In [21]:

```
# Question 1:

def split(endYear,startYear=2012):

    import pandas as pd
    import numpy as np
    from datetime import datetime

    #Importing dataset
    df_energy = pd.read_excel(r"D:\Ruchi Folder\Python\Assignment5\Energy.xlsx")


    # Column columns  "Accumulated Other Comprehensive Income (Loss)" to "Selling, Gene

    df_selected=df_energy.loc[:,"Accumulated Other Comprehensive Income (Loss)": "Selli


    date_dataf=df_energy.loc[:,"Data Date":"Fiscal Year"]


    # Joining df_Selected with df_Dates
    new_dataf=date_dataf.join(df_selected)





    if endYear==None:

           Test=new_dataf[new_dataf["Fiscal Year"]==startYear]

           Train=new_dataf[new_dataf["Fiscal Year"]!=startYear]

    elif endYear!=None:

           Test= new_dataf[(new_dataf["Fiscal Year"] >= startYear) & (new_dataf["Fiscal

           Train=new_dataf[~((new_dataf["Fiscal Year"] >= startYear) & (new_dataf["Fisc



    return(Test, Train)


split(2013)
```

```
C:\Users\Bipin\anaconda3\lib\site-packages\openpyxl\styles\stylesheet.py:221: UserWarnin
g: Workbook contains no default style, apply openpyxl's default
  warn("Workbook contains no default style, apply openpyxl's default")
```

Out[21]:

```
(      Data Date  Fiscal Year  Accumulated Other Comprehensive Income (Loss)  \
 8     20120331         2012                                        -1057.0
 9     20120630         2012                                         -779.0
 10    20120930         2012                                         -675.0
 11    20121231         2012                                         -493.0
 12    20130331         2013                                         -452.0
 ..         ...          ...                                            ...
 827   20121231         2012                                         -464.0
 828   20130331         2013                                         -382.0
```

```
829   20130630         2013                                     -254.0
830   20130930         2013                                     -205.0
831   20131231         2013                                     -204.0

     Current Assets - Other - Total  Current Assets - Total  \
8                            1421.0                  8212.0   
9                            1699.0                  7925.0   
10                           2167.0                  8157.0   
11                           2148.0                  8387.0   
12                           8448.0                 11418.0   
..                              ...                     ...   
827                           110.0                 13029.0   
828                           172.0                 15297.0   
829                           190.0                 14252.0   
830                           206.0                 13531.0   
831                           197.0                 12737.0   

     Other Long-term Assets  Non-Current Assets - Total  \
8                    3675.0                     32435.0   
9                    3417.0                     32689.0   
10                   3302.0                     34055.0   
11                   3596.0                     35054.0   
12                   3104.0                     30961.0   
..                      ...                         ...   
827                   254.0                     14194.0   
828                   320.0                     15476.0   
829                   328.0                     15443.0   
830                   320.0                     15472.0   
831                   324.0                     15648.0   

     Assets Netting & Other Adjustments  \
8                                -350.0   
9                                -305.0   
10                               -186.0   
11                               -227.0   
12                               -205.0   
..                                  ...   
827                                84.0   
828                               138.0   
829                               -30.0   
830                               -62.0   
831                               -21.0   

     Accum Other Comp Inc - Derivatives Unrealized Gain/Loss  \
8                                               -547.0         
9                                                 78.0         
10                                              -152.0         
11                                               -24.0         
12                                               -11.0         
..                                                 ...         
827                                                0.0         
828                                                4.0         
829                                                4.0         
830                                                4.0         
831                                                4.0         

     Accum Other Comp Inc - Other Adjustments  ...  \
8                                       107.0  ...   
9                                      -270.0  ...   
10                                       67.0  ...   
11                                      170.0  ...   
12                                       28.0  ...   
..                                        ...  ...   
827                                       0.0  ...   
828                                       0.0  ...   
```

```
829                                       0.0  ...
830                                       0.0  ...
831                                       0.0  ...

     Implied Option EPS Diluted  Implied Option EPS Diluted Preliminary  \
8                           0.0                                     0.0
9                           0.0                                     0.0
10                          0.0                                     0.0
11                          0.0                                     0.0
12                          0.0                                     0.0
..                          ...                                     ...
827                         0.0                                     NaN
828                         0.0                                     NaN
829                         0.0                                     NaN
830                         0.0                                     NaN
831                         0.0                                     NaN

     Implied Option EPS Basic 12MM  Implied Option 12MM EPS Basic Preliminary  \
8                              0.0                                        0.0
9                              0.0                                        0.0
10                             0.0                                        0.0
11                             0.0                                        0.0
12                             0.0                                        0.0
..                             ...                                        ...
827                            NaN                                        NaN
828                            NaN                                        NaN
829                            NaN                                        NaN
830                            NaN                                        NaN
831                            NaN                                        NaN

     Implied Option EPS Basic  Implied Option EPS Basic Preliminary  \
8                         0.0                                   0.0
9                         0.0                                   0.0
10                        0.0                                   0.0
11                        0.0                                   0.0
12                        0.0                                   0.0
..                        ...                                   ...
827                       0.0                                   NaN
828                       0.0                                   NaN
829                       0.0                                   NaN
830                       0.0                                   NaN
831                       0.0                                   NaN

     Implied Option Expense  Implied Option Expense Preliminary  \
8                       0.0                                 0.0
9                       0.0                                 0.0
10                      0.0                                 0.0
11                      0.0                                 0.0
12                      0.0                                 0.0
..                      ...                                 ...
827                     0.0                                 NaN
828                     0.0                                 NaN
829                     0.0                                 NaN
830                     0.0                                 NaN
831                     0.0                                 NaN

     Research and Development Expense  \
8                                 NaN
9                                 NaN
10                                NaN
11                                NaN
12                                NaN
..                                ...
827                               NaN
828                               NaN
```

```
829                             NaN
830                             NaN
831                             NaN

     Selling, General and Administrative Expenses
8                                           385.0
9                                           340.0
10                                         1011.0
11                                          749.0
12                                          400.0
..                                            ...
827                                         314.0
828                                         243.0
829                                         358.0
830                                         304.0
831                                         336.0

[240 rows x 362 columns],
     Data Date  Fiscal Year  Accumulated Other Comprehensive Income (Loss)  \
0     20100331         2010                                        -1749.0
1     20100630         2010                                        -1603.0
2     20100930         2010                                        -1377.0
3     20101231         2010                                        -1159.0
4     20110331         2011                                         -873.0
..         ...          ...                                            ...
839   20151231         2015                                         -318.0
840   20160331         2016                                         -318.0
841   20160630         2016                                         -321.0
842   20160930         2016                                         -327.0
843   20161231         2016                                         -234.0

     Current Assets - Other - Total  Current Assets - Total  \
0                            1502.0                  8780.0
1                            1434.0                  8296.0
2                             865.0                  8839.0
3                            1002.0                  8780.0
4                             759.0                  9436.0
..                              ...                     ...
839                           183.0                  9471.0
840                           204.0                  8097.0
841                           142.0                 10304.0
842                           176.0                  9545.0
843                           236.0                 10401.0

     Other Long-term Assets  Non-Current Assets - Total  \
0                    2568.0                     21169.0
1                    2587.0                     21204.0
2                    2742.0                     24646.0
3                    2638.0                     26616.0
4                    2602.0                     27201.0
..                      ...                         ...
839                   119.0                     33644.0
840                   886.0                     33661.0
841                   875.0                     33829.0
842                   848.0                     33748.0
843                   107.0                     34012.0

     Assets Netting & Other Adjustments  \
0                                -299.0
1                                -254.0
2                                -228.0
3                                -517.0
4                                -789.0
..                                  ...
839                               -62.0
```

```
840                             -67.0
841                            -146.0
842                            -278.0
843                            -688.0

     Accum Other Comp Inc - Derivatives Unrealized Gain/Loss  \
0                                              -1312.0
1                                              -1058.0
2                                               -962.0
3                                               -786.0
4                                               -688.0
..                                                 ...
839                                                4.0
840                                                4.0
841                                                4.0
842                                                4.0
843                                                4.0

     Accum Other Comp Inc - Other Adjustments  ...  \
0                                       -33.0  ...
1                                      -148.0  ...
2                                       -27.0  ...
3                                        11.0  ...
4                                       192.0  ...
..                                        ...  ...
839                                       0.0  ...
840                                       0.0  ...
841                                       0.0  ...
842                                       0.0  ...
843                                       0.0  ...

     Implied Option EPS Diluted  Implied Option EPS Diluted Preliminary  \
0                           0.0                                     0.0
1                           0.0                                     0.0
2                           0.0                                     0.0
3                           0.0                                     0.0
4                           0.0                                     0.0
..                          ...                                     ...
839                         NaN                                     NaN
840                         NaN                                     NaN
841                         NaN                                     NaN
842                         NaN                                     NaN
843                         NaN                                     NaN

     Implied Option EPS Basic 12MM  Implied Option 12MM EPS Basic Preliminary  \
0                              0.0                                        0.0
1                              0.0                                        0.0
2                              0.0                                        0.0
3                              0.0                                        0.0
4                              0.0                                        0.0
..                             ...                                        ...
839                            NaN                                        NaN
840                            NaN                                        NaN
841                            NaN                                        NaN
842                            NaN                                        NaN
843                            NaN                                        NaN

     Implied Option EPS Basic  Implied Option EPS Basic Preliminary  \
0                         0.0                                   0.0
1                         0.0                                   0.0
2                         0.0                                   0.0
3                         0.0                                   0.0
4                         0.0                                   0.0
..                        ...                                   ...
839                       NaN                                   NaN
```

```
840                    NaN                               NaN
841                    NaN                               NaN
842                    NaN                               NaN
843                    NaN                               NaN

     Implied Option Expense  Implied Option Expense Preliminary  \
0                      0.0                                  0.0
1                      0.0                                  0.0
2                      0.0                                  0.0
3                      0.0                                  0.0
4                      0.0                                  0.0
..                     ...                                  ...
839                    NaN                                  NaN
840                    NaN                                  NaN
841                    NaN                                  NaN
842                    NaN                                  NaN
843                    NaN                                  NaN

     Research and Development Expense  \
0                                 NaN
1                                 NaN
2                                 NaN
3                                 NaN
4                                 NaN
..                                ...
839                               NaN
840                               NaN
841                               NaN
842                               NaN
843                               NaN

     Selling, General and Administrative Expenses
0                                           559.0
1                                           576.0
2                                           608.0
3                                           805.0
4                                           760.0
..                                            ...
839                                         403.0
840                                         378.0
841                                         401.0
842                                         420.0
843                                         406.0

[604 rows x 362 columns])
```

In [37]:

```
                                         # Question 3:



import sklearn
from sklearn import linear_model
from sklearn import preprocessing
from sklearn import datasets
from sklearn import svm
import numpy as np

from sklearn import linear_model
from sklearn.model_selection import cross_val_score
from sklearn.metrics import classification_report, confusion_matrix
from sklearn.metrics import mean_squared_error, r2_score
```

In [38]:

```
#Question 3 - 1

#Load the dataset
diabets_df=datasets.load_diabetes()
```

In [39]:

```
#Display the data
diabets_df
```

Out[39]:

```
{'data': array([[ 0.03807591,  0.05068012,  0.06169621, ..., -0.00259226,
          0.01990842, -0.01764613],
        [-0.00188202, -0.04464164, -0.05147406, ..., -0.03949338,
         -0.06832974, -0.09220405],
        [ 0.08529891,  0.05068012,  0.04445121, ..., -0.00259226,
          0.00286377, -0.02593034],
        ...,
        [ 0.04170844,  0.05068012, -0.01590626, ..., -0.01107952,
         -0.04687948,  0.01549073],
        [-0.04547248, -0.04464164,  0.03906215, ...,  0.02655962,
          0.04452837, -0.02593034],
        [-0.04547248, -0.04464164, -0.0730303 , ..., -0.03949338,
         -0.00421986,  0.00306441]]),
 'target': array([151.,  75., 141., 206., 135.,  97., 138.,  63., 110., 310., 101.,
         69., 179., 185., 118., 171., 166., 144.,  97., 168.,  68.,  49.,
         68., 245., 184., 202., 137.,  85., 131., 283., 129.,  59., 341.,
         87.,  65., 102., 265., 276., 252.,  90., 100.,  55.,  61.,  92.,
        259.,  53., 190., 142.,  75., 142., 155., 225.,  59., 104., 182.,
        128.,  52.,  37., 170., 170.,  61., 144.,  52., 128.,  71., 163.,
        150.,  97., 160., 178.,  48., 270., 202., 111.,  85.,  42., 170.,
        200., 252., 113., 143.,  51.,  52., 210.,  65., 141.,  55., 134.,
         42., 111.,  98., 164.,  48.,  96.,  90., 162., 150., 279.,  92.,
         83., 128., 102., 302., 198.,  95.,  53., 134., 144., 232.,  81.,
        104.,  59., 246., 297., 258., 229., 275., 281., 179., 200., 200.,
        173., 180.,  84., 121., 161.,  99., 109., 115., 268., 274., 158.,
        107.,  83., 103., 272.,  85., 280., 336., 281., 118., 317., 235.,
         60., 174., 259., 178., 128.,  96., 126., 288.,  88., 292.,  71.,
        197., 186.,  25.,  84.,  96., 195.,  53., 217., 172., 131., 214.,
         59.,  70., 220., 268., 152.,  47.,  74., 295., 101., 151., 127.,
        237., 225.,  81., 151., 107.,  64., 138., 185., 265., 101., 137.,
        143., 141.,  79., 292., 178.,  91., 116.,  86., 122.,  72., 129.,
        142.,  90., 158.,  39., 196., 222., 277.,  99., 196., 202., 155.,
         77., 191.,  70.,  73.,  49.,  65., 263., 248., 296., 214., 185.,
         78.,  93., 252., 150.,  77., 208.,  77., 108., 160.,  53., 220.,
        154., 259.,  90., 246., 124.,  67.,  72., 257., 262., 275., 177.,
         71.,  47., 187., 125.,  78.,  51., 258., 215., 303., 243.,  91.,
        150., 310., 153., 346.,  63.,  89.,  50.,  39., 103., 308., 116.,
        145.,  74.,  45., 115., 264.,  87., 202., 127., 182., 241.,  66.,
         94., 283.,  64., 102., 200., 265.,  94., 230., 181., 156., 233.,
         60., 219.,  80.,  68., 332., 248.,  84., 200.,  55.,  85.,  89.,
         31., 129.,  83., 275.,  65., 198., 236., 253., 124.,  44., 172.,
        114., 142., 109., 180., 144., 163., 147.,  97., 220., 190., 109.,
        191., 122., 230., 242., 248., 249., 192., 131., 237.,  78., 135.,
        244., 199., 270., 164.,  72.,  96., 306.,  91., 214.,  95., 216.,
        263., 178., 113., 200., 139., 139.,  88., 148.,  88., 243.,  71.,
         77., 109., 272.,  60.,  54., 221.,  90., 311., 281., 182., 321.,
         58., 262., 206., 233., 242., 123., 167.,  63., 197.,  71., 168.,
        140., 217., 121., 235., 245.,  40.,  52., 104., 132.,  88.,  69.,
        219.,  72., 201., 110.,  51., 277.,  63., 118.,  69., 273., 258.,
         43., 198., 242., 232., 175.,  93., 168., 275., 293., 281.,  72.,
        140., 189., 181., 209., 136., 261., 113., 131., 174., 257.,  55.,
         84.,  42., 146., 212., 233.,  91., 111., 152., 120.,  67., 310.,
         94., 183.,  66., 173.,  72.,  49.,  64.,  48., 178., 104., 132.,
        220.,  57.]),
```

```
 'frame': None,
 'DESCR': '.. _diabetes_dataset:\n\nDiabetes dataset\n----------------\n\nTen baseline v
ariables, age, sex, body mass index, average blood\npressure, and six blood serum measur
ements were obtained for each of n =\n442 diabetes patients, as well as the response of
interest, a\nquantitative measure of disease progression one year after baseline.\n\n**D
ata Set Characteristics:**\n\n  :Number of Instances: 442\n\n  :Number of Attributes: Fi
rst 10 columns are numeric predictive values\n\n  :Target: Column 11 is a quantitative m
easure of disease progression one year after baseline\n\n  :Attribute Information:\n
- age     age in years\n      - sex\n      - bmi     body mass index\n      - bp      av
erage blood pressure\n      - s1      tc, T-Cells (a type of white blood cells)\n      -
s2      ldl, low-density lipoproteins\n      - s3      hdl, high-density lipoproteins\n
- s4      tch, thyroid stimulating hormone\n      - s5      ltg, lamotrigine\n      - s6
glu, blood sugar level\n\nNote: Each of these 10 feature variables have been mean center
ed and scaled by the standard deviation times `n_samples` (i.e. the sum of squares of ea
ch column totals 1).\n\nSource URL:\nhttps://www4.stat.ncsu.edu/~boos/var.select/diabete
s.html\n\nFor more information see:\nBradley Efron, Trevor Hastie, Iain Johnstone and Ro
bert Tibshirani (2004) "Least Angle Regression," Annals of Statistics (with discussion),
407-499.\n(https://web.stanford.edu/~hastie/Papers/LARS/LeastAngle_2002.pdf)',
 'feature_names': ['age',
  'sex',
  'bmi',
  'bp',
  's1',
  's2',
  's3',
  's4',
  's5',
  's6'],
 'data_filename': 'C:\\Users\\Bipin\\anaconda3\\lib\\site-packages\\sklearn\\datasets\\d
ata\\diabetes_data.csv.gz',
 'target_filename': 'C:\\Users\\Bipin\\anaconda3\\lib\\site-packages\\sklearn\\datasets
\\data\\diabetes_target.csv.gz'}
```

In [40]:

```python
                                                  #Question3 - 2:
new_df= diabets_df.data[:,np.newaxis,2]

atest=new_df[-20:]
atrain=new_df[:-20]

btest=diabets_df.target[-20:]
btrain=diabets_df.target[:-20]
```

In [41]:

```python
                                                  #Question3 - 3:


var_lm=linear_model.LinearRegression()

var_lm.fit(atrain,btrain)
model_y= var_lm.predict(atest)
model_y
coef=var_lm.coef_
ms_error=mean_squared_error(btest,model_y)
r_square=r2_score(btest,model_y)
print(coef)
print(ms_error)
print(r_square)
```

```
[938.23786125]
2548.0723987259703
0.47257544798227136
```

In [46]:
```
                                                            #Question3 - 4:

var_fold=10

new_var_lm=var_lm.fit(atrain,btrain)

rang=list(range(1,11))

#use For Loop
for res_data in rang:

    res_data=cross_val_score(new_var_lm,atrain, btrain,cv=10)
    #print the result
    print(res_data)
```

```
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
[0.28527158 0.05133253 0.24112318 0.44536036 0.23478297 0.38300126
 0.43946172 0.25480749 0.32934734 0.31765983]
```

In [48]:
```
                                                #Question3 - 5

from sklearn.ensemble import RandomForestRegressor
from sklearn.metrics import accuracy_score

regressor = RandomForestRegressor(max_depth = 7,random_state = 0)
```

In [49]:
```
regressor.fit(atrain,btrain)
pred = regressor.predict(atest)
print(pred)
```

```
[296.05572177  95.45286735 203.13858856  86.22366823 134.69925732
 172.44011759 266.00002655 113.6590721  102.92884903  97.98467848
 181.16902649  84.89712172 178.36582945 110.96031125  89.13732091
 234.07248578 111.99641223 111.99641223 187.42421128  98.03151117]
```

In [50]:
```
regressor.score(atrain,btrain)
```

Out[50]:
```
0.5264240887187687
```

In [51]:
```
                                                #Question3 - 6
```

```
from sklearn.model_selection import GridSearchCV
```

In [52]:
```
hy_parameters = {'max_depth': [None,7,4],'min_samples_split':[2,10,20]}
```

In [53]:
```
grid_GBR = GridSearchCV(estimator = regressor,param_grid = hy_parameters,cv = 2,n_jobs
```

In [54]:
```
grid_GBR.fit(atrain,btrain)
```

Out[54]:
```
GridSearchCV(cv=2, estimator=RandomForestRegressor(max_depth=7, random_state=0),
             n_jobs=1,
             param_grid={'max_depth': [None, 7, 4],
                         'min_samples_split': [2, 10, 20]})
```

In [56]:
```
print('Results')
print('\nThe best estimator:',grid_GBR.best_estimator_)
print('\nThe best score :',grid_GBR.best_score_)
print('\nThe best parameter:',grid_GBR.best_params_)
```

```
Results

The best estimator: RandomForestRegressor(max_depth=4, min_samples_split=20, random_stat
e=0)

The best score : 0.2854161850966666

The best parameter: {'max_depth': 4, 'min_samples_split': 20}
```

In [57]:
```
                                                          # Question 2



import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import datetime as dt
```

In [59]:
```
data = pd.read_excel('D:\Ruchi Folder\Python\Assignment5\ResearchDatasetV2.0.xlsx',pars
```

In [63]:
```
data['Date'] = pd.to_datetime(data['Date'], format='%Y%m%d')
data.describe()
```

Out[63]:

	Signal	ClosePrice
count	667.000000	667.000000
mean	5.166802	163.169369
std	23.392809	39.210384
min	-3.802670	127.495000
25%	3.418083	140.880000

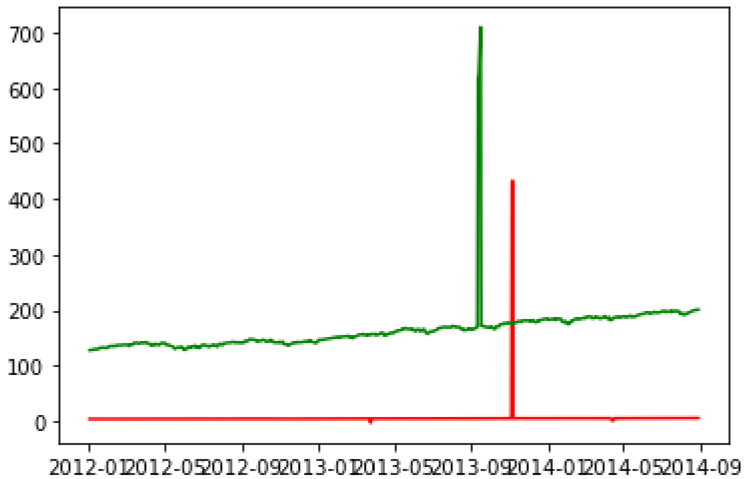
	Signal	ClosePrice
50%	3.893689	159.750000
75%	4.408313	181.500000
max	432.961165	710.310000

In [62]:

```
y = data['Signal'].tolist()
x = data['ClosePrice'].tolist()
```

In [64]:

```
plt.plot(data['Date'].tolist(), y, color='r', label='signal')
plt.plot(data['Date'].tolist(), x, color='g', label='closeprice')
plt.show()
```



In [65]:

```
data['signal_return'] = data['Signal'].pct_change()
data['price_return'] = data['ClosePrice'].pct_change()
data.describe()
```

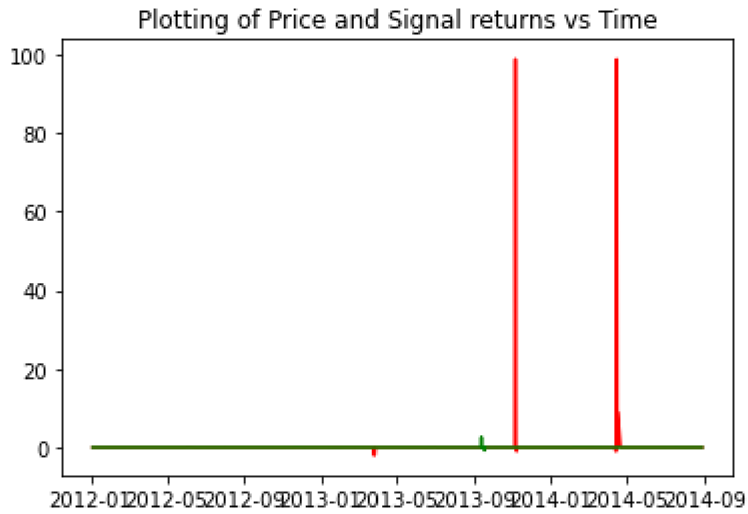
Out[65]:

	Signal	ClosePrice	signal_return	price_return
count	667.000000	667.000000	666.000000	666.000000
mean	5.166802	163.169369	0.301753	0.003760
std	23.392809	39.210384	5.421266	0.107373
min	-3.802670	127.495000	-2.003073	-0.759161
25%	3.418083	140.880000	-0.003357	-0.003199
50%	3.893689	159.750000	0.000111	0.000727
75%	4.408313	181.500000	0.005180	0.005070
max	432.961165	710.310000	98.796977	2.653778

In [67]:

```
plt.plot(data['Date'].tolist(), data['signal_return'], color='r', label='signalreturn')
plt.plot(data['Date'].tolist(), data['price_return'], color='g', label='closepriceretur
```

```
plt.title('Plotting of Price and Signal returns vs Time')
plt.show()
```

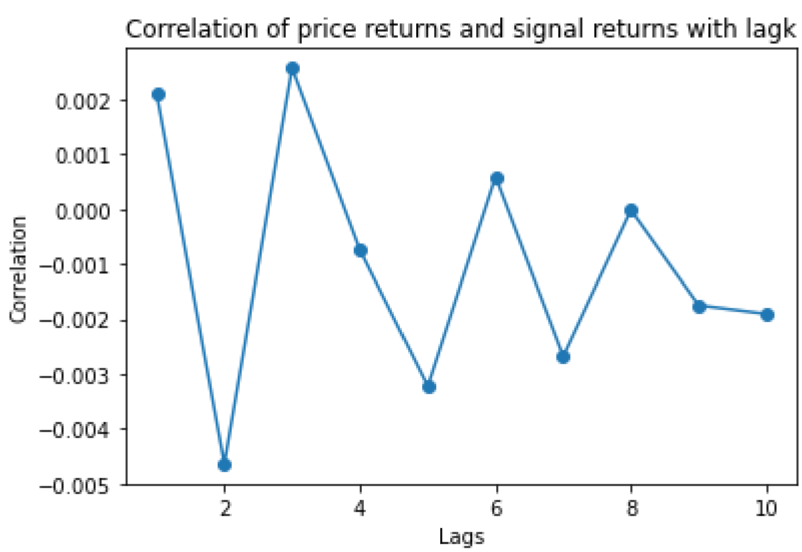


In [68]:

```
def cross_correlate(x, y, lag):
    return x.corr(y.shift(lag))
```

In [70]:

```
cross_corrs = []
for lag in range(1, 11):
    cross_corrs.append(cross_correlate(data['price_return'],
                                       data['signal_return'], lag))
plt.plot(range(1,11), cross_corrs, marker='o')
plt.xlabel('Lags')
plt.ylabel('Correlation')
plt.title('Correlation of price returns and signal returns with lagk')
plt.show()
```



In [72]:

```
grp_1 = []
grp_1_idx = []
grp_2 = []
grp_2_idx = []
grp_3 = []
```

```
grp_3_idx = []
grp_4 = []
grp_4_idx = []
for index in range(len(data['signal_return'])):
    if data['signal_return'][index] < -0.01:
        grp_1.append(data['signal_return'][index])
        grp_1_idx.append(index)
    elif data['signal_return'][index] >= -0.01 and data['signal_return'][index] < 0:
        grp_2.append(data['signal_return'][index])
        grp_2_idx.append(index)
    elif data['signal_return'][index] >= 0 and data['signal_return'][index] < 0.01:
        grp_3.append(data['signal_return'][index])
        grp_3_idx.append(index)
    elif data['signal_return'][index] >= 0.01:
        grp_4.append(data['signal_return'][index])
        grp_4_idx.append(index)
grp_1_idx = [x+3 for x in grp1_idx if x+3 < len(data.index)]
grp_2_idx = [x+3 for x in grp2_idx if x+3 < len(data.index)]
grp_3_idx = [x+3 for x in grp3_idx if x+3 < len(data.index)]
grp_4_idx = [x+3 for x in grp4_idx if x+3 < len(data.index)]

#calculate the average return
average_returns = [np.mean(data['price_return'][grp_1_idx]),
                   np.mean(data['price_return'][grp_2_idx]),
                   np.mean(data['price_return'][grp_3_idx]),
                   np.mean(data['price_return'][grp_4_idx])]
grps = ['<-0.01', '-0.01:0', '0:0.01', '>0.01']
```

In [73]:

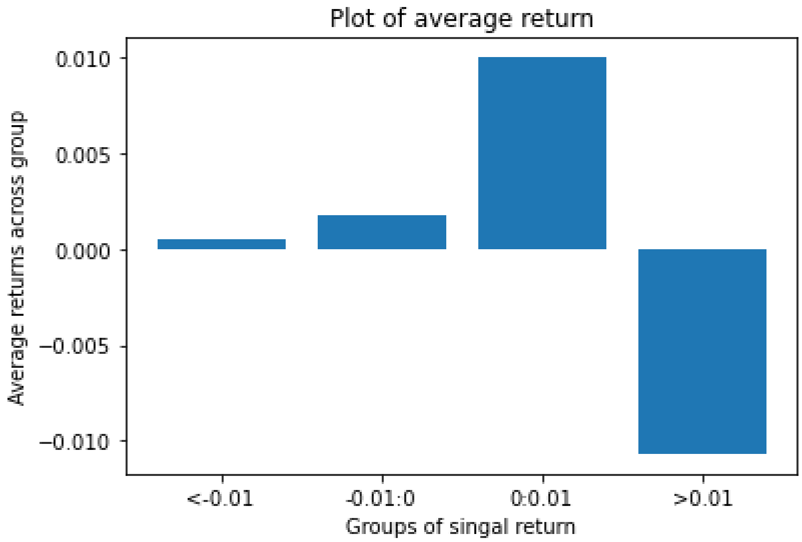
```
average_returns
```

Out[73]:

```
[0.0005273010550047138,
 0.0017594901175337868,
 0.009988932451912276,
 -0.010687383162841551]
```

In [78]:

```
plt.bar(grps, average_returns)
plt.title('Plot of average return')
plt.xlabel('Groups of singal return')
plt.ylabel('Average returns across group')
plt.show()
```

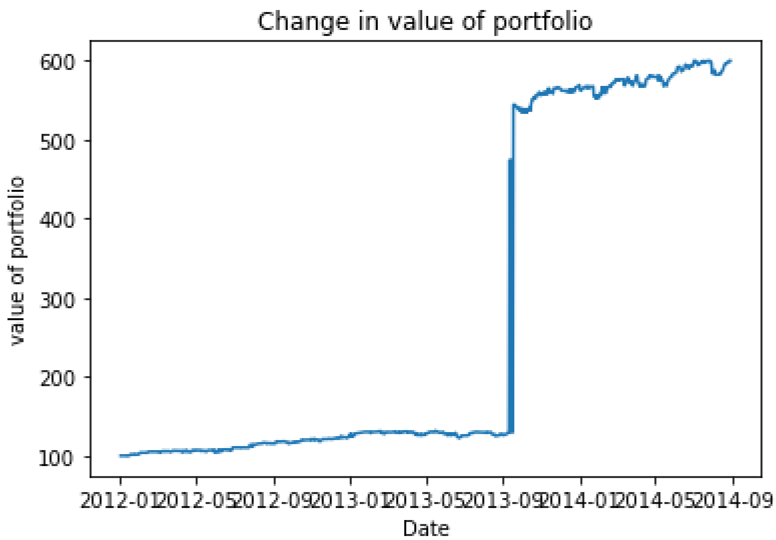


In [ ]:

```
value = [100]
investment_available = 100
data['buy_signal'] = pd.Series(np.zeros(len(data.index)))
data['sell_signal'] = pd.Series(np.zeros(len(data.index)))
invested = False
for i in range(len(data.index)):

    if data['buy_signal'][i] == 1:
        shares = investment_available/data['ClosePrice'][i]
        invested = True

    if data['sell_signal'][i] == 1:
        investment_available = shares * data['ClosePrice'][i]

    if data['signal_return'][i] < 0 and invested == False:
        data['buy_signal'][i+3] = 1


    if invested == False:
        value.append(value[i-1])

    if invested == True:
        value.append(shares*data['ClosePrice'][i])

    if data['sell_signal'][i] == 1:
        invested = False

    if data['signal_return'][i] > 0.01 and invested == True:
        data['sell_signal'][i+3] = 1
```

In [79]:

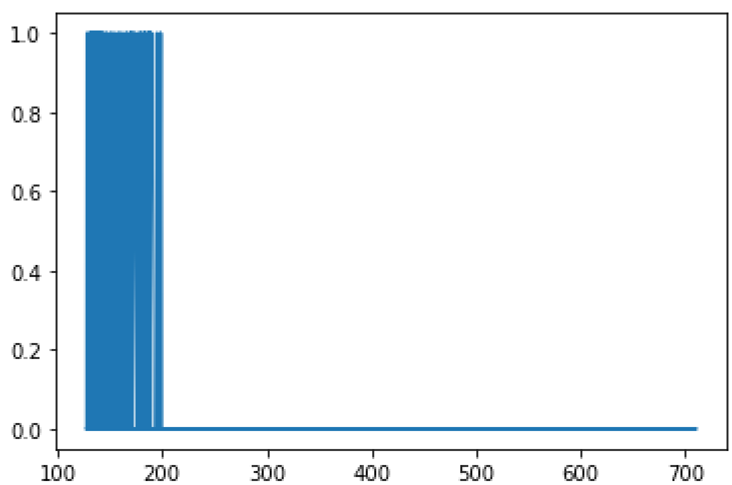
```
plt.plot(data['Date'], value[1:])
plt.xlabel('Date')
plt.ylabel('value of portfolio')
plt.title('Change in value of portfolio')
plt.show()
```



In [80]:

```
total_return = value[-1]/value[0] -1
total_return
plt.plot(data['ClosePrice'], data['buy_signal'])
risk_free = 0.01 #1% rate of return on risk free asset
portfolio_return = total_return # Calculated above
std = np.std(pd.Series(value).pct_change())
sharpe = (total_return - risk_free) / std
sharpe
```

Out[80]: 25.612455172694524



In [ ]: