Going back in Time to invest in a S&P 500 stock

A Report by Daniel Gaza

Which companies would provide me with the greatest return if I could travel back in time to 2014? That is the one thought that hasn't left my mind since I bought this time machine from a pawn shop on top of the mountains and came across this S&P 500 stock data on Maven Analytic playground. Thank goodness I'm a data analyst whose job it is to find answers.

Introduction

Maven Analytic has issued yet another challenge, this time to S&P 500 companies. The analysis will culminate in a one-page report containing visualizations, insights, and recommendations.

The purpose of this report is to show the entire process of creating the one-page report as well as other important aspects of the analysis that could not be included in the one-page report.

The analysis's goal is to answer questions about volatility and must-traded stocks in the dataset from 1/2/2014 to 29/12/2017. Oh, let me end the suspense: based on the analysis, I was able to determine which stocks I should buy in 2017 in order to profit. Do you have a time machine and want this amazing investment tip? Stick with me until the end.

Everyone would have wanted to invest in companies like Google, Amazon, Apple, and others when they were first listed on the stock market if they knew how much they are worth today; if you couldn't or did not invest then, this is an opportunity for Maven Analytic to invest in the next best thing. Keep reading until the end to find out about that potential stock.

Objective of Analysis

Maven Analytics will want to know the answers to these questions so that they can decide which company stock to invest in. Maven Analytic is establishing retirement funds for her employees, with the overall goal of investing in long-term stocks that will generate profits for her employees. When we first enter the stock market as investors, we want to understand stock behavior so that we can profit from our investments in long-term, low-risk stocks.

Maven Analytics is interested in learning the answers to these various questions.

1. On which date in the sample did the most overall trading volume occur? Which two stocks were most actively traded on that date?

2. Which day of the week has the highest volume? lowest?

3. When did Amazon (AMZN) experience the greatest volatility, as measured by the difference between the high and low prices?

4. Which stock would you invest in if you could go back in time and invest from 1/2/2014 to 12/29/2017? What percentage gain would you expect?

Data

Maven Analytic provided the data for this analysis; this is one of many competitions held by Maven on a monthly basis to help students and analysts develop their skills while also providing a platform to showcase their abilities.

From 2014 to 2017, historical stock market data for current S&P 500 companies. Each record represents a single trading day and contains the ticker symbol, volume, high, low, open, and close prices.

To make the analysis process easier and simpler, I **added** volatility, day of week, and profit columns to the date.

Tools used

Used SQL server is adding new columns and analyzing data in order to answer the necessary questions. PowerBI is being used to create visuals for easy understanding by end users, as well as a single-page report.

Data cleaning and Wrangling

Maven Analysis did an excellent job of cleaning the data before uploading it, so no cleaning or restructuring was required; this realization came after going through the data.

The datatypes were automatically assigned when the data was imported to SQL server; however, I had to re-assign appropriate datatypes to all columns so that we did not encounter errors when performing analysis.

```
--Assigning appropriate DataTypes
ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN date DATE;

ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN open_ FLOAT;

ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN high_ FLOAT;

ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN low FLOAT;

ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN close_ FLOAT;

ALTER TABLE [S&P 500 Stock Prices].[dbo].[Stock_Prices]
ALTER COLUMN volume FLOAT;
```

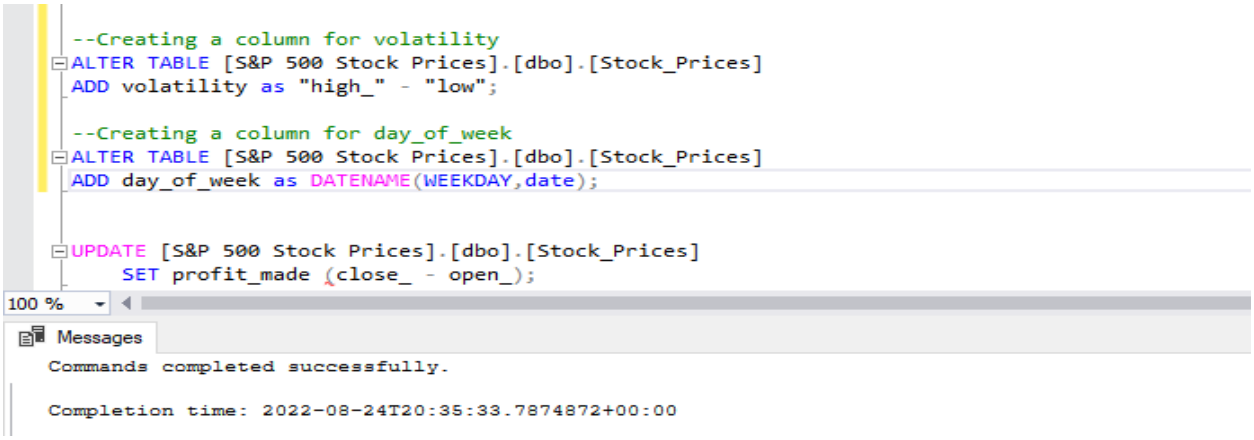
```
100 %
Messages
Commands completed successfully.

Completion time: 2022-08-24T20:35:33.7874872+00:00
```

Queries used in re-assigning of datatype to their original type.

Created new columns,

- **Volatility**: calculated as the difference between the high and low prices of each stock in the dataset.
- **Day of week:** the day of the week of each stock price is extracted from the date column**.**

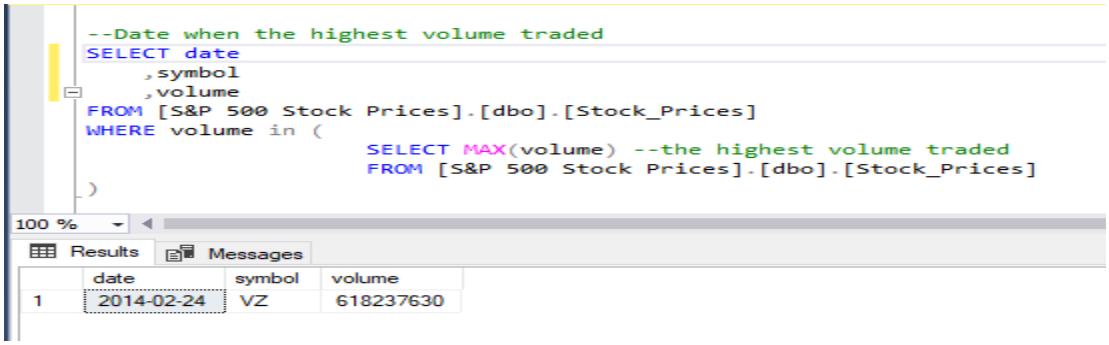


Queries were used in creating the new columns.

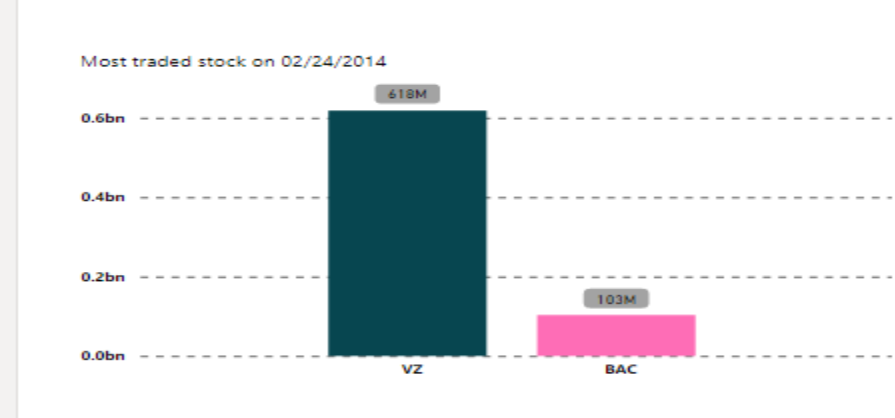
Exploratory Data Analysis

1. Date with the largest overall trading volume and two stocks traded most.

Maven Analytics will want to know which date has the most trading transactions, as well as which stocks traded the most on that day. We at Maven Analytic are attempting to understand the stock market and trading patterns in order to determine the best way to enter the market without losing money, which is the retirement plan set up for her employees.



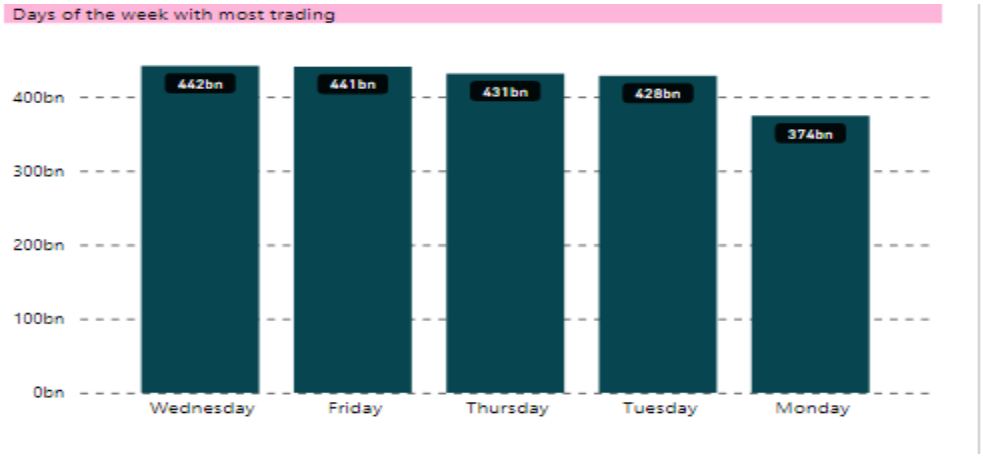
We can see from the query used to return the date with the most trading transactions that 24/02/2014 had the most transactions. Identifying the stock that traded the most on that day can help us gain a better understanding of the stocks.



On 02/24/2014, the most traded stocks were VZ (Verizon Communications Inc) and BAC (Bank of America Corp).

2. On which day of the week does volume tend to be highest? lowest?

A stock trader understands that knowing when to buy and sell is critical to her success. Using the data, we'd like to know which day of the week traders trade the most, so we can know which days we should be watching the stock market closely so we don't miss a profitable opportunity or when to avoid a loss.



We can see from the column chart above that trading was highest on Wednesday and lowest on Monday from 1/2/2014 to 12/29/2017. What we can deduce from this is that when stock prices are either dropping or rising on specific days, we should know traders are either buying or dumping stocks.

3. On which date did Amazon (AMZN) see the most volatility, measured by the difference between the high and low prices?

Volatility represents the size of an asset's price swing around the mean price; it is a statistical measure of its return dispersion. Standard deviation or variance is used to calculate volatility. Higher volatility indicates that the value of a security can potentially be spread out over a wider range of values. This means that the security's price can change dramatically in either direction in a short period of time. More information on volatility can be found here. To determine how frequently the prices of S&P 500 companies' stocks rise and fall, volatility was calculated using the standard derivation from the difference between high and low prices.

```
SELECT STDEVP(volatility) AS Volatility_Rate
FROM [S&P 500 Stock Prices].[dbo].[Stock_Prices]
WHERE symbol = 'AMZN'

--Date Amazon saw the most volatility
SELECT date
      ,volatility
FROM [S&P 500 Stock Prices].[dbo].[Stock_Prices]
WHERE volatility = (
                        select max(volatility)
                            from [S&P 500 Stock Prices].[dbo].[Stock_Prices]
                            where symbol = 'AMZN')
```

100 %

Results | Messages

	Volatility_Rate
1	6.934058291128 78

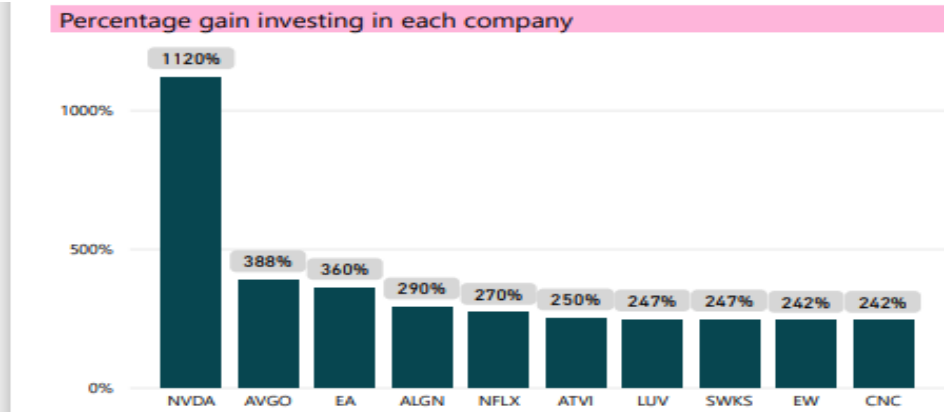
	date	volatility
1	2017-06-09	85.99

AMZN (Amazon) rising and falling volatility is 6.934. Importantly, on September 6, 2017, Amazon (AMZN) experienced its highest volatility, resulting in its highest prices ever on that date. Imagine buying AMZN stock when the market opened and selling when it closed. Profits would have been enormous, with profits exceeding 600%.

4. If you could go back in time and invest in one stock from 1/2/2014 – 12/29/2017, which would you choose? What % gain would you realize?

Remember the plan with the time machine to tell Maven analytics which stock would have made the most money if invested in 2014-01-02. This result can be obtained by calculating the percentage gain from each stock in the dataset by subtracting the purchase price from the present value of the stock, dividing the result by the purchase price, and multiplying by it 100. More information on calculating stock gains can be found in this investopedia article. The purchase

prices will be the close prices as of 2014-01-02, and the current price will be the close price as of 2017-12-29.



Insights

These are the following insights gotten from this analysis

- VZ (Verizon Communications Inc) and BAC (Bank of America Corp) were the most traded stocks on 02/24/2014, which also happened to be the day with the most trades in the data set.
- What we can conclude from this is that when stock prices fall or rise on specific days, we can assume traders are either buying or selling stocks. With additional research, investors can determine whether to sell or hold their stocks.
- Amazon's (AMZN) volatility was around 6.9%, and the AMZN stock experienced the most volatility on September 6, 2017, when Amazon prices fluctuated the most.
- The stocks with the highest prices as of 12/29/2017 would not have yielded the greatest profit if invested in 2014. PCLN (The Priceline Group now Booking Holdings) had the highest stock price in 2017, but if invested in 2014, it would have made 51%, while NVDA (NVIDIA Corporation) has the highest potential gain of 1120.05%.

Recommendations

- From the S&P stock table contenting stocks from 2014 to 2017, NVDA (NVIDIA Corporation) delivered the highest returns if invested in 2014, with a gain of 1120.05%.