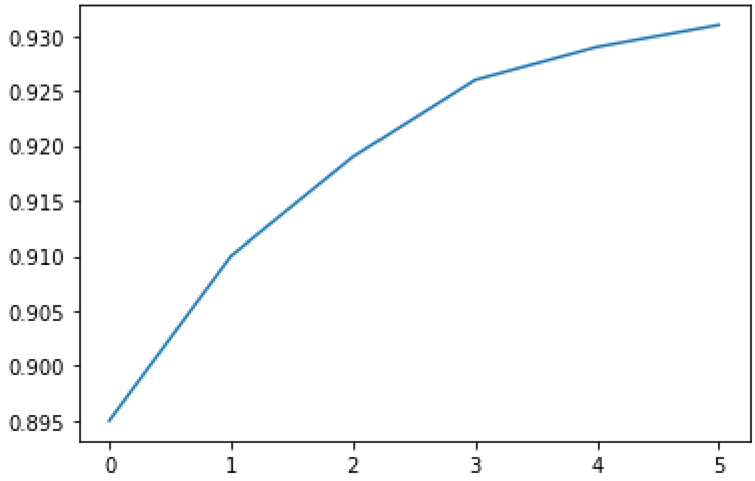
```
import matplotlib.pyplot as plt
import numpy as np
```

Line Charts

A Line chart is a graph that represents information as a series of data points connected by a straight line. In line charts, each data point or marker is plotted and connected with a line or curve.

```
samples = [0.895, 0.91, 0.919, 0.926, 0.929, 0.931]
plt.plot (samples)
```

```
[<matplotlib.lines.Line2D at 0x7f2ee2a009d0>]
```



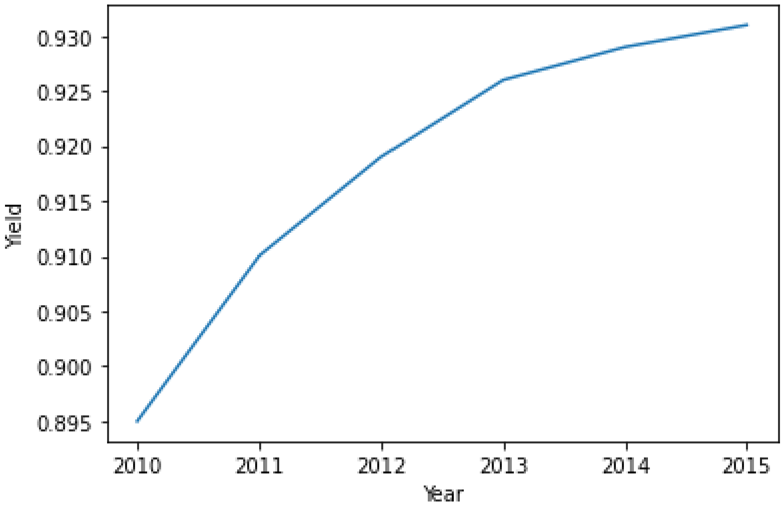
```
# we can add the x-axis values.
years = [2010, 2011, 2012, 2013, 2014, 2015]
samples = [0.895, 0.91, 0.919, 0.926, 0.929, 0.931]
plt.plot(years, samples)
```

```
[<matplotlib.lines.Line2D at 0x7f2ee2cc12d0>]
```

Add labels to the axes.

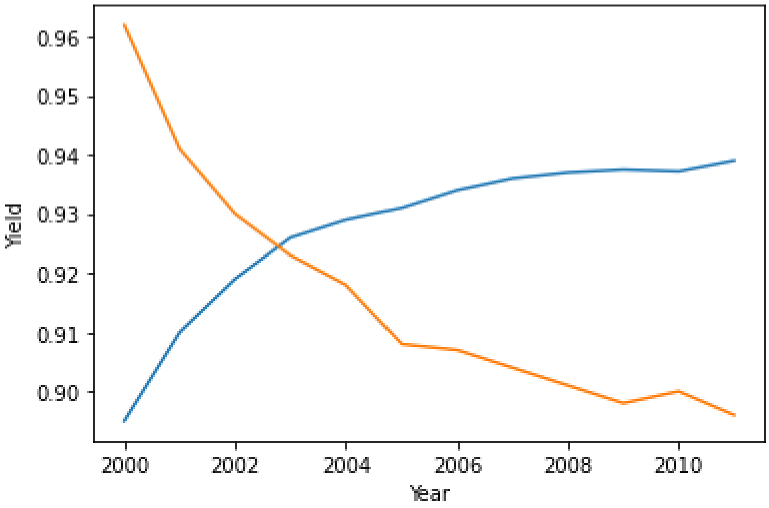
```
plt.xlabel('Year')
plt.ylabel('Yield')
plt.plot(years, samples)
```

```
[<matplotlib.lines.Line2D at 0x7f2ee2c6d2d0>]
```



Plot multiple datasets on the same graph to compare multiple objects on the same graph.

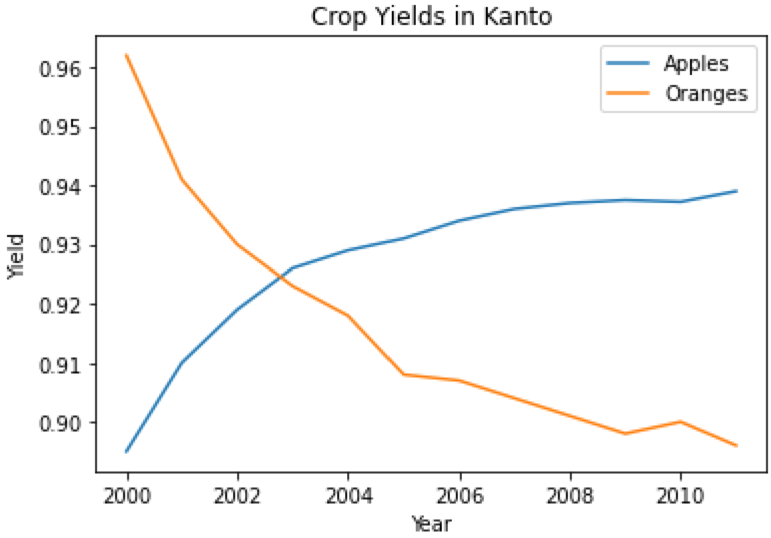
```
years = range (2000, 2012)
apples = [0.895, 0.91, 0.919, 0.926, 0.929, 0.931, 0.934, 0.936, 0.937, 0.9375, 0.9
oranges = [0.962, 0.941, 0.930, 0.923, 0.918, 0.908, 0.907, 0.904, 0.901, 0.898, 0.

plt.plot(years, apples)
plt.plot (years, oranges)
plt.xlabel ('Year' )
plt.ylabel ('Yield')
plt.show()
```



Add legend and title to graph

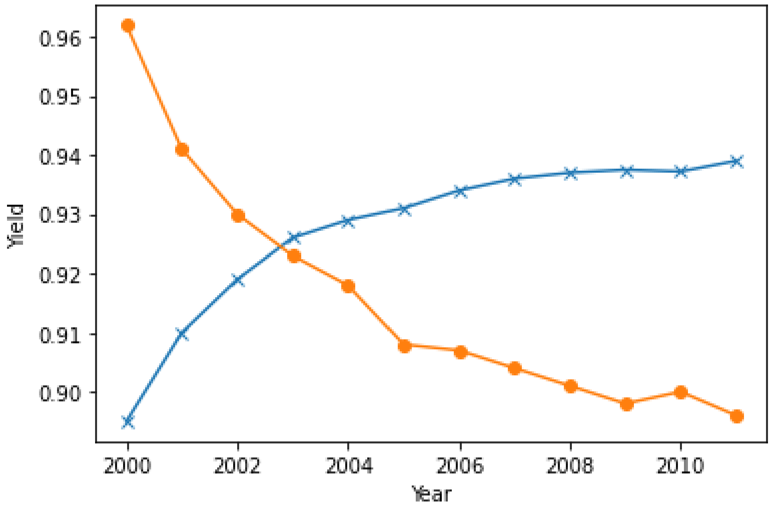
```
plt.plot(years, apples)
plt.plot (years, oranges)
plt.xlabel ('Year' )
plt.ylabel ('Yield')

plt.title("Crop Yields in Kanto")
plt. legend (['Apples', 'Oranges'])
```

```
    <matplotlib.legend.Legend at 0x7f2ee2c0ba90>
```



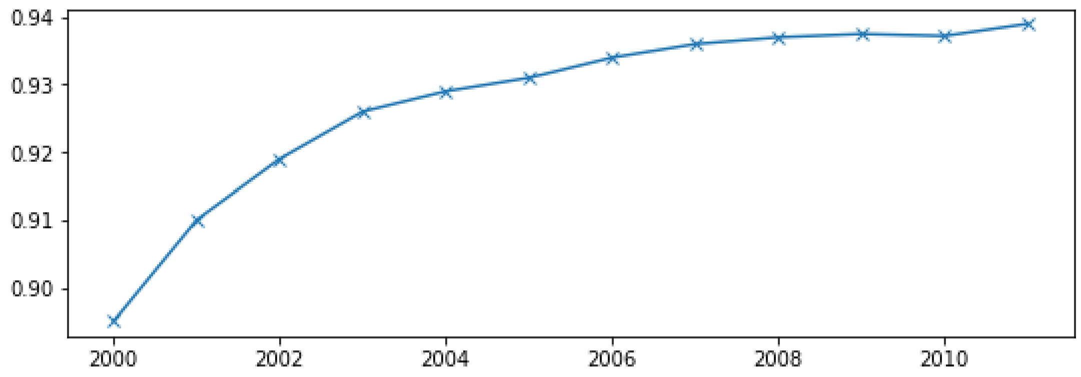
Use markers to show each data point on our graph.

```
plt.plot(years, apples, marker='x')
plt.plot (years, oranges, marker='o')
plt.xlabel ('Year' )
plt.ylabel ('Yield')
plt.show()
```



Use the plt.figure function to change the size of the figure.

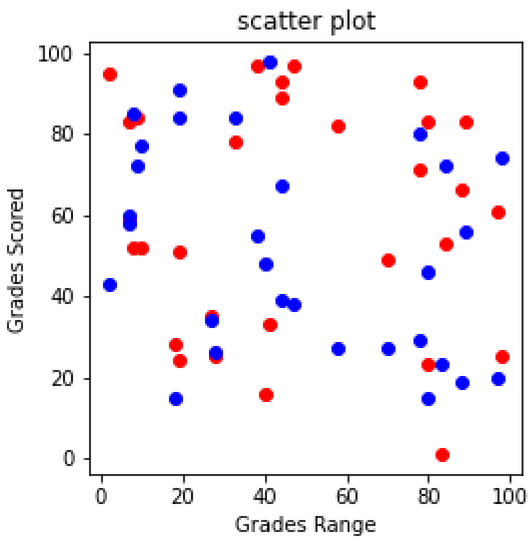
```
plt.figure(figsize=(9,·3))
plt.plot(years, apples, marker='x')
plt.show()
```



ScatterPlot

Scatter plots are used to plot data points on horizontal and vertical axis in the attempt to show how much one variable is affected by another. Each row in the data table is represented by a marker the position depends on its values in the columns set on the X and Y axes. A third variable can be set to correspond to the color or size of the markers, thus adding yet another dimension to the plot.

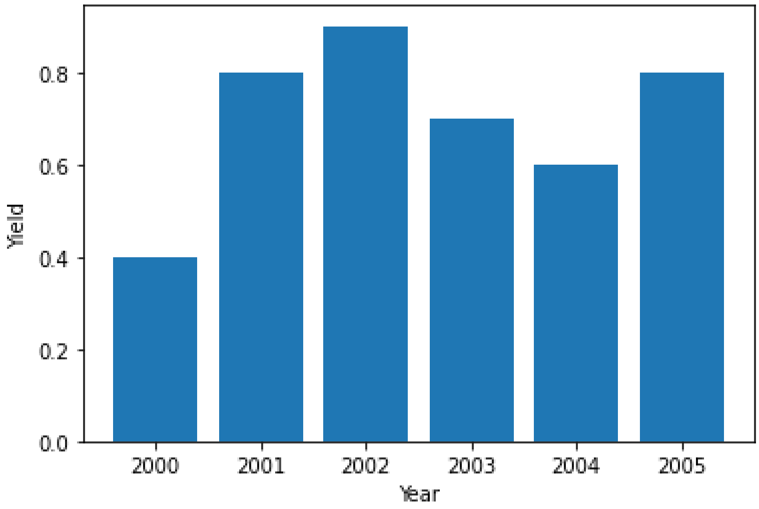
```
girls_grades = np.random.randint(100, size=30)
boys_grades = np.random.randint(100, size=30)
grades_range = np.random.randint(100, size=30)
fig=plt.figure(figsize=(3, 3))
ax=fig.add_axes([0,0,1,1])
ax.scatter(grades_range, girls_grades, color='r')
ax.scatter(grades_range, boys_grades, color='b')
ax.set_xlabel('Grades Range')
ax.set_ylabel('Grades Scored')
ax.set_title('scatter plot')
plt.show()
```



Bar Graphs

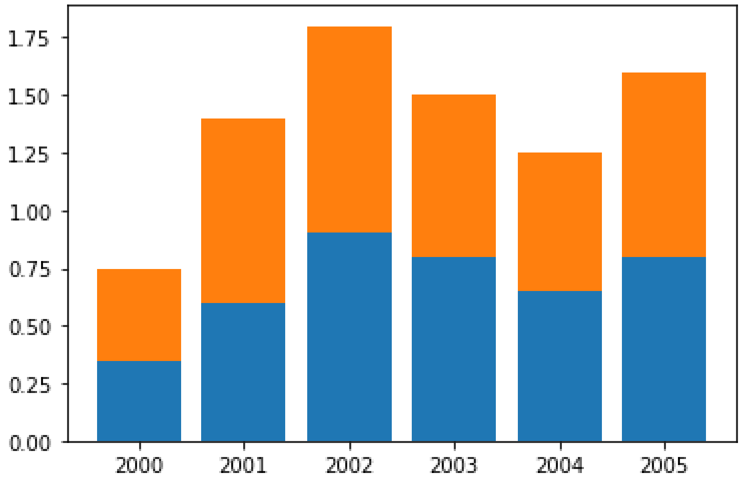
A bar graph plots data with the help of bars, which represent value on the y-axis and category on the x-axis.

```
years = range (2000, 2006)
oranges = [0.4, 0.8, 0.9, 0.7, 0.6, 0.8]

plt.xlabel('Year')
plt.ylabel('Yield')
plt.bar(years, oranges)
plt.show()
```



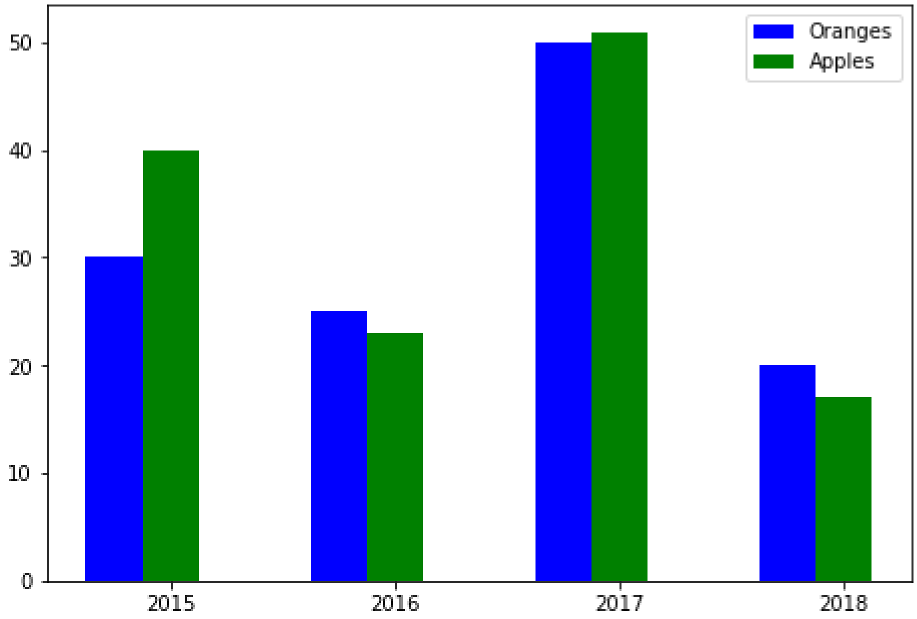
We can also stack bars on top of each other to plot many object.

```
apples = [0.35, 0.6, 0.9, 0.8, 0.65, 0.8]
plt.bar (years, apples)
plt.bar (years, oranges, bottom=apples)
plt.show()
```



When comparing several quantities and when changing one variable, we might want a bar chart where we have bars of one color for one quantity value. We can plot multiple bar charts by playing with the thickness and the positions of the bars. The data variable contains three series of four values.

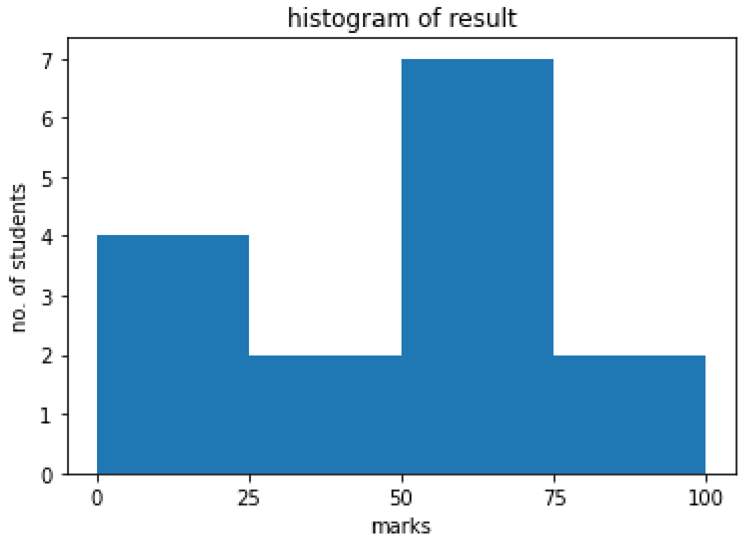
```
data = [[30, 25, 50, 20],
   [40, 23, 51, 17]]
X = np.arange(4)
fig=plt.figure()
ax=fig.add_axes([0,0,1,1])
ax.bar(X + 0.00, data[0], color = 'b', width = 0.25)
ax.bar(X + 0.25, data[1], color = 'g', width = 0.25)
ax.set_xticks([0.25,1.25,2.25,3.25])
ax.set_xticklabels([2015,2016,2017,2018])
ax.legend(labels=['Oranges','Apples'])
plt.show()
```



Histogram

A histogram is an accurate representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable. It is a kind of bar graph.

```
from matplotlib import pyplot as plt
import numpy as np
fig,ax=plt.subplots(1,1)
a = np.array([72,87,5,73,56,73,55,54,11,20,51,5,79,31,27])
ax.hist(a, bins = [0,25,50,75,100])
ax.set_title("histogram of result")
ax.set_xticks([0,25,50,75,100])
ax.set_xlabel('marks')
ax.set_ylabel('no. of students')
plt.show()
```



Exercise

1. Read iris.csv

1. Plot a simple scatter of 2 features of the iris dataset.

   X label = sepal_length

   Y label = sepal_width

2. Read the sepal_length and show it using the histogram with bin = 20.

2. Read company_sales_data.csv

1. Read total profit of all months and using a line plot to show it.
2. Read all month of toothpaste sales data and show it using a scatter plot.
3. Read facecream and facewash product and using the bar chart to show them.*italicized text*

3. Find the frequency of each word in a given text (at least 300 words), and show 30 most common word using line plot.

✓ 0s completed at 11:06 AM ● ✕