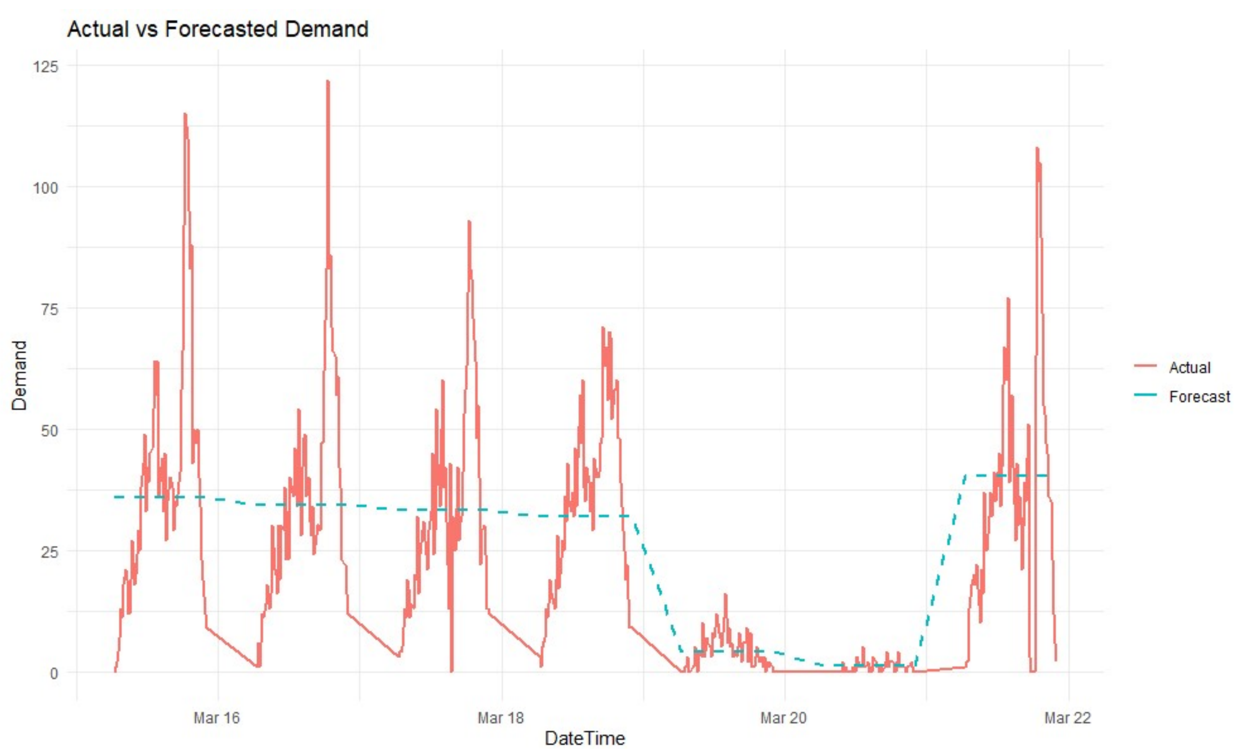
Technical Summary:

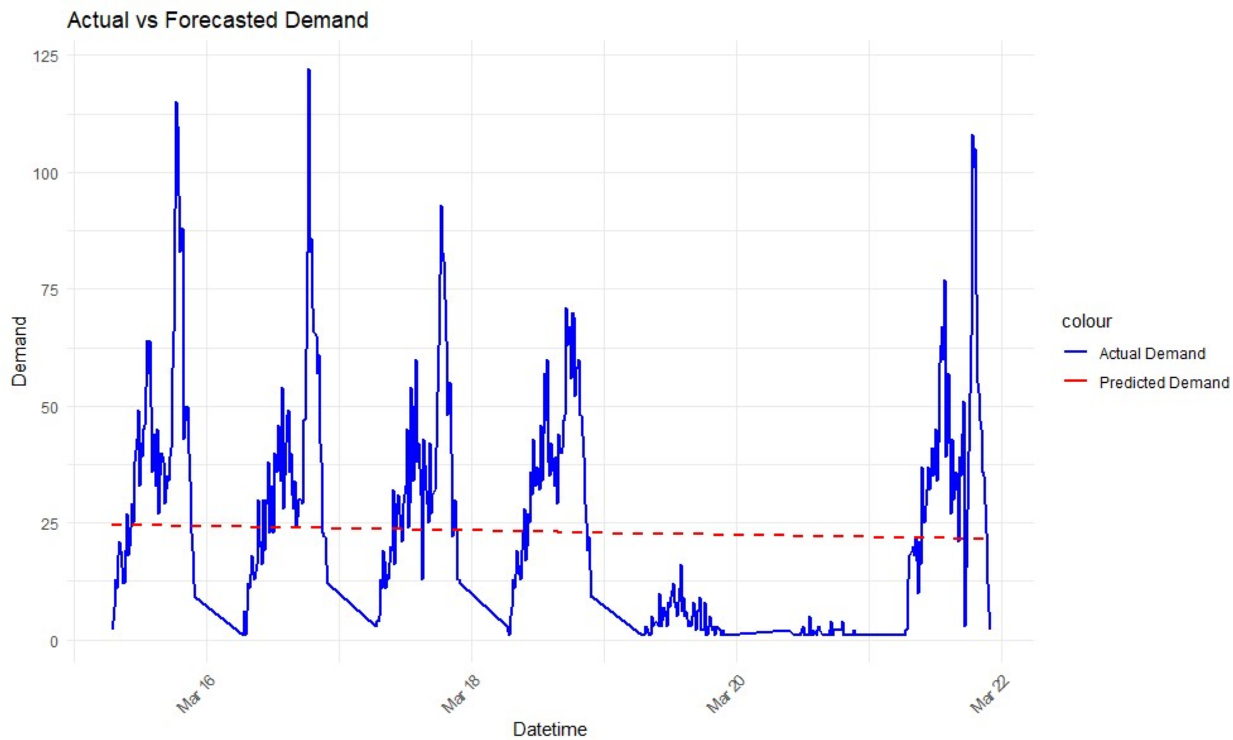
The Naive forecast as a benchmark.

- **Mean Absolute Error (MAE):** The average absolute error between the forecasted demand and the actual demand is approximately 12.66. This value indicates the average magnitude of errors in the forecasts, regardless of direction.
- **Root Mean Squared Error (RMSE):** The square root of the average squared differences between the forecasted demand and the actual demand is approximately 18.99. RMSE gives more weight to larger errors and is useful for understanding the magnitude of error.
- **Filtered Mean Absolute Percentage Error (MAPE):** After filtering out instances where the actual demand is zero to avoid division by zero issues, the MAPE is approximately 127.59%. This high value suggests that the forecast errors are large relative to the actual demand values, indicating significant discrepancies in the forecast accuracy.



Validation data of the Naive Forecast.

Linear Regression Model:



Comparing the performance metrics of the naive forecast and the linear regression model reveals interesting insights:

- **Naive Forecast Metrics:**
  - MAE: 12.6568765072166
  - RMSE: 18.9888449798315
  - Filtered MAPE: 127.594538853236
- **Linear Regression Model Metrics (Filtered for Demand > 0):**
  - MAE: 18.6840522177133
  - RMSE: 24.1018505149405
  - MAPE: 314.343608815863

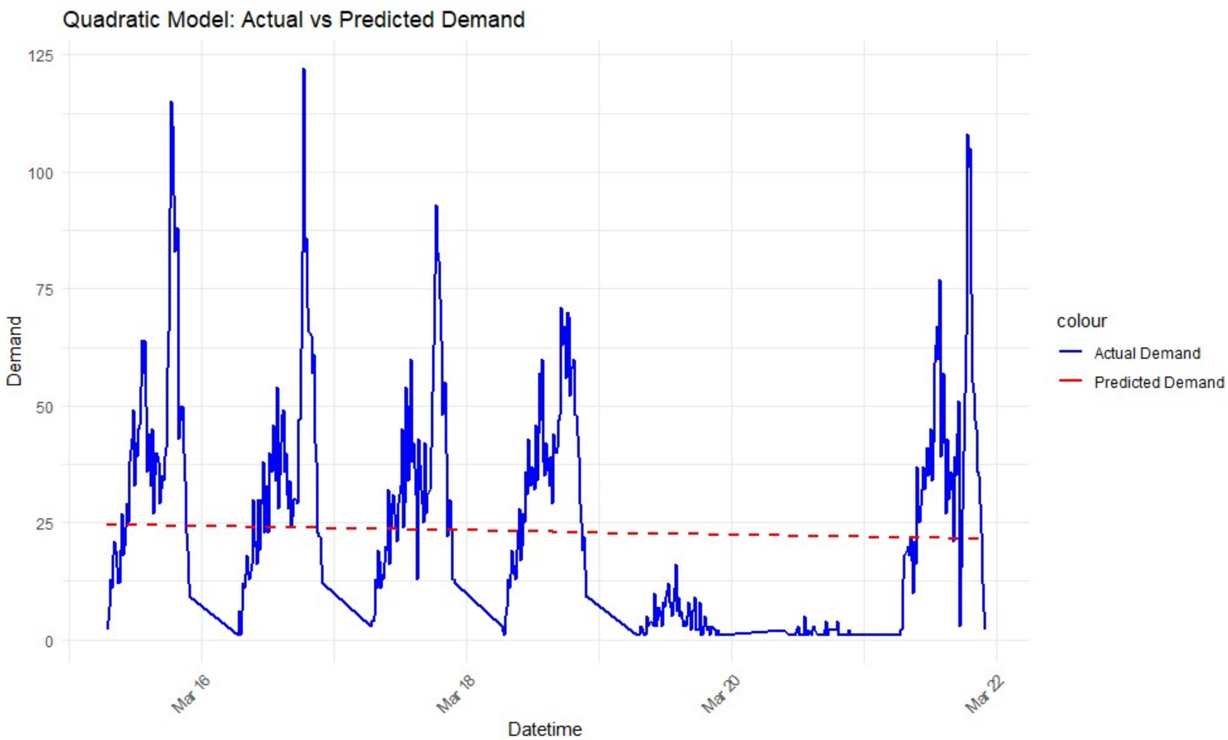
The naive forecast demonstrates lower MAE and RMSE values compared to the linear regression model, suggesting that it was more accurate in this specific case. The naive method is simpler and uses less information (only the previous period's demand) to make forecasts, which might have been sufficiently effective given the dataset's characteristics.

The MAPE values, especially, highlight a significant difference in percentage error between the two methods. The linear regression model's filtered MAPE is substantially higher, indicating that, on average, the linear model's predictions deviate more from the actual values in relative terms than the naive forecast.

Given these results, the naive forecast may be considered more reliable for this specific forecasting task.

The main takeaway is that the linear model does a poor job of capturing the weekend data in comparison to the weekday data. It attempts to model something in the middle without capturing any of the trend or seasonality in the data.

Quadratic model:



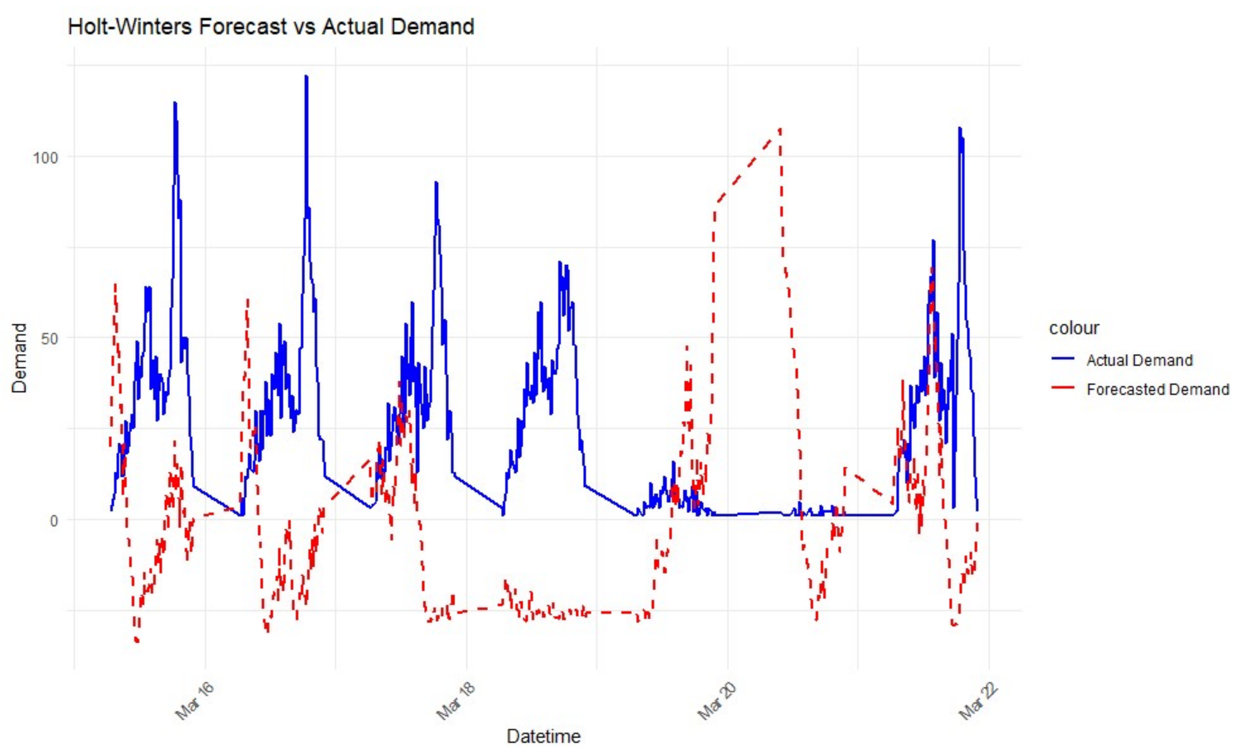
Summary of Error Metrics

- **Naive Forecast**:
  - MAE: 12.66
  - MAPE: 127.59
  - RMSE: 18.99
- **Linear Model**:
  - MAE: 18.68
  - MAPE: 314.34
  - RMSE: 24.10
- **Quadratic Model**:
  - MAE: 18.68
  - MAPE: 314.34
  - RMSE: 24.10

The naive forecast outperforms both the linear and quadratic models in terms of MAE and RMSE. This suggests that seasonally adjusted data of the demand in the Naive model is more accurate than using linear or quadratic regression models to predict future demand for this specific dataset.

The error metrics for the linear and quadratic models are identical, indicating that the quadratic term did not provide additional predictive power. This similarity in performance can be attributed to the quadratic term being undefined (NA) due to singularities. Essentially, the model couldn't find a significant quadratic relationship in the data, leading to the quadratic model collapsing back into a linear model.

Holt Winters Method:



Holt-Winters Model

- **MAE:** 39.73
- **MAPE:** 394.02%
- **RMSE:** 49.05

Seasonal Naive Forecast

- **MAE:** 12.66
- **MAPE:** 127.59%
- **RMSE:** 18.99

Linear Model

- **MAE:** 18.68
- **MAPE:** 314.34%
- **RMSE:** 24.10

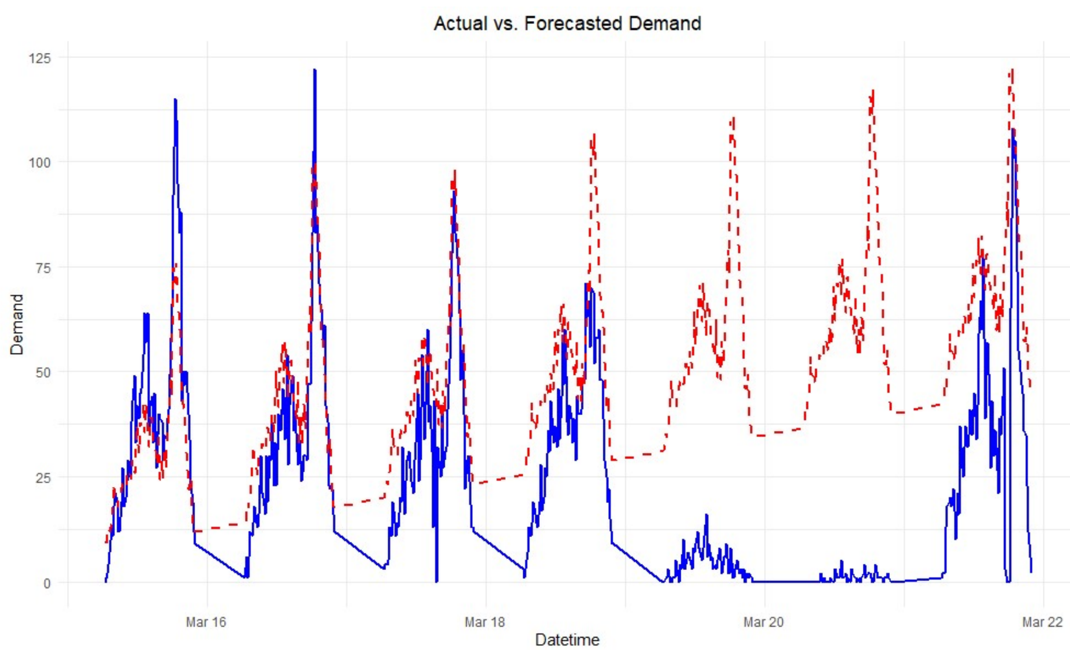
Quadratic Model

(Since the Quadratic model performed similarly to the Linear model due to the singularity issue, its metrics are akin to the Linear model's metrics.)

The Holt-Winters method did not outperform the Seasonal Naive Forecast. Since the data is not perfectly seasonal within a 7 day period and the usage is significantly greater during the week and less during the weekend, the forecast does a poor job in capturing the seasonality. I tried a few different frequencies and still had poor results vs the naive model.

Arima

- **MAE:** 28.92261
- **Adjusted MAPE:** 824.6387%
- **RMSE:** 37.73276
- **MASE:** 4.376186



The Arima model does an excellent job of forecasting the weekdays and not the weekends. The Naive model does an excellent job of forecasting the weekends and not the weekdays. I spent a couple of hours trying to combine the weekdays Arima model with the weekend naive model and was unable to get the code to work. The reason for this is the 'seasonality' of the extreme usage (variability) on weekdays or workdays and the low variability of the weekend days which the naive model captured well due to its ability to forecast recent data well.

Due to the obvious strengths and weaknesses of opposing models, the error metrics do not represent the final model well.