**SEM -2 MINI PROJECT**

**BITCOIN VALUE PREDICTION**



**BY:**

KAVYA.P (1833022)

PRIYADHARSHINI.R (1833039)

PRANEETHA.T.R (1833036)

SARATHRAM.M (1833049)

SRIDHAR.S (1833051)

**GUIDE:**

Mrs.RADHIKA

**INTRODUCTION:**

A cryptocurrency is a digital currency that is created and managed through the use of advanced encryption techniques known as cryptography. Cryptocurrency made the leap from being an academic concept to (virtual) reality with the creation of Bitcoin in 2009. The fact that the cryptocurrency works without the regulation from a single administrator or central bank makes it a decentralized currency. Ever since its creation , this has grown popularity across the globe and has been accepted as a medium of exchange by many merchants. The peer to peer network enables users to directly transfer money from one person to another without the involvement of any intermediary. This method of transfer has enabled transactions to be much cheaper compare to any other means of money transfers. However there have been many speculations about the vulnerability of bitcoin. Some economic analysts predict a big change in crypto is forthcoming as institutional money enters the market. So here we have predicted the bitcoin value of forthcoming years using regression. Since the volatility of the currency also raises the question of whether bitcoin is worthy of being a real currency and whether it is safe. To better understand the safety of bitcoins, we bring it out with swot analysis.

**TRANSACTIONS:**

The block chain is a shared publicledger on which the overall Bitcoin network relies. All confirmed transactions are included in this block chain. It allows Bitcoin wallets to calculate their spendable balance so that new transactions can be verified thereby ensuring the bitcoins are actually owned by the spender which is managed through the advanced encryption technique to identify its internal strengths and weaknesses, as well as its external opportunities and threats.s known as cryptography. A transaction is a transfer of value between Bitcoin wallets which gets included in the block chain. Bitcoin wallets keep a secret piece of data called a private key which is used to sign transactions, providing a mathematical proof that they have come from the owner’s wallet. The signature also prevents the transaction from being altered by anybody once it has been issued. All transactions that are being broadcasted to the network will be confirmed within 10-20 minutes, through a process called mining. Mining is a distribute consensussystem that is used to confirm pending transactions by including them in the block chain and it also enforces a chronological order in the block chain, protects the neutrality of the network, and allows different computers to agree on the state of the system.

**OUR PROJECT:**

Our project is to predict the bitcoin market price for future years.We have collected the data of the previous years from the KAGGLE website from which we are going to predict the bitcoin market price by taking in the attribute such as market price. We have framed the SWOT analysis to identify its internal strengths and weaknesses, as well as its external opportunities and threats for bitcoin investments.

And moreover we are going to use the following statistical methods-

We find relationship between time period and the bitcoin market price with the previous years data using CORRELATION , so that we can conclude how the bitcoin market price varies in the subsequent years,

REGRESSION shows a relationship between an independent variable i.e. date and a dependent variable i.e. the bitcoin market price,

To cross check the obtained results,whether it is accurate or not,HYPOTHESIS TEST is made using Z test

We put the data in the datasturucture tool called trees and used binary search,which we are going to implement in PYTHON language.

**SWOT ANALYSIS:**

It is a study undertaken to identify its internal strengths and weaknesses, as well as its external opportunities and threats. As far as investments in bitcoin concerns, it is essential to look over at these areas.

**STRENGTHS:**

They follow peer-to-peer transactions. Users can make transactions without involving an intermediary since the transactions are recorded in a public ledger. The security offered by bitcoins also play an major role. Large amount of returns is possible in bitcoins. They cannot be tracked or stolen. It makes bitcoin a tax-free currency. Even if the economy crashes, bitcoin can survive.

**WEAKNESS**:

The volatility of bitcoins is most worrying one. It is a decentralized monetary system and has no regulation. The price of bitcoins can neither be determined nor controlled. Crippling slow transactions and accessibility loss.

**OPPORTUNITY:**

The world is heading to a cashless economy where you don’t need wallet and some carry material. Many global vendors like amazon have accepting bitcoin as their payment money. The major opportunity perhaps lies in the Blockchain technology itself. Bitcoin is one of thousand potentials of the Blockchain technology.

**THREATS:**

Bitcoin’s trackless encryption can attract criminals & money launderers to use their money in illegal way. As number of people will adopt it as a currency, number of criminal acts will increase. Blockchain technology will inspire many financial institutes, but at same time cyber attackers will also be looking for ways to hack into the system by deciphering the long codes. Defending itself against cyber theft is going to be probably the biggest hurdle for bitcoin in the coming years.

**DATA STRUCTURE USED**:

The given data is implemented in TREE and searching is done using BINARY SEARCH TREE.

**Binary Search Tree** is a node-based binary tree data structure which has the following properties:

* The left subtree of a node contains only nodes with keys lesser than the node’s key.
* The right subtree of a node contains only nodes with keys greater than the node’s key.
* The left and right subtree each must also be a binary search tree.
* Whenever an element is to be searched, start searching from the root node. Then if the data is less than the key value, search for the element in the left subtree. Otherwise, search for the element in the right subtree.

## **STATISTICAL TOOL:**

**CORRELATION:**

{Correlation is a statistical measure that indictes the extent to which two or more variables fluctuate together \.....heredisplaystyle r\_{xy}\quad {\overset {\underset {\mathrm {def} }{}}{=}}\quad {\frac {\sum \limits \_{i=1}^{n}(x\_{i}-{\bar {x}})(y\_{i}-{\bar {y}})}{(n-1)s\_{x}s\_{y}}}={\frac {\sum \limits \_{i=1}^{n}(x\_{i}-{\bar {x}})(y\_{i}-{\bar {y}})}{\sqrt {\sum \limits \_{i=1}^{n}(x\_{i}-{\bar {x}})^{2}\sum \limits \_{i=1}^{n}(y\_{i}-{\bar {y}})^{2}}}},}…

Here our attributes are date and bitcoin market price .A correlation coefficient is a statistical measure of the degree to which changes to the value of one variable predict change to the value of another and its formula is:

**r=N(∑XY)−(∑X)(∑Y)/([N∑X2−(∑X)2][N∑Y2−(∑Y)2])^1/2**

where,

r-correlation coefficient

X-date

Y-bitcoin price value

we get the correlation co efficient value as **0.9** which means that they both are correlated and means that its increasing market price depends on the duration.

**REGRESSION:**

Linear regression is an important tool in analytics. The technique uses statistical calculations to plot a trend line in a set of data points. The trend line give the future market price of bitcoins. Linear regression shows a relationship between an independent variable i.e. date and a dependent variable i.e. the bitcoin market price,

**Regression Equation:  (y) = a + b(x)**

b= regression coefficient

a= intercept point of the regression line which is in the y-axis.

N= Number of values or elements

X=date

Y= bitcoin market price

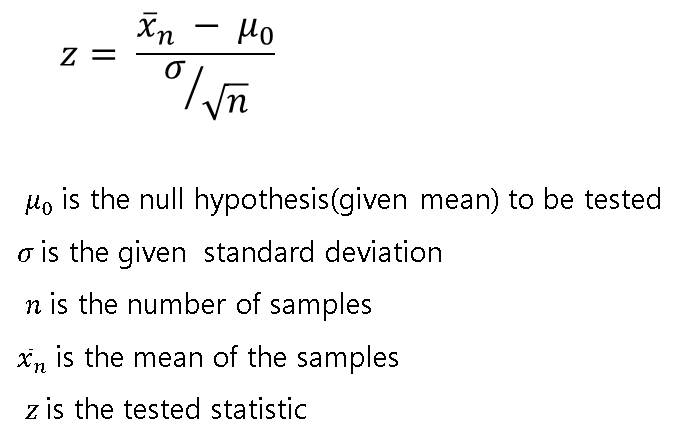
The fitted equation is:

**y = -1906.9949793051371 + 69.05222695005884x**

**HYPOTHESIS TESTING:**

Hypothesis testing is used to infer the result of a hypothesis performed on sample data from a larger population.The test tells the analyst whether or not his primary hypothesis is true.Statistical analysts test a hypothesis by measuring and examining a random sample of the population being analysed.

Z-TEST is a statistical tool used to determine whether two population means are different when the variances are known and the sample size is large.the formula for Z-TEST is calculated by deducting population mean from the X-VALUE and then the result is divided by the population standard deviation.Z-TEST for single mean is used to test a hypothesison a specific value of the population mean.



Null hypothesis:

The predicted value for a given year is Incorrect

Alternative hypothesis:

The predicted value for a given year is correct

It is one tailed with **5%** level of significance

Zcal= x (value obtained from regression)

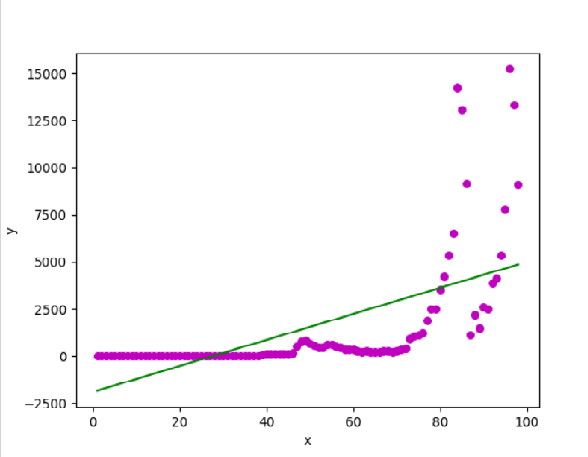
Ztab= 1.645

If (Zcal > Ztab)

Hence we Reject null hypothesis,i.e.the predicted value is correct.

**GRAPH:**

Here the graph represents the increase in bitcoin market value over the years .



**CODE:**

from prettytable import PrettyTable

import matplotlib.pyplot as plt

import seaborn as sns

import numpy as np `

import pandas as pd

from scipy import stats

from sklearn.linear\_model import LinearRegression

x = PrettyTable()

import csv

with open('bitcoindataset1.csv', newline='') as csvfile:

data = list(csv.reader(csvfile))

year=[]

yr=[]

price=[]

n=len(data)

#print("Length of the given data:",n)

class Node:

def \_\_init\_\_(self,yr,avg):

self.left = None

self.right = None

self.yr = yr

self.avg = avg

def insert(self,yr,avg):

if self.yr:

if yr < self.yr:

if self.left is None:

self.left = Node(yr,avg)

else:

self.left.insert(yr,avg)

elif yr > self.yr:

if self.right is None:

self.right = Node(yr,avg)

else:

self.right.insert(yr,avg)

else:

self.yr = yr

def search(self,key):

x.field\_names = [data[0][1],data[0][3]]

if key < self.yr:

return self.left.search(key)

elif key > self.yr:

return self.right.search(key)

else:

x.add\_row([ self.yr,self.avg])

return x

for i in range(1,n):

year.append(float(data[i][0]))

yr.append(float(data[i][1]))

price.append(float(data[i][3]))

root=Node(year[1],price[1])

for i in range(0,n-1):

root.insert(yr[i],price[i])

def menu():

print("\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* MAIN MENU \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

print()

choice = int(input(""" Select a number from the given menu

1: Show data

2: Central tendency values of price

3: Correlation and Regression

4: Prediction

5: Search

6: Graph

7: Quit

Please enter your choice: """))

while(1):

if choice == 1:

showdata()

menu()

elif choice ==2:

centraltendency()

menu()

elif choice ==3:

correg()

menu()

elif choice==4:

prediction()

menu()

elif choice==6:

main()

elif choice==7:

print(" \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* THANK YOU \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*")

exit(1)

elif choice==5:

search()

menu()

else:

print(" You are only allowed to select the given options ")

print(" Please don't exceed limit ")

print(" ")

menu()

def showdata():

x.field\_names = (data[0][1], data[0][2], data[0][3])

for i in range (1,n):

x.add\_row([data[i][1], data[i][2], data[i][3]])

print(x)

print("-------------------------------------------------------------------------------")

def search():

a=data[1][1]

b=data[n-1][1]

find\_yr=float(input("Enter a year between {}-{}:".format(a,b)))

print(root.search(find\_yr))

print("-------------------------------------------------------------------------------")

def centraltendency():

mean=np.mean(price)

median=np.median(price)

print("Mean:",mean)

print("Median:",median)

print("-------------------------------------------------------------------------------")

def estimate\_coef(x, y):

n = np.size(x)

m\_x, m\_y = np.mean(x), np.mean(y)

SS\_xy = np.sum(y\*x) - n\*m\_y\*m\_x

SS\_xx = np.sum(x\*x) - n\*m\_x\*m\_x

b\_1 = SS\_xy / SS\_xx

b\_0 = m\_y - b\_1\*m\_x

return(b\_0, b\_1)

print("-------------------------------------------------------------------------------")

def correg():

cov=0.0

mx=0.0

my=0.0

mx2=0.0

b=0.0

a=0.0

mx=np.mean(year)

my=np.mean(price)

for i in range(len(price)):

cov=cov+((year[i]-mx)\*(price[i]-my))

mx2=mx2+((year[i]-mx)\*\*2)

b=cov/mx2

a=my-b\*mx

print("y = {} + {}x".format(a,b))

print("-------------------------------------------------------------------------------")

def prediction():

cov=0.0

mx=0.0

my=0.0

mx2=0.0

b=0.0

a=0.0

y=0.0

mx=np.mean(year)

my=np.mean(price)

for i in range(len(price)):

cov=cov+((year[i]-mx)\*(price[i]-my))

mx2=mx2+((year[i]-mx)\*\*2)

b=cov/mx2

a=my-b\*mx

yy=int(input("Enter the year : "))

mm=int(input("Enter the month: "))

ye=((yy-2010)\*12)+mm

y=a+b\*ye

#print(y,a,b)

print("The predicted price of the year {} is {}".format(yy,y))

print("-------------------------------------------------------------------------------")

print(" ")

print(" ------------------------ HYPOTHESIS TESTING ------------------------ ")

sd=np.std(price)

obs=(n-1)

n2=obs\*\*0.5

z=abs(y-my)/(sd/n2)

#print(z)

print("we have taken the 5% level of significance")

print(" ")

print("z-table=1.645")

print(" ")

print("The Hypothesis value is: %f"%(z))

print(" ")

val=1.645

if val>z:

print("We ACCEPT the Null Hypothesis")

print("The predicted value is Wrong")

else:

print("We REJECT the Null Hypothesis")

print("The predicted value is Correct")

print("-------------------------------------------------------------------------------")

def plot\_regression\_line(x, y, b):

plt.scatter(x, y, color = "m",

marker = "o", s = 30)

y\_pred = b[0] + b[1]\*x

plt.plot(x, y\_pred, color = "g")

plt.xlabel('x')

plt.ylabel('y')

plt.show()

print("-------------------------------------------------------------------------------")

menu()

def main():

data=pd.read\_csv('bitcoindataset1.csv')

print("\n",data)

x = data.Slno.values

y = data.Average.values

b = estimate\_coef(x, y)

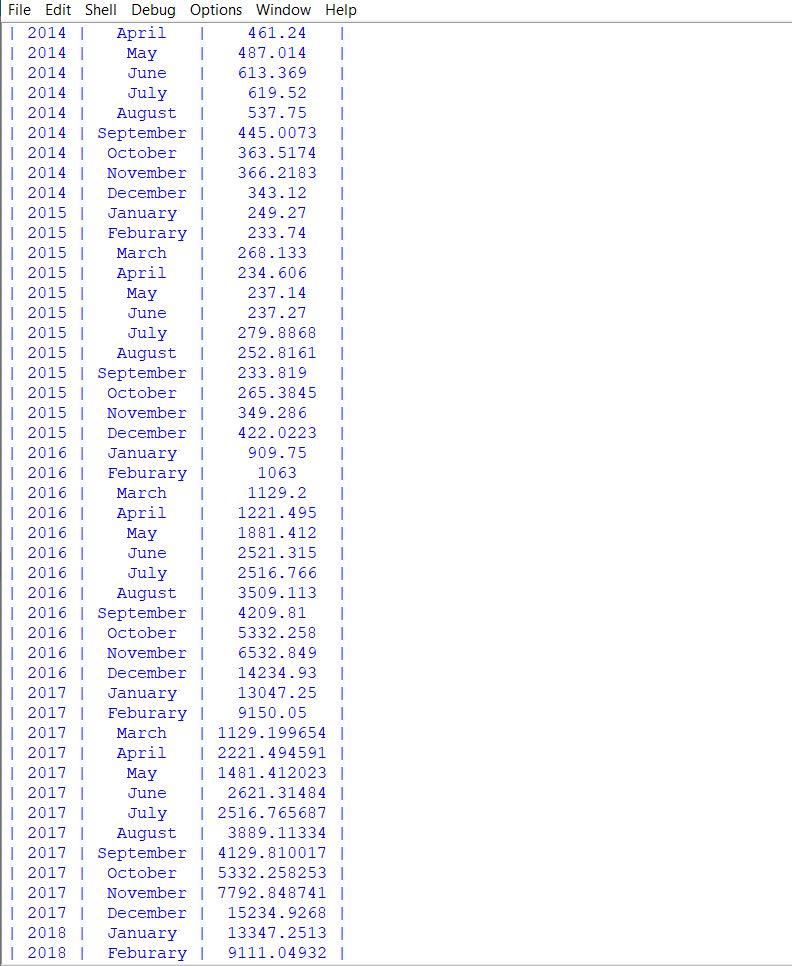
print("Estimated coefficients:\nb\_0 = {} \ \nb\_1 = {}".format(b[0], b[1]))

plot\_regression\_line(x, y, b)

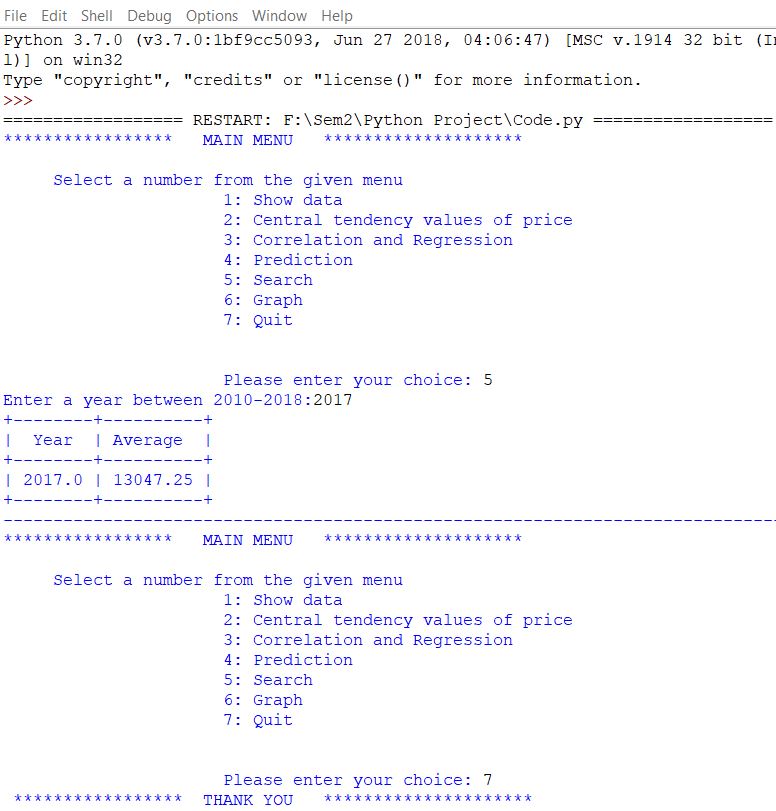
print("-------------------------------------------------------------------------------")

menu()

**OUTPUT:**







##### **CONCLUSION:**

The bitcoin market for the upcoming 6 years are being predicted using regression in python and is graphically represented as:

As per the latest report of RBI, the following is allocation to different asset class.

1.Real estate 84%

2. Precious metals gold-silver etc., 11%.

3.Financial asset (non tangible) 5%whichincluded fixes deposits/insurances/PPF/Post/Mutual funds etc.

Mutual fund industry witnessed 20,000 crore of fund flow and market went all time high. Just imagine if 1% of total asset shifts from physical to financial investment then the amount of wealth would be created will be mind blowing.

Every form of investment has its own strength and shortcomings. Since bitcoin having a strong foundations, these can overcome disadvantages. Making it as a part of the mainstream financial system would have to satisfy very divergent. system would have to satisfy very divergent criteria . These investment medium happens once in a generation.. Some countries like USA has started to accept payment through Bitcoins .Increase in number of users can make it an official currency which would make bitcoin to mark itself it's position in future.