Kehua Chu, UID: 806153163

.

1.) Pull in Data and Convert ot Monthly

In [1]:

```
import yfinance as yf
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [10]:

```
apple_data = yf.download(’AAPL’)
df = apple_data.resample(”M”).last()[[”Adj Close”]]
```

```
[*********************100%%**********************]  1 of 1 completed
```

2.) Create columns.

- Current Stock Price, Difference in stock price, Whether it went up or down over the next month, option premium

In [11]:

```
# Difference in stockprice, to the next period
df[’Diff’] = df.diff().shift(-1)
```

In [13]:

```
# Target up or down
df[’Target’] = np.sign(df[’Diff’])
```

In [26]:

```
# Option Premium
df[’Premium’] = 0.08 * df[’Adj Close’]
```

In [27]:

```
# if we got Falso Positive, -100 diff  -> we really don’t want
df
```

Out[27]:

Date	Adj Close	Diff	Target	Premium	Predictions
1980-12-31	0.117887	-0.020296	-1.0	0.009431	-1.0
1981-01-31	0.097591	-0.006045	-1.0	0.007807	-1.0
1981-02-28	0.091546	-0.006909	-1.0	0.007324	-1.0
1981-03-31	0.084637	0.013386	1.0	0.006771	1.0
1981-04-30	0.098023	0.016409	1.0	0.007842	1.0
...	...	...	...	...	...
2023-05-31	176.778061	16.675507	1.0	14.142245	1.0
2023-06-30	193.453568	2.473389	1.0	15.476285	1.0
2023-07-31	195.926956	-8.304138	-1.0	15.674156	-1.0
2023-08-31	187.622818	-16.638077	-1.0	15.009825	-1.0
2023-09-30	170.984741	-0.439423	-1.0	13.678779	-1.0

514 rows × 5 columns

3.) Pull in X data, normalize and build a LogReg on column 2

In [15]:

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn import metrics
```

In [28]:

```
X = pd.read_csv("Xdata.csv", index_col="Date", parse_dates=["Date"])
```

In [29]:

```
y = df.loc[:"2023-09-30","Target"].copy()
df = df.loc[:"2023-09-30",:].copy()
```

In [30]:

```
logreg = LogisticRegression()

logreg.fit(X, y)

y_pred = logreg.predict(X)
```

4.) Add columns, prediction and profits.

In [31]:

```
df['Predictions'] = y_pred
```

In [45]:

```
df['Profits'] = 0.

# True positives

# df.loc[(df['Predictions'] == 1) & (df['Target'] == 1),'Profits' ] = df.loc[(df['Predictions']
df.loc[(df['Predictions'] == 1) & (df['Target'] == 1),'Profits' ] = df['Premium'] # rows, coma,

# False positives
df.loc[(df['Predictions'] == 1) & (df['Target'] == -1),'Profits' ] = 100*df['Diff'] + df['Prem

# True Negatives
# = 0

# False Negatives
# = 0
```

In [46]:

```
df
```

Out[46]:

	Adj Close	Diff	Target	Premium	Predictions	Profits
Date						
1980-12-31	0.117887	-0.020296	-1.0	0.009431	-1.0	0.000000
1981-01-31	0.097591	-0.006045	-1.0	0.007807	-1.0	0.000000
1981-02-28	0.091546	-0.006909	-1.0	0.007324	-1.0	0.000000
1981-03-31	0.084637	0.013386	1.0	0.006771	1.0	0.006771
1981-04-30	0.098023	0.016409	1.0	0.007842	1.0	0.007842
...	...	...	...	...	...	...
2023-05-31	176.778061	16.675507	1.0	14.142245	1.0	14.142245
2023-06-30	193.453568	2.473389	1.0	15.476285	1.0	15.476285
2023-07-31	195.926956	-8.304138	-1.0	15.674156	-1.0	0.000000
2023-08-31	187.622818	-16.638077	-1.0	15.009825	-1.0	0.000000
2023-09-30	170.984741	-0.439423	-1.0	13.678779	-1.0	0.000000

514 rows × 6 columns

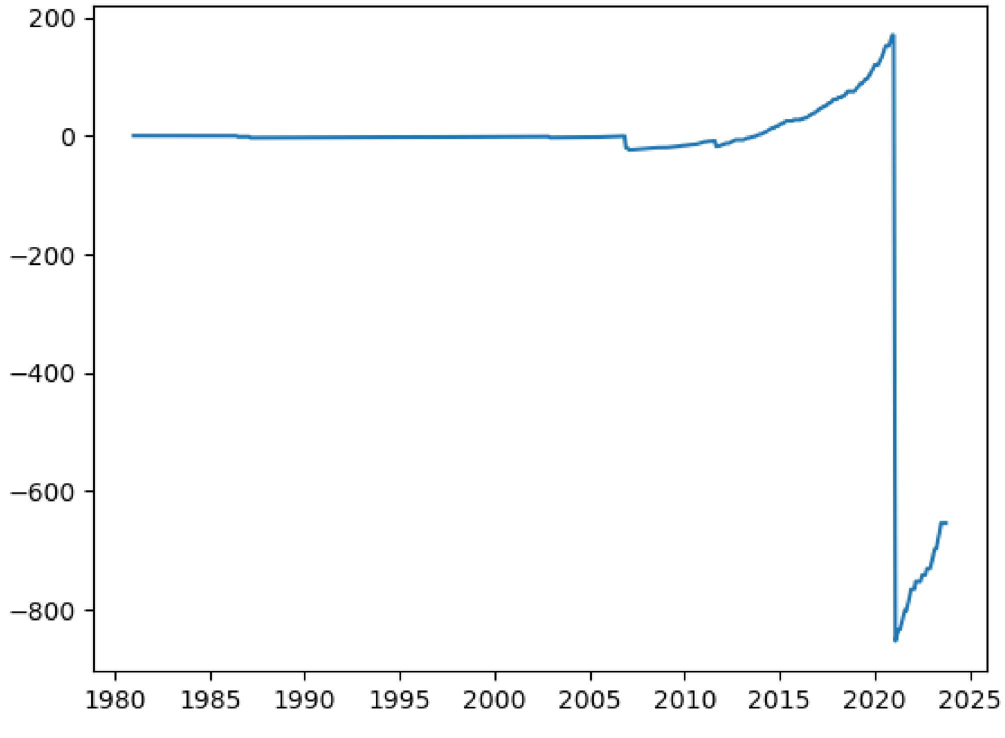
5.) Plot profits over time

In [47]:

```
plt.plot(np.cumsum(df['Profits']))
plt.show
```

Out[47]:

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



In [ ]:

```
# pandemic >> huge loss
# then, picking up small returns
```

5.5.) Your skills from MQE to help Mr.Lius ventures

Mr. Lius's ventures have some many business segments including content creastor selling personal brands, data colelction to make new features like trading strategies, data analysis realted to blockchain and so on. Several skills from MQE may help his busniess. Firstly, using the data collection, washing and analysing skill, we can help his venture with data analysis works no matter which industry the service is aiming at. Secondly, since we have strong storytelling skills, we can help customers, who are not familiar with trading and data, to know this ventures' service better.