**Project Overview**

This project aims to analyze the viewership data for the Super Bowl, with a focus on understanding the potential impact of Taylor Swift's presence at the 2024 Super Bowl, where her boyfriend Travis Kelce, a player for the Kansas City Chiefs, was participating.

**Data Source**

The data used in this analysis is obtained from Wikipedia and includes information such as the year, winner, runner-up, margin of victory, overtime status, broadcast network, ratings, viewership numbers (main and alternative casting), and total viewership.

**Assumptions**

1. Factors like population growth, the number of devices available for viewing, and the specific teams participating in the Super Bowl finals are assumed to have no impact on the viewership numbers for 2024.
2. Taylor Swift's presence at the 2024 Super Bowl is considered the primary factor that could potentially influence viewership numbers.

**Data Cleaning**

Several data cleaning steps were performed to prepare the data for analysis:

1. **Separating Winner and Runner-Up Columns**: The initial data had the winner and runner-up teams combined in a single column. This was split into two separate columns using the "Text to Columns" feature in Excel and the delimiter "--".
2. **Quantifying Margin and Overtime**: Two new columns, "Margin" and "OT" (Overtime), were created. The margin of victory was manually entered, and the overtime status was determined using the formula =IF(ISNUMBER(SEARCH("OT",$D2)), "1", "0"), where $D2 represents the cell containing the original game information.
3. **Cleaning Viewership Data**: The "Main Vwrs" column contained the letter "M" for millions, which was removed using the "Text to Columns" feature. The "Alt Cast" column had a mix of "M" for millions and "K" for thousands, which was handled using the formula =IF(RIGHT($J2,1)="M", VALUE(LEFT($J2, LEN($J2)-1)), IF(RIGHT($J2,1)="K", VALUE(LEFT($J2, LEN($J2)-1))/1000, VALUE($J2)/1000000)), where $J2 represents the cell containing the original alternative casting viewership data.

**Analysis**

Two forecasting methods, ARIMA (Autoregressive Integrated Moving Average) and Holt's method, were used to predict the viewership for the 2024 Super Bowl.

1. **ARIMA Method**:
   * Definition: ARIMA is a powerful forecasting model that can handle a wide range of time series patterns, including trends, seasonality, and non-stationarity.
   * Reason for Use: The ARIMA model was chosen because it can capture complex patterns and handle non-stationarity in the data, which is often present in real-world time series.
   * Forecast for 2024: The ARIMA model forecasted a viewership of 123.41 million for the 2024 Super Bowl.
2. **Holt's Method**:
   * Definition: Holt's method is an exponential smoothing technique that can handle trends in the data.
   * Reason for Use: The Holt's method was selected because it is effective for data exhibiting a linear trend, which seems to be the case for the Super Bowl viewership data based on historical trends.
   * Forecast for 2024: The Holt's method forecasted a viewership of 122.91 million for the 2024 Super Bowl.

Both forecasting methods were also implemented in Tableau, and the results aligned with the Python-based forecasts.

**Conclusion**

The forecasted viewership numbers for the 2024 Super Bowl, ranging from 122.91 million to 123.41 million, show a potential spike compared to recent years. This spike could be attributed to the presence of Taylor Swift at the event, as she attended to support her boyfriend Travis Kelce, who played for the Kansas City Chiefs.

While Taylor Swift did not perform during the halftime show, her presence and the associated media attention could have generated increased interest and excitement among her fans, potentially contributing to the anticipated surge in viewership.

It is important to note that the assumptions made in this analysis exclude the impact of factors such as population growth, the number of devices available for viewing, and the specific teams participating in the Super Bowl finals. These assumptions were made to isolate the potential effect of Taylor Swift's presence on the viewership numbers.

However, it is possible that other factors, such as the quality of the game, overall marketing and promotion efforts, and the teams' popularity, could also influence viewership numbers. Further analysis and data collection, including factors like social media engagement, ticket sales, and merchandise sales related to Taylor Swift's attendance, could provide additional insights into the extent of her influence on the Super Bowl viewership numbers.