

**Project 3**

**Forecasting Financial Time Series**

Sankalp Biswal

College of Professional Studies, Northeastern University Boston

ALY6050: Introduction to Enterprise Analytics

Professor Roy Wada

March 12, 2024

**Introduction**

This report examines time series forecasting for Apple Inc. (AAPL), Honeywell (HON), and dry wine prices using exponential smoothing and auto.arima(). Through statistical tests and evaluation of forecasting errors, we aim to assess the effectiveness of these methods.

**Part - 1**

1. **Simple line plot for Apple(AAPL) and Honeywell(HON) stock prices in USD**

*Fig. 1 - Line plot for AAPL and HON*

**Interpretation for Fig. 1**

* **Apple Inc. (AAPL):** The orange line representing Apple Inc. shows an overall upward trend over the time period shown. The stock price rises significantly from around $40 in early 2019 to nearly $200 by the end of the period shown. The line does show some fluctuations, which could be considered short-term irregular behavior or volatility. However, from this plot, there does not seem to be a clear seasonal pattern, as the fluctuations do not appear at regular intervals.
* **Honeywell Inc. (HON):** The blue line for Honeywell Inc. also shows an upward trend, although it is less pronounced than that of Apple Inc. The stock price starts at around $140 and ends at slightly above $160. Similar to Apple, there is volatility throughout the period, but no discernible seasonality.

1. **Performed exponential smoothing to forecast 253rd period of both AAPL and HON.**

*Table 1 – Mean absolute deviation and Mean absolute percentage deviation for different value of alpha for Apple and Honeywell stock prices*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | AAPL (Apple Inc.) | | | | HON (Honeywell) | | | |
| Alpha (α) | 0.15 | 0.35 | 0.55 | 0.75 | 0.15 | 0.35 | 0.55 | 0.75 |
| MAD | 3.82 | 2.60 | 2.25 | 2.16 | 5.15 | 4.14 | 3.70 | 3.50 |
| MAPE | 3.87% | 2.47% | 2.08% | 1.97% | 3.06% | 2.35% | 2.03% | 1.90% |

**Interpretation for Table 1**

* The most accurate forecast for Apple stock is achieved with an alpha value of 0.75, having the lowest **MAPE** of 1.97%.
* Similarly, for Honeywell stock, the most accurate forecast is also achieved with an alpha value of 0.75, having the lowest **MAPE** of 1.90%.

**Explanation**

* Alpha (α) is a smoothing factor that controls the rate at which the influence of the observations at prior time steps decays exponentially. A higher alpha discounts older observations faster. The most accurate α values for both stocks are the highest ones (0.75), indicating that giving more weight to the most recent observations has resulted in more accurate forecasts for these particular stocks.

***Q. Why* α *= 0.75 gave the most accurate results?***

* I feel this could be because recent trends and patterns are more indicative of future movements than older data for these stocks. Stock prices are often influenced by recent events, such as economic news, company performance, and market sentiment, which would be captured more effectively by a higher alpha that places more emphasis on recent observations.

1. **Performed adjusted exponential smoothing to forecast 253rd period of both AAPL and HON.**

Table 2 - *Mean absolute percentage deviation for alpha = 0.55 and different values of beta for Apple and Honeywell stock prices*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | AAPL (Apple Inc.) | | | | HON (Honeywell) | | | |
| Beta(β) | 0.15 | 0.25 | 0.45 | 0.85 | 0.15 | 0.25 | 0.45 | 0.85 |
| MAPE | 1.97% | 1.95% | 1.94% | 1.98% | 1.98% | 1.94% | 1.89% | 1.85% |

**Interpretation for Table 2**

* The most accurate forecast for AAPL is achieved with a β value of 0.45, as it has the lowest **MAPE** of 1.94%.
* The most accurate forecast for HON is achieved with a β value of 0.85, with the lowest **MAPE** of 1.85%.

**Explanation:**

* The different optimal values of β for AAPL and HON suggest that the importance of the trend component in their time series forecasts is different. For AAPL, a moderate value of β (0.45) is optimal, suggesting that incorporating a moderate level of the trend from the most recent observations leads to the best forecasting accuracy. This could mean that while the trend is significant for forecasting AAPL's stock, it does not change so rapidly that the most recent trends always dominate.
* For HON, the higher β value (0.85) is optimal, indicating that the recent trend is very significant in predicting future values. A higher β suggests that the stock price of HON is more influenced by the direction in which it has been moving most recently, and thus, recent changes in the trend should be heavily accounted for.

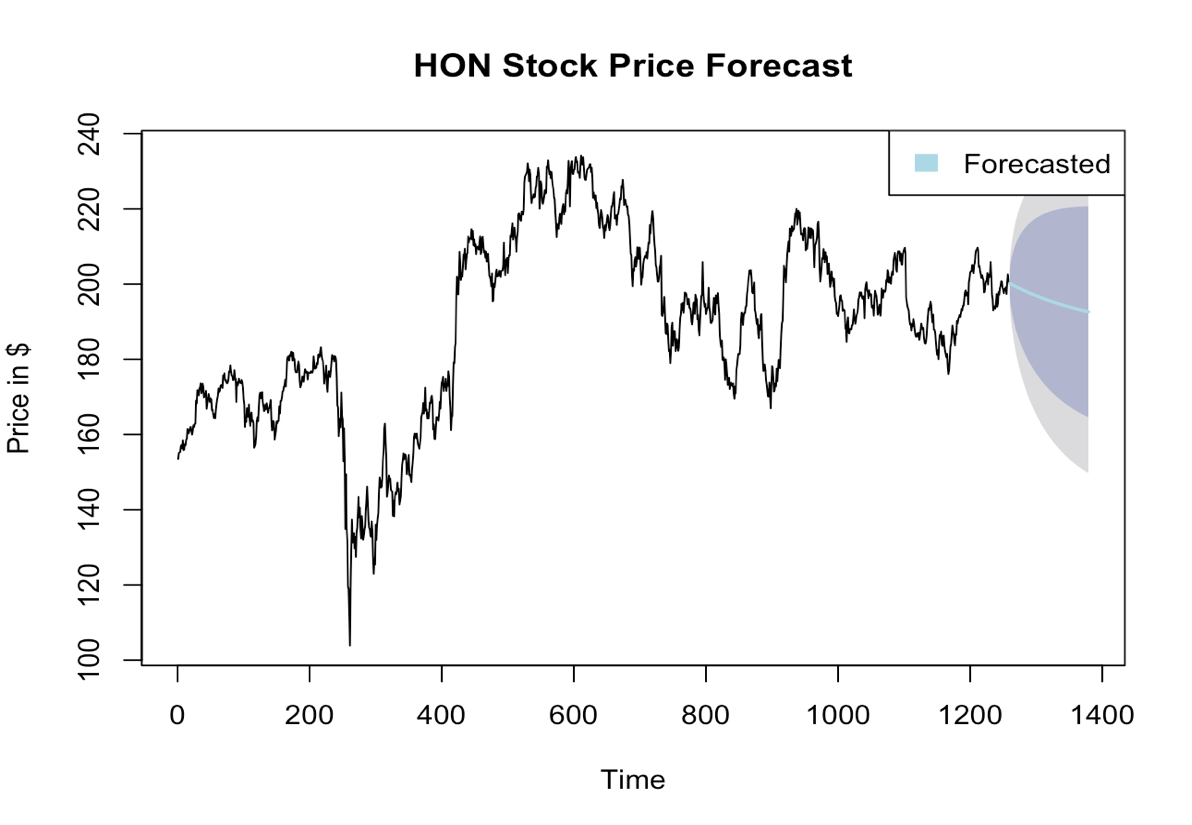
***Q. Why the above β values gave the most accurate results?***

* Stock price movements can be influenced by a lot of factors including company performance, investor sentiment, and market conditions. These factors can change rapidly, making recent trends a strong indicator of future movements, which is possibly why a higher β yields a more accurate forecast for HON. For AAPL, the optimal β suggests that while recent trends are important, there may also be a more stable, long-term trend that should not be disregarded in forecasts.

**Part -2**

1. **Plotted the actual series + forecasts up to 120 months for Honeywell and Apple stock price using ARIMA**

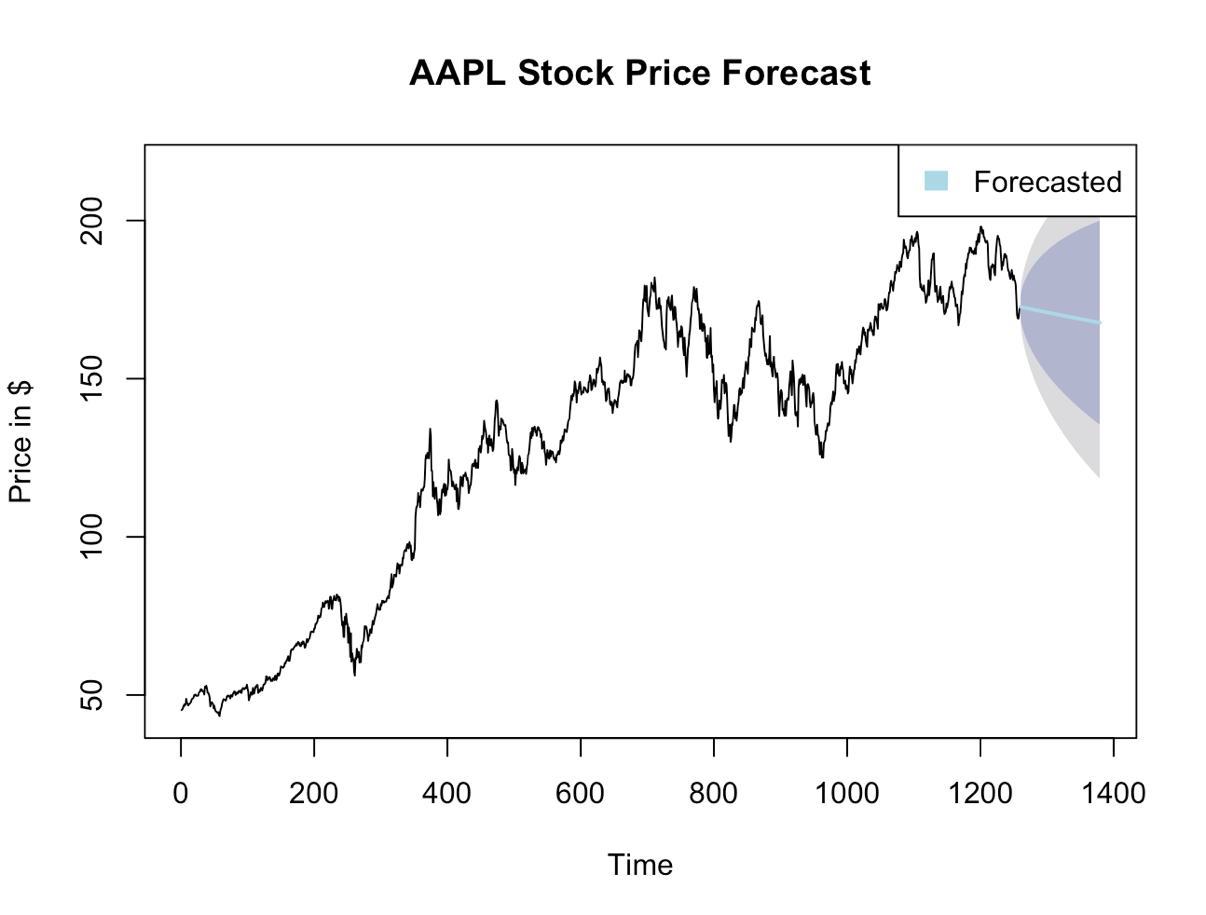
*Fig. 2 – Forecasting for Honeywell stock using ARIMA*



**Interpretation for Fig. 2**

The graph displays the historical and forecasted stock price over time. The black line represents the stock's historical price, showing high volatility with an upward trend. The forecasted period is highlighted by a blue line, starting where the historical data ends, indicating that the price might continue to experience fluctuations. The shaded area suggests a range of possible future prices, capturing the uncertainty of the forecast. As time progresses, this area widens, reflecting increased uncertainty in the longer term.

*Fig. 3 – Forecasting for Apple stock using ARIMA*



**Interpretation for Fig. 3**

The graph illustrates the historical price movement and future forecast of apple stock price. The black line traces the stock's historical price, showcasing an overall increasing trend with periods of volatility. The forecasted price, depicted by a blue line, starts at the conclusion of the historical data, suggesting a potential continuation of previous volatility within a certain range. The shaded area represents the forecast's confidence interval, which grows wider with time, implying greater uncertainty in the forecast the further out we look. There's no strong indication of a decisive trend in the forecasted period.

1. **Testing for stationarity using ADF (Augmented Dickey – Fuller) Test on both the**

**stocks**

**Hypothesis –**

* ***Null Hypothesis (H0***) : Data is not stationary
* ***Alternate Hypothesis (H1)*** : Data is stationary

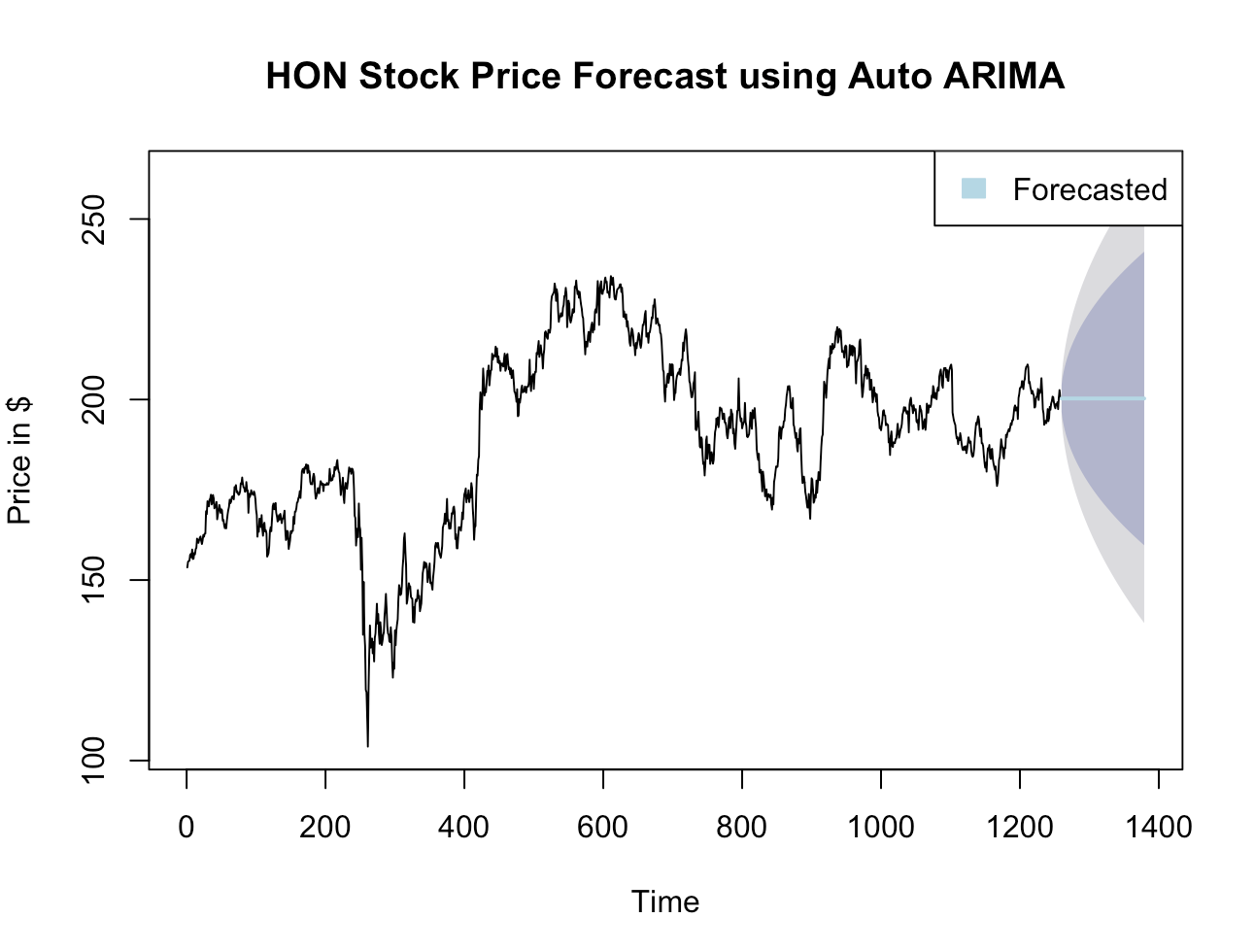
**Interpretation**

1. **Honeywell -** We observe that **p-value = 0.4921** i.e > **0.05** (Significance level). We **fail to** **reject** the Null Hypothesis. Therefore, we don’t have sufficient evidence to say that the time series is stationary
2. **Apple -** We observe that **p-value = 0.3937** i.e > **0.05** (Significance level). We **fail to** **reject** the Null Hypothesis. Therefore, we don’t have sufficient evidence to say that the time series is stationary

Therefore, we conclude that both the time series are **non – stationary**

1. **Plotted the actual series + forecasts up to 120 months for Honeywell and Apple stock price using Auto - ARIMA**

*Fig. 4 – Forecasting for Honeywell stock using Auto-ARIMA*

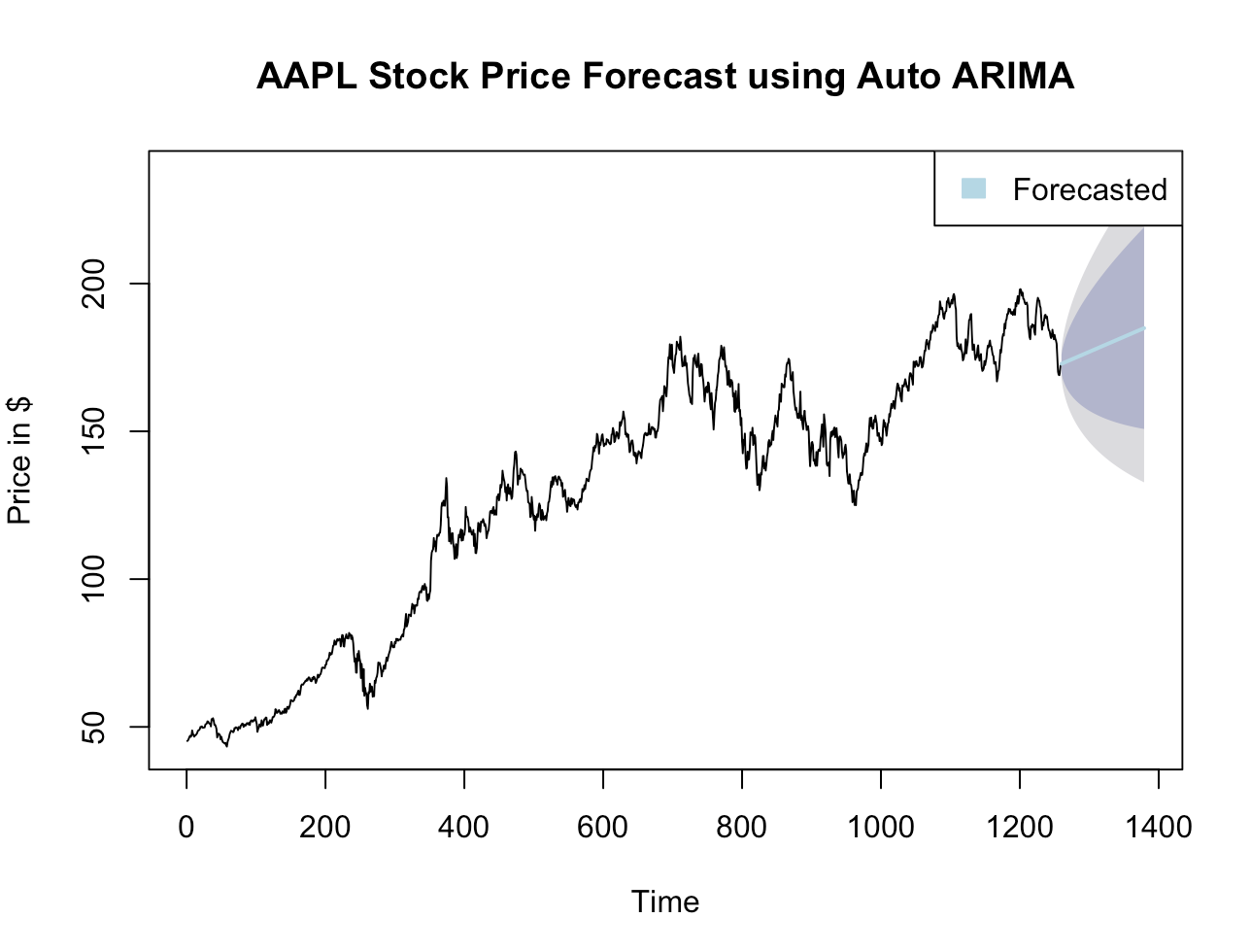
****

**Model used -** ARIMA(0,1,0) or Random Walk

**Interpretation for Fig. 4**

The historical data shows significant volatility with an overall uptrend, followed by stabilization. The forecast, marked by the blue line, begins at the end of the historical data and indicates a slight upward trend within the forecast horizon. The shaded area around the forecast represents the confidence interval, which increases over time, suggesting growing uncertainty as the forecast extends further into the future. The model seems to predict that the stock price will likely continue to fluctuate while maintaining the level reached at the end of the historical period.

*Fig. 5 – Forecasting for Apple stock using Auto-ARIMA*

****

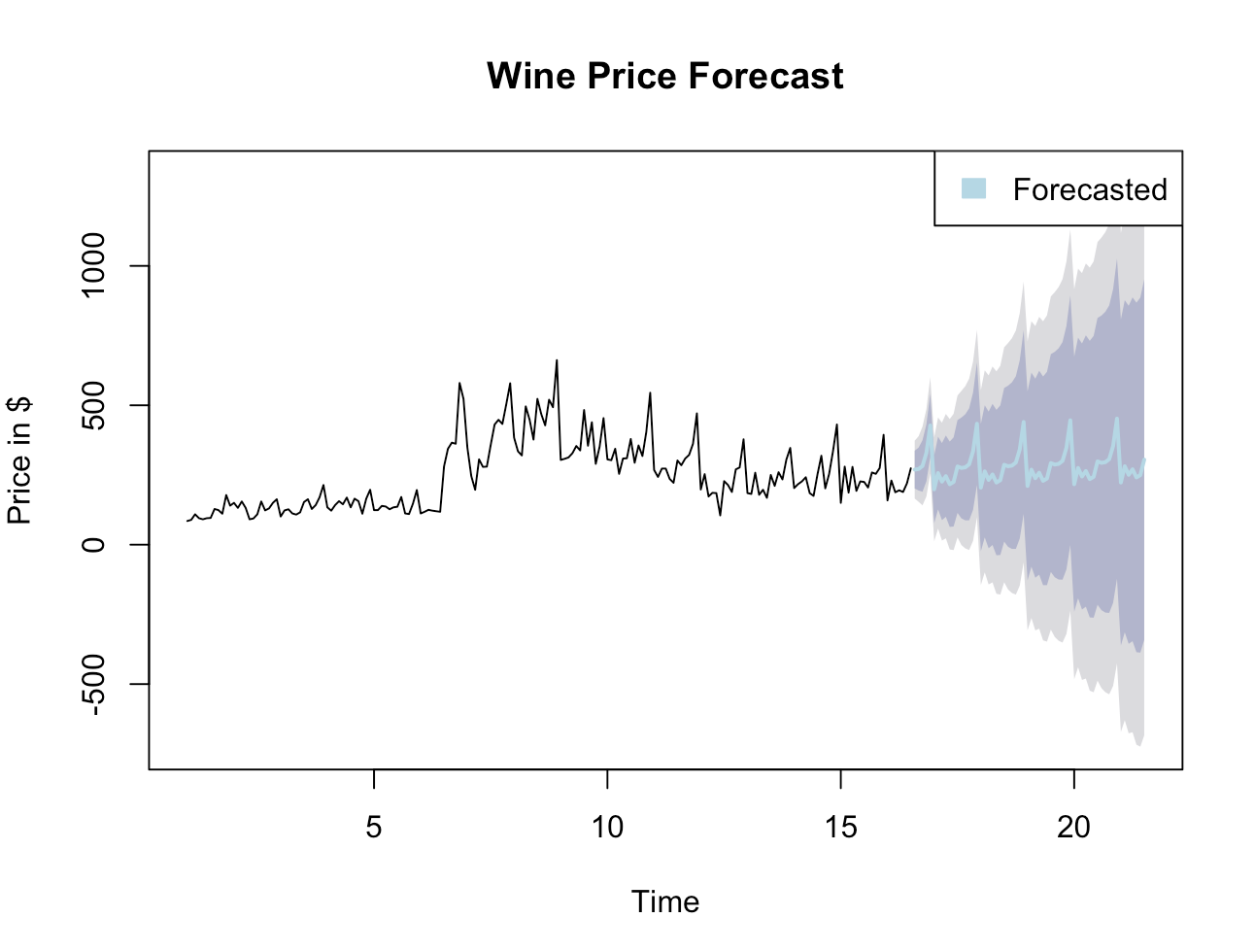
**Model used -** ARIMA(0,1,0) or Random Walk

**Interpretation for Fig. 5**

The historical data, represented by the black line, exhibits a strong upward trend followed by more recent fluctuations. The blue line indicates the projected future price, suggesting a trend continuation with possible slight upward momentum. The forecast's confidence interval, depicted by the shaded area, widens over time, which signifies increasing uncertainty in the price prediction as it moves further into the future.

1. **Plotted the actual series + forecasts up to 120 months for Wine price using Auto - ARIMA**

*Fig. 6 – Forecasting for Wine price using Auto-ARIMA*

****

**Model used -** ARIMA(0,1,1)(2,0,0)[6] or Seasonal Arima

**Interpretation for Fig. 6**

The graph provides a forecast for wine prices over a specified time frame. The historical price data shows variability without a clear long-term trend. The forecast, indicated by a blue line, suggests that prices will continue to fluctuate, but with a potential downward bias as indicated by the central tendency of the forecasted line. The confidence intervals, shown as progressively darker shades of grey, expand significantly as time goes on, which indicates a high degree of uncertainty in the forecasted prices. The negative values in the forecast could suggest a potential issue with the underlying forecasting model, as negative prices are not feasible in the context of wine pricing.

***Q. Why do you get different conclusions? What is the preferred method? What did you  
learn? How would you improve on these analysis?***

**Differences in Conclusions:**

* The stock price forecasts result in a smoother prediction path with less volatility in the forecasted values. This could be due to the inherent properties of the financial markets where past prices are a strong indicator of future prices, despite the random day-to-day fluctuations.
* The wine price forecast results in a more volatile and uncertain prediction path. This might be because the wine market could be influenced by more complex seasonal patterns and external factors that are not captured as smoothly by the model.

**Preferred Method:**

* For HON and AAPL, a non-seasonal auto.arima() model seems suitable as it appropriately captures the random walk nature of stock prices without the need for complex seasonal components.
* For the wine prices, a seasonal ARIMA model is preferred if the data indeed shows strong seasonality. However, the appearance of negative values and high forecast uncertainty suggests that either the model is not well-fitted or that the data contains anomalies that need to be addressed.

**Learnings:**

* Model selection should be considered based on the data's characteristics. Auto.arima() can effectively select an appropriate model for stock prices, but seasonal effects in other markets like wine require additional attention.
* The quality of the input data is critical, as anomalies or non-standard data patterns can lead to unreliable forecasts.

**Improvements:**

* Further diagnostic checks should be conducted on the wine price data to understand the negative forecast values and high uncertainty. Transformation or anomaly detection might be required.
* Additional data preprocessing steps such as de-trending or de-seasonalizing could be performed before fitting the models.
* For stock prices, considering a model that accounts for potential structural breaks could provide more realistic forecasts.

**Conclusion:**

The analysis reveals that auto.arima() effectively models the random walk nature of AAPL and HON stock prices, while the dry wine prices require a seasonal ARIMA approach. The results highlight the importance of recent trends in forecasting stock prices and suggest that further model adjustments are needed for the wine prices due to their volatile nature. This study underscores the need for careful model selection based on data characteristics to improve forecast accuracy.

**References:**

1. Prabhakaran, S. (2022, April 4). *Augmented Dickey Fuller Test (ADF Test) &#8211; Must Read Guide*. Machine Learning Plus. <https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/>
2. May, L. (2022, January 5). *Comparing Holt-Winters exponential smoothing and ARIMA models for time series analysis*. Medium. <https://medium.com/@lawrence.may/comparing-holt-winters-exponential-smoothing-and-arima-models-for-time-series-analysis-659d6f7738c1>