## **Introduction**

In the world of business, nearly all executives and managers regularly want to know what’s going to happen next in their business. A quick search for “business trends” on Google yields over 3 billion results, the top results being list-style articles of trends for different sectors. With this interest in the future, being able to use analysis tools to forecast is a key skill to develop.

A forecast can be defined as “look[ing] at how hidden currents in the present signal possible changes in direction for companies, societies, or the world at large” [1]. This must be done in a logical, explainable manner, in order to engage the critical thinking skills of the consumer of the forecast, who will make a decision based on the forecast information.

Forecasting is not just for the business realm - it is regularly used by individuals in their personal lives when they make decisions about anything that persists into the future. Actions such as purchasing a home, starting a small business, and accepting a job offer all involve some level of forecasting, whether of not that forecast is analytically rigorous. For this case study, we will demonstrate how time series analysis can be used to apply mathematical rigor to a question of personal interest: the stock price of an investment.

**Background**

Time series data are data collected over time, often in equal time increments. The data only makes sense in the order it is collected, referred to as serial correlation. This means it cannot be randomly sampled like other data types because it is not independent of the other observations in the dataset. Specific types of analysis are required to manage these unique characteristics.

In order to use these analysis tools, several things must be true about the time series. First, it must be stationary. A stationary time series has a mean, variance and covariance that do not change with time. The main ways to confirm that these conditions are true are by visual inspection and by using the Dickey-Fuller Test, a statistical test where if the null hypothesis is rejected, the time series is stationary. When a dataset does not meet these criteria, transformations must be done to make the time series stationary [2].

In order to stationarize a time series, there are several tactics. If the variance is not consistent over the time series, the data can be transformed to bring the data into comparable scale using techniques such as log transformations. If the mean is not consistent, indicating a trend, detrending may be useful in order to bring the time series to stationary. This typically involves subtracting out a model of the trend. The model of the trend can be generated taking averages over aggregated time periods, taking smoothed rolling averages, or by fitting a regression or polynomial model [3].

When a time series shows a repeatable pattern that follows the calendar, this is called seasonality. This also needs to be removed from the time series in order to make it stationary. One way of managing seasonality, particularly if there is also trending present, is to use differencing. Differencing looks at the difference between subsequent observations in the data, rather than the actual observed values. There are also technical techniques to decompose a time series into its trend, season and residual components, and the residual component is used, if it is stationary [3].

Once a time series has been stationarized, time series analysis can be conducted. The method used in this case study is called ARIMA, which stands for AutoRegressive Integrated Moving Averages. This method brings together two main components: auto-regression and moving averages. Auto-regression is the dependence of the next value on previous values in the time series. Different degrees of auto-regression indicate how far back each value looks for its dependence [2]. The moving average component of the model represents the dependency between the next value and the residual errors of the moving average of the previous terms in the sequence. The Integrated component of the ARIMA model is a built-in differencing that aims to complete the stationarization step as part of the overall model. In cases where data is seasonal, this integrated differencing may not be enough, and manual differencing and transformation is necessary [4].

The ARIMA model is tuned using three parameters, one corresponding to each component. The number of terms to look back (also known as the order of lags) is notated by *p.* The number of times the observed data are differenced (also known as the degree of differencing) is the term *d.* Finally, the number of terms to look back for the moving average calculations (also known as the order of the moving average) is denoted *q*. Ideal values of *p* and *q* can be determined by looking at the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots [3]. After a time series has been differenced, the point where the functions in the plots first cross the upper confidence interval is the ideal number of lags for *p* in the PACF plot, and *q* in the ACF plot[4].

Once an ARIMA model is tuned, in order to use it to make a forecast, it must be taken back to its original format. This means adding reversing any transformations or adding back in any trend or seasonality components that were removed in the process of stationarizing the time series [3].

## **Method**

## **Results**

## **Conclusions**

## **Future Work**

## 

## **References**

1. P. Saffo, "Six Rules for Effective Forecasting", Harvard Business Review, 2007. [Online]. Available: https://hbr.org/2007/07/six-rules-for-effective-forecasting.
2. T. Srivastava, "A Complete Tutorial on Time Series Analysis and Modelling in R", Analytics Vidhya, 2015. [Online]. Available: https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/.
3. [3]A. Jain, "Complete guide to Time Series Forecasting (with Codes in Python)", Analytics Vidhya, 2016. [Online]. Available: https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/.
4. [4]J. Chen, "Autoregressive Integrated Moving Average (ARIMA)", Investopedia, 2019. [Online]. Available: https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp. [Accessed: 26- Jun- 2019]

For coding and how to use the ARIMA in Python:

<https://www.analyticsvidhya.com/blog/2016/02/time-series-forecasting-codes-python/>

For how to interpret and understand what’s going on in an ARIMA:

<https://www.analyticsvidhya.com/blog/2015/12/complete-tutorial-time-series-modeling/>

## 

## **Appendix - Code**