

Project Workflow Summary

Monday 4 hours – Setup & Data Collection

I created a public GitHub repository for this research project because it is easier for me to track all my work, and I uploaded everything related to the stock market there as well. I also uploaded my Rstudio file and the early code that I wrote, structured folders for data, analysis scripts, and output.

I installed the required R packages: tidyverse, tidyquant, lubridate, and zoo. Tidyverse, tidyquant, and lubridate packages help make stock market data analysis easier. tidyverse is used to clean, filter, and visualize data. tidyquant allows you to pull stock prices and calculate things like returns or moving averages. lubridate makes it simple to work with dates, such as sorting data by month or year. Together, they help organize, analyze, and understand stock trends more effectively.

Downloaded CSV files of Apple, Microsoft, Nvidia, and Google from Yahoo Finance, filtering only the years 2000 – 2023, and I also uploaded all the datasets to my GitHub repository.

Tuesday (4.5 hours) – Data Cleaning & EDA

Imported CSVs into R and converted Date columns to the proper date format. Got a few syntax errors because the datasets stored the date in 2 different ways (mm-dd-year and mm/dd/year). I also filtered the dataset for years 2000 to 2023 because 2 datasets have data from 1984, and there is a dataset that only started in 2000. I decided to filter all datasets to be from the year 2000 until 2023 only.

I created 4 line plot visualizations of Close prices for each company and also created a dashboard using the Tableau application. I computed daily returns using $\text{Close}/\text{lag}(\text{Close}) - 1$

Wednesday (3 hours) – Volatility & Correlation Analysis

I started by writing R code in RStudio to calculate the standard deviation of daily returns for each stock as a measure of risk, selected only the Date and Return columns, joined all datasets using `full_join()` by date, and generated a correlation matrix to study inter-stock dependencies.

Thursday (3 hours) – Beta & Risk Summary

I chose Apple as a market benchmark because of its dominant position in the worldwide tech industry. I computed beta for Microsoft, Nvidia, and Google using $\text{cov}(\text{stock}, \text{Apple}) / \text{var}(\text{Apple})$ and created a consolidated risk summary table including standard deviation and beta values.

Friday (4 hours) – Tableau Visualizations

Imported each stock CSV into Tableau, created individual line charts of closing price over time.

And also designed a 2x2 dashboard comparing all four companies.

Saturday (4 hours) – Report Compilation & GitHub Update

I spent 4 hours compiling all my analysis, charts, and summary tables into a single RMarkdown report, which I then exported as an HTML file. After that, I uploaded the final version of my code, cleaned datasets, the README file, and the output files to my GitHub repository. I also made sure to write clear commit messages and organize everything into appropriate folders to keep the project well-structured and easy to navigate.

Sunday (3.5 hours) – Reflection & Submission

I completed the weekly activity log and inserted all the activities that I did in the template given in Canvas, wrote the final reflection and conclusions, and also reviewed and proofread the report, committed “Final Version” to GitHub.

Some of the Key Findings this week:

- *Apple rose from \$0.23 (2003) to \$198.11 (2023); steady growth fueled by product innovation.*
- *Microsoft rebounded from a 2009 low of \$15.15 to \$377.85 in 2023 due to cloud expansion.*
- *Nvidia showed the steepest growth from \$0.61 (2002) to \$950.02 (2024), reflecting the AI boom.*
- *Google peaked at \$2,996.77 in 2021 after strong digital ad and cloud performance.*
- *Nvidia had the highest volatility; Google showed the least correlation with the others.*