

```

1 import quandl
2 import numpy as np
3 from sklearn.naive_bayes import GaussianNB
4 from sklearn.model_selection import train_test_split
5 from sklearn import preprocessing
6 from sklearn.metrics import accuracy_score, confusion_matrix
7

```

```

1 df = quandl.get('WIKI/AMZN').
2 df.head()

```



| | Open | High | Low | Close | Volume | Ex-Dividend | Split Ratio | Adj. Open | Adj. High | Adj. Low | Adj. Close |
|------------|-------|-------|-------|-------|-----------|-------------|-------------|-----------|-----------|----------|------------|
| Date | | | | | | | | | | | |
| 1997-05-16 | 22.38 | 23.75 | 20.50 | 20.75 | 1225000.0 | 0.0 | 1.0 | 1.865000 | 1.979167 | 1.700000 | 1.700000 |
| 1997-05-19 | 20.50 | 21.25 | 19.50 | 20.50 | 508900.0 | 0.0 | 1.0 | 1.708333 | 1.770833 | 1.666667 | 1.666667 |
| 1997-05-20 | 20.75 | 21.00 | 19.63 | 19.63 | 455600.0 | 0.0 | 1.0 | 1.729167 | 1.750000 | 1.666667 | 1.666667 |

```

1 df['Profit'] = np.where(df['Open']<df['Close'], 1, 0)
2 df = df.drop(columns=['Ex-Dividend', 'Split Ratio'])

1 X = df.drop(columns=['Profit']).

1 y = df['Profit'].

1 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3).

1 model = GaussianNB()
2 model.fit(X_train, y_train)
3 y_pred = model.predict(X_test)
4 print('Accuracy Score: ',model.score(X_test, y_test))
5 print('Confusion Matrix', confusion_matrix(y_pred, y_test))

```



```

Accuracy Score: 0.5085714285714286
Confusion Matrix [[660 642]
                  [132 141]]

```