Wilshire 5000 Index VS Shanghai Composite Index

Sohail Ahmad, Kenneth Romero,
Jeffrey Umanzor, Joquanna Scott,
Priti Saha, Aquib Zaman

12/16/24

Team One

AGENDA

01 **BACKGROUND**
Introduction
Problem Statement

02 **Methodology**
UML Diagrams for
Workflow & Uses Cases

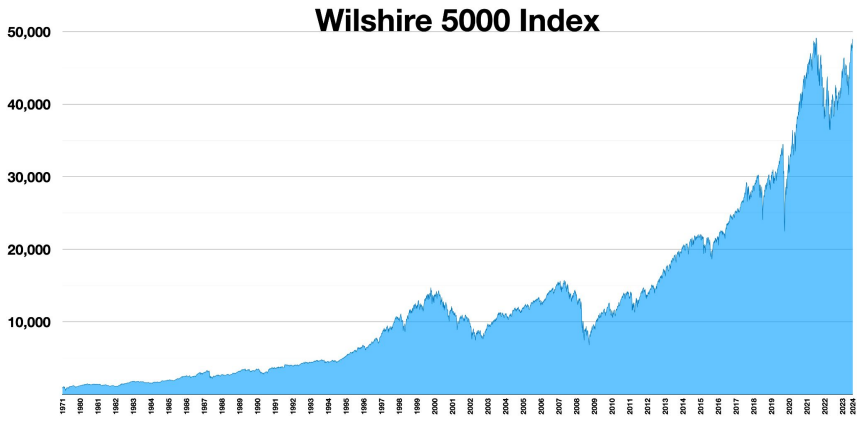
03 **ETL**
Dataset Logic
ETL Process

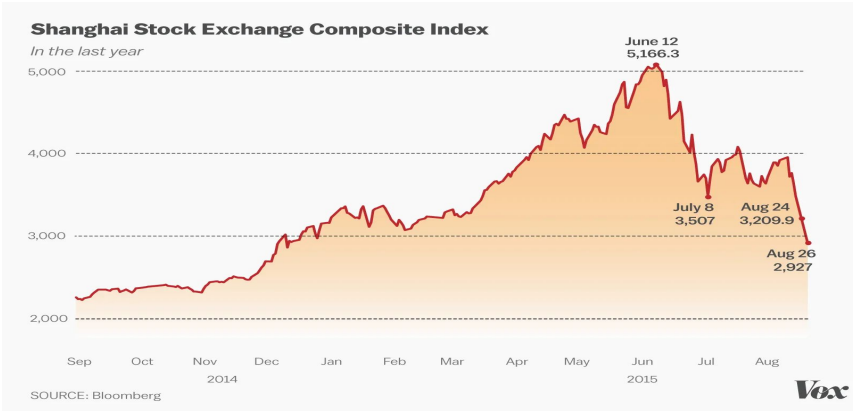
04 **Dashboard Showcase**
Tableau Dashboard
Showcase

05 **Predictive Analysis**
Discussing Findings

06 **Conclusion**
Q&A

Introduction & Problem Statement





How do broad market indices, like the Wilshire 5000 Index (US) and the Shanghai Composite Index (China), perform during global economic downturns, and what factors contribute to their resilience or vulnerability in such events?

The Wilshire 5000 Index is basically a snapshot of the entire U.S. stock market. It includes all the publicly traded companies in the U.S. that have easily available price data, so it covers everything from big companies to tiny ones. It's often called the total stock market index because it gives a good overall view of how the market is doing. The Shanghai Composite Index, on the other hand, tracks the stocks listed on the Shanghai Stock Exchange in China. It includes a mix of big state-owned companies and private businesses, giving a sense of how the Chinese market and economy are performing. Both are great for understanding the bigger picture in their respective markets.

Wilshire 5000 Index (US) and the Shanghai Composite Index (China)

The Wilshire 5000 Total Market Index is one of the broadest measures of the U.S. stock market, designed to represent the performance of nearly every publicly traded company headquartered in the United States. Despite its name, the index doesn't always have 5,000 components its number of constituents fluctuates as companies enter or exit the market. It includes all stocks with readily available pricing data, spanning large-cap, mid-cap, small-cap, and even micro-cap stocks. The Wilshire 5000 covers a wide range of industries and sectors, making it a key indicator of the overall health and direction of the U.S. equity market. The Wilshire 5000 captures the performance of not just the major players but also smaller companies, offering a more comprehensive view. It's often used by analysts, investors, and economists to gauge market trends and evaluate the broader economy state.

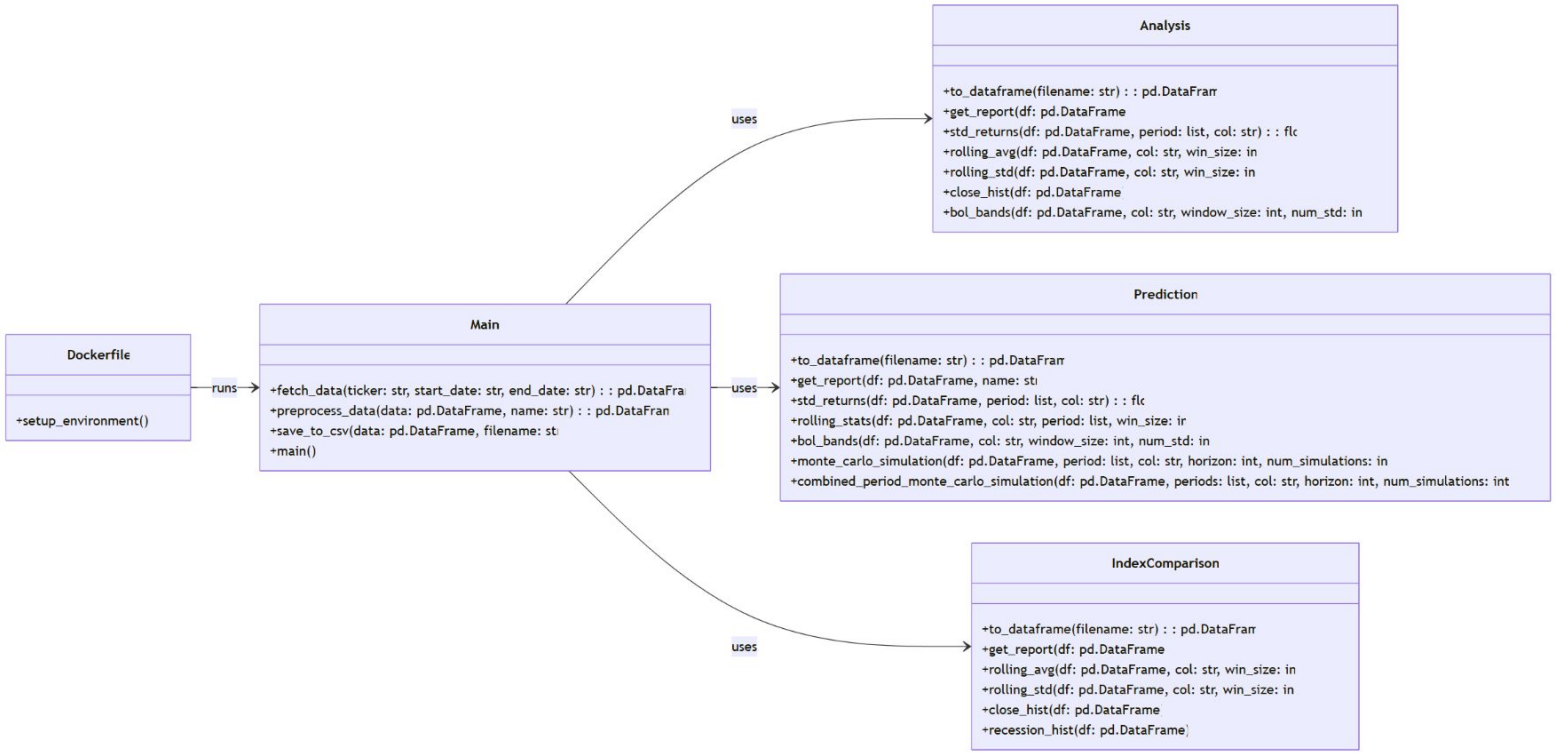
The Shanghai Composite Index is a key benchmark for China's stock market, tracking all A-shares and B-shares listed on the Shanghai Stock Exchange. It includes a wide range of companies, from large state-owned enterprises to private firms, giving insight into the overall performance of the Chinese economy. The index is heavily influenced by domestic investors and government policies, making it a useful gauge of market sentiment and economic trends in China. It's a popular reference point for understanding the health of China's financial markets.

Introduction

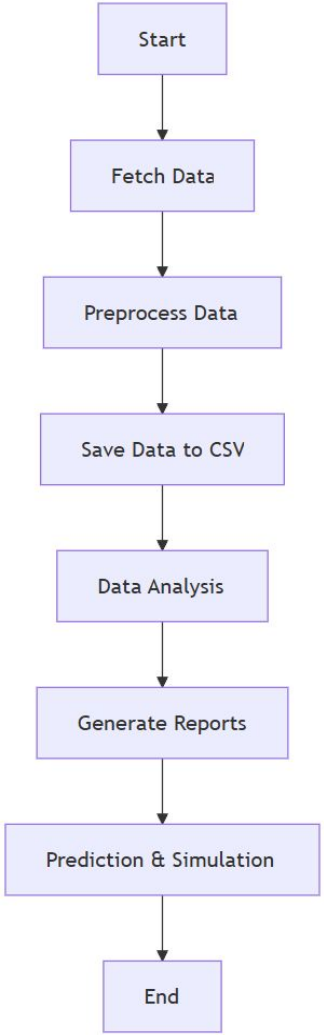
Background/Context:
Analyzing stock market volatility helps us uncover patterns in market dynamics, particularly how investor behavior and market stability are influenced during economic fluctuations. The Wilshire 5000 Index, as a comprehensive measure of the U.S. stock market, and the Shanghai Composite Index, representing the Chinese market, serve as excellent benchmarks for studying these trends. Together, these indices reflect broader economic shifts in two of the world's largest economies. This project will utilize the yfinance API to collect historical data on both indices, ensuring data is properly organized and cleaned for reliable analysis. Leveraging Python libraries for data visualization and statistical modeling, we aim to identify and compare volatility trends across these markets. By exploring how they respond to economic challenges, this study provides valuable insights for academic research and practical strategies for investors managing market uncertainty.

Problems We Found Answers To:
After analyzing the data, we aim to address key questions about stock market volatility during economic downturns. Specifically, we will compare the volatility levels of the Wilshire 5000 Index and the Shanghai Composite Index to understand how these markets differ in their responses to economic crises. We'll explore patterns in price fluctuations for both indices and their relationships with broader economic indicators. Additionally, we plan to investigate the recovery speed of each index following downturns, identifying the factors that influence these recovery paths. We will also analyze which economic events or conditions have the most significant impact on the volatility of each index and assess how these effects vary between the U.S. and Chinese markets. Finally, through statistical modeling, we aim to provide predictive insights into future volatility trends, offering valuable guidance for investors navigating uncertain market conditions.

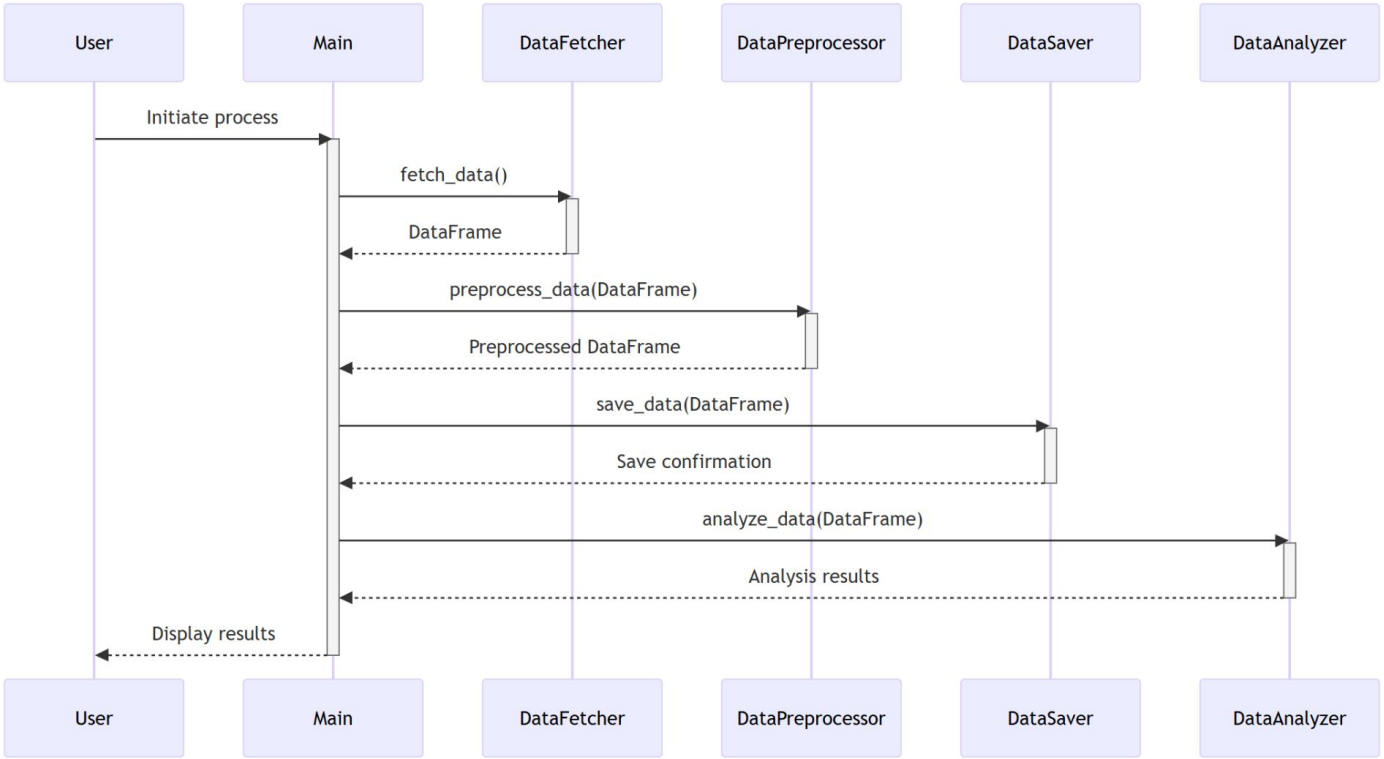
Class Diagram



Workflow Diagram



Sequence Diagram



ETL Process

- Extraction: Data was extracted from Yahoo Finance using the **yfinance** module of Python
- Transform:
  - Missing data values were imputed using pandas **ffill()** and **bfill()** for comprehensive filling
  - Normalized and Daily Return prices were calculated during preprocessing and added to the dataset
    - Normalized: normalization of Adjusted Close price data
    - Daily Return: percent change of the Adjusted Close prices
- Load:
  - For initial EDA operations, the dataset was loaded into Python dataframes.
  - For dashboarding purposes, data was held in csv files before being loaded into Tableau for analysis.

Dataset Logic

- Final Structure of both DataFrames:

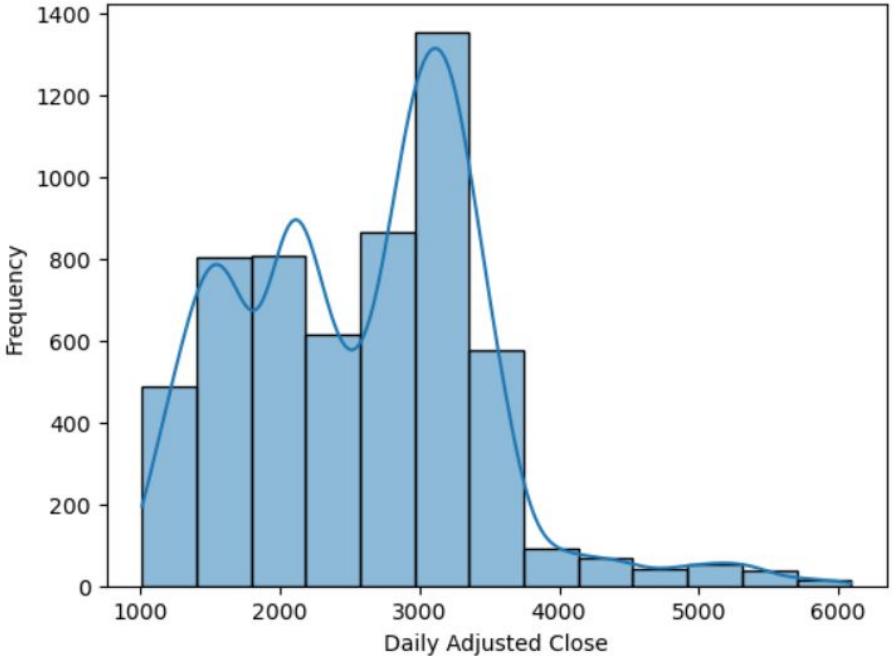
```
Data columns (total 9 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   Adj Close      5810 non-null   float64
 1   Close          5810 non-null   float64
 2   High           5810 non-null   float64
 3   Low            5810 non-null   float64
 4   Open           5810 non-null   float64
 5   Volume         5810 non-null   float64
 6   Normalized     5810 non-null   float64
 7   Daily Return   5809 non-null   float64
 8   Index          5810 non-null   object
dtypes: float64(8), object(1)
```

Initial Findings

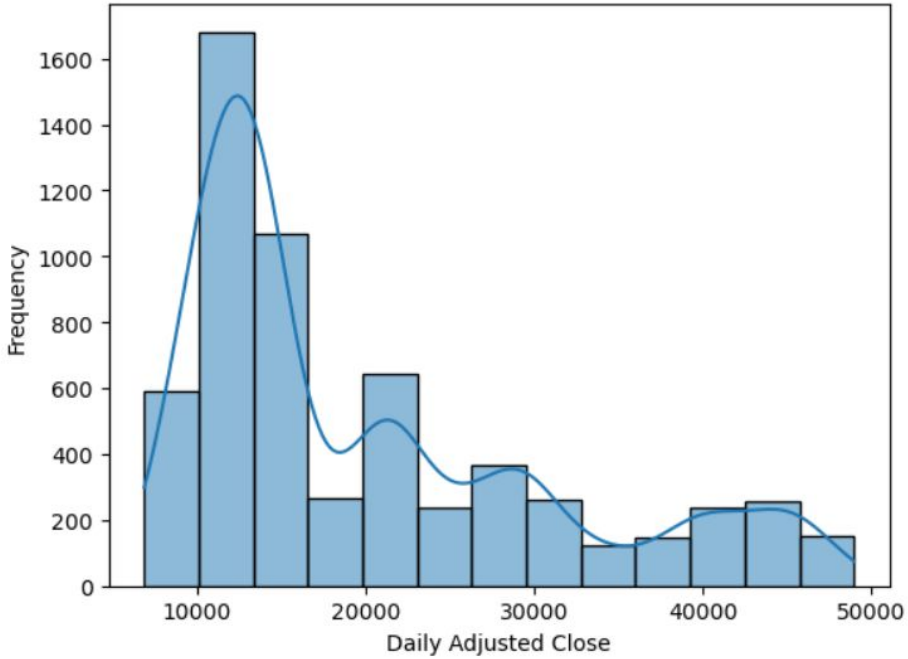
- Total dataset comparison:
  - Shanghai - red,
  - Wilshire - green

Initial Findings (cont.)

- Adjusted Close Distribution for both indices



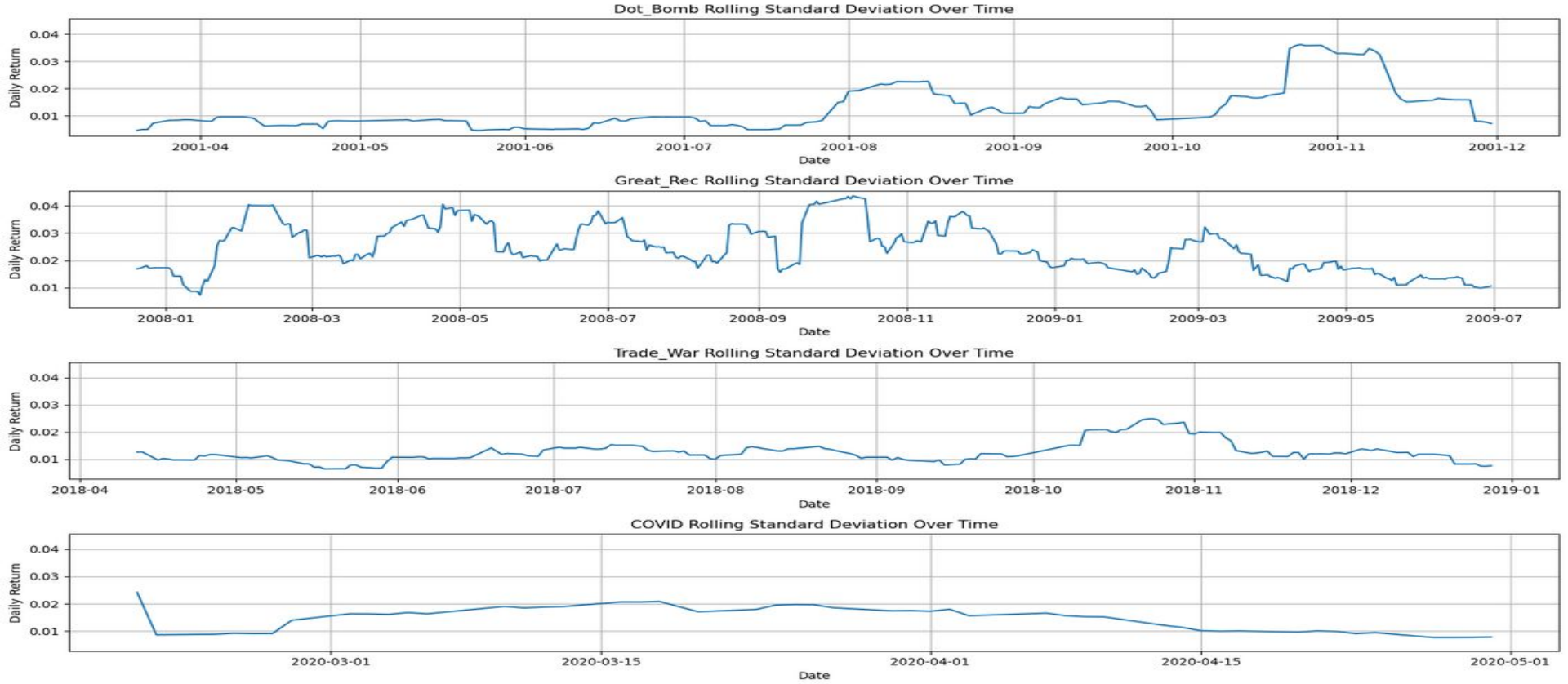
Shanghai Composite



Wilshire 5000

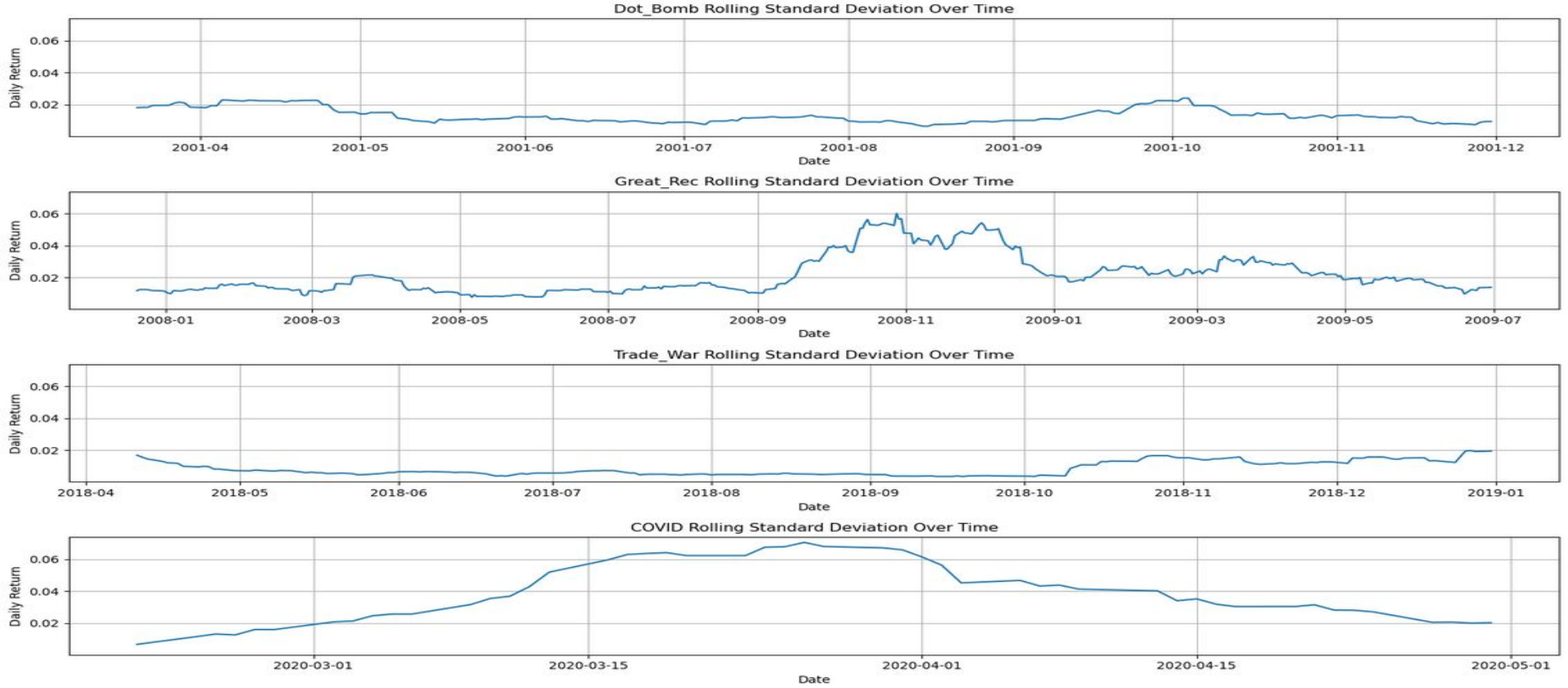
Initial Findings (cont.)

- **Shanghai** Daily Return Rolling Standard Deviation



Initial Findings (cont.)

- **Wilshire** Daily Return Rolling Standard Deviation



Dashboard Showcase

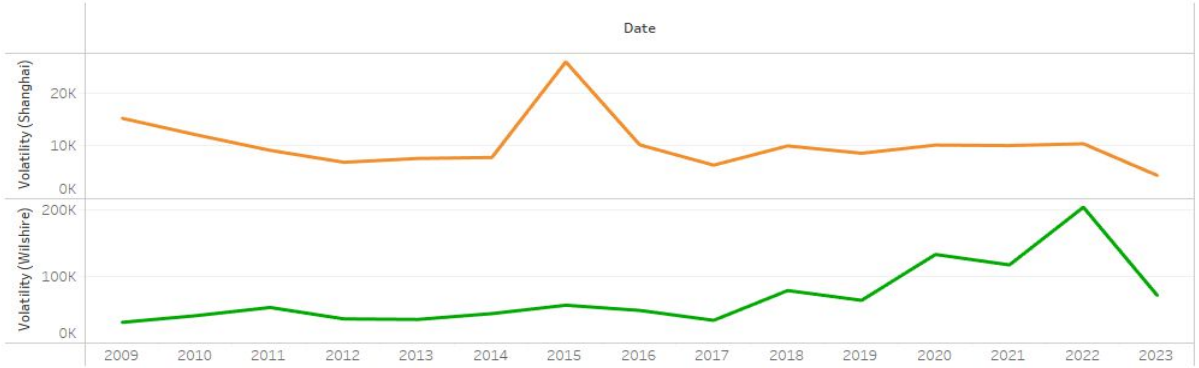
Overview | COVID-19 | China-United States Trade War | Great Recession | Dot-Com Bomb

Daily Return

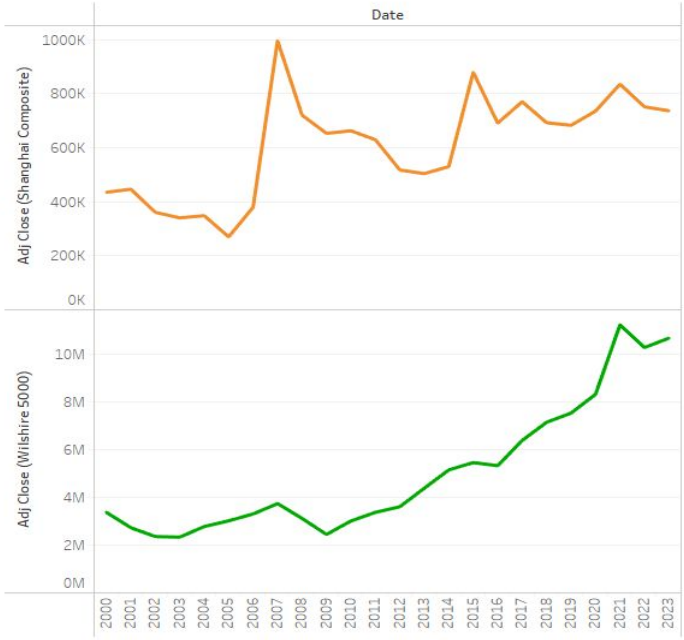
Index (Shangai Composite)
Shanghai Composite

Index (Wilshire 5000)
Wilshire 5000

Volatility Over Time

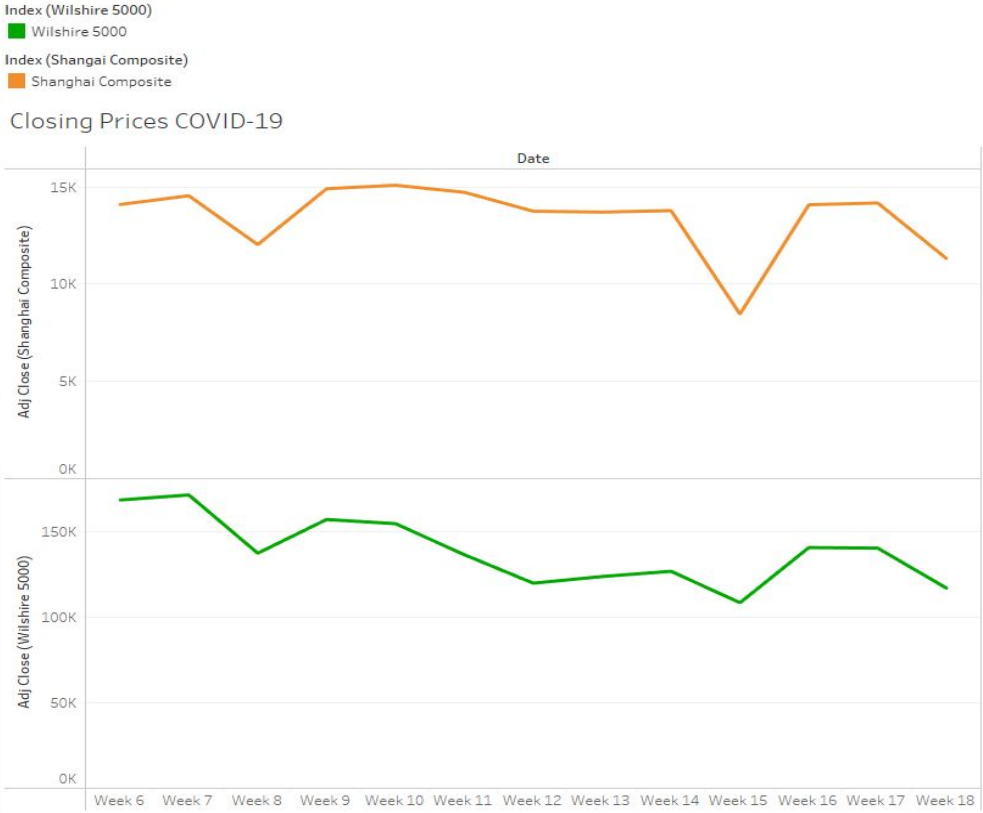


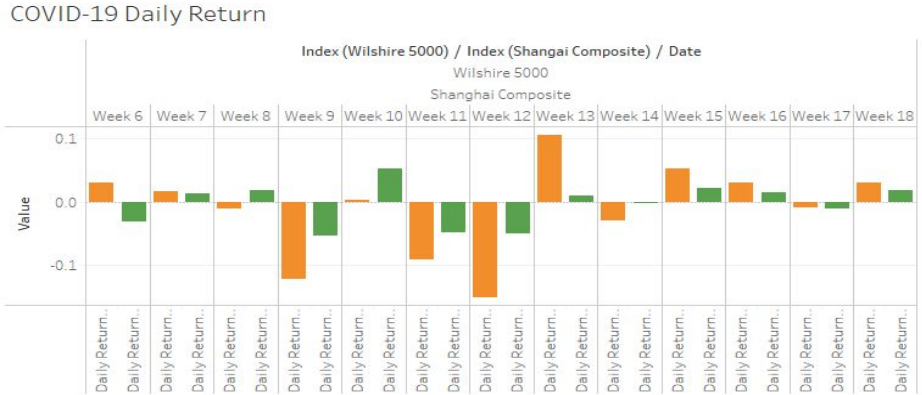
Closing Prices



Dashboard Showcase (cont.)

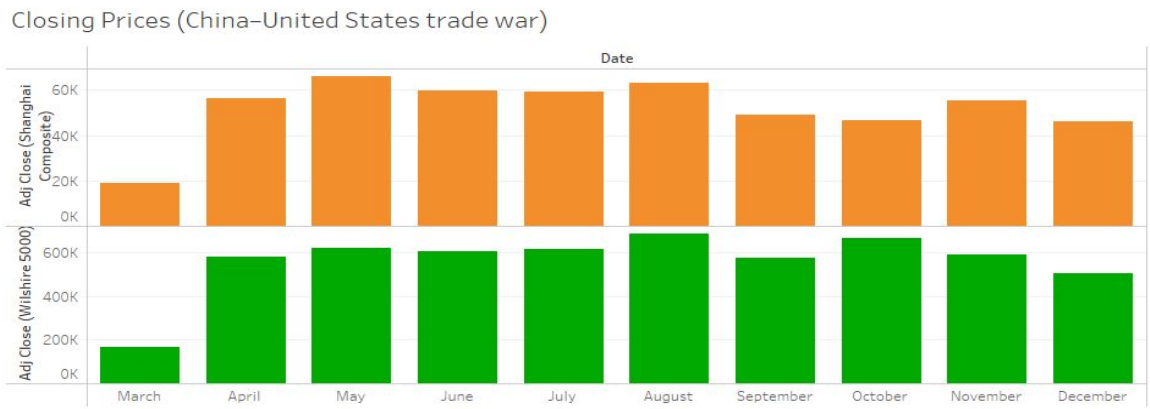
Overview | COVID-19 | China-United States Trade War | Great Recession | Dot-Com Bomb

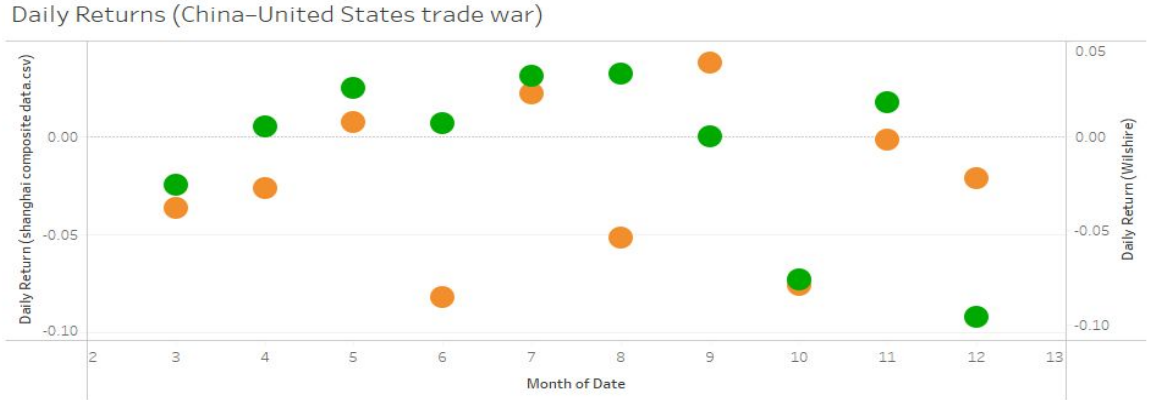




Dashboard Showcase (cont.)

Overview | COVID-19 | China-United States Trade War | Great Recession | Dot-Com Bomb

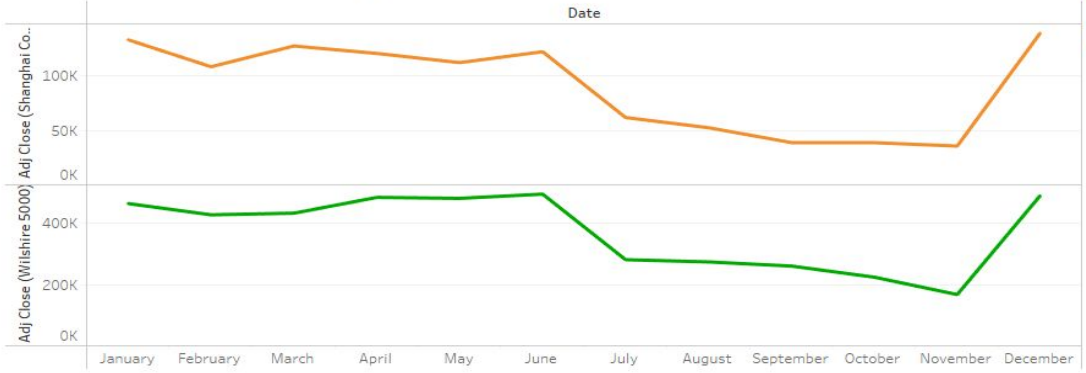




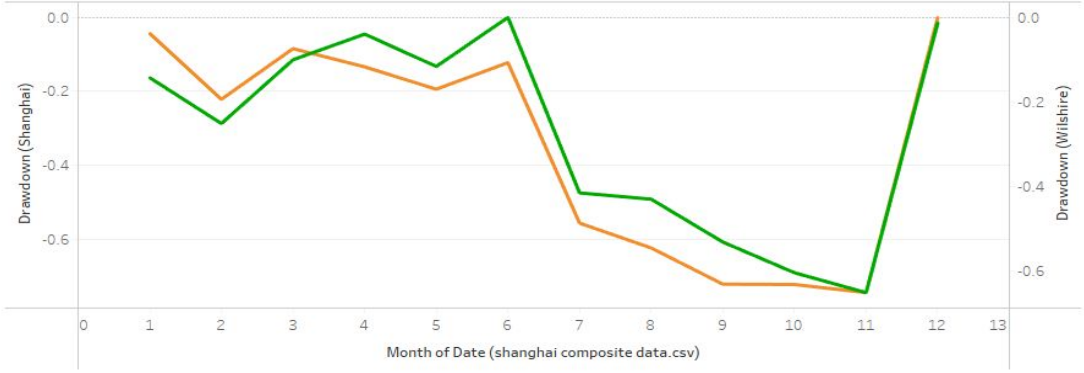
Dashboard Showcase (cont.)

Overview | COVID-19 | China-United States Trade War | **Great Recession** | Dot-Com Bomb

Closing Prices (Great Recession)

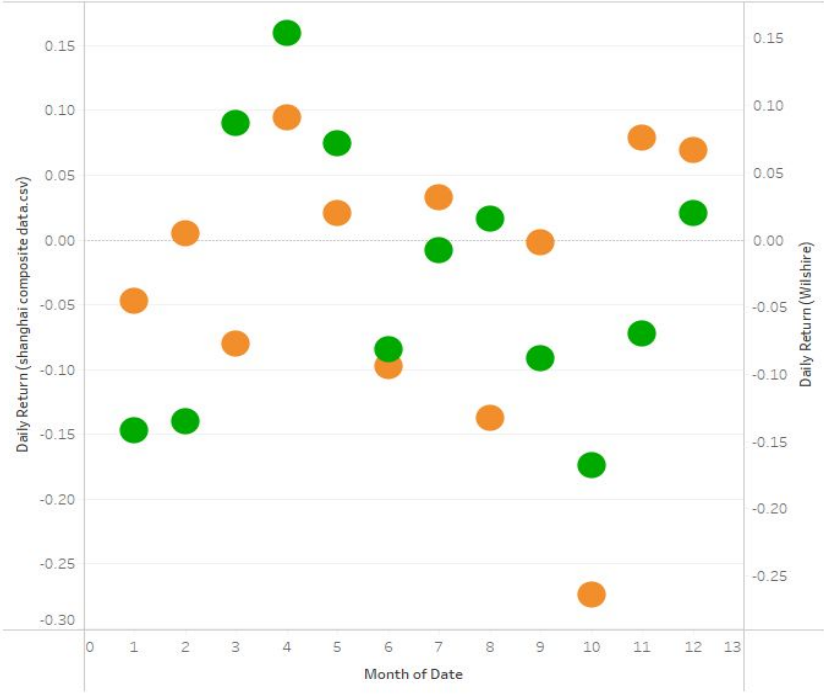


Drawdown (Great Recession)

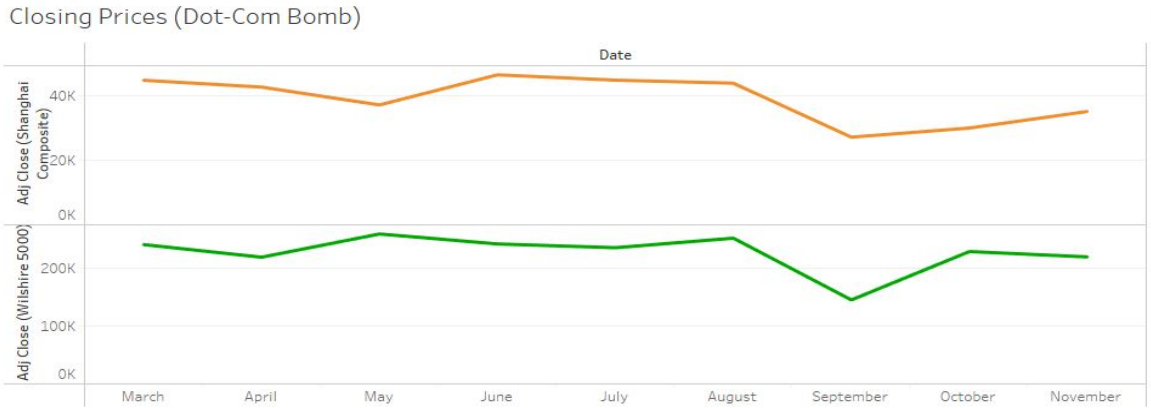


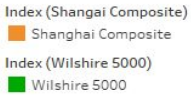
Index (Shangai Composite)
Shanghai Composite
Index (Wilshire 5000)
Wilshire 5000

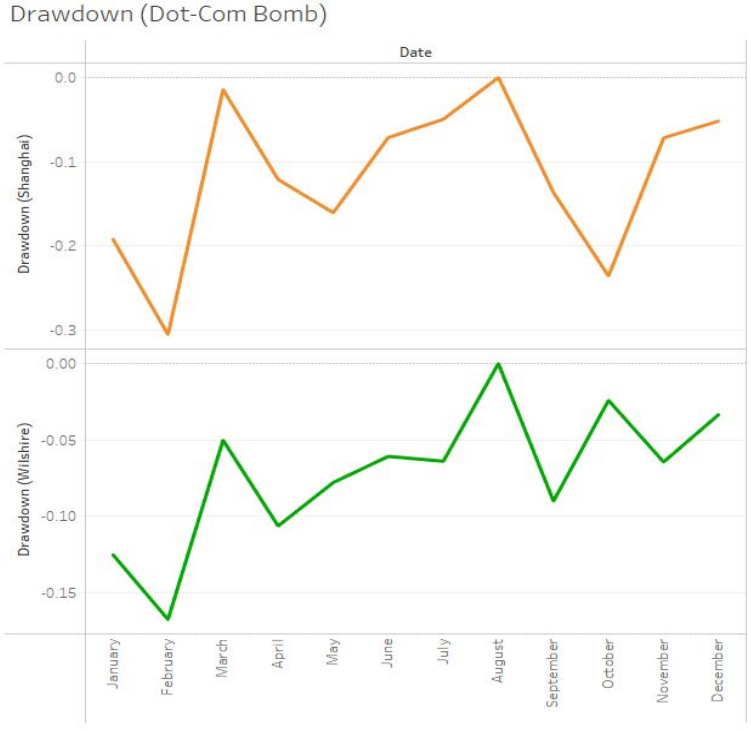
Daily Return (Great Recession)

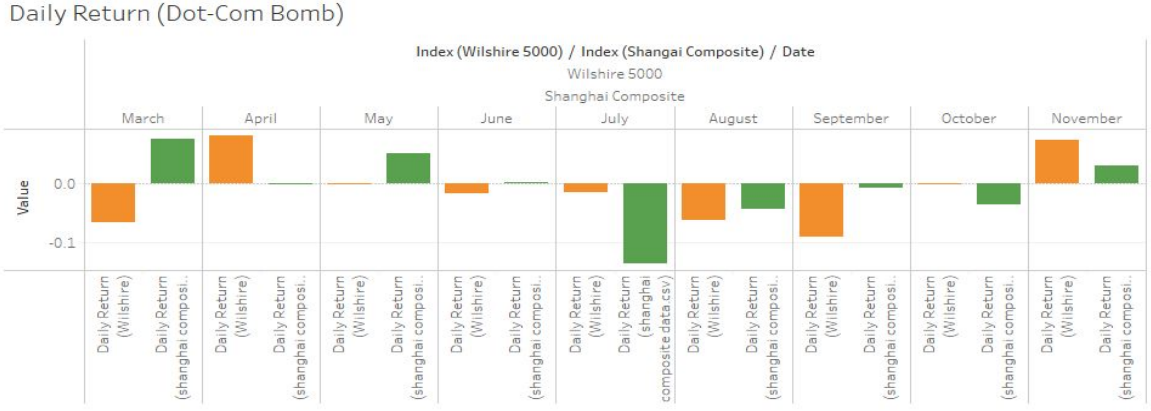


Dashboard Showcase (cont.)



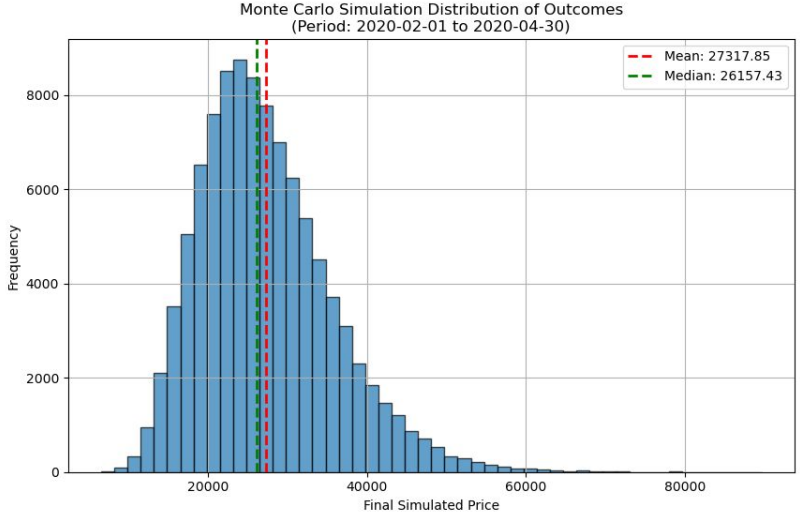






Predictive Analysis

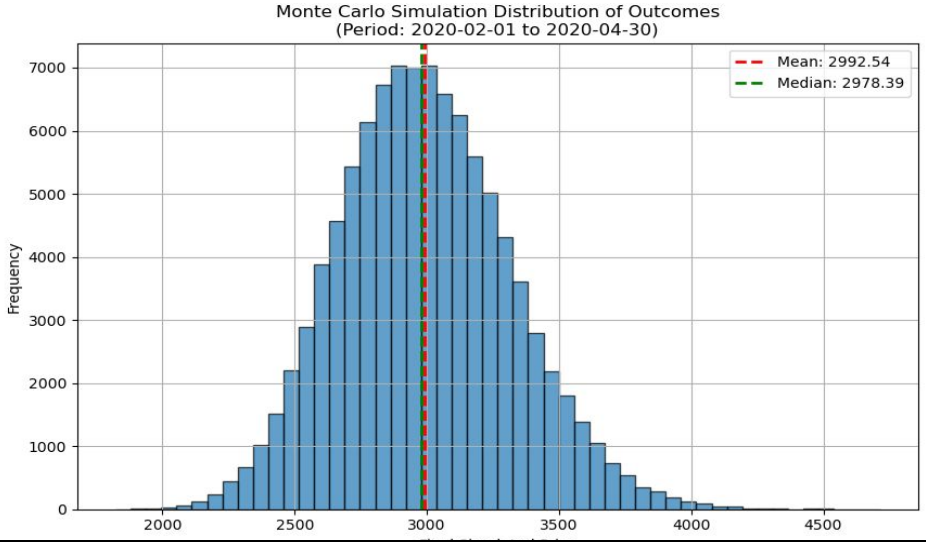
Monte Carlo Simulation (Wilshire 5000) - COVID-like scenario:




```
Monte Carlo Simulation Summary (2020-02-01 to 2020-04-30):
Mean final price: 27317.85
Median final price: 26157.43
Std Dev of final price: 8414.84

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.12%
Median of daily returns (averaged over 60 days): -0.04%
Average daily returns standard deviation: 3.90%
```

Monte Carlo Simulation (Shanghai Composite) - COVID-like scenario:

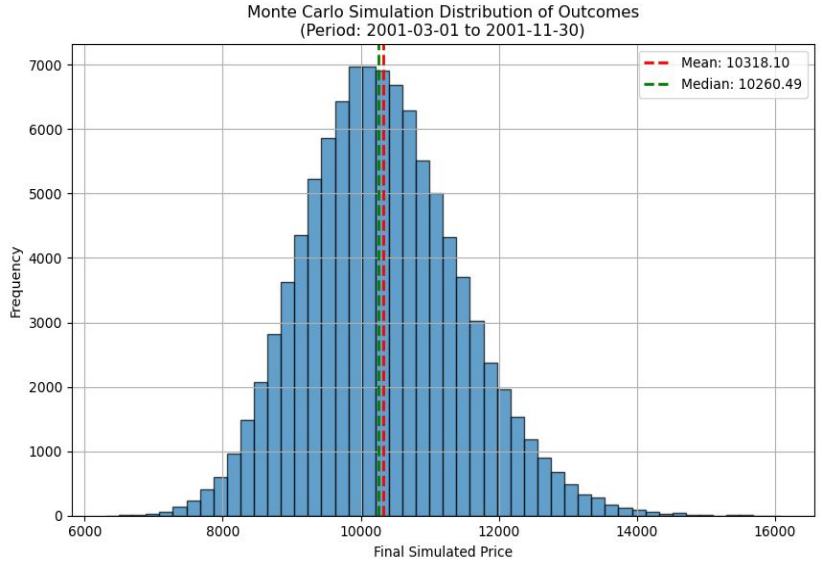



```
Monte Carlo Simulation Summary (2020-02-01 to 2020-04-30):
Mean final price: 2992.54
Median final price: 2978.39
Std Dev of final price: 325.77

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): 0.08%
Median of daily returns (averaged over 60 days): 0.17%
Average daily returns standard deviation: 1.40%
```

Predictive Analysis (cont.)

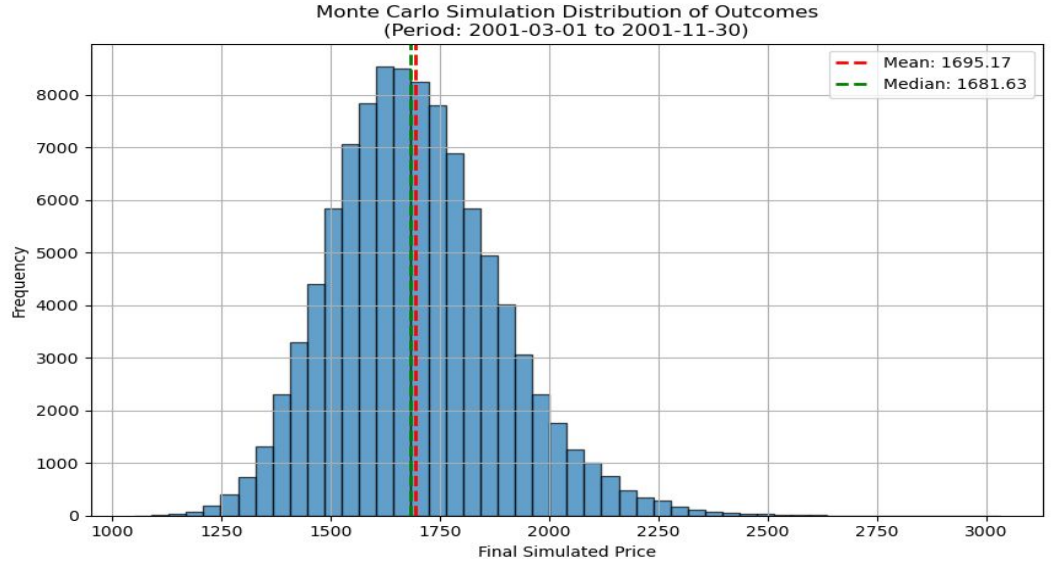
Monte Carlo Simulation (Wilshire 5000) – Dot-Com Bubble scenario:




```
Monte Carlo Simulation Summary (2001-03-01 to 2001-11-30):
Mean final price: 10318.10
Median final price: 10260.49
Std Dev of final price: 1133.16

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.03%
Median of daily returns (averaged over 60 days): -0.05%
Average daily returns standard deviation: 1.41%
```

Monte Carlo Simulation (Shanghai Composite) – Dot-Com Bubble scenario:

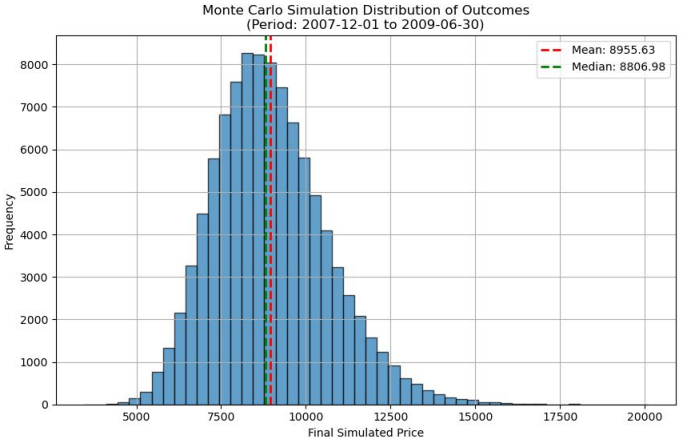



```
Monte Carlo Simulation Summary (2001-03-01 to 2001-11-30):
Mean final price: 1695.17
Median final price: 1681.63
Std Dev of final price: 191.46

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.05%
Median of daily returns (averaged over 60 days): 0.02%
Average daily returns standard deviation: 1.44%
```

Predictive Analysis (cont.)

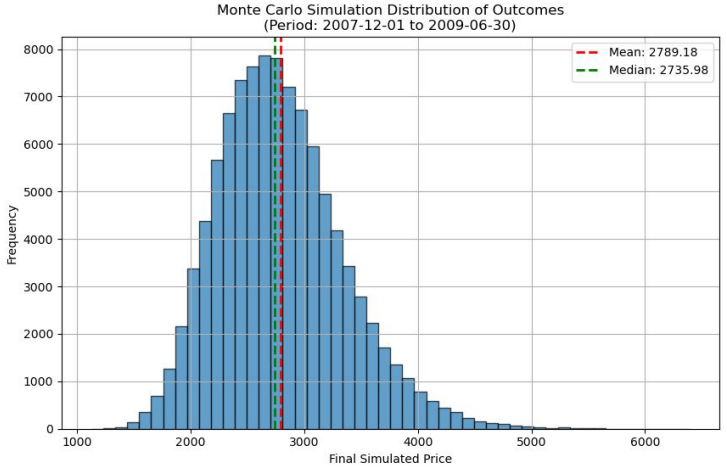
Monte Carlo Simulation (Wilshire 5000) – Great Recession scenario:




```
Monte Carlo Simulation Summary (2007-12-01 to 2009-06-30):
Mean final price: 8955.63
Median final price: 8806.98
Std Dev of final price: 1679.59

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.09%
Median of daily returns (averaged over 60 days): -0.01%
Average daily returns standard deviation: 2.40%
```

Monte Carlo Simulation (Shanghai Composite) – Great Recession scenario:

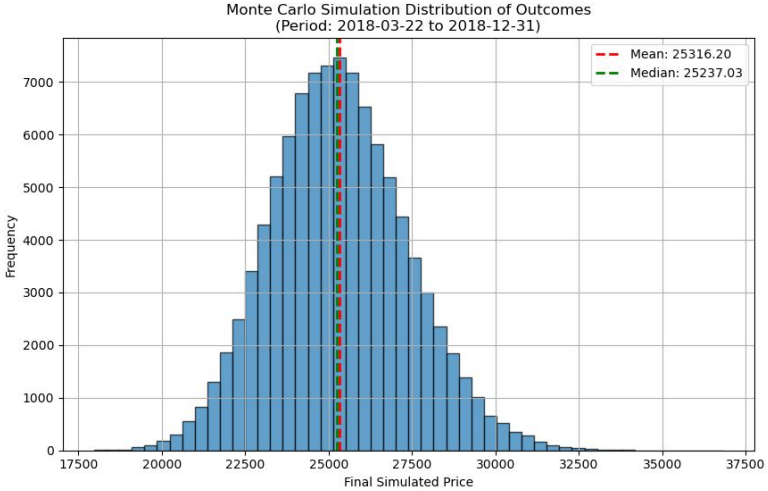



```
Monte Carlo Simulation Summary (2007-12-01 to 2009-06-30):
Mean final price: 2789.18
Median final price: 2735.98
Std Dev of final price: 556.02

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.10%
Median of daily returns (averaged over 60 days): -0.00%
Average daily returns standard deviation: 2.55%
```

Predictive Analysis (cont.)

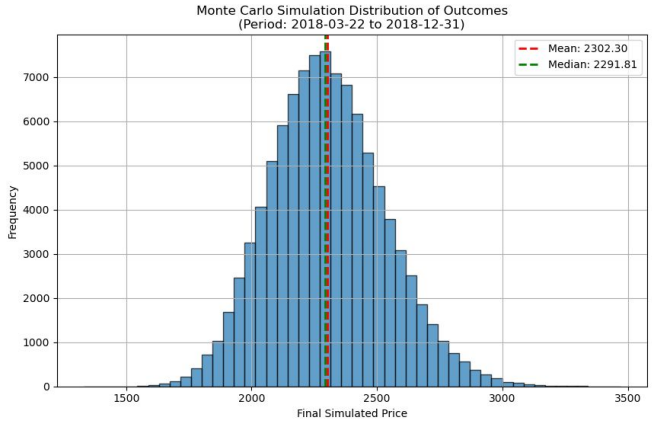
Monte Carlo Simulation (Wilshire 5000) – Trade War scenario:



Monte Carlo Simulation Summary (2018-03-22 to 2018-12-31):
Mean final price: 25316.20
Median final price: 25237.03
Std Dev of final price: 2053.18

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.03%
Median of daily returns (averaged over 60 days): 0.04%
Average daily returns standard deviation: 1.04%

Monte Carlo Simulation (Shanghai Composite) – Trade War scenario:

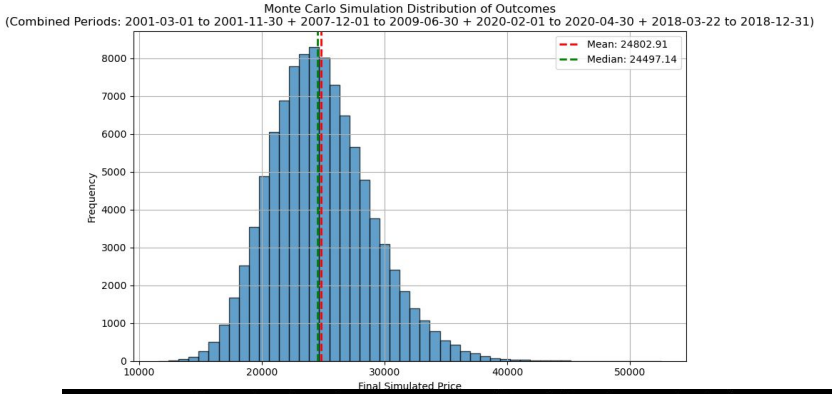


Monte Carlo Simulation Summary (2018-03-22 to 2018-12-31):
Mean final price: 2302.30
Median final price: 2291.81
Std Dev of final price: 229.95

Daily Returns Summary Across All Simulations:
Mean of daily returns (averaged over 60 days): -0.13%
Median of daily returns (averaged over 60 days): -0.18%
Average daily returns standard deviation: 1.28%

Predictive Analysis (cont.)

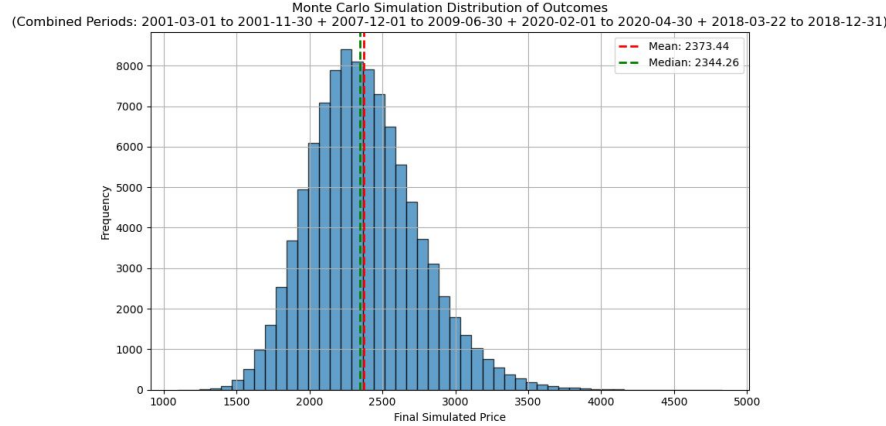
Monte Carlo Simulation (Wilshire 5000)




```
Monte Carlo Simulation Summary (Combined Periods):
Mean final price: 24802.91
Median final price: 24497.14
Std Dev of final price: 4101.76

Daily Returns Summary Across All Simulations (Combined):
Mean of daily returns (averaged over 60 days): -0.06%
Median of daily returns (averaged over 60 days): 0.01%
Average daily returns standard deviation: 2.13%
```

Monte Carlo Simulation (Shanghai Composite)




```
Monte Carlo Simulation Summary (Combined Periods):
Mean final price: 2373.44
Median final price: 2344.26
Std Dev of final price: 372.38

Daily Returns Summary Across All Simulations (Combined):
Mean of daily returns (averaged over 60 days): -0.08%
Median of daily returns (averaged over 60 days): -0.05%
Average daily returns standard deviation: 2.01%
```

THANK YOU

Any Questions?