***1)TO OPEN A CSV FILE***

!pip install pandas

import pandas as pd

tesla=pd.read\_csv('Tesla.csv - Tesla.csv.csv')

print(tesla)

***2)TO DO MATHEMATIC CALCULATIONS***

!pip install numpy

import numpy as np

a=np.array([[30,40],[50,60]])

b=np.array([[50,40],[30,20]])

np.add(a,b)

output:-

array([[80, 80],

[80, 80]])

*3)DATA VISUALIZATION*

import matplotlib.pyplot as plt

import numpy as np

x=np.array(["A","B","C"])

y=np.array([10,20,30])

plt.xlabel("x values")

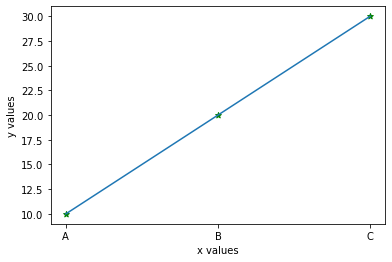
plt.ylabel("y values")

plt.scatter(x,y,marker="\*",color="green")

plt.plot(x,y)

plt.show()

output:-



*4)DATA VISUALIZATION IN CSV FILES*

import matplotlib.pyplot as plt

import numpy as np

import pandas as pd

df=pd.read\_csv("diabetes.csv")

plt.xlabel("glucose")

plt.ylabel("insulin level")

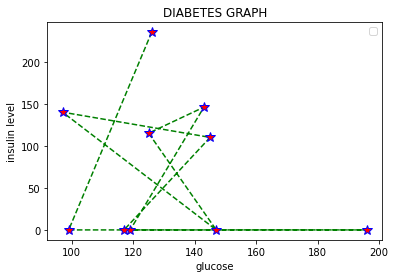
plt.plot(df["Glucose"][20:30],df["Insulin"][20:30],"g\*--",ms=10,mfc="red",mec="blue")

plt.legend()

plt.title("DIABETES GRAPH")

plt.show()

output:-



*A)BAR GRAPH*

plt.bar(df["Age"][0:25],df["Outcome"][0:25])

plt.show()

output:-

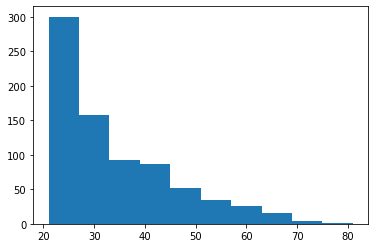


*B)HISTOGRAM*

plt.hist(df["Age"])

plt.show()

output:-

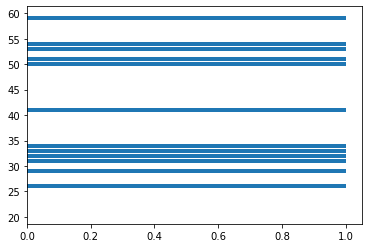


*C)HORIZONTAL BAR GRAPH*

plt.barh(df["Age"][0:25],df["Outcome"][0:25])

plt.show()

output:-



*D)PIE CHART*

a=df.loc[df["Age"]==50].count()[0]

b=df.loc[df["Age"]<50].count()[0]

c=df.loc[df["Age"]>50].count()[0]

z=[a,b,c]

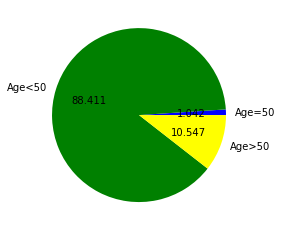
l=["Age=50","Age<50","Age>50"]

c=["blue","green","yellow"]

plt.pie(z,labels=l,autopct="%.3f",colors=c)

plt.show()

output:-



i)df['status'].unique()

output:-

array(['Placed', 'Not Placed'], dtype=object)

x=df[df['status']=='Placed'].count()[0]

y=df[df['status']=='Not Placed'].count()[0]

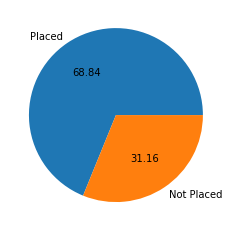
z=[x,y]

l=['Placed','Not Placed']

plt.pie(z,labels = l,autopct='%1.2ff')

plt.show()

output:-



*5)TO CSV*

c={'name': ['usa', 'india', 'german', 'rwanda'],

'codes': [1, 91, 49, 250],

'language': ['english', 'kannada', 'spanish', 'african'],

'food': ['sandwich', 'ragi muddhe', 'burger', 'pizza'],

'sports': ['baseball', 'gilli dandhu', 'cricket', 'swimming']}

df=pd.DataFrame(c)

df.to\_csv('country data.csv')

*CORELATION*

import pandas as pd

import numpy as np

import seaborn as sns

import sklearn

import matplotlib.pyplot as plt

d =pd.read\_csv('bangalore.csv')

df=d.corr()

p=df["price"]

p #price increases for bath,total\_sqft\_int,balocny

#price decreases for -ve values(ratings)

a=d.loc[d["price"]==100].count()[0]

b=d.loc[d["price"]<100].count()[0]

c=d.loc[d["price"]>100].count()[0]

z=[a,b,c]

l=["house price=100","house price<100","house price>100"]

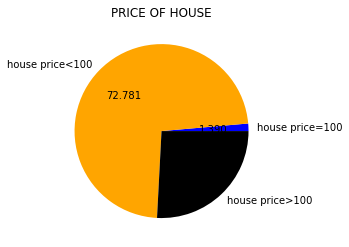
c=["blue","green","yellow"]

plt.pie(z,labels=l,autopct="%.3f",colors=c)

plt.title("PRICE OF HOUSE")

plt.show()

output:-



*MACHINE LEARNING*

Types of AI

It predicts the o/p based on past experiences and while predicting it automatically learns the data from trained data.

Ex: Email spam, Medical diagnosis, Salary prediction, Image recognition, Product recommendation

Train data=well known data both i/p and o/p values

Test data=only i/p values

Types of ML

1.Supervised learning

2.Un Supervised learning

3.Semi- Supervised learning

from sklearn.linear\_model import LinearRegression

reg=LinearRegression()

reg.fit(df[[‘YearsExperience’]],df[‘Salary’])

#prediction

reg.predict([[1.1]])#x

reg.coef\_#m

reg.intercept\_#c

#mx+c

df1=pd.read\_csv(‘Salary\_test.csv’)

salary=reg.predict(df1)

salary

*LINEAR REGRESSION*

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

df=pd.read\_csv("Train.csv")

df1=df.drop(['Item\_Identifier','Outlet\_Identifier','Outlet\_Establishment\_Year','Outlet\_Size','Outlet\_Location\_Type','Outlet\_Type'],axis=1)

from sklearn.preprocessing import LabelEncoder

l=LabelEncoder()

f=l.fit\_transform(df1['Item\_Fat\_Content'])

p=l.fit\_transform(df1['Item\_Type'])

df1['Fat']=f

df1['Product']=p

df2=df1.drop(['Item\_Fat\_Content','Item\_Type'],axis=1)

df2[0:3]

output:-

Item\_Weight

Item\_Visibility

Item\_MRP

Item\_Outlet\_Sales

Fat

Product

#split the data

x=df2.drop("Item\_Outlet\_Sales",axis=1)

x[0:2]

output:-

| **Item\_Weight** | **Item\_Visibility** | **Item\_MRP** | **Fat** | **Product** |
| --- | --- | --- | --- | --- |

y=df2['Item\_Outlet\_Sales']

from sklearn.model\_selection import train\_test\_split

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2)#test size=20%

from sklearn.linear\_model import LinearRegression

reg = LinearRegression()

reg.fit(x\_train,y\_train)#target variable

sales=reg.predict(x\_test)

plt.plot(x\_test,sales)

plt.show()

output:-

