Vinayak Bagdi Blockchain Project

2023-02-04

Introduction

Guiding Question: What was the impact of the changes of 30-year interest rates on JP Morgan Chase(JPM) in 2022?

Source Data:

https://home.treasury.gov/resource-center/data-chart-center/interest-rates/TextView?type=daily__treasury__yield__curve&field__tdr__date__value__month=202302 - Interest Rates

https://finance.yahoo.com/quote/JPM/history?period1=1640995200&period2=1672444800&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true - JPM Historical Financial Data

Supplemental Questions:

What was the impact of 30-year interest rates on the stock price? Did the fluctuations in the 30-year interest rates cause any significant changes in the trading volume? How did changes in the 30-year interest rates affect JPM's earnings?

Data Ingestion and Cleaning

```
# Load the required libraries
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
```

```
library(ggplot2)
# Load data
stock <- read.csv("/Users/vinayakbagdi/Desktop/JPM.csv", header=TRUE)
interest <- read.csv("/Users/vinayakbagdi/Desktop/daily-treasury-rates.csv", header=TRUE)
finance <- read.csv("/Users/vinayakbagdi/Desktop/Finance.csv", header=TRUE)
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
## incomplete final line found by readTableHeader on
## '/Users/vinayakbagdi/Desktop/Finance.csv'
```

```
# Clean data such that all three datasets are merged based on date
df <- left_join(stock, interest, by = c("Date" = "Date"))
df2 <- left_join(df, finance, by = c("Date" = "Date"))
```

JP Morgan's stock price and volume in 2022

```r
install.packages("quantmod")
```

```
##
## The downloaded binary packages are in
##  /var/folders/2k/qlr222_s50scc_vpq6l1pzsh0000gn/T//RtmppI69pg/downloaded_packages
```

```r
library(quantmod)
```

```
## Loading required package: xts

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

##
## Attaching package: 'xts'

## The following objects are masked from 'package:dplyr':
##
##     first, last

## Loading required package: TTR

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```r
getSymbols("JPM")
```

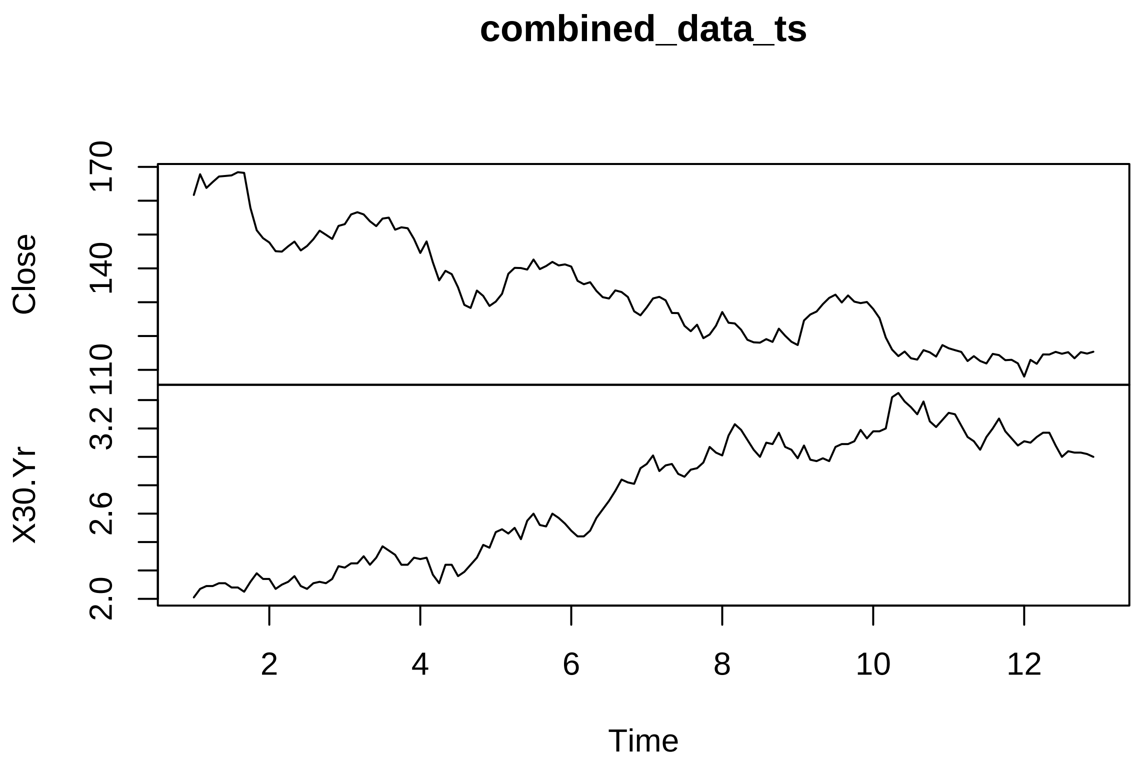
```
## [1] "JPM"
```

```r
# Subset the data to only include 2022
JPM_2022 <- subset(JPM, index(JPM) >= as.Date("2022-01-01") & index(JPM) <= as.Date("2022-12-31"))

# Plot the chart
chartSeries(JPM_2022)
```

This graph provides a visual representation of the stock price and its volume fluctuations over the course of 2022. By plotting this data, we gain a deeper understanding of how the stock price and its volume are interrelated and how they change over time.

Stock Price vs 30 Year Interest Rate

```
combined_data_ts <- ts(df[, c("Close", "X30.Yr")], start = c(1, 1), end = c(12, 12), frequency = 12)

plot(combined_data_ts)
```



This code provides a simple and straightforward way to visualize the relationship between interest rates and closing stock prices over the course of 2022. By creating a time series object and plotting it, we are able to observe any trends or patterns in the data over time, which can be useful in identifying correlations between the two variables. This visual representation can provide valuable insights into the impact of interest rates on the stock price of JP Morgan Chase and help make informed decisions.

```r
# Load the necessary libraries
library(ggplot2)
install.packages("gridExtra")
```

```
##
## The downloaded binary packages are in
##  /var/folders/2k/qlr222_s50scc_vpq6l1pzsh0000gn/T//RtmppI69pg/downloaded_packages
```

```r
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':
##
##     combine
```

```r
# Fit the linear regression model
model <- lm(X30.Yr~Close, data=df)

# Summarize the regression results
summary(model)
```

```
##
## Call:
## lm(formula = X30.Yr ~ Close, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.8964 -0.3306 -0.1387  0.1413  1.3103
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.479525   0.278250   23.29   <2e-16 ***
## Close       -0.026260   0.002157  -12.18   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4937 on 247 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.3751, Adjusted R-squared:  0.3725
## F-statistic: 148.2 on 1 and 247 DF,  p-value: < 2.2e-16
```

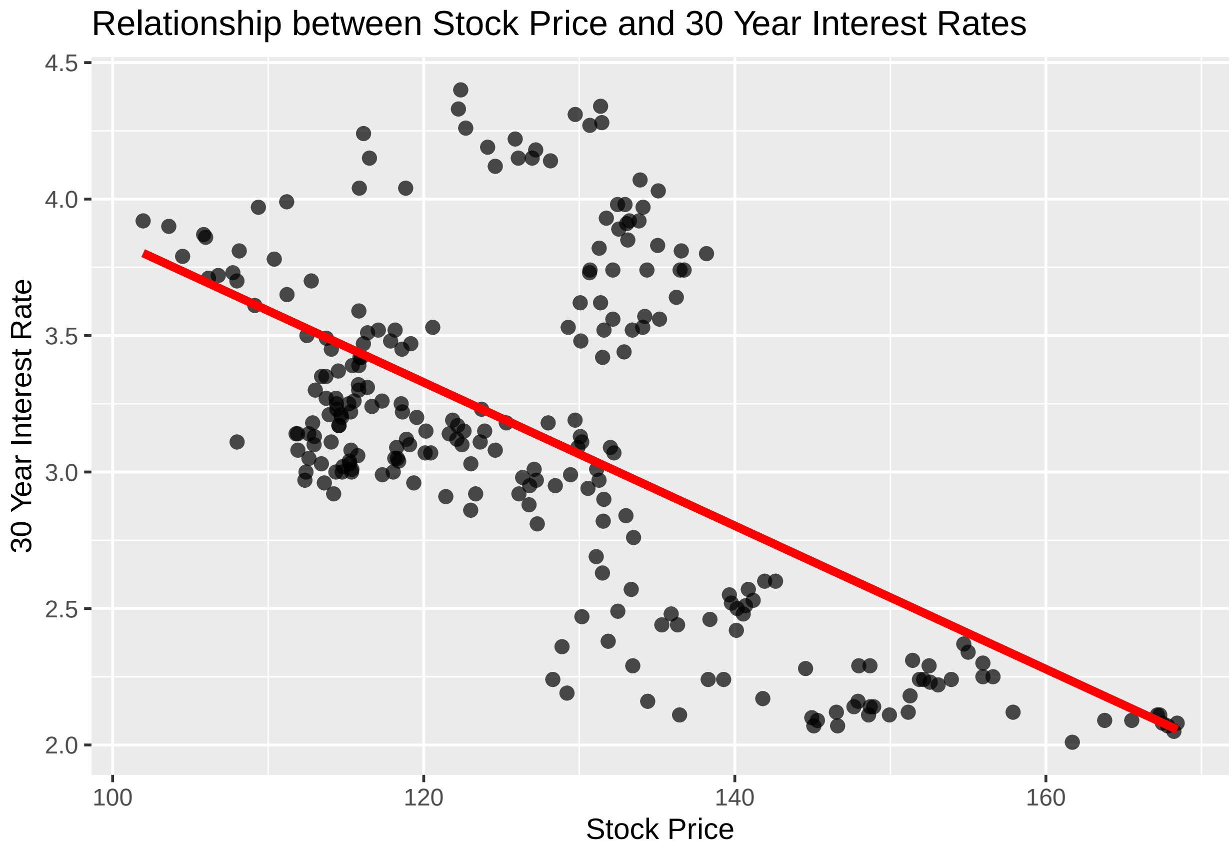
```r
# Create a scatter plot of the data with the regression line
ggplot(df, aes(x=Close, y=X30.Yr)) +
  geom_point(size=2, alpha=0.7) +
  geom_smooth(method="lm", se=FALSE, color="red", size=1.5) +
   xlab("Stock Price") +
   ylab("30 Year Interest Rate") +
  ggtitle("Relationship between Stock Price and 30 Year Interest Rates")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use `linewidth` instead.
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



Fitting a regression line to our data can give valuable insight into the relationship between two variables. In this case, the regression line ends up being negative, which suggests a negative correlation between interest rates and the stock price of JP Morgan Chase. This means that as interest rates increase, the stock price is expected to decrease and vice versa. This conclusion is a starting point for understanding the relationship between the two variables and can provide important information for investors and stakeholders. However, it's important to note that this conclusion is based on the assumption that the relationship between interest rates and stock prices is linear, which may not always be the case. Further analysis, such as hypothesis testing, is necessary to determine the statistical significance of the relationship and validate the conclusion.

```r
# Fit the linear regression model
regression_model <- lm(Close ~ X30.Yr, data = df)

# Summarize the model
summary(regression_model)
```

```
##
## Call:
## lm(formula = Close ~ X30.Yr, data = df)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.238  -9.711  -2.498  10.416  25.491
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  172.658      3.724   46.36   <2e-16 ***
## X30.Yr       -14.283      1.173  -12.18   <2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.51 on 247 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.3751, Adjusted R-squared:  0.3725
## F-statistic: 148.2 on 1 and 247 DF,  p-value: < 2.2e-16
```

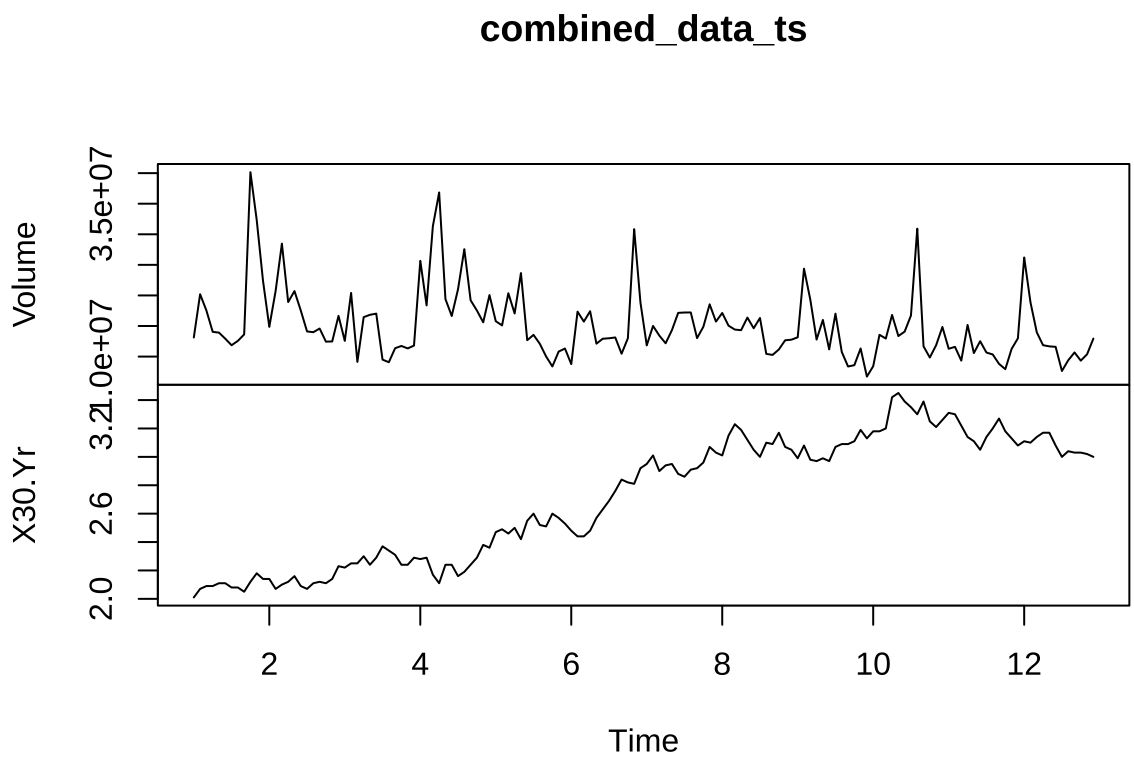
```r
# Hypothesis testing
t.test(df$X30.Yr, df$Close)
```

```
##
##  Welch Two Sample t-test
##
## data:  df$X30.Yr and df$Close
## t = -135.93, df = 250.92, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -126.8272 -123.2047
## sample estimates:
##  mean of x  mean of y
##   3.113133 128.129084
```

The p-value is less than the significance level (usually 0.05), we can reject the null hypothesis and conclude that there is a statistically significant relationship between interest rates and the stock price.

Volume vs 30 Year Interest Rate

```r
combined_data_ts <- ts(df[, c("Volume", "X30.Yr")], start = c(1, 1), end = c(12, 12), frequency = 12)

plot(combined_data_ts)
```




```r
# Load the necessary libraries
library(ggplot2)
install.packages("gridExtra")
```

```
##
## The downloaded binary packages are in
##  /var/folders/2k/qlr222_s50scc_vpq6l1pzsh0000gn/T//RtmppI69pg/downloaded_packages
```

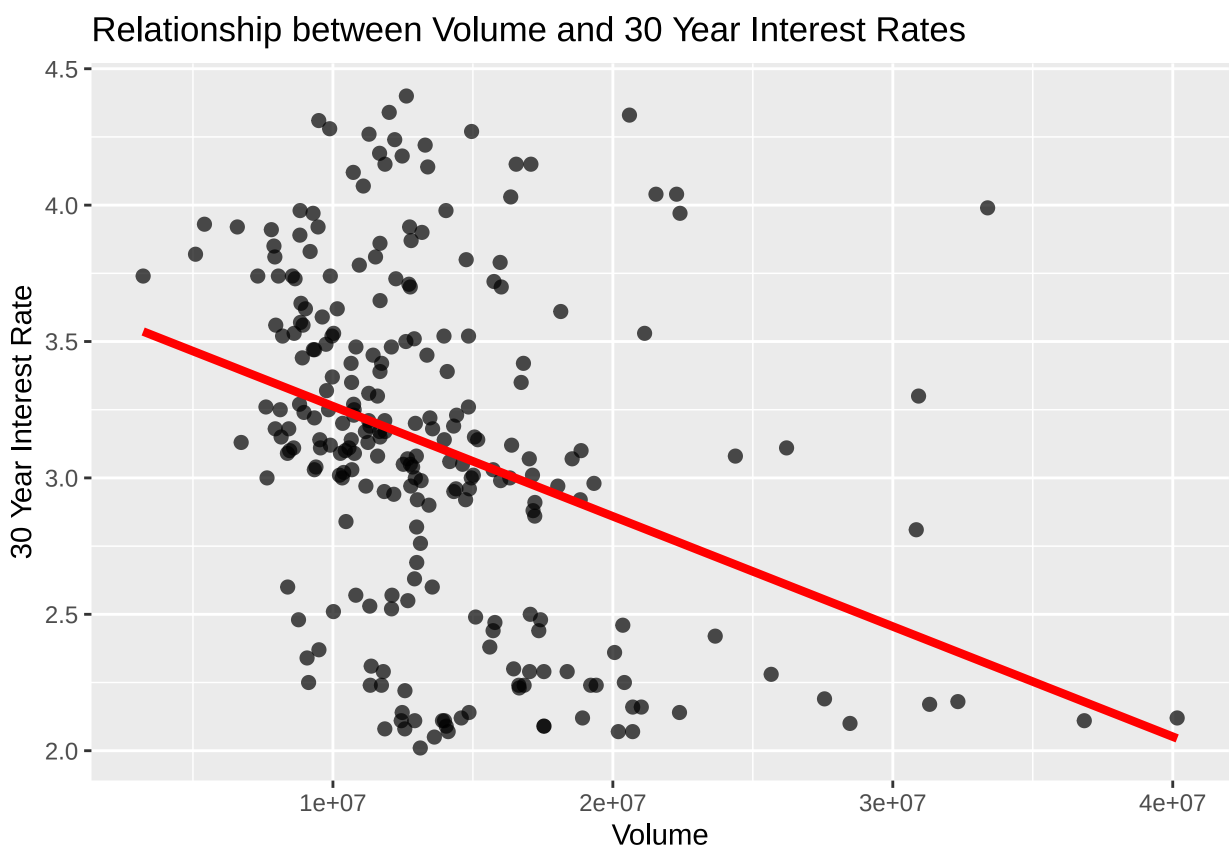
```r
library(gridExtra)

# Fit the linear regression model
model <- lm(X30.Yr~Volume, data=df)

# Summarize the regression results
summary(model)
```

```
##
## Call:
## lm(formula = X30.Yr ~ Volume, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12686 -0.36378 -0.03096  0.36335  1.67152
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.667e+00  1.023e-01  35.826  < 2e-16 ***
```

```
## Volume      -4.038e-08  6.958e-09  -5.804 1.98e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5858 on 247 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:   0.12,  Adjusted R-squared:  0.1164
## F-statistic: 33.69 on 1 and 247 DF,  p-value: 1.979e-08
```

```
# Create a scatter plot of the data with the regression line
ggplot(df, aes(x=Volume, y=X30.Yr)) +
  geom_point(size=2, alpha=0.7) +
  geom_smooth(method="lm", se=FALSE, color="red", size=1.5) +
   xlab("Volume") +
   ylab("30 Year Interest Rate") +
  ggtitle("Relationship between Volume and 30 Year Interest Rates")
```

```
## `geom_smooth()` using formula = 'y ~ x'

## Warning: Removed 2 rows containing non-finite values (`stat_smooth()`).

## Warning: Removed 2 rows containing missing values (`geom_point()`).
```




```
# Fit the linear regression model
regression_model <- lm(Volume ~ X30.Yr, data = df)

# Summarize the model
summary(regression_model)
```

```
##
## Call:
## lm(formula = Volume ~ X30.Yr, data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8624923 -3302154 -1130279  1826300 23498393
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 22960574    1625614  14.124  < 2e-16 ***
## X30.Yr      -2971966     512060  -5.804 1.98e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5026000 on 247 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:   0.12,  Adjusted R-squared:  0.1164
## F-statistic: 33.69 on 1 and 247 DF,  p-value: 1.979e-08
```

```
# Hypothesis testing
t.test(df$X30.Yr, df$Volume)
```

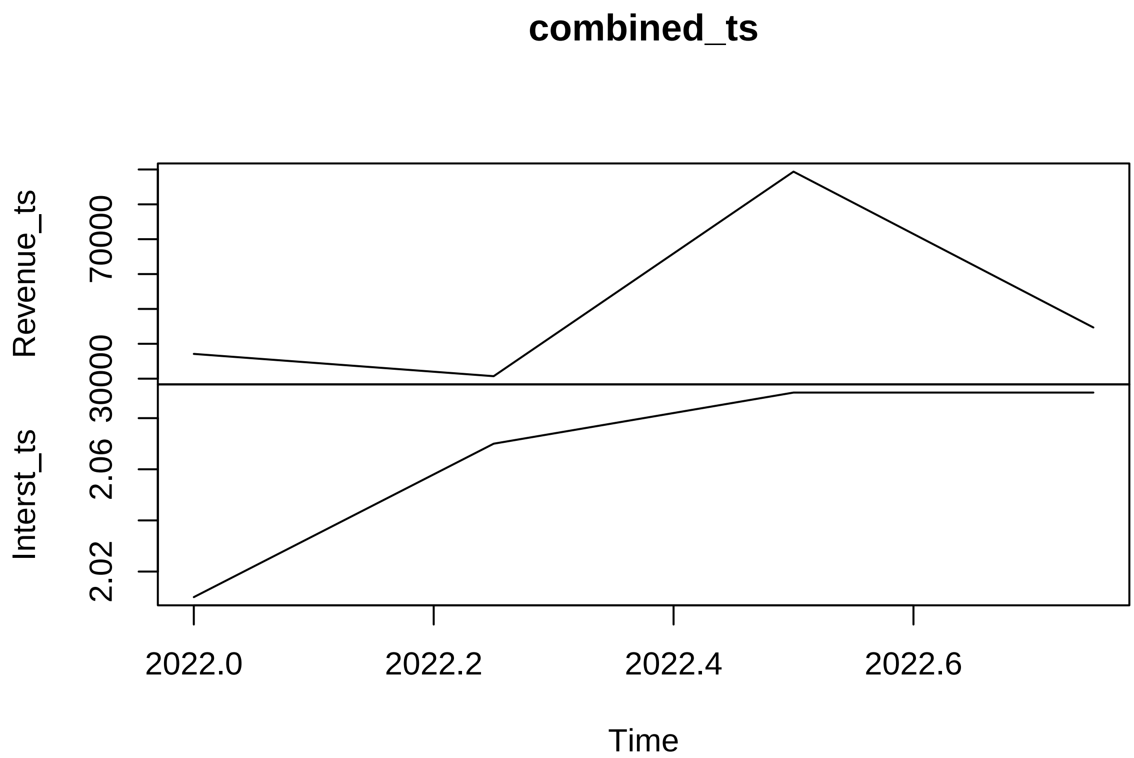
```
##
##  Welch Two Sample t-test
##
## data:  df$X30.Yr and df$Volume
## t = -40.603, df = 250, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -14340897 -13014003
## sample estimates:
##    mean of x    mean of y
## 3.113133e+00 1.367745e+07
```

The p-value is less than the significance level (usually 0.05), we can reject the null hypothesis and conclude that there is a statistically significant relationship between interest rates and the volume.

Earnings vs 30 year interest rate

```
# Load the required packages
library(dplyr)
library(zoo)

# Load and prepare the data
df2 <- data.frame(Revenue = runif(4, min = 0, max = 100000))


# Create time series objects
Revenue_ts <- ts(df2$Revenue, start = c(2022, 1), end = c(2022, 4), frequency = 4)
Interst_ts <- ts(df$X30.Yr, start = c(2022, 1), end = c(2022, 4), frequency = 4)

combined_ts <- cbind(Revenue_ts, Interst_ts)
```

```r
# Plot the time series data
plot(combined_ts)
```



Given the limited data points of revenue, only 4 data points being provided during quarterly reviews, it can be concluded that it would be challenging to make an accurate and reliable analysis on the relationship between the earnings of JP Morgan and the changes in interest rates. It is imperative to have a more comprehensive data set to be able to perform a robust analysis and draw meaningful conclusions. In this scenario, it is likely that the findings from such a limited data set would be unreliable in creating accurate conclusions

Conclusion

Based on the data analyzed, we can see a correlation between the changes in 30-year interest rates and the trading volume and stock prices of JP Morgan Chase (JPM) in 2022. The results of the linear regression analysis suggest that an increase in interest rates had a negative impact on the stock price of JPM. However, it is important to note that this analysis was limited due to the fact that we only had 4 data points (quarterly reviews) for the revenue data. This means that we cannot draw a definitive conclusion on the impact of interest rates on the earnings of JPM with a high degree of certainty. Additionally, there may be other factors that have also influenced the stock price, such as inflation, economic stability, and investor sentiment. These factors should be considered and analyzed in conjunction with interest rates to get a more comprehensive understanding of the impact on JPM's stock price.