**STOCK PRICE PREDICTION**

**Supraja Ravipati**

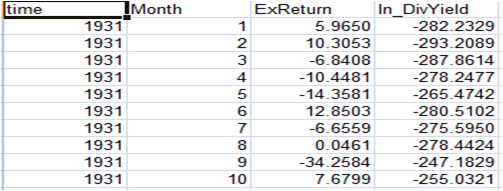
**Raamkumar Thiyagrajan**

**Sanjay Jain**

**Shashank Goswami**

**Introduction:**

Our data has monthly data of stock prices from year 1931 to 2002. The size of the dataframe is 864x4. We are going to make predictions on Excess returns with Dividend yield as additional variable.



**Excess Returns:** Excess return, also known as alpha, is a measure of how much a fund has under or outperformed the benchmark against which it is compared. This important financial return metric allows investors to compare sets of funds against each other, in order to see which fund has generated greater excess returns. Excess return can be positive (denoting outperformance relative to the benchmark) or negative (indicating underperformance). It is a measure of the portion of a fund’s return which is not explained by overall market returns. As such, an excess return analysis can help determine whether outperformance is the result of a portfolio manager’s skill, or simply the result of movements in stock markets. Similarly, for alternative indices, excess return can gauge the quality of the index’s strategy and underlying stock-selection rules.

**In a nutshell: Greater Excess Returns -> Better the stock to invest**

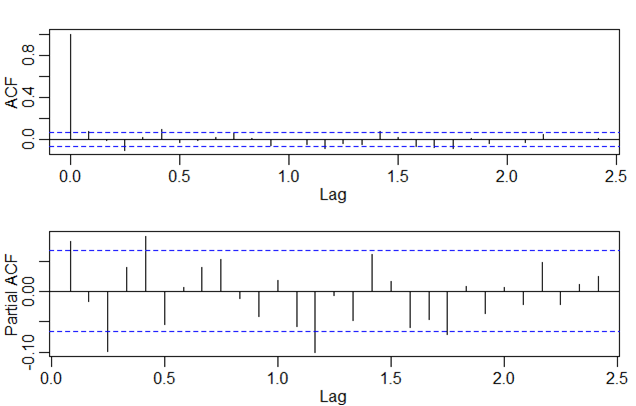
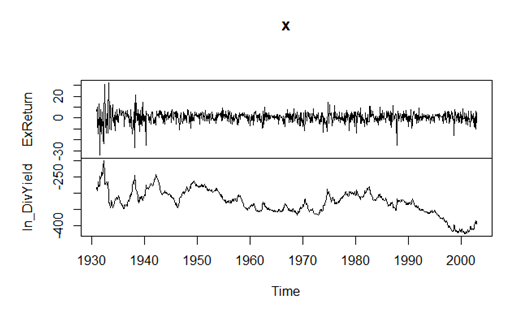
Hence, predicting how excess returns is performing, in a way helps us evaluates the performance of the stock.

**Dividend yield:** Dividend yield is a method used to measure the amount of cash flow you're getting back for each dollar you invest in an equity position.

Suppose company ABC's stock is trading at $20 and pays yearly dividends of $1 per share to its shareholders. Also, suppose that company XYZ’s stock is trading at $40 and also pays annual dividends of $1 per share. Company ABC’s dividend yield is 5% (1 ÷ 20), while XYZ’s dividend yield is only 2.5% (1 ÷ 40). Assuming all other factors are equivalent, an investor looking to use their portfolio to supplement their income would likely prefer ABC's stock over that of XYZ, as it has double the dividend yield.

**Dividend Yield is an additional information that helps us in predicting the excess returns**

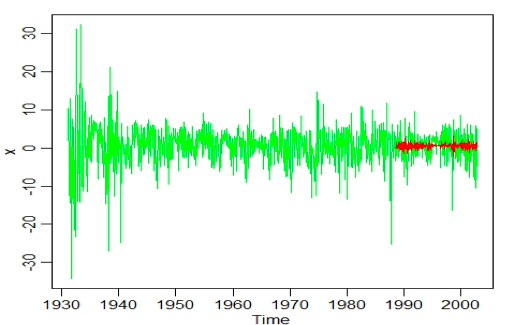
**Data Preprocessing:**

Converted the data set into a time series data. Performed log transformation and difference in the ExReturns and looked at the acf and pacf plot. But the acf and pacf didn’t seem to improve even after transformation. So, the model fit was carried out just with the feature ExReturns over time. Fitted a SARIMA model and observed the standardized residuals plot and accordingly fine-tuned for our model for best fit and predictions

**Analyses and Results:**

**SARIMA without additional variables**

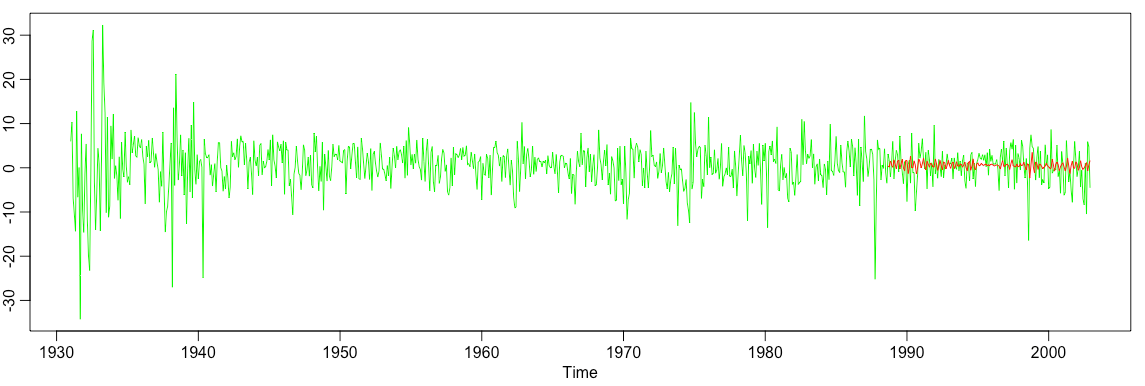
SARIMA models are very much suitable for time series analysis. Our stock price data gave really good results with SARIMA for ExReturn prediction. As a first step towards our SARIMA model, we performed analysis on the stock data by plotting acf and pacf of ExReturn over time. Initially we hypothesized AR(2) model for our data. For prediction, we split our data into 80% for training the model and 20% for testing the model. We ran the model with AR(2), and performed one step ahead prediction. Based on the MSPE we fine-tuned our model and arrived at SARIMA(2,0,2)x(0,0,0) model. This model gave us really good MSPE score.



“The MSPE is” 1.531.

**SARIMA with additional variables**

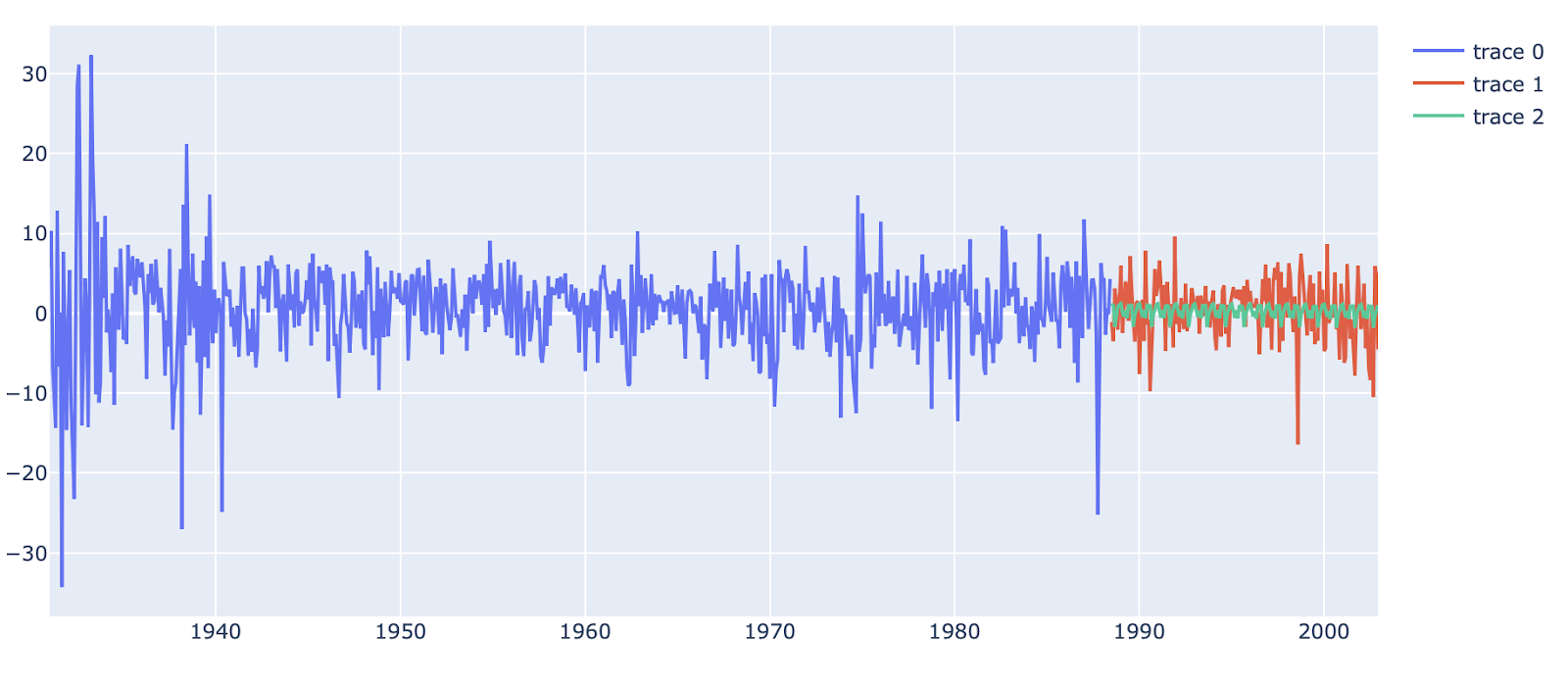
After performing univariate analysis with SARIMA, we proceeded with SARIMA models with additional variables, Dividend yield was taken as an additional variable. To proceed with the analysis, ACF, PACF plots of dividend yield were plotted, which displayed clear trends in dividend yield. The new variable was inculcated in the SARIMA model. 80-20 train test split was done on the additional variable and passed into the required arguments of SARIMA model. Model AR(2) was initially picked, however for choosing the best model auto arima function was used which gave SARIMA(2,0,2)x(0,0,0) as the best model, with low MSPE value.



“The MSPE is” 1.407.

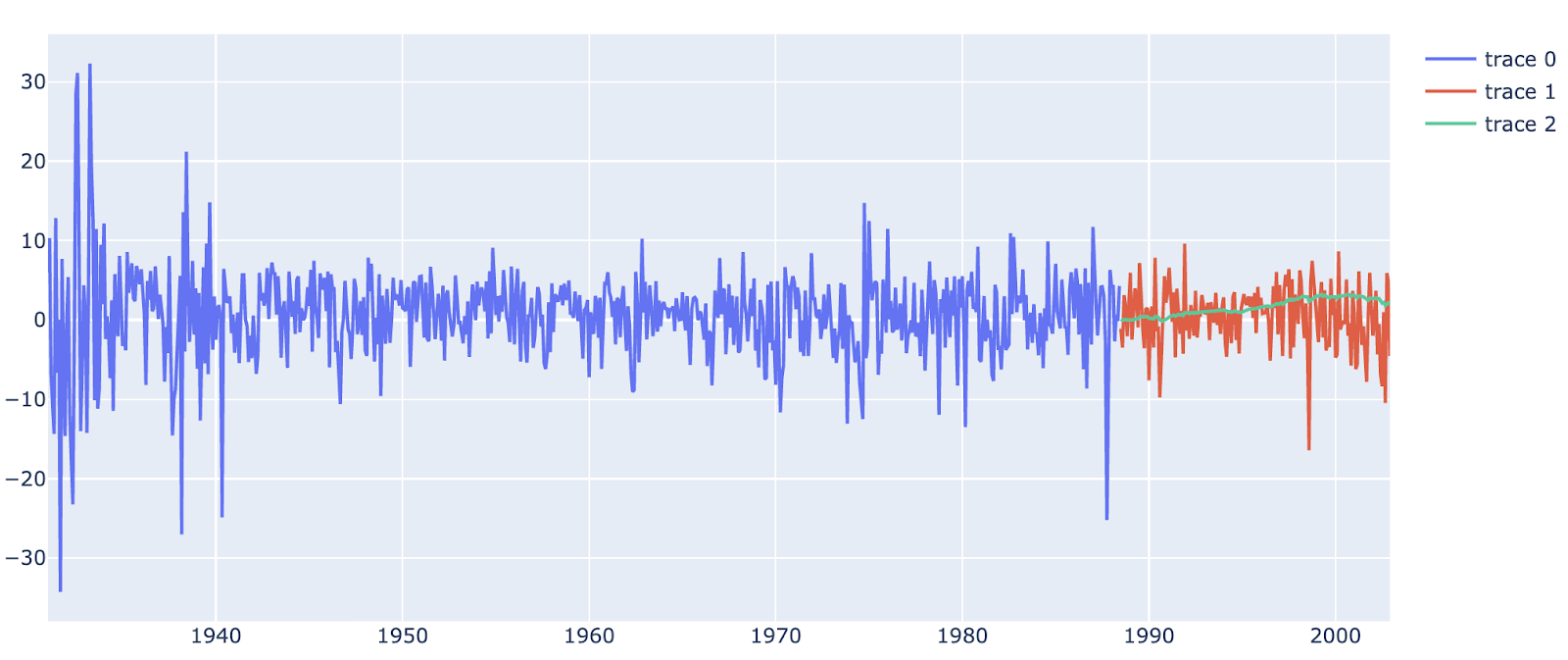
**Prophet**

Prophet is a forecasting model developed by Facebook. It is based on additive model approach where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. Prophet accurately captures trend and seasonality in the data. Prophet is flexible enough to make predictions with and without additional variables. Therefore, predictions were made for Excess return standalone and again with dividend yield as an additional variable. The predictions made without additional variable surpassed SARIMA model predictions, while in case of prediction with additional variable it lagged behind SARIMA model.



**Prediction without additional variable.**

“The MSPE is” 1.392.



**Prediction with additional variable**

“The MSPE is” 3.963.

**Discussion:**

Can make better predictions/estimations by

1. Adding more additional variables
2. Subsetting the data; taking relatively recent years data as the trends/seasonality a 100 years ago would not be as relevant to the stock price in 2002.
3. By using advanced deep learning techniques such as LSTM

**Responsibilities:**

Ramkumar, Supraja worked together mainly on Data Preprocessing, simple SARIMA model together. Shashank mainly focused on modelling SARIMA with additional variables. Sanjay focused working on Prophet.

**References:** [**https://www.quantilia.com/excess-returns/**](https://www.quantilia.com/excess-returns/) **,** [**https://www.thebalance.com/understanding-dividend-yield-3140782**](https://www.thebalance.com/understanding-dividend-yield-3140782)