

General Instructions for All Projects:

- Out of the three projects, please select two to work on. We recommend considering Project 3 if possible.
- During the course of these projects, should you encounter any questions or uncertainties, no one will be available to provide direct answers. In such cases, please make your best assumptions, justify them, and proceed with your work.

Project 1: Predicting the S&P/Case-Shiller U.S. National Home Price Index

Objective: Predict the future values of the S&P/Case-Shiller U.S. National Home Price Index using relevant indicators from the FRED database, ensuring no look-ahead bias in the data.

Instructions:

Data Collection: Use the FRED API to collect the historical data for the S&P/Case-Shiller U.S. National Home Price Index. You may choose any other indicators from FRED that you believe are relevant to your model. Ensure you abide by the guidelines provided in the FRED's documentation, especially regarding look-ahead information.

Data Preprocessing: Clean the dataset by handling any missing values, outliers, or any anomalies. Properly align the dates to ensure that no look-ahead bias is introduced.

Feature Engineering: Extract or engineer features from the data that may have predictive power.

Model Selection: Choose an appropriate machine learning or statistical model for the time series forecasting task.

Model Training and Validation: Split the data into a training and validation set. Train the chosen model using the training set and validate its performance using the validation set.

Evaluation: Evaluate the model's accuracy using appropriate metrics such as RMSE, MAE, etc.

Report: Summarize the findings, mentioning the indicators used, the reasoning behind the choice of the model, and the results achieved. Please also attach your code.

Project 2: Predicting Correlation of Selected Stocks with the SPY Index

Objective: Predict the correlation of GOOG, AMZN, JPM, GME, and XOM with the SPY index for the next two months using only historical prices.

Instructions:

Data Collection: Download the historical price data for the stocks GOOG, AMZN, JPM, GME, XOM, and the SPY index from Yahoo Finance or other sources. Ensure to account for corporate adjustments.

Data Preprocessing: Clean the data by handling any missing values, outliers, or anomalies. Check for corporate adjustments and account for them.

Model Selection: Choose an appropriate model to predict the future correlations. Consider models suitable for time series forecasting.

Model Training and Validation: Train the chosen model using historical correlations and validate its performance.

Evaluation: Assess the model's predictive accuracy using suitable metrics.

Report: Present your findings detailing the approach, choice of model, and achieved results. Please attach your code.

Project 3: Idea Generation for Price Prediction Signal for a Macro Asset or ETF

Objective:

Conceptualize a novel price prediction signal for a chosen macro asset or broad-based index ETF. This project focuses on idea generation, understanding the underlying rationale, and potential data-driven insights that could back up the proposed signal. While the main emphasis is on the conceptual framework, implementation or preliminary testing is a valuable addition.

Begin by selecting a major macro asset or broad-based index ETF of interest. Based on your research, conceptualize a potential price prediction signal. This could be rooted in technical analysis, fundamental analysis, or any unconventional sources or patterns you've identified. Articulate the reasoning behind your proposed signal. Why do you believe it might have

predictive power? Are there any historical events or data-driven insights that might support your idea?

Detail your thought process in conceptualizing the signal, the research conducted, and the rationale behind its potential effectiveness. If you opted for preliminary testing, discuss your data, findings and the implications they might have for the feasibility of your signal.