

# HW1

Noah Love

1/18/2021

## HW1

This homework is to meant for you to brush up some prerequisites. If some of these topics are new to you, feel free to ask on Piazza how to approach these. Context - Is gold price inversely related to the market?

There's a belief that the money from the stock market will escape to gold when the stock market is not doing well. The demand for gold and the expectations for the market are often reflected in the pricing of the assets, i.e. high demand yields high gold prices and upward expectations also lead to higher stock prices.

## Q1

Please use the 'TIME\_SERIES\_WEEKLY' API listed on Alpha Vantage to get the weekly time series data for

'V00': an arbitrarily chosen ETF that tracks the market

'GDXJ': an arbitrarily chosen ETF for gold

For this problem, simply show the code for your query and print out the number of weeks of data for each time series. Your API key should NOT appear in your solutions but the URL you're using and the query should be shown.

Hint:

You will need to claim a free API key before you can query data

The functions in httr should be helpful, here is some sample code if you have not done so before.

```
#library(httr)
# url_V00 <- "https://www.alphavantage.co/query?function=TIME_SERIES_WEEKLY&symbol=V00&apikey=SEDREZCZE"
# url_GDXJ <- "https://www.alphavantage.co/query?function=TIME_SERIES_WEEKLY&symbol=GDXJ&apikey=SEDREZCZE"
#
# responce <- GET(url = url_V00, query = params)

function_name <- "TIME_SERIES_WEEKLY"
stock_ticker <- "V00"
my_data_type <- "csv"
output_size <- "full"
api_call <- paste0("https://www.alphavantage.co/query?function=",
                   function_name,
                   "&symbol=",
                   stock_ticker,
                   "&outputsized=",
                   output_size,
                   "&apikey=",
                   api_key,
                   "&datatype=",
                   my_data_type)
```

```
VOO <- read.csv(url(api_call))
```

```
head(VOO)
```

```
##      timestamp  open    high    low  close  volume
## 1 2021-01-20 347.80 353.7100 346.36 352.83 6354363
## 2 2021-01-15 347.36 350.3183 343.53 345.37 15438912
## 3 2021-01-08 345.02 350.6190 335.37 350.42 20904599
## 4 2020-12-31 341.74 344.3700 340.87 343.69 10833270
## 5 2020-12-24 336.83 340.3800 334.11 339.16 11429244
## 6 2020-12-18 338.76 342.5000 334.89 340.64 17013034
```

*#For VOO*

```
function_name <- "TIME_SERIES_WEEKLY"
stock_ticker <- "GDXJ"
my_data_type <- "csv"
output_size <- "full"
api_call <- paste0("https://www.alphavantage.co/query?function=",
                  function_name,
                  "&symbol=",
                  stock_ticker,
                  "&outputsize=",
                  output_size,
                  "&apikey=",
                  api_key,
                  "&datatype=",
                  my_data_type)
```

```
GDXJ <- read.csv(url(api_call))
```

```
head(GDXJ)
```

```
##      timestamp  open  high    low  close  volume
## 1 2021-01-20 50.44 52.34 49.55 52.06 15001327
## 2 2021-01-15 52.60 53.41 49.33 49.35 38498209
## 3 2021-01-08 57.09 58.55 52.53 53.66 53354228
## 4 2020-12-31 54.45 55.12 52.74 54.24 25206958
## 5 2020-12-24 53.59 54.75 51.63 53.35 25876383
## 6 2020-12-18 50.80 55.96 49.77 54.42 40719212
```

## Q2

Please plot the close price for VOO against the different weeks and overlay the regression line for this scatter plot.

You do not need to label your week index but the prices should be labeled.

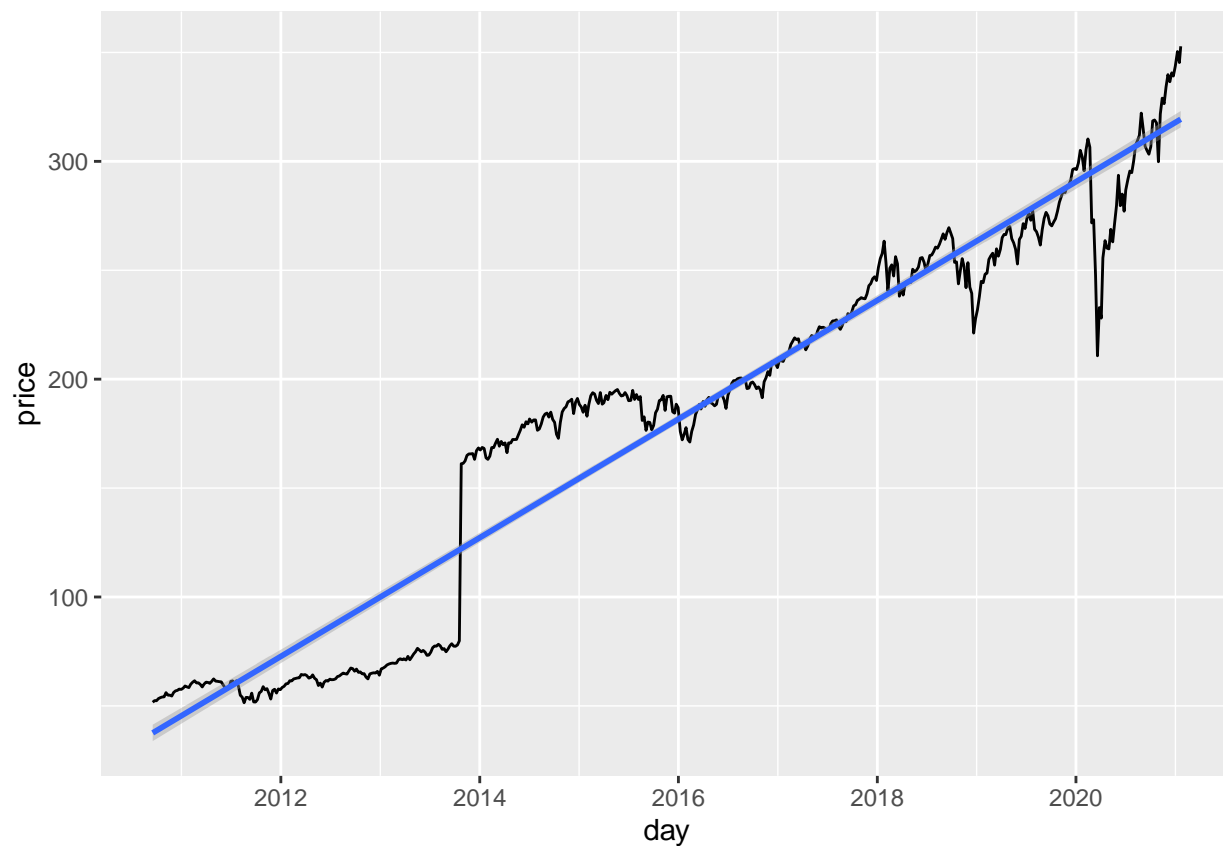
```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
```

```
##      filter, lag
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
V00$timestamp <- as.Date(V00$timestamp)

ggplot(V00, aes(x = timestamp, y = close)) +
  geom_line() +
  xlab("day") +
  ylab("price") +
  geom_smooth(method = "lm", formula = y~x)
```



### Q3

Please plot the residuals from the regression in Q2 against the close price of GDXJ.

label your axes with units. Your title should include the correlation value, rounded to the nearest hundredth. Please show the code that demonstrates your decision on merging the 2 time series.

```
linear <- lm(V00$close ~ V00$timestamp)

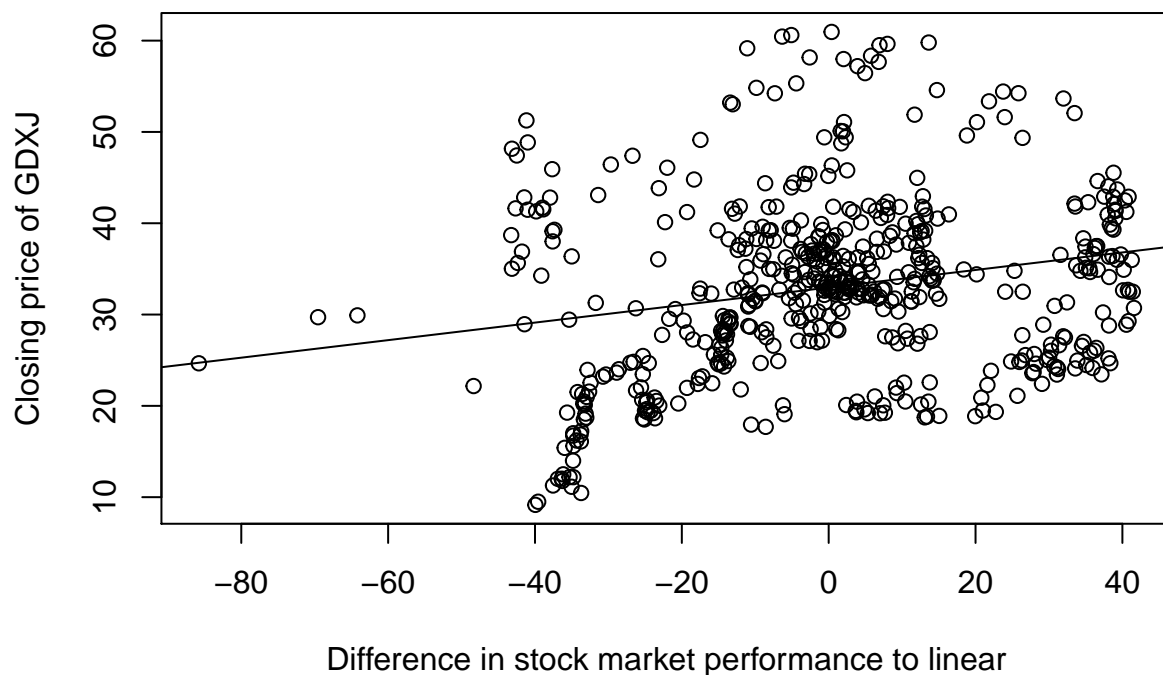
subset_GDXJ <- GDXJ[1:length(linear$residuals),]
#subset_GDXJ

gold_vs_stock <- lm(subset_GDXJ$close ~ linear$residuals)
```

```
summary(gold_vs_stock)
```

```
##
## Call:
## lm(formula = subset_GDXJ$close ~ linear$residuals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -19.9637  -6.6311  -0.2735   5.2986  28.1383
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    32.96080    0.40657   81.071 < 2e-16 ***
## linear$residuals  0.09605    0.01822    5.272 1.96e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.457 on 539 degrees of freedom
## Multiple R-squared:  0.04903,    Adjusted R-squared:  0.04727
## F-statistic: 27.79 on 1 and 539 DF,  p-value: 1.959e-07
plot(x = linear$residuals, y = subset_GDXJ$close,
     xlab="Difference in stock market performance to linear",
     ylab = "Closing price of GDXJ",
     main = "Correlation coeff: 0.049",
     abline(lm(subset_GDXJ$close ~ linear$residuals)))
```

**Correlation coeff: 0.049**



#### Q4

Relying only on the scatter plot, would you say the belief between gold and the market is supported or rejected? Please explain.

Relying on the scatter plot and the line I plotted, gold is not inversely related. In fact, when the market is beating its performance, on average gold is closing higher as well. Also, when the stock market is doing really poorly (compared to the linear regression) gold also closes really low as seen by the strong tail around -40.