Name Haojuan He

GSI Yikun (Alison) Jiang.

Prof. Evgeniya Duzhak
ECO 140
Fall 2020

Homework 1

1/ Let A,B,C be annual returns on Apple INC, government Bonds and CityGroup INC stocks respectively. The table below describes the probability distributions for these stock returns based on the state of the economy.

	A	*p(A)*	*B*	*P(B)*	*C*	*P(C)*
Recession	0.04	0.2	0.09	0.2	-0.2	0.2
Normal	0.22	0.6	0.1	0.6	0.1	0.6
Expansion	0.3	0.2	0.11	0.2	0.4	0.2

a) Are A, B and C random variables? Yes

(Yes) No (circle the correct answer)

b) Is it (discrete) or continuous probability distribution? (circle the correct answer)

discrete.

c) For each random variable, compute the expected value, variance, standard deviation. **Comment** on the results for **each** stock. (**Show** your work.)

$E[A] = \sum_{i=1}^{n} A_i * P(A_i) = 0.04*0.2 + 0.22*0.6 + 0.3*0.2$

$= 0.2$

$Var[A] = \sum_{i=1}^{n} (A_i - \mu_A)^2 * P(A_i) = (0.04-0.2)^2*0.2 + (0.22-0.2)^2*0.6 + (0.3-0.2)^2*0.2.$

$= 0.00736$

$sd[A] = \sqrt{Var(A)} = \sqrt{0.00736}$

≈ 0.086

$E[B] = \sum_{i=1}^{n} B_i * P(B_i) = 0.09*0.2 + 0.1*0.6 + 0.11*0.2$

$= 0.1$

$Var[B] = \sum_{i=1}^{n} (B_i - \mu_B)^2 * P(B_i) = (0.09-0.1)^2*0.2 + (0.1-0.1)^2*0.6 + (0.11-0.1)^2*0.2.$

$= 0.00004$

sd[B] = $\sqrt{Var(B)} = \sqrt{0.00004}$

≈ 0.00632

E[C] = $\sum_{i=1}^{n} C_i * P(C_i) = -0.2*0.2 + 0.1*0.6 + 0.4*0.2$

$= 0.1$

Var[C] = $\sum_{i=1}^{n} (C_i - \mu_C)^2 * P(C_i) = (-0.2-0.1)^2*0.2 + (0.1-0.1)^2*0.6 + (0.4-0.1)^2*0.2$

$= 0.036$

sd[C] = $\sqrt{Var(C)} = \sqrt{0.036}$

≈ 0.1897

Comment: the expected value of the annual return means on average, how much you can earn/lost from the bond. The variance means how much will the annual return differ in three economic situation (how much/likely your annual return is going to be different than the on-average annual return).

d) Suppose you have a portfolio consisting of $200 in Apple INC stocks, $200 in government bonds and $100 in CitiGroup INC stocks. Calculate the expected annual return (note that the answer should be in % units):

$E[\text{annual return}] = E\left[\frac{\#\text{change in stock}}{\#\text{original amount}}\right] = E\left[\frac{\$(A)*A + \$(B)*B + \$(C)*C}{200+100+200}\right]$

$= \frac{1}{500} * \{200*E(A) + 100*E[B] + 200*E[C]\}$

$= \frac{1}{500} * \{200*0.2 + 100*0.1 + 200*0.1\} = 70/500 = 0.14$

⇒ thus the expected annual return is 0.14.

e) Given that A, B and C are independent, calculate the variance and standard deviation of the **returns** to that portfolio.(*Hint*: Calculate Var(0.4A+0.4B+0.2C))

⇒ $Var[\text{returns}] = Var[0.4A + 0.4B + 0.2C]$

$= 0.4^2 Var[A] + 0.4^2 Var[B] + 0.2^2 Var[C]$ ← independent A&B&C.

$= 0.4^2 * 0.00736 + 0.4^2 * 0.00004 + 0.2^2 * 0.036$

$= 0.002624$

⇒ $SD[\text{returns}] = \sqrt{Var(\text{returns})} = \sqrt{0.002624} \approx 0.0512$

⇒ thus, $Var[\text{returns}] = 0.002624$ & $SD[\text{returns}] = \sqrt{0.002624}$

2. The file **gold_sp500.xls** contains a table with weekly data for the prices of gold and the S&P 500 index for the period 1985 to 2020.

(a) What is the type of this data set?

panel data.

(b) Make a time series plot of the data. Make sure it looks professional. Include labels, title, units of measurement. Interpret the plot. Attach the printout.

(c) Pick a stock of a company that has a ticker that starts with the same letter as your last name (e.g. Duzhak = DIS for Walt Disney company). Find and download weekly data for your stock over the same period as the rest of the observations in your data set. Add it to the existing data file.

⇒ Home Depot Inc.

(d) Using this data, estimate the mean, standard deviation, skewness, and kurtosis of gold, S&P 500 and your stock prices for ***each*** decade. Put your answers into the "Answer_Table" table below. **Interpret the results**.

Create a new data series, *ln*(*Gold), ln(S&P500) and ln(Your stock*), by taking a natural log of every observation. Then make the new series, measuring a percentage change in the price of a commodity/security. For example, the percentage change in the price of gold for a given month can be found according to the formula $\Delta ln(Gold) = ln(Gold_t) - ln(Gold_{t-1})$. Call these variables *rGold, rSP500, r"Your stock name"*.

(e) Find the mean, standard deviation, skewness, and kurtosis for the created returns for each decade. Put the answers into the table. **Comment on your results.**

(f) How do these results compare to the ones obtained in part (d)? Are they very different? Why?

(g) Make a histogram of *S&P500* returns with 51 bins. Does it look normally distributed? Use the first 4 moments of the distribution to argue your case.

(h) Find the *correlation* between *Gold, S&P500*, and *Your stock* for each decade. Repeat the same calculations for the returns. Put your answers into the table. Interpret your findings.

(i) Is Gold a good hedge against the market portfolio? Why? Would you rather use the results you obtained for prices or for the returns to back up your answer? Explain. Hint: A hedge against the S&P 500 is an asset that appreciates in value at times when the S&P 500 performs poorly.

(j) Find the correlations between the returns on *S&P500* and *Gold* in 2020. Interpret the coefficient and provide plausible explanation.

correlation = 0.63694416

(k) How does *your stock* compare to the other assets? Use the results you obtained in (c) – (g) to answer this question.

ANSWER_TABLE

	80's(1985-1990)		90's (1990-2000)		00's (2000-2010)		10's (2010-2019)	
	Mean	St.dev	Mean	St.dev	Mean	St.dev	Mean	St.dev
Gold	391.3	51.98313	359.4	39.3819	519.5	240.8231	1345	179.7541
SP500	262.1	49.23802	655.6	322.828	1187.1	197.612	1953	584.5916
DIS	0.7435	0.4404583	14.316	12.34572	36.78	9.405198	107.03	59.45448
	Mean	Sd	Mean	sd	Mean	sd	Mean	sd
rGold	0.0010833	0.0171882	-0.000692	0.0140548	0.002630	0.0115715	0.000511 2	0.0168379
rSP500	0.0024938	0.0229168	0.0028479	0.0185831 4	-0.000166	0.0279678	0.0019788 2	0.0194328
rDIS	0.006129	0.05528	0.00692	0.040388	-0.001533 5	0.053834	0.0038422 8	0.0282178
Corr(G,SP)	level	Return	level	return	level	return	level	return
	0.5014399	0.2726279	-0.804862	-0.8022968	0.0972805 5	-0.638183	-0.343439	-0.3190966.
Corr(G,DIS)	-0.089412	-0.13031	0.03764	0.05404	-0.018167	-0.037842	-0.048974	0.0164664.

80's

skewness = -0.03

kurtosis = -1.163107

S = 0.0167892

K = -0.846363

S = 0.9583374

K = 0.00443437

90's

S = -0.715120

K = -0.552019

skewness = 0.9837261

kurtosis = -0.3579949

S = 1.817662

K = 2.858235

00's

S = 0.764602

K = -0.6900582

S = -0.1490936

K = -0.8512976

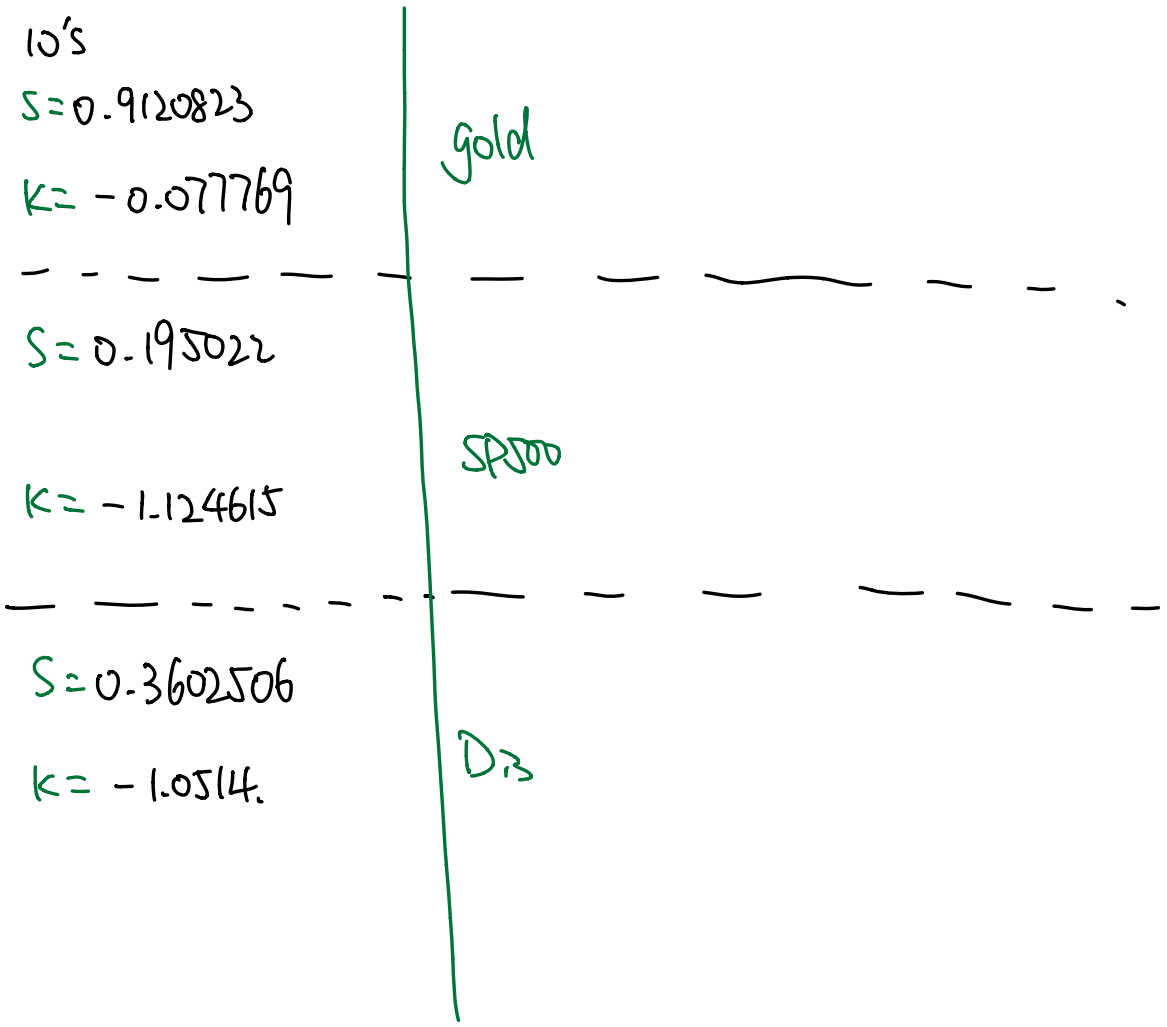
S = 0.37561

K = -0.1680135.

gold

SP500

DIS


10's
S=0.9120823
K=-0.077769
gold
S=0.195022
SP500
K=-1.124615
S=0.3602506
K=-1.0514.
D3

80	90	00	10
S = 0.0282178	S = 1.63302	S = -0.458193	S = -0.5234106.
K = 3.182651	K = 18.78426	K = 2.76225	K = 2.267616
S = -1.47545	S = -0.04499939	S = -0.8605483	S = -0.5851427
K = 5.635581	K = 0.7649965	K = 7.017354	K = 1.926442
S = -0.470913	S = -0.123665	S = -0.5077196	S = -0.2704096
K = 2.812268	K = 1.020691	K = 5.738275	K = 1.273214

⇒ skewness & kurtosis for the rGold, rSP500 and rDis.

g).

⇒ No, the histogram does not look normally distribute.
This is because normal distribution needs to be symmetric and the skewness of the distribution needs to be 0. Also, the kurtosis of the distribution needs to have equal left tail and right tail.

⇒ from the histogram, we can clearly tell that the distribution is not symmetric and the skewness is not 0.

⇒ since the histogram distribution is not symmetric, we can conclude that the S&P500 is not normally distributed.

ECON140

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com (http://rmarkdown.rstudio.com).

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
data1 <- read.csv('gold_sp500.csv')
mystock <- read.csv('HD_hd.csv')
modify <- mystock[c("Date","Close")]
data1$HD_price <- modify$Close
head(data1,10)
```

```
##    observation_date year   gold  sp500 HD_price
## 1         1/28/1985 1985 301.75 178.63 0.359396
## 2          2/4/1985 1985 303.49 182.19 0.351166
## 3         2/11/1985 1985 301.66 181.60 0.403292
## 4         2/18/1985 1985 305.90 179.36 0.411523
## 5         2/25/1985 1985 299.31 183.23 0.433471
## 6          3/4/1985 1985 288.53 179.10 0.414266
## 7         3/11/1985 1985 288.82 176.53 0.414266
## 8         3/18/1985 1985 290.54 179.04 0.378601
## 9         3/25/1985 1985 317.93 180.66 0.375857
## 10         4/1/1985 1985 325.82 179.03 0.356653
```

```
decade1_80 <- subset(data1, year<1990)
decade1_90 <- subset(data1, year>1989 & year<2000)
decade1_00 <- subset(data1, year>1999 & year<2010)
decade1_10 <- subset(data1, year>2009 & year<2020)
summary(decade1_80$gold)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   288.5   343.4   392.9   391.3   436.2   490.1
```

```
sd(decade1_80$gold)
```

```
## [1] 51.98313
```

```
#install.packages("moments")
library(moments)
#install.packages("e1071")
library(e1071)
```

```
##
## Attaching package: 'e1071'
```

```
## The following objects are masked from 'package:moments':
##
##     kurtosis, moment, skewness
```

```
log80 <- log(decade1_80[,c("gold","sp500","HD_price")])
log90 <- log(decade1_90[,c("gold","sp500","HD_price")])
log00 <- log(decade1_00[,c("gold","sp500","HD_price")])
log10 <- log(decade1_10[,c("gold","sp500","HD_price")])
rGold80 <- diff(log80$gold)
rGold90 <- diff(log90$gold)
rGold00 <- diff(log00$gold)
rGold10 <- diff(log10$gold)
rSP500_80 <- diff(log80$sp500)
rSP500_90 <- diff(log90$sp500)
rSP500_00 <- diff(log00$sp500)
rSP500_10 <- diff(log10$sp500)
rHD80 <- diff(log80$HD_price)
rHD90 <- diff(log90$HD_price)
rHD00 <- diff(log00$HD_price)
rHD10 <- diff(log10$HD_price)
data2020 <- subset(data1, year=2020)
log2020 <- log(data2020[,c("gold","sp500","HD_price")])
sp500_2020 <- log2020$sp500
gold_2020 <- log2020$gold
summary(rGold90)
```

```
##       Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
## -0.0569606 -0.0073377 -0.0004415 -0.0006921  0.0052941  0.1372879
```

```
summary(rGold10)
```

```
##       Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
## -0.0928958 -0.0099553  0.0012019  0.0005112  0.0114232  0.0551407
```

```
sd(rGold80)
```

```
## [1] 0.0171882
```

```
sd(rGold90)
```

```
## [1] 0.01405479
```

```
sd(rGold00)
```

```
## [1] 0.02115715
```

```
sd(rGold10)
```

```
## [1] 0.01683979
```

```
sd(rSP500_80)
```

```
## [1] 0.0229186
```

```
sd(rSP500_90)
```

```
## [1] 0.01858314
```

```
sd(rSP500_00)
```

```
## [1] 0.02789678
```

```
sd(rSP500_10)
```

```
## [1] 0.01943208
```

```
mean(rSP500_90)
```

```
## [1] 0.002824788
```

```
mean(rSP500_00)
```

```
## [1] -0.0005165902
```

```
mean(rSP500_10)
```

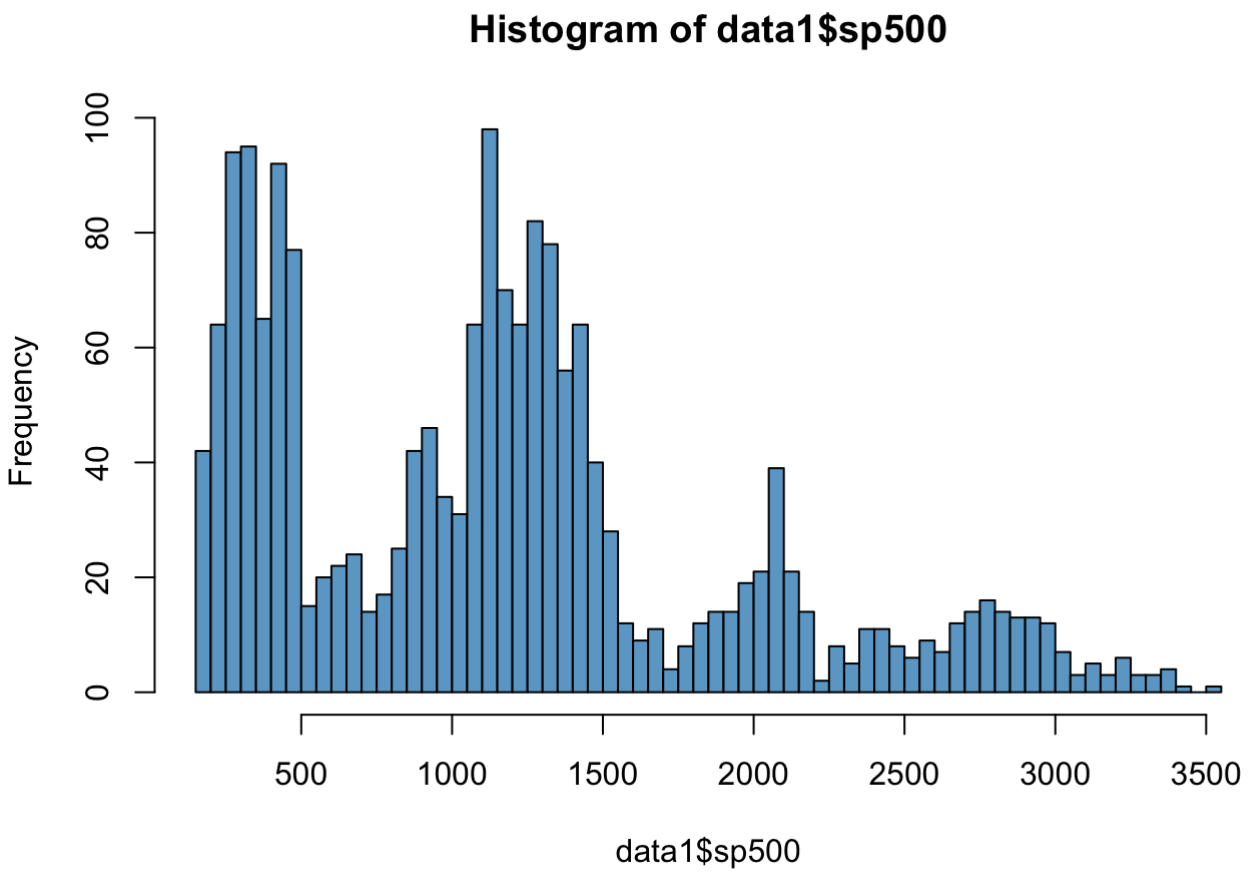
```
## [1] 0.001978816
```

```
cor(sp500_2020,gold_2020)
```

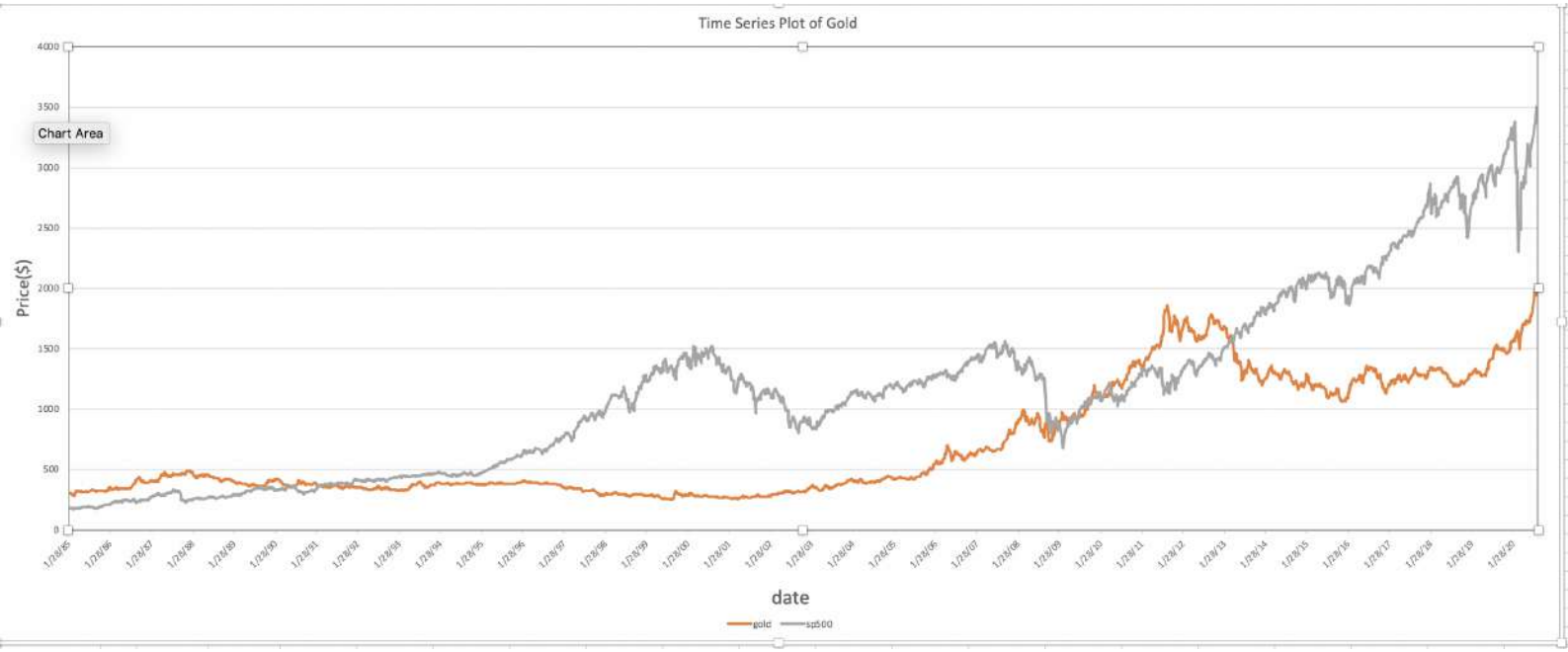
```
## [1] 0.6369416
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.


Time Series Plot of Gold
Chart Area
Price($)
4000
3500
3000
2500
2000
1500
1000
500
0
1/28/85
1/28/86
1/28/87
1/28/88
1/28/89
1/28/90
1/28/91
1/28/92
1/28/93
1/28/94
1/28/95
1/28/96
1/28/97
1/28/98
1/28/99
1/28/00
1/28/01
1/28/02
1/28/03
1/28/04
1/28/05
1/28/06
1/28/07
1/28/08
1/28/09
1/28/10
1/28/11
1/28/12
1/28/13
1/28/14
1/28/15
1/28/16
1/28/17
1/28/18
1/28/19
1/28/20
date
gold
sp500

2.

d).

The mean of gold, SP500, and Home Depot stock price indicates the average of the stock price in this decade. The standard deviation of those data means how much each weekly daily is different from the average decade price. Also, the skewness of the variable indicates the symmetric of the variable and the kurtosis indicates how heavy the tail is for this variable. We can see that the gold price in the ’80s has a negative skewness and a negative kurtosis. This indicates that the gold price in the ’80s is more likely to be lower than the mean 80’s gold price (fatter left tail). According to this characteristic, we can then conclude from the data that gold prices in the ’90s are more like to be lower than the average price and have a fatter left tail. On the contrary, although the gold price in the 00’s and 10’s still have fatter left tails, they are more likely to be higher than the decade average gold price. For SP500, 80’s, 90’s, and 10’s SP500 distribution have fatter left tails and are more likely to be higher than the decade average. Meanwhile, the 00’s SP500 has a fatter left tail and is more likely to be lower than the decade average SP500. Last but not the least, for our Home Depot stock price, in the 80s and 90s, the stocks are more likely to be higher than the decade average and are distributed more toward the right tail. In terms of the 00s and 10s, Home Depot Inc’s stock is still more likely to be higher than the decade average but has fatter left tails, (more distributed toward the lower price range).

e).

The mean of gold, rSP500, and rDIS indicate on average how much the stock price is increase/decrease weekly. (positive value means increase while a negative value means decrease). The standard deviation of those variable means how much each weekly increase/decrease are likely to be different from the decade mean weekly price change. The skewness means how the variable distribution is skewed, a negative skewness means the data are skewed to the left (more likely to be lower than average) while a positive skewness means the variable is skewed to the right (more likely to be higher than average). The kurtosis of the data means how the tail is distributed for this variable, a negative kurtosis means a fatter left tail while a positive kurtosis means a fatter right tail.

f).

Compare to the data we obtain in part d, those new variables rGold, rSP500, rDis are very different. To begin with, the mean of those data means very different things. While the mean only reflects the actual average price in the decade, the new variable mean indicates the average weekly change of the price in the decade. So different meaning in mean also indicates the different meaning of the standard deviation, the standard deviation for those new variables means how different are those weekly price change in a decade. That why the variable in part d and those new variables have very different skewness and kurtosis value. (Not only the difference in amount but even in sign). Those differences in the data are mainly caused by the difference in what those variables are measuring in the stock market.

g).

No, the histogram does not look normally distributed.

This is because normal distribution needs to be symmetric and the skewness of the distribution needs to be 0. Also, the kurtosis of the distribution needs to have equal left tail and right tail. From the histogram, we can clearly tell that the distribution is not symmetric and the skewness is not 0. Since the histogram distribution is not symmetric, we can conclude that the S&P500 is not normally distributed.

h).

The correlation coefficient reflects how one variable will change when the other variable change. A positive correlation means two variables change in the same direction while a negative correlation means two variables change in the opposite direction. According to those characteristics, from the correlation coefficient, we observed, in the 80s: 1.If gold price increase, the SP500 increases. 2.If the gold price increase, the Home Depot stock price decrease. 3. If the difference in weekly gold price increase, the difference in weekly SP500 increase. 4. If the difference in weekly gold price increase, the difference in weekly Home Depot stock price decrease. Then in the 90s: 1.If gold price increase, the SP500 decreases. 2.If the gold price increase, the Home Depot stock price decrease. 3. If the difference in weekly gold price increase, the

difference in weekly SP500 increase. 4. If the difference in weekly gold price increase, the difference in weekly Home Depot stock price decrease. In the 00s: 1.If gold price increase, the SP500 increases. 2.If the gold price increase, the Home Depot stock price decrease. 3. If the difference in weekly gold price increase, the difference in weekly SP500 increase. 4. If the difference in weekly gold price increase, the difference in weekly Home Depot stock price decrease. In the 10s: 1.If gold price increase, the SP500 decreases. 2.If the gold price increase, the Home Depot stock price decrease. 3. If the difference in weekly gold price increase, the difference in weekly SP500 decrease. 4. If the difference in weekly gold price increase, the difference in weekly Home Depot stock price decrease.

i).

As far as I concern, I think gold is a good hedge against the market portfolio. From the chart, we can see that the S&P500 decrease from the 80s to the 90s decade. However, the price in gold doesn't decrease much despite the huge drop in SP500. Despite the poor performance in SP500, the gold price still hold relatively high, this is why I think gold is a good hedge against the market portfolio.

However, I do rather use the results for the returns to back up my answers. This is because the returns directly show the average weekly change in SP500 performance as well as the average change in gold. Since returns are all about "change, it is a more upfront reflection of how the market is doing as well as how the gold price is fluctuating. A negative returns mean the marketing is performing poorly or the price is constantly

decreasing while a positive returns means the market is performing well or the price is constantly increasing. From the chart, we can easily observe that in both times where the returns for SP500 is negative (poor market performance), the gold price is still increasing. Thus, we can conclude that gold still appreciates in values at time when the S&P500 performs poorly. That's why using returns, it's easier for us to show that Gold is a good hedge against the market portfolio.

j.)

The correlations between the returns of S&P500 and Gold in 2020 is 0.6369416. This is a positive correlation, it means in 2020, that while S&P500 increases, the Gold price will increase as well. This two variables change in both direction probably because, in 2020, which both the value of every currency vary a lot due to the sudden attack of COVID19, gold became a relatively stable asset with stable value. So what some of the large companies will do is that they will put their profit into the gold market in order to avoid currency devaluation. As a result, when all the companies are doing well (which rSP500) increase, the demand for gold increase and thus the gold price increase (rGold) as well. That's why in 2020, the return of S&P500 and Gold have a positive correlation coefficient.

k).

Compare to the other asset, I think Home Depot Inc, the stock I picked, is a pretty good asset to invest in. Since we already clarify that why the return variables are

more efficient when we are evaluating the asset, we can see that Home Depot Inc has more advantages compared to other assets from those returns variable. We can see that not only the mean weekly price is increasing each decade the weekly price change of the stock is increasing pretty constantly throughout the years. (except the 00s decade). When we look at the most current 10s decade, we can see that the Home Depot have the highest weekly average price increase out of all the demonstrated asset. This means that if we invest in Home Depot, we will receive the highest average return from our investment. As a result, from all those aspects, I am convinced that Home Depot Inc is a pretty good asset to invest in compare to the other asset.