

# Case Study Rubric

DS 4002 – Instructor: Professor Alonzi

Due: To Be Determined

## Submission Format

- Github Repository – contains code, necessary data, output, and final write-up

## Individual Assignment

Preparatory Assignments – none

## Why am I doing this?

This will be an opportunity in order to showcase your technical skills as well as interpret crucial times series results. The case study will present a real-life scenario to one you may undertake in the future or in any data science or statistical courses or in a professional setting.

## What am I going to do?

The Github repository for this case study can be found at: <https://github.com/ashritakodali/Time-Series-Analysis-Case-Study>. In addition, the Github repository for the original project can be found: <https://github.com/briseek/DS4002-Stocks-Time-Series>. You will obtain data from the S&P 500 website and conduct a times series analysis in order to predict future returns. Specifically, you will first explore the data and conduct a times series analysis. You will explore your response variable and see if any transformations will be required. Afterwards, you will build several SARIMA and GARCH models and determine which ones meet model diagnostics. Finally, you will predict future forecasts and determine which model is the best at predicting the S&P 500. The deliverables will include the following:

- One page document detailing the process and reflecting on said process
- Well Document R Markdown Script
- A folder with the necessary output
- A folder with the original and modified data
- A folder containing the data appendix

## Tips for Success:

- Make your variables easy to interpret and easy to call back upon
- Make sure you understand the model diagnostics and SARIMA and GARCH model assumptions before building them
- Familiarize yourself with R and how to format a R Markdown script

## How will I know I have Succeeded?

Category	Details
Formatting	<ul style="list-style-type: none"><li>● One Github Repository (submitted via link on Canvas)<ul style="list-style-type: none"><li>○ Create a new Github repository for this assignment and name it "CS2 Predicting the S&amp;P 500"<ul style="list-style-type: none"><li>■ This repository will contain the following:<ul style="list-style-type: none"><li>● README.md</li><li>● LICENSE.md</li><li>● SCRIPTS FOLDER<ul style="list-style-type: none"><li>○ Contains the scripts used to conduct the times series analysis</li></ul></li></ul></li></ul></li></ul></li></ul>

	<ul style="list-style-type: none"> <li>● OUTPUT FOLDER <ul style="list-style-type: none"> <li>○ Contains the images of the important output from the EDA and the model building process</li> </ul> </li> <li>● DATA FOLDER <ul style="list-style-type: none"> <li>○ INITIAL FOLDER <ul style="list-style-type: none"> <li>■ Contains the raw data from the S&amp;P 500 website</li> </ul> </li> <li>○ FINAL FOLDER <ul style="list-style-type: none"> <li>■ Contains the modified final data</li> </ul> </li> </ul> </li> <li>● REFERENCES.md</li> <li>● final_analysis.pdf</li> </ul>
Written Portion	<ul style="list-style-type: none"> <li>● Reflect on the project as a whole and what implications it has <ul style="list-style-type: none"> <li>○ Discuss the problem and explain why it is important <ul style="list-style-type: none"> <li>■ Explain the goal and the problem</li> <li>■ Provide context as to why this problem is important</li> </ul> </li> <li>○ Explain the general modeling approach and why it can be useful <ul style="list-style-type: none"> <li>■ Explain the general steps used to build the models</li> <li>■ Justify why these steps were taken</li> </ul> </li> <li>○ Discuss the results of entire process and what <ul style="list-style-type: none"> <li>■ Explain and present your results</li> <li>■ Interpret the findings in a broader context (make sure to relate it back to the goal)</li> </ul> </li> </ul> </li> <li>● Provide a reflection on the entire process <ul style="list-style-type: none"> <li>○ Provide a positive critique of what you liked</li> <li>○ Provide a negative critique or a suggestion that you would have liked to see</li> </ul> </li> </ul>
Code	<p>A well documented R Markdown file will be created that will execute your times series analysis. In the code, you must include the following:</p> <ul style="list-style-type: none"> <li>● The Exploratory Data Analysis and Transformations of the Data</li> <li>● The SARIMA models</li> <li>● The GARCH models</li> <li>● Summary Performance Metrics of Each Model</li> <li>● Comments throughout the file that explain each step and your thought process</li> </ul>
References	<p>In the REFERENCES.md, you must include the following:</p> <ul style="list-style-type: none"> <li>● Use IEEE documentation to cite any sources you used</li> <li>● Use IEEE documentation to cite any external help you received (ie: ChatGPT)</li> </ul>

Acknowledgments: Thank you to Professor Alonzi for providing the Rubric Structure!