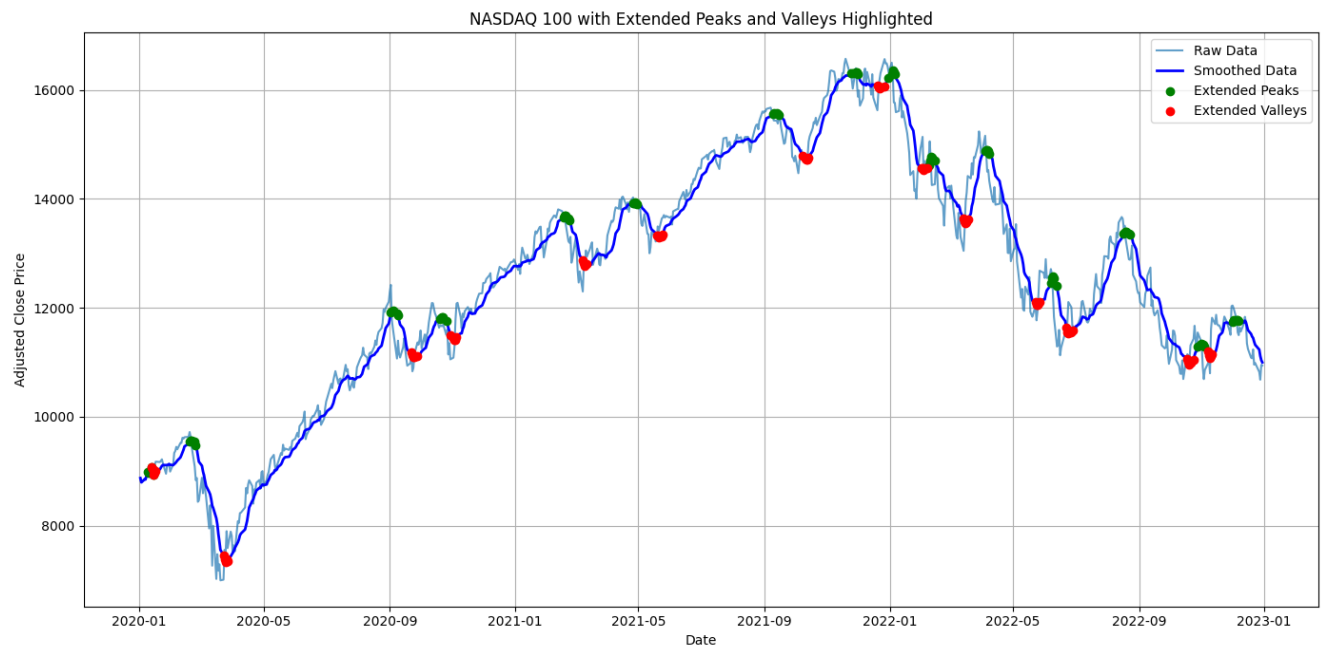


# NASDAQ 100 Trend Classification: Machine Learning's Approach to Identifying Major Turning Points

The objective of this case study is to design and implement a predictive classification model that uses broad economic and market indicators to forecast significant peaks and valleys in the NASDAQ 100 index from 2007 to 2023.



## Context

The stock market is influenced by a myriad of factors, ranging from macroeconomic indicators to investor sentiment. Identifying peaks and valleys, or turning points, in major stock market indices can provide valuable insights for investors and policy makers. This project seeks to harness the predictive power of broad economic and leading market indicators to forecast these turning points in the NASDAQ 100.

## Criteria for Success

- Accurately classify data points into peaks (1), valleys (3), or neither (2) with a high level of accuracy, precision, and recall.
- Develop a model that can predict these classifications using the provided leading market indicators with a significant level of confidence.
- Offer insights into which leading indicators are most influential in predicting the NASDAQ 100's peaks and valleys.

## Scope of Solution Space

The model will use the smoothed NASDAQ 100 data to train and test its predictions. The scope will include preprocessing of data, feature engineering from leading market indicators, model selection, training, testing, and evaluation.

- **Model Development:** Our solution involves developing a classification model, likely employing algorithms such as decision trees, random forests, or gradient-boosted trees.
- **Feature Importance:** Post-model development, we will analyze the importance of each leading market indicator in predicting the peaks and valleys.

## Constraints

- **Data Availability:** Continuous and consistent data for all selected market indicators within the defined time frame.
- **Data Quality:** The accuracy and reliability of external data sources can impact predictions.
- **Computation:** Depending on the complexity of the model, computational power could be a limiting factor.
- **Overfitting:** A model that's too complex might overfit to the training data and perform poorly on unseen data.

## Stakeholders

- **Investors:** Individuals or entities interested in the performance of the NASDAQ 100.
- **Financial Analysts:** Professionals tasked with analyzing stock market trends and providing recommendations.
- **Policy Makers:** Officials who might use this model to understand the health of the stock market and possibly inform economic policies.

## Data Sources

*NASDAQ 100 data:* Obtained from Yahoo Finance using the yfinance library.

*Leading Market Indicators:* Data on DOW transport index, CPI and core inflation, VIX index, S&P price to earnings ratio, the 10-year and 2-year treasury yield, oil prices, interest rates, and unemployment would be sourced from financial databases and platforms like the Federal Reserve Economic Data (FRED) platform.