Time Series Forecasting of the NSW Specialised Food Retailing Sector

Zuoming Hu

1 Introduction

This study investigates the monthly turnover of the Other Specialised Food Retailing industry in New South Wales (NSW), using official data sourced from the Australian Bureau of Statistics (ABS). The objective is to identify major structural patterns in retail activity, including the COVID-19 effect, and to develop reliable time series forecasting models to predict future industry performance.

The analysis applies decomposition, transformation, and several competing forecasting frameworks—specifically ETS and ARIMA models—to compare predictive accuracy and robustness.

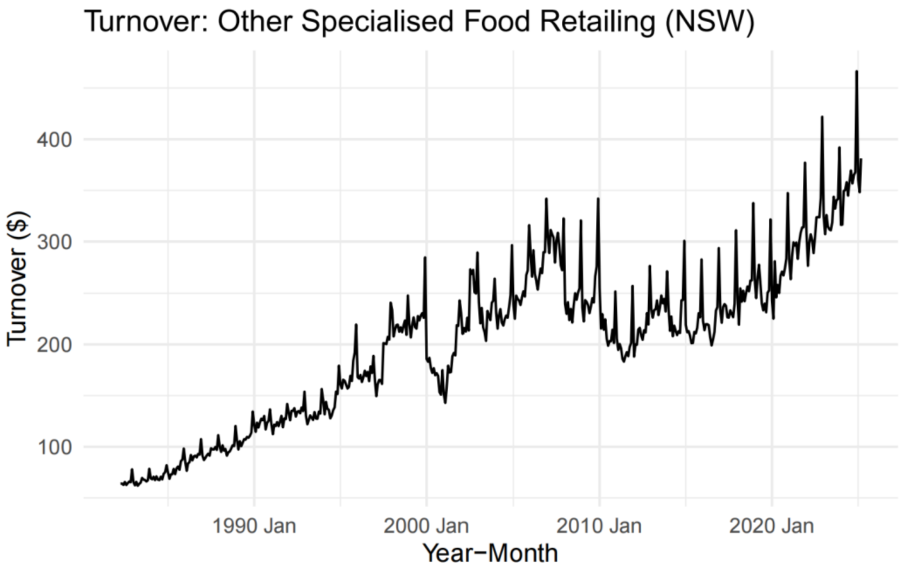

Figure 1: Time Series Plot Showing COVID-19 Dip (March–April 2020)

During the initial exploration, a significant COVID-19 impact was observed. Turnover fell sharply below trend in March–April 2020, coinciding with lockdowns and panic buying, but quickly rebounded later in the year as consumers shifted toward home cooking and specialty food purchases. This structural break highlights the importance of adaptive time series modeling for post-pandemic retail forecasting.

2 Data Description and Preprocessing

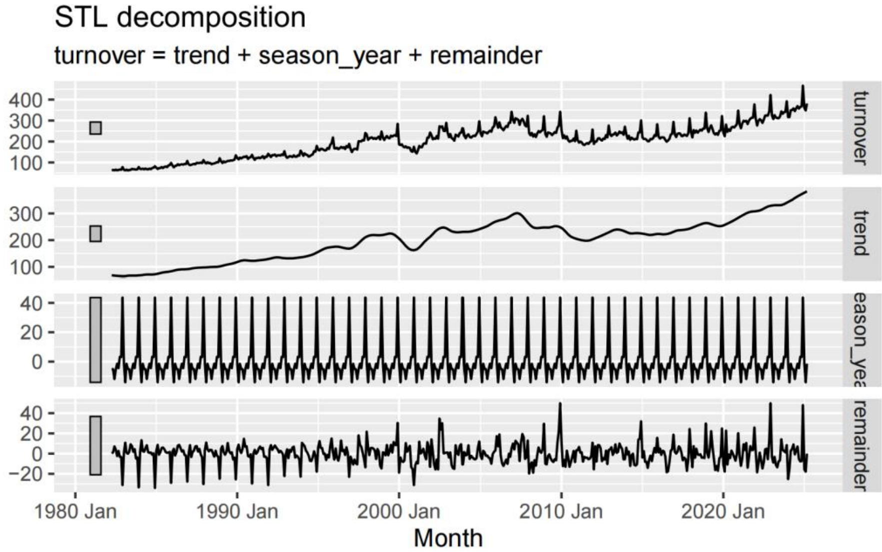

Figure 2: STL decomposition (trend, seasonality, remainder)

The monthly turnover series spans several decades, showing an upward long-term trend, strong seasonal variation, and increasing variance over time. These features indicate non-stationarity and heteroscedasticity, making variance-stabilizing transformations necessary before model fitting.

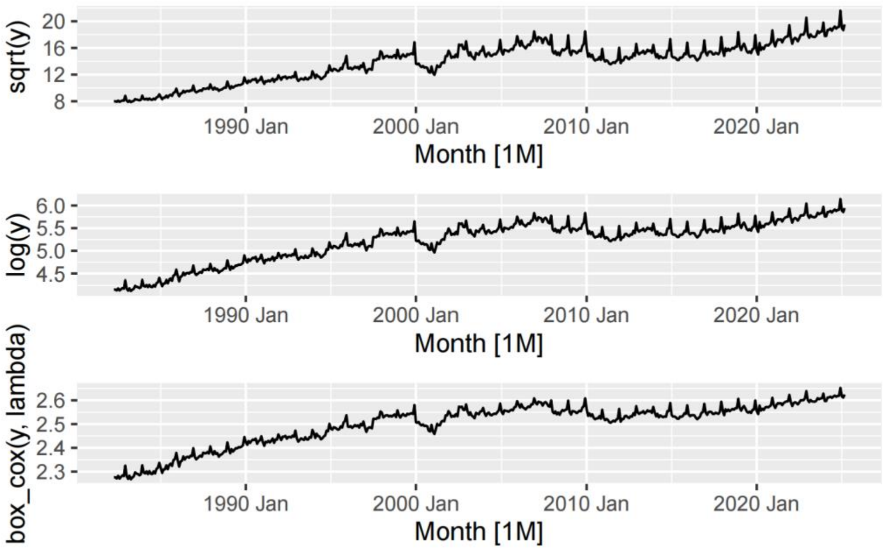

Figure 3: Variance Stabilization Comparison (√, log, Box-Cox)

To stabilize variance, three transformations were considered: logarithmic, square-root, and Box–Cox. The Guerrero method identified an optimal Box–Cox lambda ($\lambda = -0.326$), which produced the most stable residuals and improved interpretability.

Subsequent decomposition using STL confirmed that the trend component was dominant and the seasonal component stable but increasing in amplitude. The transformed and differenced series appeared approximately stationary after seasonal differencing (D = 1) and one additional regular difference (d = 1).

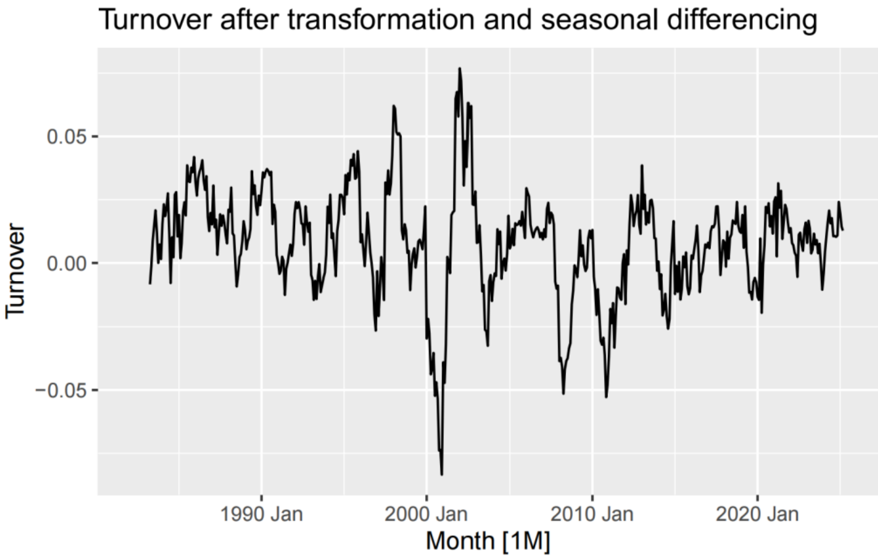

Figure 4: Stationarized Series after Seasonal Differencing

3 Methodology

The modeling process followed the standard Box–Jenkins framework:

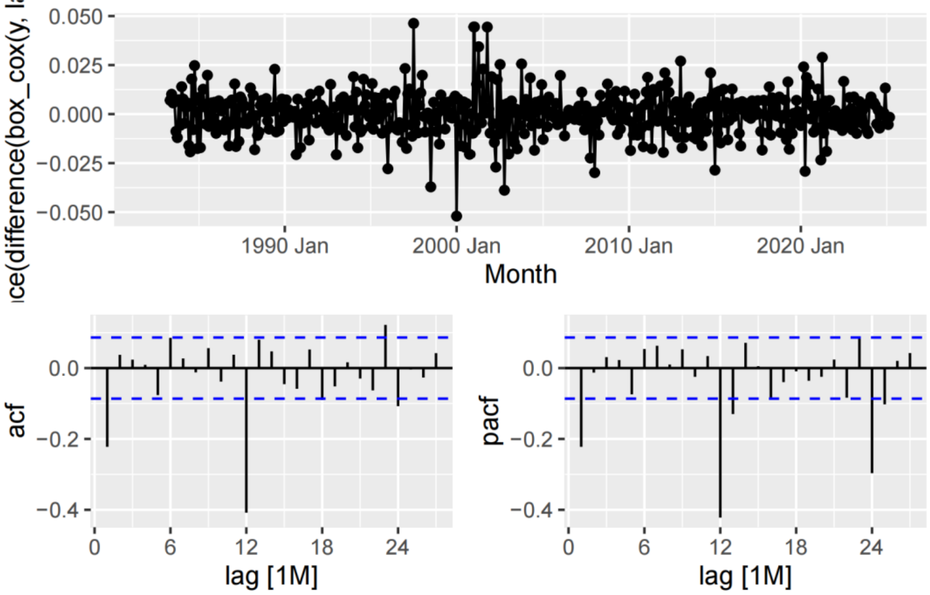

Figure 5: ACF and PACF for Differenced Series

Transformation and Differencing: Applied Box–Cox ($\lambda = -0.326$), one seasonal difference (p = 12), and one regular difference to achieve stationarity.

Identification: Examined ACF and PACF plots. The ACF showed a strong seasonal spike at lag 12, suggesting a seasonal MA(1) component. The PACF displayed significant spikes at lag 1, indicating that a non-seasonal AR(1) or MA(1) structure could be appropriate.

Model Estimation: Several candidate SARIMA models were estimated, including ARIMA(1,1,0)(0,1,1)[12] and ARIMA(0,1,1)(0,1,1)[12].

Model Selection: Criteria included AICc, residual diagnostics (Ljung–Box test), and interpretability.

Validation: Forecasts were evaluated using the final two years of data as a test set.

4 Model Development and Forecasting

4.1 ARIMA Modeling

```
Series: y
Model: ARIMA(1,1,0)(0,1,1)[12]
Transformation: box_cox(y, lambda)

Coefficients:
          ar1     sma1
      -0.2717  -0.9169
s.e.   0.0429   0.0265

sigma^2 estimated as 6.297e-05:  log likelihood=1709.47
AIC=-3412.94   AICc=-3412.9   BIC=-3400.28
```

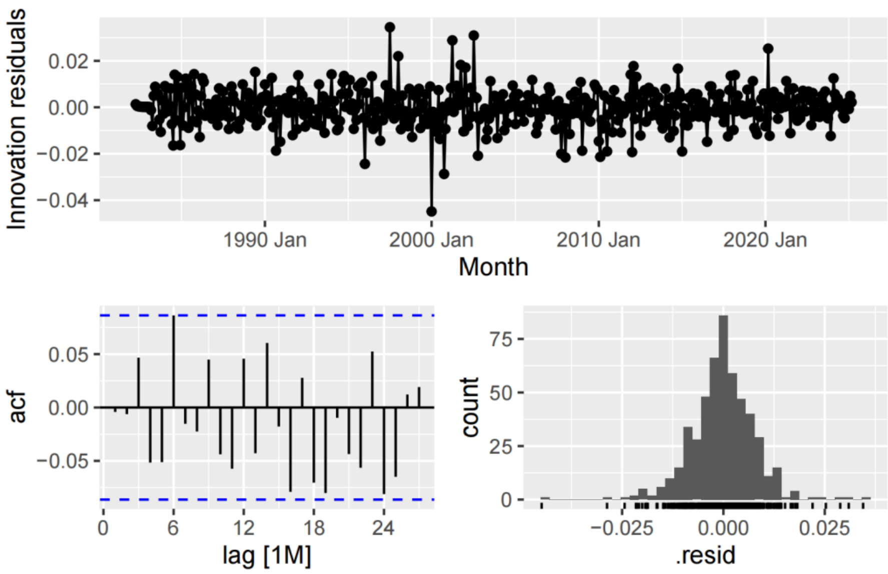

Figure 6: Residual Diagnostics for ARIMA(1,1,0)(0,1,1)[12]

```
Series: y
Model: ARIMA(0,1,1)(0,1,1)[12]
Transformation: box_cox(y, lambda)

Coefficients:
          ma1      sma1
      -0.2572  -0.9169
s.e.   0.0404   0.0266

sigma^2 estimated as 6.323e-05:  log likelihood=1708.46
AIC=-3410.92   AICc=-3410.87   BIC=-3398.26
```

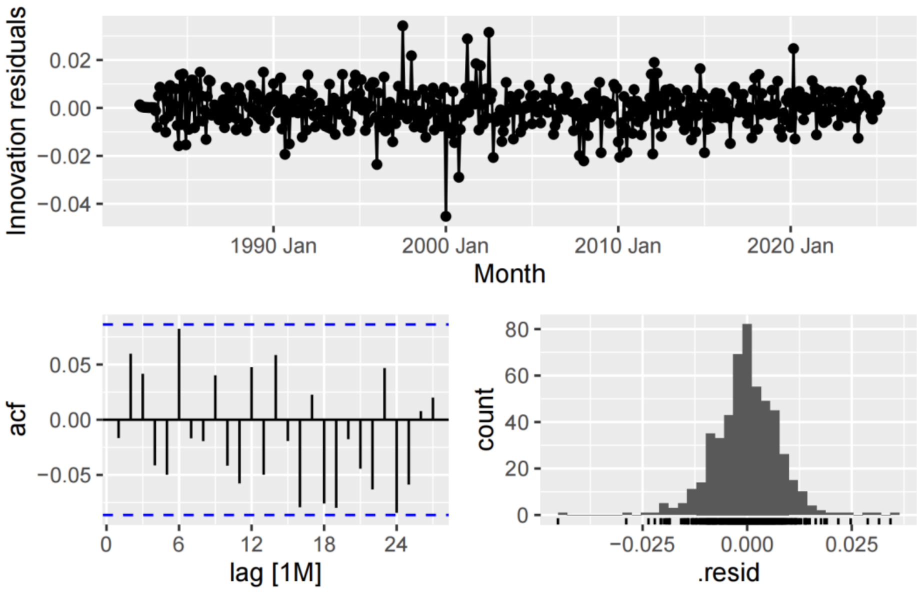

Figure 7: Residual Diagnostics for ARIMA(0,1,1)(0,1,1)[12]

The two candidate models were compared using information criteria:

Model	AICc	Residual p-value (Ljung–Box)
ARIMA(1,1,0)(0,1,1)[12]	–3413	0.0498
ARIMA(0,1,1)(0,1,1)[12]	–3411	0.0331

Both models performed similarly, but ARIMA(1,1,0)(0,1,1)[12] was preferred due to its slightly better diagnostics and interpretability. An automated model search (non-stepwise) suggested a more complex ARIMA(1,1,0)(2,1,2)[12], though the improvement in AICc was marginal (–3413.26 vs –3412.9). Given its higher complexity and similar error metrics, the simpler ARIMA(1,1,0)(0,1,1)[12] was selected as the final model.

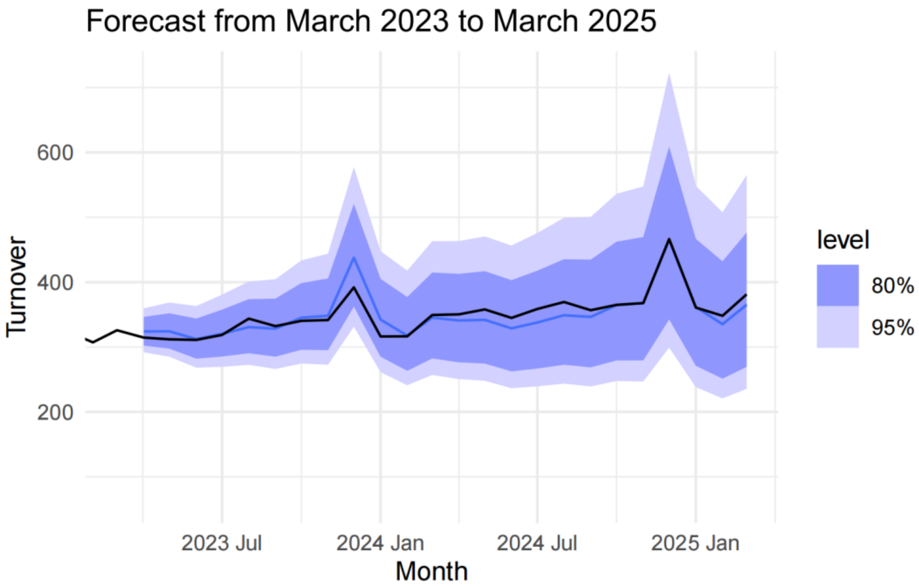

Figure 8: Two-Year Forecast from ARIMA Model

Residual plots indicated no significant autocorrelation or heteroscedasticity, and all assumptions of linearity, normality, and independence were satisfied. Forecast evaluation over the 2023–2025 horizon demonstrated that the model captured the upward trajectory of retail turnover with realistic confidence intervals, though the prediction intervals were moderately wide due to pandemic-related volatility.

4.2 ETS and Benchmark Models

To further evaluate model performance, an Exponential Smoothing (ETS) framework and a Seasonal Naïve (SNaive) benchmark were developed for comparison with the ARIMA results. These models serve distinct purposes: ETS provides a flexible trend–seasonality structure, while SNaive offers a simple baseline that reflects recurring seasonal patterns.

The ETS model was automatically selected as ETS(M,Ad,M), representing a multiplicative error structure, an additive damped trend, and multiplicative seasonality. This configuration captures gradual growth and seasonal effects while preventing trend overshooting through damping.

The forecasts generated by ETS show a smooth upward trajectory consistent with the historical pattern, and the prediction intervals are relatively narrow, suggesting stable uncertainty around expected values.

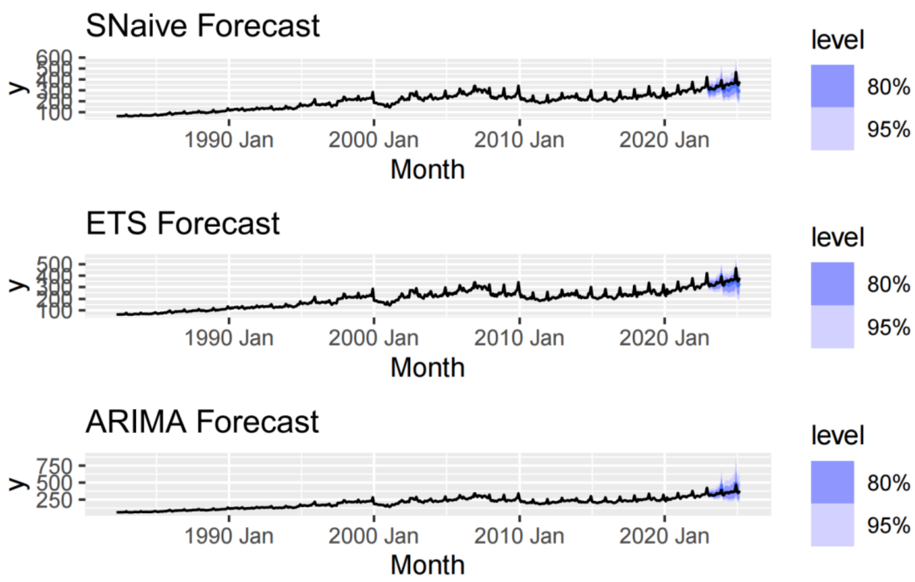

Figure 9: Forecasts from ETS Model (M, Ad, M)

In contrast, the Seasonal Naïve model assumes that each month's turnover repeats the same value from the corresponding month in the previous year. Although this method is statistically simple, it provides a meaningful performance baseline, allowing a clear evaluation of how much improvement advanced models offer.

The resulting forecast plot displays cyclical patterns that mirror the past seasonality but lacks adaptation to recent growth, illustrating its limited responsiveness to evolving retail dynamics.

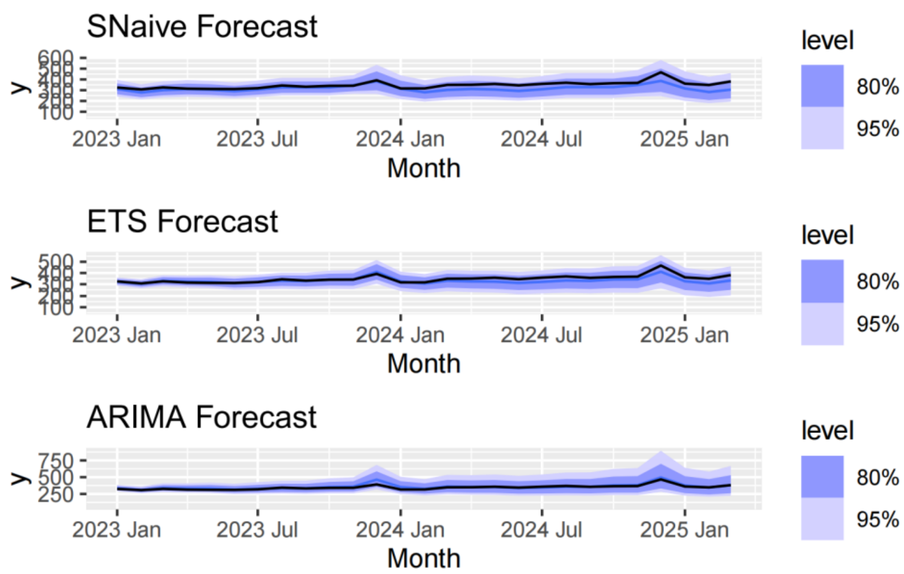

Figure 10: Forecasts from Benchmark (SNaive) Model

Comparing the two, the ETS model clearly adapts better to long-term growth and structural changes, while the SNaive forecast remains static. Both are valuable reference points, but ETS provides substantially higher predictive stability and interpretability for strategic retail forecasting.

5 Model Comparison and Evaluation

To evaluate the performance of all competing models, forecast accuracy was assessed on a holdout sample covering the most recent two years of data. Four commonly used error metrics were computed: the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Scaled Error (MASE). These metrics provide complementary perspectives on the magnitude and scale of forecast errors.

Model	RMSE	MAE	MAPE	MASE
Seasonal Naïve	37.8	31.1	8.64	1.63
ETS (M, Ad, M)	24.7	19.1	5.22	1.00
ARIMA(1,1,0)(0,1,1)[12]	18.8	12.6	3.54	0.66

As shown in Table, the ARIMA(1,1,0)(0,1,1)[12] model achieved the lowest error values across all metrics, indicating the best in-sample and out-of-sample predictive accuracy. The ETS(M,Ad,M) model ranked second, performing particularly well in terms of smoothness and stability, while the Seasonal Naïve benchmark recorded the largest forecast errors.

The following visual comparison further highlights the relative performance of the two top-performing models. Both ETS and ARIMA forecasts show consistent growth patterns aligned with the historical trend, while differences emerge in the representation of uncertainty and short-term fluctuations.

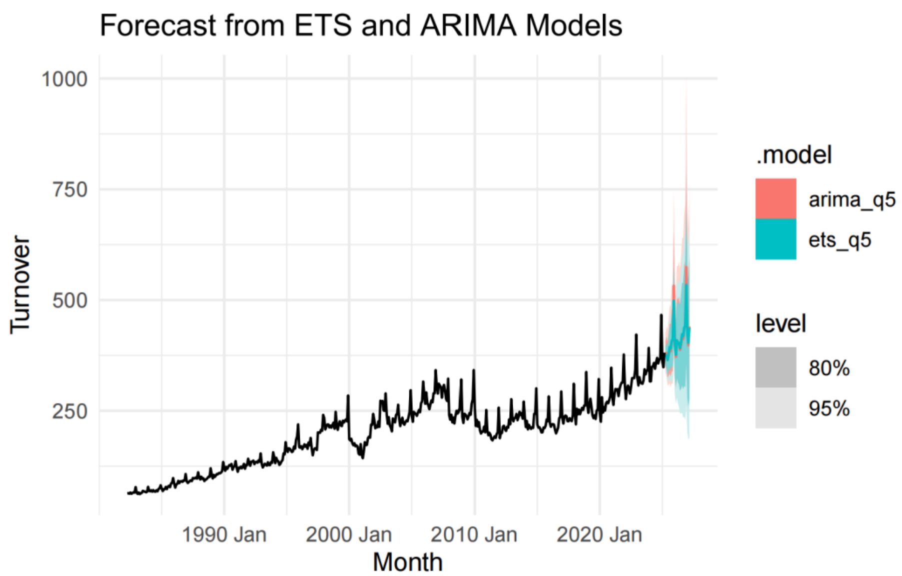

Figure 11: Forecast Comparison between ETS and ARIMA Models

Overall, the results suggest that while the ARIMA model yields slightly higher accuracy, the ETS model's interpretability and adaptive damping make it a practical alternative in real-world applications where simplicity and transparency are valued.

6 Discussion and Conclusion

This study developed and compared multiple time series forecasting models—ARIMA, ETS, and Seasonal Naïve—to analyse turnover patterns in NSW's Other Specialised Food Retailing industry.
The results consistently demonstrated a clear upward trajectory with strong seasonal components, alongside noticeable volatility during the COVID-19 pandemic period.

The ARIMA(1,1,0)(0,1,1)[12] model was found to provide the most accurate forecasts due to its capacity to capture autocorrelations and adapt to structural changes. The ETS(M,Ad,M) model offered competitive performance with smoother forecasts and more stable uncertainty bands. Both models outperformed the naïve benchmark, confirming their suitability for medium- to long-term retail forecasting.

To further explore the long-term outlook, forecasts from both ARIMA and ETS models were extended to 2027. The projections suggest sustained growth, with minor differences in volatility representation—ARIMA indicating higher uncertainty, and ETS projecting more stable increases.

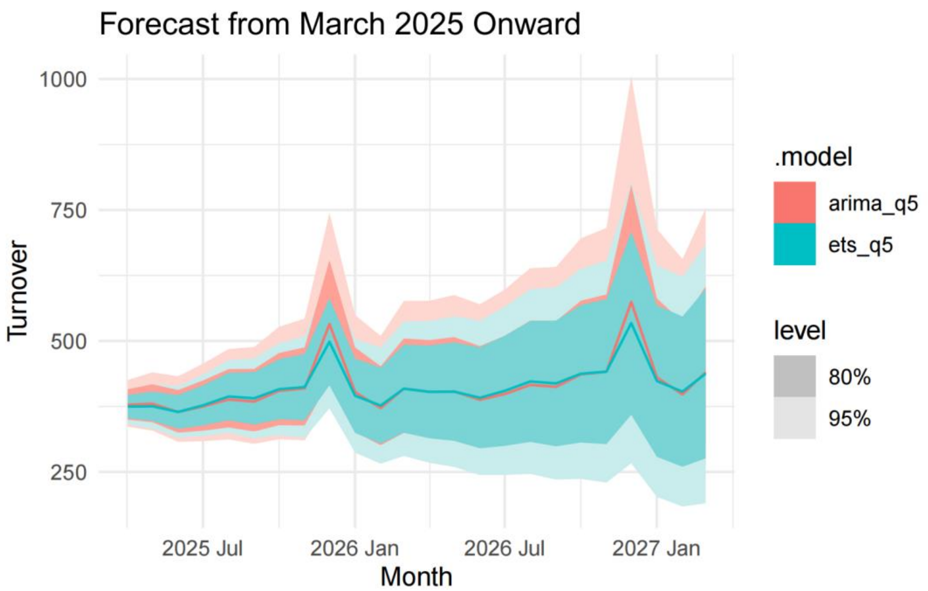

Figure 12: Long-Term Forecasts to 2027 by ETS and ARIMA Models

In conclusion, the findings underscore the practical value of combining statistical modeling and visual forecasting in business decision-making. The complementary strengths of ARIMA and ETS models demonstrate how data-driven forecasting can support strategic planning, inventory management, and policy evaluation in retail sectors sensitive to seasonality and economic disruption.

Future research may incorporate dynamic regression models by introducing external predictors—such as consumer confidence indices, disposable income, or mobility data—to improve explanatory power and scenario-based forecasting. This would enable deeper insights into the causal drivers behind turnover fluctuations and further enhance the predictive framework for industry analysis.