# Stock Price Prediction in the Presence of High Volatility, Missing Data, and Extreme Events

## Evaluating the Performance of Different Prediction Models and Data Preparation Techniques

Your Name

Date

## Introduction

In this presentation, we will explore the challenges in stock price prediction, particularly in the presence of high volatility, missing data, and extreme events. We will discuss the significance of addressing these challenges for investors and financial institutions and present the research questions we aim to answer.

## Data Preparation

We will explain the process of preparing the four versions of the dataset for each time series:

* Original data with missing values removed and outliers not identified
* Original data with missing values removed and outliers identified
* Imputed data with outliers not identified
* Imputed data with outliers identified

We will also describe the data imputation techniques used, such as linear interpolation and rolling mean, and explain the methods used for outlier detection, for example, clustering-based methods.

## Data Resampling

In this section, we will describe the rationale behind resampling the time series to daily frequency, including capturing potential daily and weekly periodicities and computational efficiency. We will explain the resampling methods used, such as resampling to daily frequency using the mean of the original data.

## Benchmark Model

We will explain the random walk simulation as a benchmark for comparison. We will describe how the random walk model is used to simulate the unpredictability of stock prices in the short term and discuss the importance of comparing the performance of prediction models against this benchmark.

## Hypotheses

In this part of the presentation, we will list the different hypotheses to be tested:

* H1: ts1 can predict ts1
* H2: ts2 can predict ts2
* H3: ts1 can predict ts2
* H4: ts2 can predict ts1

We will explain the motivation behind testing these hypotheses, including investigating the relationships between the two time series and their predictability.

## Prediction Models

We will introduce the chosen prediction models (e.g., ARIMA, GARCH, LSTM, etc.), describe the key characteristics of each model and why they were chosen for the study, and mention that these models will be applied to test the hypotheses on the four versions of the dataset.

## Model Evaluation

In this section, we will describe the evaluation metrics (MSE, MAE, R-squared) and their significance in evaluating model performance. We will explain the process of comparing the performance of different models on different dataset versions, highlighting the importance of understanding the impact of data preparation techniques on prediction accuracy. We will also discuss the investigation of the impact of data imputation techniques, such as linear interpolation and rolling mean, on the accuracy of stock price predictions. Additionally, we will assess the effectiveness of different outlier detection techniques in improving the model's robustness to extreme events.

## {Insert Results and Graphs}

In this part of the presentation, we will show the results of the experiments, including graphs and tables that illustrate the performance of the prediction models on different dataset versions and under various hypotheses. We will discuss the key insights and findings derived from the results.

## Results and Conclusions

In the final part of the presentation, we will summarize the findings, highlighting the most effective techniques for handling missing data, outliers, and challenges of stock price prediction in the presence of high volatility and extreme events. We will provide insights into the relationships between the two time series and their predictability. We will discuss the limitations of the study and potential avenues for future research, such as exploring alternative data imputation and outlier detection techniques, or testing additional prediction models in the context of stock price prediction with high volatility and extreme events.