Name Haojuan He
Berkeley ID 3033721461.

Prof. *Evgeniya Duzhak*
ECON 140 Fall 2020

Homework 4

[55 points] 1. At the end of each semester students evaluate the quality of their professors. Among many questions, students rate Professor's overall quality and perceived easiness of the Professor's class. We obtained the sample of 10 data points on the scores of Professor's perceived easiness (5 = Easy 1 = Hard) (X) and their overall rating(5=the Best 1 = the worst) (Y). Some people might argue that Professors that teach easier classes get higher evaluation scores, to find out whether it is in fact true use the following data to answer the questions below.

Y	2.7	3.4	3.8	2.6	4.4	4.8	2.5	1.4	4.5	4.9
X	2.1	2.3	2.7	3.2	3.8	3.6	2.2	1.7	4.4	4

Note: a series of massive sunspot activities in September caused major magnetic interferences that ruined all electronics, including your laptops. It left you with no choice but to complete the following tasks by hand.

(a) Calculate the following (show your work!): $\bar{X}, \bar{Y}, \sum_{i=1}^{10}(X_i - \bar{X})(Y_i - \bar{Y}), \sum_{i=1}^{10}(X_i - \bar{X})^2$

$\bar{X} = (2.1+2.3+2.7+3.2+3.8+3.6+2.2+1.7+4.4+4)/10$
$= 30/10$
$= 3$

$\bar{Y} = (2.7+3.4+3.8+2.6+4.4+4.8+2.5+1.4+4.5+4.9)/10$
$= 35/10$
$= 3.5$

$\sum_{i=1}^{10}(X_i - \bar{X})(Y_i - \bar{Y}) = -0.8*-0.9+(-0.1)*(-0.7)+(-0.3)*0.9+0.2*(-0.9)$
$+(0.8)*0.9+1.3*0.6+(-1)*(-0.8)+(-2.1)*(-1.3)+1*1.4+1.4*1$
$= 0.72+0.07+(-0.09)+(-0.18)+0.72+0.78+0.8+2.73+1.4+1.4$

$= 8.35.$

$\sum_{i=1}^{10}(X_i - \bar{X})^2 = (-0.9)^2+(-0.7)^2+(0.3)^2+(+0.2)^2+(0.8)^2+0.6^2+(-0.8)^2+(-1.3)^2$
$+1^2+1.4^2$
$= 0.81+0.49+0.09+0.04+0.64+0.36+0.64+1.69+1.96$

$= 7.72$

(b) Assume that these observations were generated by the following model:

$$Y = \beta_0 + \beta_1 X + u, \text{ where } u \text{ is the random error term}$$

Compute OLS estimators of β_0 and β_1 using the formulas derived in class

$\hat{\beta}^{OLS}_1 = \widehat{cov}(X,Y) / Var(X) = \Sigma(X_i - \bar{X})(Y_i - \bar{Y}) / \Sigma(X_i - \bar{X})^2$

$= 8.35 / 7.72$

≈ 1.08

$\hat{\beta}^{OLS}_o = \bar{Y} - \hat{\beta}_1 * \bar{X}$

$= 3.5 - 1.08 * 3$

$= 0.26$

(c) Provide an interpretation for the slope and the intercept that you estimated.

[2 p.] $\widehat{\beta_0}$ $\Rightarrow$ the rating a prof will get if the class is extremely hard $(X=0)$ is 0.26

[3 p.] $\widehat{\beta_1}$ $\Rightarrow$ each decreasing level of difficulty will increase the class rating by 1.08 points

(d) Calculate the predicted rating, $\hat{Y}$ and the estimated error $\hat{u}_i = Y_i - \hat{Y}_i$ for all $i=1,\ldots,10$

#	1	2	3	4	5	6	7	8	9	10
$\hat{Y}_i$	2.528	2.744	3.176	3.716	4.364	4.148	2.636	2.096	5.012	4.48
$\hat{u}_i$	0.172	0.656	0.624	-1.116	0.036	0.652	-0.136	-0.696	-0.512	0.32.

$\Rightarrow \hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

$= 0.26 + 1.08 * X_i$

$\Rightarrow \hat{u}_i = Y_i - \hat{Y}_i.$

(d) The Chair of the Economics Department honors the best teachers with an award. To choose the best candidate, she looks at end-of-term faculty evaluations by students and picks the professors with the highest scores. Someone points out that professors teaching easy classes tend to get higher evaluations and the Chair needs to adjust for the class easiness. Given the regression model you used above, propose a way to recognize the teachers that outperform the others.

⇒ since $\hat{\beta}_1$ is positive, it is true that there's a positive relation btw the easiness of the class and the higher rating

⇒ so simply use the rating might be bias for those who teach difficult classess

⇒ a good way to fix this is that since we already know how much estimated rating of the teacher will increase, with each increase unit of easyness of the class

⇒ we can set an universal avg difficulty of every class to k.

⇒ the adjust the rating of each class considering the difference btw the actual difficulty and our set value k.

⇒ adjusted rating = original rating − 1.08 * (actual difficulty − k).

⇒ then we can use this adjusted rating to recognize the teachers that out perform others.

2. Your manager gives you a task of evaluating the performance of a particular stock by estimating stock's beta. A stock's (or portfolio's) β (beta) is a measure of the volatility of that stock relative to the volatility of the market—therefore, the market has a β equal to one. For individual securities, β can take on any value, negative or positive.

Some basic facts about β:

- If an individual security has $\beta = 1$, then that security tends to move in line with the market.
- If $\beta = 0$ the returns on the security are uncorrelated with returns of the market.
- If $0 < \beta < 1$ the security tends to move in the same direction as the market, but is less volatile than the market.
- If $\beta > 1$ the security tends to move in the same direction as the market, but is more volatile than the market.
- If $\beta < 0$ the security tends to move in the opposite direction of the market. Above average returns on the market would tend to be associated with below average returns on the security and vice versa.

A stock's β is calculated from an OLS regression of the daily returns of the stock on the daily returns of the market. If R_{st} is the return of the stock on day t, R_{mt} is the return of the market at time t and R_f is a risk-free return, then we fit the linear regression model

$$R_{st} - R_f = \alpha_s + \beta_s (R_{mt} - R_f) + u_{st}$$

and use the estimate of the slope as the β for the stock.

In this exercise, you will calculate β for "YOUR stock" using all available data from September 2, 2019 to present. We will use the returns of the S&P 500 to serve as our "market returns" in the calculation and the *average return* on three-month treasury bonds as a risk-free return. The steps below walk you through the calculation.

YOUR stock - is any stock that starts with the 1st letter of your Last name

For example: For Evgeniya Duzhak

"YOUR stock" is *DIS* for Walt Disney

(a) Find and download the daily three-month treasury yields and closing values of the S&P 500 and YOUR stock for the time period starting Sep 2, 2019 and ending Sep 1, 2020. Import both data sets into a spreadsheet. What website did you use to download the data? What is the value of R_f?

(b) Calculate the daily percentage returns for the S&P 500 and YOUR stock. The daily return for day t is:

$$\ln(v_t) - \ln(v_{t-1})$$

where v_t is the closing value of the security on day t. What was the daily percentage return of the S&P 500 on March 4, 2020?

(c) Create a new data series measuring excess market return $(R_{mt} - R_f)$ and excess return on YOUR stock $(R_{St} - R_f)$. Write down the formula or command you used.

(d) Produce a scatterplot of the excess return of YOUR stock versus the excess market returns on S&P500. Comment on the strength, direction, and form (e.g., linear, quadratic, non-linear) of the scatterplot. Copy and paste the scatterplot into your document.

In [133]:
```
import numpy as np
import pandas as pd
df_3MT = pd.read_csv('3MTY.csv')
#df_3MT = df_3MT[['Date','Price']]
#df_3MT['N_Date'] = df_3MT['Date'][::-1]
#df_3MT['N_Price'] = df_3MT['Price'][::-1]
df_SP500 = pd.read_csv('SP500_1.csv')
df_HD = pd.read_csv('HD_stock.csv')
df_HD = df_HD[['Date','Close']]
df_SP500 = df_SP500[['Date','Close']]
#temp_numb = [float(i) for i in df_3MT['DTB3']]
df_3MT = df_3MT.iloc[::-1]
```

In [134]:
```
#R_mt = (np.diff(df_SP500['Close']) / df_SP500['Close'][:-1]).values
#bond_return = (np.diff(df_3MT['Price']) / df_3MT['Price'][:-1]).values
R_f = np.mean(df_3MT['Price'])/100
R_f
#df_SP500['SP500'][1]
```

Out[134]:
```
0.008996105610561057
```

website for three month trajecotry: [https://fred.stlouisfed.org/series/DTB3 (https://fred.stlouisfed.org/series/DTB3)](https://fred.stlouisfed.org/series/DTB3) website for closing value of SP500: [https://fred.stlouisfed.org/series/SP500 (https://fred.stlouisfed.org/series/SP500)](https://fred.stlouisfed.org/series/SP500) website for stock price of Home Depot Inc:[https://finance.yahoo.com/quote/HD/history?period1=475718400&period2=1589241600&interval=1wk&filter=history&frequency=1wk (https://finance.yahoo.com/quote/HD/history?period1=475718400&period2=1589241600&interval=1wk&filter=history&frequency=1wk)](https://finance.yahoo.com/quote/HD/history?period1=475718400&period2=1589241600&interval=1wk&filter=history&frequency=1wk) the risk free return R_f is 0.008996105610561057

In [135]:
```
R_mt = np.diff(np.log(np.array(df_SP500['Close'])))
R_HD = np.diff(np.log(np.array(df_HD['Close'])))
df_SP500.loc[df_SP500['Date'] == '2020-03-04']
R_mt[126]
```

Out[135]:
```
-0.034510782514423965
```

The daily percentage return of SP500 on March 4, 2020 is -0.034510782514423965, which is -3.45%

In [136]:
```
exc_mt = R_mt - R_f
exc_hd = R_HD - R_f
```

In [137]:
```
data = {'excessive mt': exc_mt,'excessive HD':exc_hd}
df_excessive = pd.DataFrame(data)
df_excessive
```

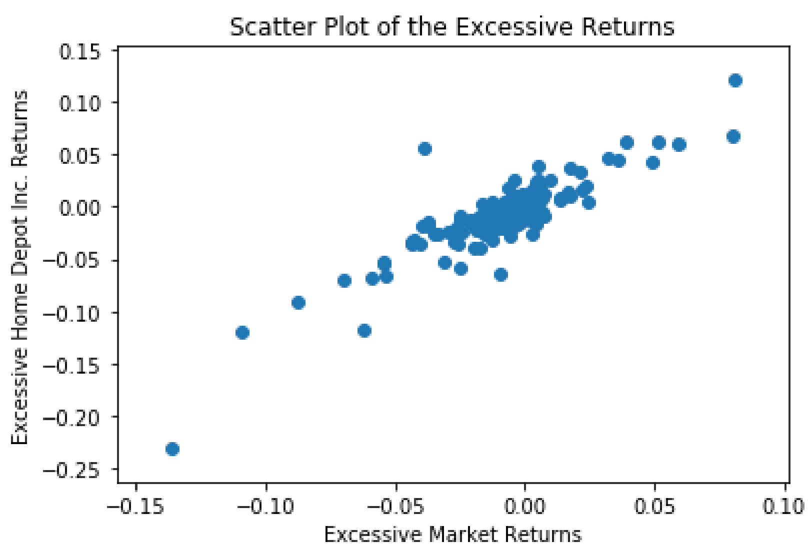
Out[137]:

	excessive mt	excessive HD
0	0.001788	-0.008639
1	0.003930	0.008692
2	-0.008086	0.003981
3	-0.009090	-0.001496
4	-0.008674	-0.008438
...	...	...
247	0.001148	0.011072
248	-0.007324	-0.020365
249	-0.002286	-0.017136
250	-0.011193	-0.013372
251	-0.001499	-0.005844

252 rows × 2 columns

In [138]:
```
import matplotlib.pyplot as plt
plt.scatter(exc_mt,exc_hd)
plt.title('Scatter Plot of the Excessive Returns')
plt.xlabel('Excessive Market Returns')
plt.ylabel('Excessive Home Depot Inc. Returns')
```

Out[138]: `Text(0, 0.5, 'Excessive Home Depot Inc. Returns')`



In [139]:
```
from scipy.stats import pearsonr
corr, p_value = pearsonr(exc_mt, exc_hd)
print(corr)
```

```
0.8695887144053438
```

The excessive market return and excessive Home Depot Inc. return have a pretty strong positive relationship. As ones increase, the other tends to increase as well (change in the same direction). The scatter plot also seems to have a linear relationship betwwen this two variables.

In [ ]: