2024 North Carolina Back-to-School Retail Outlook: Predictions and Analysis

NC Retail Merchants Association

P.O. Box 1030, Raleigh, NC 27602

E-mail: info@ncrma.org

Contact: (919) 832-0811

June 2024

Table of contents

List of Figures

List of Tables

1 Executive Summary

1. This report analyzes North Carolina's monthly sales tax revenue data to identify retail sales patterns and forecast 2024 back-to-school sales volume.
2. Historical trends indicate that retail sales reach their highest peak in December, with a smaller peak in June.
3. Sales volumes tend to be lower in January and February.
4. The report reveals a consistent upward trend in retail sales over the past decade.
5. Total retail spending during the back-to-school season (June-September) 2024 is estimated to be around $84 billion, up from $81.67 billion in 2023.
6. Projections for the June-September 2024 back-to-school season suggest an average monthly growth of about 2.6 percent compared to last year.
7. June 2024 forecasts anticipate a 2.6 percent increase, while July 2024 forecasts predict a 2.7 percent uptick compared to July 2023.
8. August 2024 is projected to see a 1.8 percent increase, and September 2024 forecasts a 3.3 percent uptick from the previous year.
9. These findings offer valuable insights for stakeholders seeking to understand North Carolina's retail sales dynamics.

2 Introduction

As summer draws to a close, the retail world gears up for one of its busiest periods: back-to-school season. This time of year sparks a surge in consumer activity, shaping everything from sales patterns to advertising approaches. It's not just about buying pencils and notebooks; it's a significant economic and social event for families and retailers. For families, it's all about getting ready for the upcoming academic year. That means stocking up on school supplies, new clothes, and household essentials. For retailers, it's a critical time for revenue generation and strategic planning. They must ensure they have the right products in stock, plan their marketing efforts, and offer attractive promotions to capture a significant share of consumer spending.

Accurate forecasting is key for retailers to navigate this period successfully. By understanding and predicting back-to-school shopping trends, they can tailor their inventory, pricing, and marketing strategies accordingly. This enables them to stock the right products in the right quantities, avoid stockouts or excess inventory, and maximize profitability. Underestimating demand can lead to missed sales opportunities while overestimating can result in unnecessary stocking costs. Ultimately, forecasting empowers the retail industry to prepare, adapt, and capitalize on the opportunities presented by this important annual event. Furthermore, accurate forecasting also improves supply chain management, including inventory and workforce planning.

This report analyzes and predicts retail sales in North Carolina for the 2024 back-to-school season. Firstly, it delves into historical sales data to uncover patterns, trends, and other relevant features. Understanding these aspects not only aids in comprehending the data but also enhances forecasting accuracy. The retail sales data reveals several notable features. Firstly, there is a distinct upward trend over time, indicating a consistent increase in sales. Despite this long-term growth, sales fluctuate year to year. Additionally, the data shows a clear seasonal pattern, with higher sales in December due to the holiday season, followed by sharp declines in January and February as consumer spending decreases post-holidays. Accurately forecasting this series requires accounting for these long-term, yearly, and seasonal patterns.

The comparison of back-to-school season sales data from June to September 2023 against the same period in 2022 reveals that while there was some positive growth in June and July, sales declined in August and September. June 2023 showed the highest growth rate at approximately 1.1 percent, followed by July with around 0.8 percent. However, sales dropped by 0.9 percent and 1.8 percent in August and September, respectively. On average, sales declined by about 0.2 percent per month during the season in 2023. Over the past four years, 2023 experienced the lowest growth during this season. These growth rates show a sharp decline from 2022, which saw an average monthly growth of 3.4 percent. These growth rates indicate decreased consumer spending in 2023, likely due to high inflation rates that erode the real value of consumer wealth.

This report employs various industry-standard univariate models to assess their forecasting precision. By comparing the performance of these models, it aims to identify the most reliable forecasting approach. Subsequently, this report utilizes the best-performing forecasting model to predict retail sales for the 2024 back-to-school season, providing valuable insights for retail industry stakeholders and decision-makers. The projections in this report indicate that the above trends are expected to change in 2024, with retail sales anticipated to increase by approximately 2.6 percent per month from June to September 2024.

3 Data

This report utilizes monthly data on North Carolina taxable sales sourced from the North Carolina Department of Revenue for the period January 2012 to January 2024. These data are derived from reports and payments submitted by taxpayers and are categorized based on sales and use tax registrations.[1] A time plot of this data is depicted in Figure 1 below. However, we need to preprocess this data to enhance forecast accuracy.

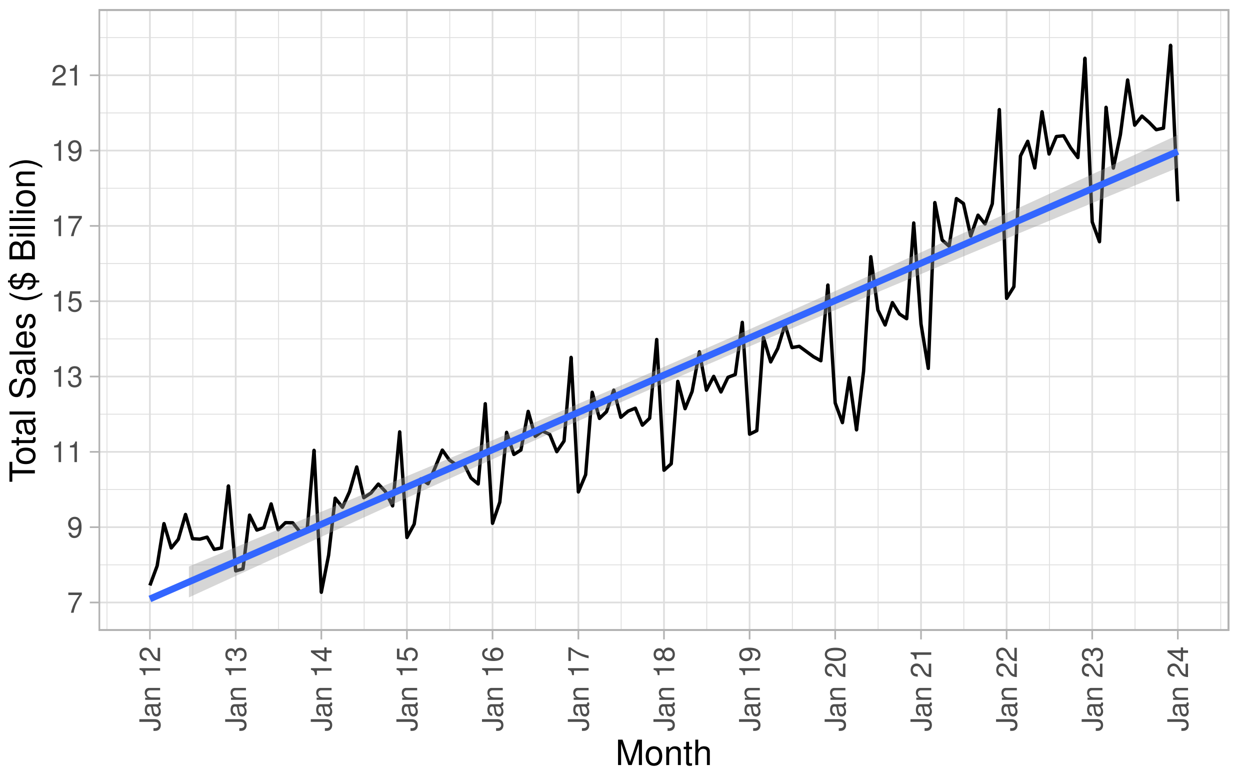
Preparing time series data is crucial for accurate forecasting. This involves comprehending the data, examining its features, and making necessary adjustments. The primary objective is to enhance forecast accuracy. Initially, inflation adjustments and mathematical transformations are applied to prepare the data for improved forecast accuracy. Subsequently, various data visualization techniques, including time plots, seasonal plots, and decomposition analysis, are employed to grasp fundamental data characteristics. These preprocessing steps make the data better suited for forecasting, enabling more precise predictions for informed decision-making.

3.1 Inflation Adjustment

Inflation-adjusting sales data is a critical step for accurate and reliable forecasting. It is important to account for inflation to avoid distorted perceptions of historical sales trends. As inflation erodes the real value of money over time, unadjusted sales figures may overstate sales growth. By applying inflation adjustments, we gain a clearer and more realistic understanding of the data, enabling us to make informed projections. Therefore, this report incorporates inflation-adjusted data into the forecasting models as a fundamental necessity and a best practice. Figure 2 presents the new inflation-adjusted data.

[1]Additional details of this data, its composition, and tax categories can be found at https://www.ncdor.gov/news/reports-and-statistics/monthly-sales-and-use-tax-statistics.

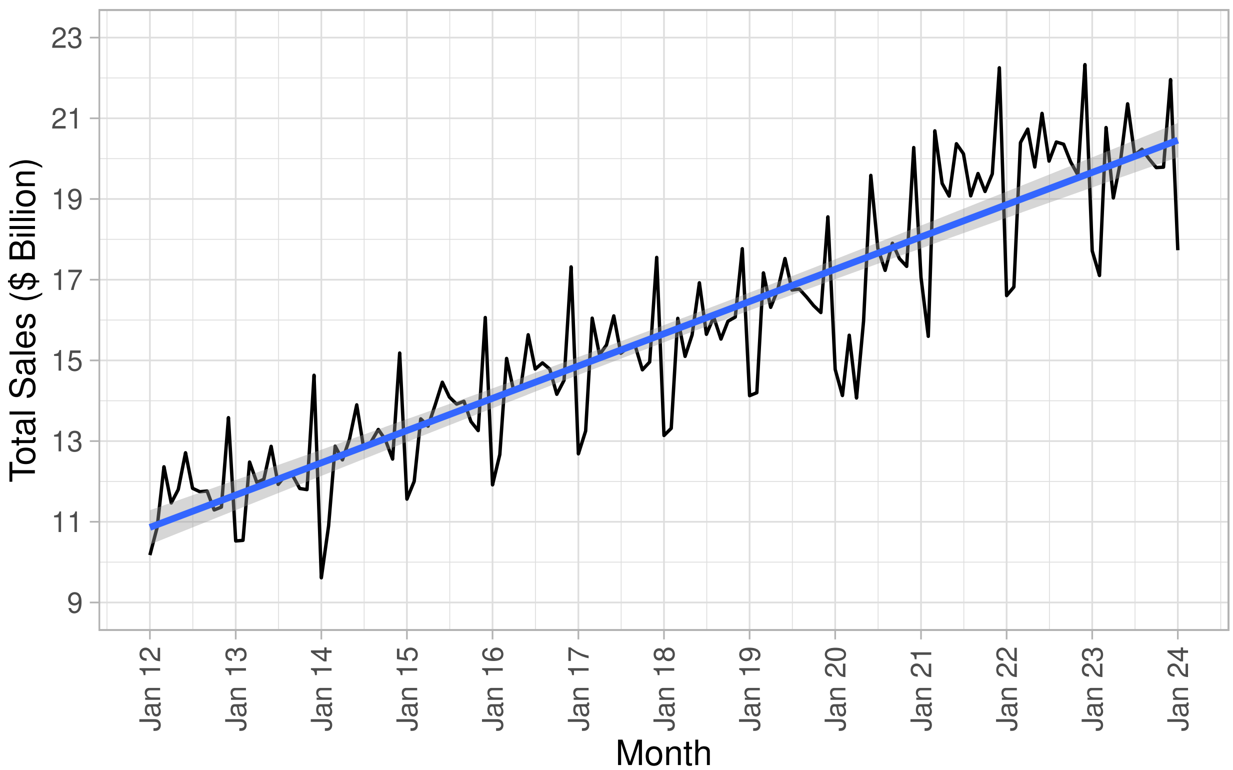
Figure 1: NC Monthly Unadjusted Sales ($ Billion) (Jan 2012 - Jan 2024)



3.2 Mathematical Adjustment

Mathematical transformation of data before forecasting is imperative for several reasons. Firstly, it simplifies the time series, making it more amenable to accurate predictions. When historical data is adjusted, it often results in a smoother and more consistent series. Secondly, if the original data exhibits a level of variation that is proportional to the level of the series, a mathematical transformation becomes invaluable. In the case of our inflation-adjusted data, it is evident that there is a noticeable increase in variation in the later years compared to earlier ones. This discrepancy can lead to inaccurate forecasts if unadjusted. A mathematical transformation can homogenize the variation across the entire series, rendering the forecasting model simpler and more reliable. In essence, this process not only enhances the predictive power of our models but also ensures that forecasts are based on more stable and consistent data. Therefore, we apply a Box-Cox transformation to our data. We can clearly see the improvement in the

Figure 2: Inflation Adjusted Monthly Sales ($ Billion) (Jan 2012 - Jan 2024)

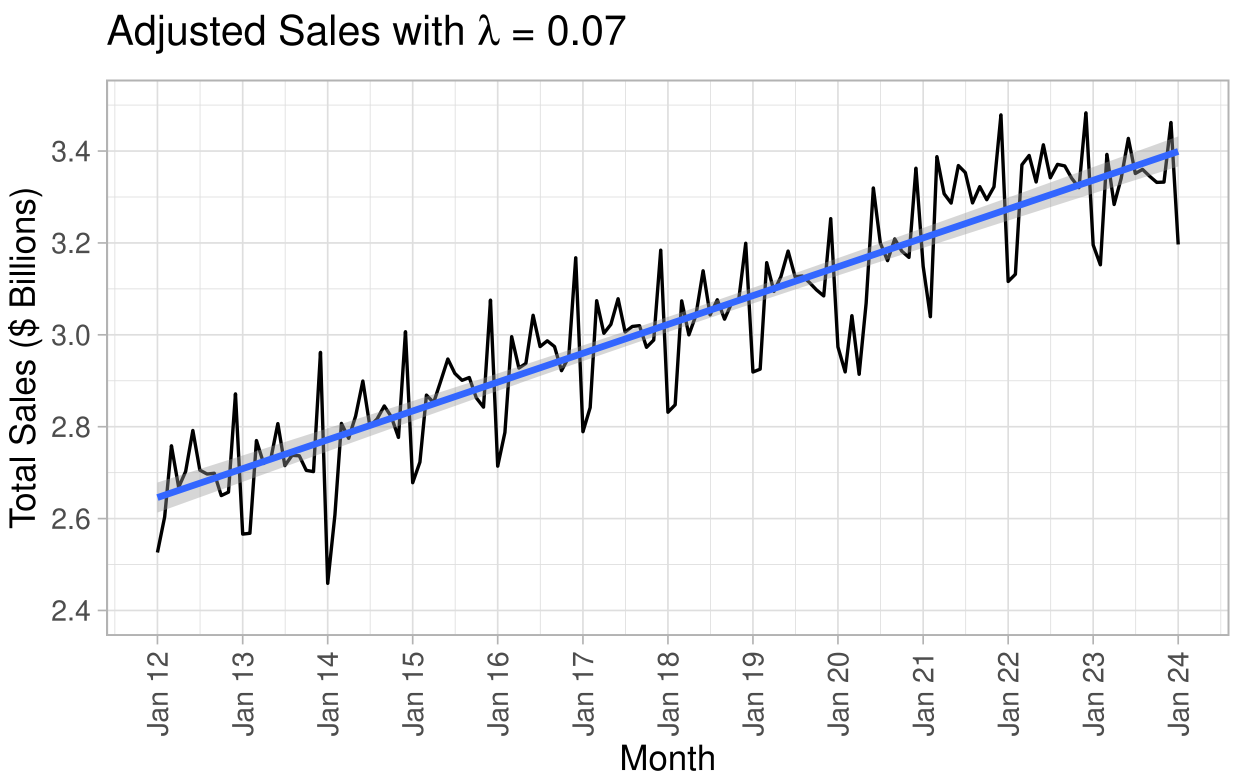


stability of the series in Figure 3.

3.3 Features

The first thing to do in any data analysis task is to plot the data. Graphs enable many features of the data to be visualized, including patterns, unusual observations, changes over time, and relationships between variables. The features that are seen in plots of the data must then be incorporated, as much as possible, into the forecasting methods to be used.

Figure 3: Box-Cox Transformation of Inflation Adjusted Sales



3.3.1 Time Plots

Figure 1 to Figure 3 display the retail data in time plots, showing observations against their respective times. These plots reveal several notable features of the data.

- **Long-Term Growth:** A distinct upward trend is evident, indicating a consistent increase in retail sales over time. Despite fluctuations, there seems to be a general trend of growth in retail sales over the years, indicating a potential overall expansion in the economy or changes in consumer behavior and purchasing power over time.
- **Yearly Fluctuations:** There are fluctuations in sales from year to year. Some years, such as 2023, show a general increase in sales compared to previous years, while others, such as 2020, show a dip, possibly due to external factors like economic downturns and Covid-19.

- **Seasonal Trends:** There appears to be a seasonal pattern in the data, with certain months consistently showing higher sales compared to others. For example, December tends to have higher sales. This could be attributed to increased spending during the holiday season. However, the post-holiday season in January and February indicates a sharp decline in sales, likely due to reduced consumer spending after the holiday season.

3.3.2 Seasonal Plots

Figure 4 displays a seasonal plot of the data. A seasonal plot presents data against the individual "seasons" during which they were observed, allowing for overlap between the data from each season. This type of plot enhances the visibility of underlying seasonal patterns, which is particularly useful for detecting changes in patterns over time. Examining this plot, a notable observation is the significant spike in sales consistently occurring in December each year. Additionally, the plot highlights an unusually high number of sales in June 2020.

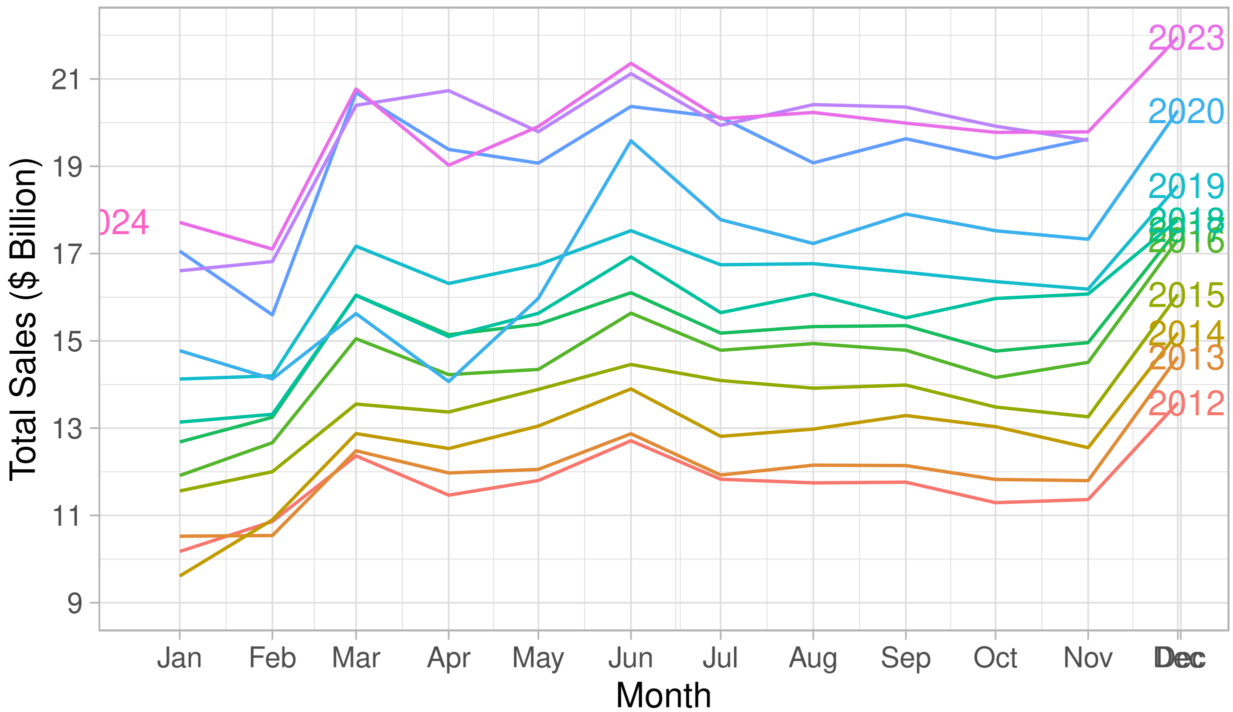
Figure 5 presents an alternative seasonal plot format. In this graph, the data for each season are aggregated into separate mini-time plots. Blue horizontal lines mark the means for each month. Analyzing this plot, it becomes evident that, on average, sales tend to peak in December and June. Conversely, the plot illustrates that sales typically dip post-holiday season, indicating lower sales volumes in January and February.

3.3.3 Decomposition

Time series data can manifest diverse patterns, and dissecting it into distinct components, each representing a specific underlying pattern category, is beneficial. Typically, we categorize a time series into three components:

- A trend-cycle component.
- A seasonal component.
- A remainder component (which encompasses any other variations in the time series)

Figure 4: Seasonal Plot 1: Retail Trade Sales in North Carolina (Jan 2012 - Jan 2024)



This decomposition enhances comprehension of the time series and improves forecast accuracy. Figure 6 decomposes the retail sales after making inflation and mathematical adjustments to the data. The top panel of this graph displays the adjusted and transformed data. The bottom three panels depict each component separately. We can reconstruct the data shown in the top panel by adding these components together. Once more, the second panel reveals a clear trend. In the third panel, the seasonal component remains relatively constant over time, indicating a similar pattern between consecutive years and even distant years. The bottom panel illustrates the remainder component, representing what remains after subtracting the seasonal and trend-cycle components from the data.

Figure 5: Seasonal Plot 2: Retail Trade Sales in North Carolina (Jan 2012 - Jan 2024)

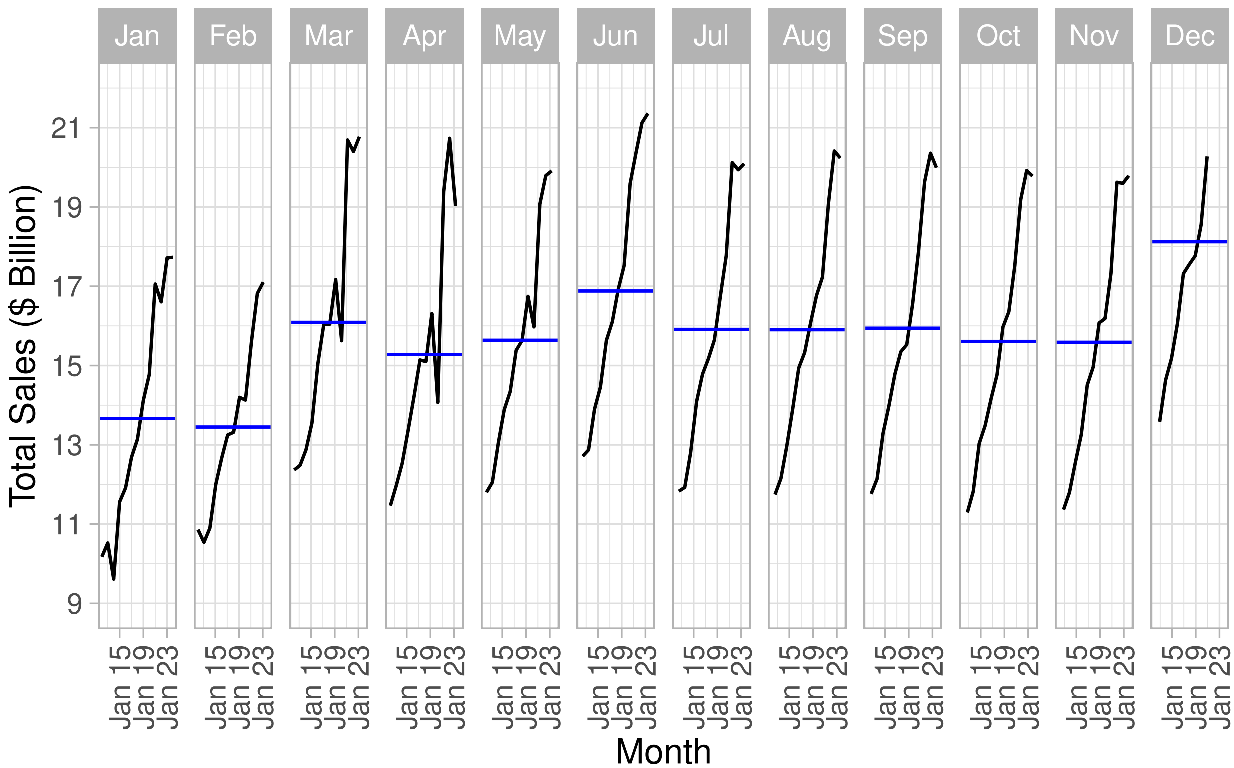


Figure 6: Decomposition of Inflation Adjusted Transformed Monthly Retail Sales

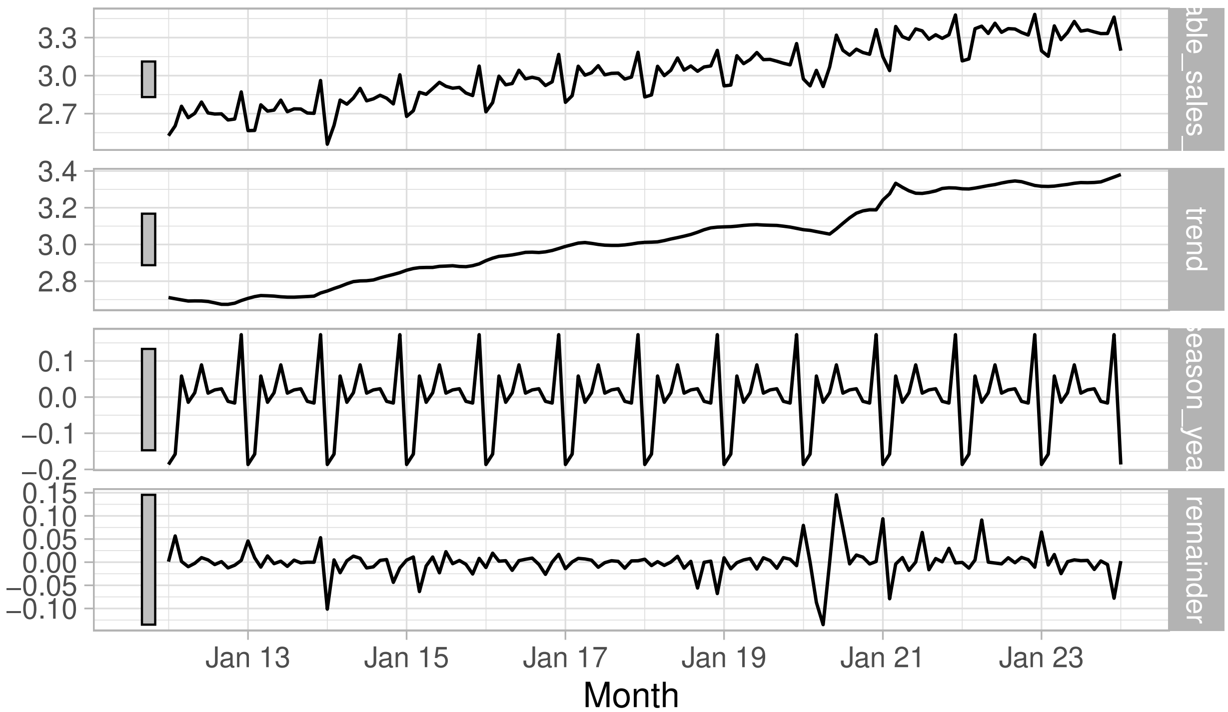


Table 1: Retail Sector Composition

Retail Sector	**Composition**
Apparel Group	Boot and shoe stores, clothing stores, clothing and accessory stores, etc.
Automotive Group	Motor vehicle dealers, service stations, garages, automotive supply stores, recappers, and repairers, and manufactured home (mobile home) dealers, etc.
Food Group	Bakeries, grocery stores, meat markets, vending machine operators, restaurants, cafeterias, grills, nightclubs, etc.
Furniture Group	Furniture stores, household appliance dealers and repair services, antique dealers, interior decorators, etc.
General Merchandise Group	Department stores, drugstores, farm implement and supply stores, general stores, hardware stores, jewelry stores, industrial machinery and supply dealers, flea markets, sporting goods stores, toy shops, variety stores, pawn shops, road building equipment and supply dealers, etc.
Lumber & Building Material Group	Sheet metal shops, steel fabricators, building hardware and machine stores, building material dealers, electrical, plumbing, and heating supply dealers, etc.
Unclassified Group	Beauty and barber shops, bookstores, coal and wood dealers, feed stores, florists, funeral homes, photographers, laundries, hospitals, and various other businesses offering a wide range of community services.

4 Forecasting Back-to-School Retail Sales

4.1 Training vs. Test Sets

Splitting data into training and test sets is a critical step in developing accurate forecasting models. This division allows us to assess a model's performance on unseen data, which simulates real-world scenarios. By training a model on a subset of the data and testing it on another, we can gauge its ability to forecast beyond the information it was trained on. This practice helps select the most appropriate forecasting model, as it provides an unbiased evaluation of predictive power of a model. Moreover, it guards against overfitting, a common pitfall where a model becomes overly complex and fits the training data too closely, leading to poor performance on new, unseen data. This separation ensures that our model captures underlying patterns without getting bogged down by noise or idiosyncrasies in the training set, ultimately leading to more reliable and robust forecasts.

This report uses data from January 2012 to January 2023 as a training set. Then, it assesses the prediction accuracy of various forecasting models using the test set spanning from February 2023 to January 2024, latest available data month. It then uses the best-selected forecasting model to predict retail sales for the back-to-school season, spanning June to September 2024.

4.2 Fitting the Forecasting Methods on the Training Set

In this step, this report employs multiple industry-standard forecasting methods on the training set to select the best model. Choosing the appropriate method based on the specific features of the time series data is crucial, as using an unsuitable method can result in inaccurate forecasts. Hence, considering the particular time series features of retail sales data explored above, this report utilizes the following forecasting methods.

- Benchmark: Seasonal Naive
- Dynamic Regression Model

- Holt-Winters' with Additive Damped Trend and Additive Seasonality (Ad, A)
- Holt-Winters' with Additive Damped Trend and Multiplicative Seasonality (Ad, M)
- Innovations State Space Models
- Neural Network Model
- Prophet Model Additive Seasonality
- Prophet Model Multiplicative Seasonality
- SARIMA (Seasonal AutoRegressive Integrated Moving Average Model)
- Simple Exponential Smoothing
- Theta Method with Additive Seasonality
- Theta Method with Multiplicative Seasonality

Moreover, employing multiple forecasting methods on the same time series and averaging the resulting forecasts is highly effective in improving out-of-sample forecast accuracy. This method, proposed by Bates and Granger in 1969, has consistently shown enhanced forecast accuracy. Numerous studies have confirmed that combining forecasts, even through simple averaging, consistently leads to better forecasting performance. Therefore, this report also evaluates the test set accuracy of all possible combinations of the top three models to determine if forecast accuracy could be improved by averaging these models.

4.3 Testing Out of Sample Forecast Accuracy

Figure 7 presents the forecast accuracy comparison across all the models. Clearly, most of the models are able to capture the patterns in the original series (black line). Similarly, Table 2 presents several out-of-sample forecasting accuracy results across all models. The accuracy measures in this table include Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Mean Absolute Scaled Error (MASE), and Root Mean Squared Scaled Error (RMSSE). This table also includes four ensembled models, calculated using all possible combinations of the top models indicated by an asterisk (*). A good forecast method produces the lowest out-of-sample error. We select our preferred forecasting model based on RMSE. Based on these out-of-sample accuracy measures, the Seasonal Naive method performs the best, followed by the Theta method with additive seasonality and the Holt-Winters method with a damped trend and additive seasonality.

Figure 7: Out of Sample Forecast Accuracy Comparison

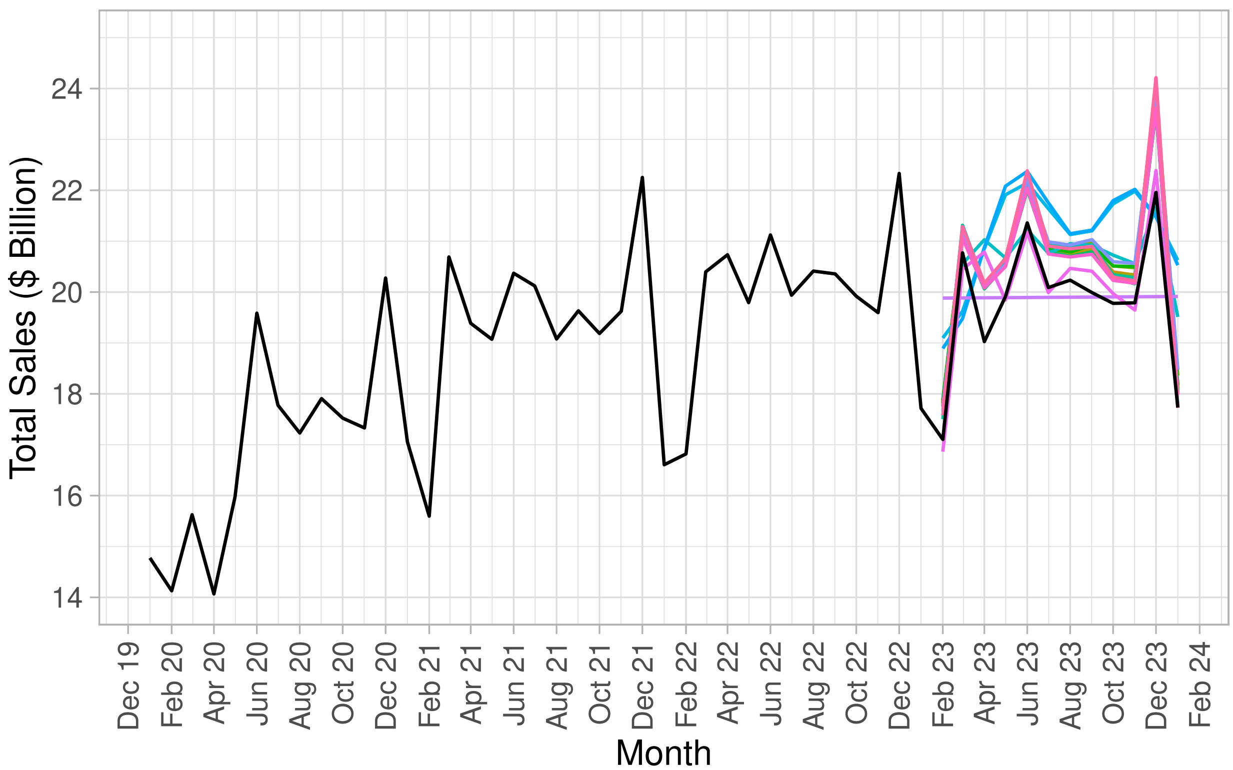


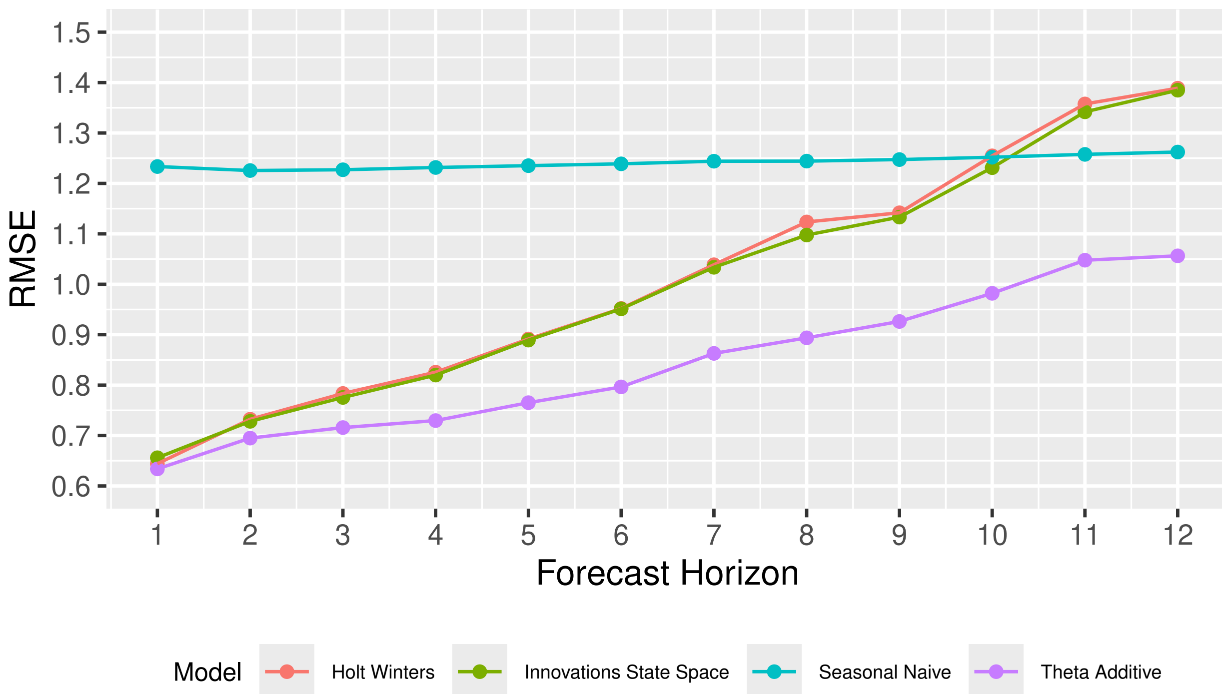
Table 2: Out of Sample Forecast Accuracy Comparison

Model	RMSE	MAPE	MASE	RMSSE
Seasonal Naive	0.561	1.747	0.371	0.462
Theta (Additive)*	0.751	3.275	0.708	0.618
Ensemble 1	0.756	3.334	0.720	0.622
Holt-Winters (Ad, A)*	0.761	3.392	0.731	0.626
Ensemble 2	0.802	3.592	0.776	0.660
Ensemble 3	0.824	3.693	0.798	0.678
Ensemble 4	0.829	3.751	0.809	0.682
State Space Model*	0.899	4.110	0.887	0.740
Holt-Winters (Ad, M)	0.923	3.847	0.844	0.760
SARIMA	0.938	4.325	0.933	0.772
Dynamic Regression	0.941	4.304	0.928	0.774
Theta (Multiplicative)	0.950	3.882	0.852	0.782
Neural Network	0.986	4.288	0.890	0.811
Simple Exponential Smoothing	1.308	4.829	1.001	1.077
Prophet (Additive)	1.712	8.186	1.707	1.409
Prophet (Multiplicative)	1.738	8.368	1.752	1.430

As a final test, this report examines the cross-validation of the top models by evaluating their ability to forecast sales one to twelve months into the future. Cross-validation is a more sophisticated approach than using simple training/test sets. In this procedure, there are multiple test sets, each consisting of a single observation. The corresponding training set includes only observations that occurred before the observation formed the test set, ensuring that no future observations are used to construct the forecast. The forecast accuracy is computed by averaging the errors over all test sets.

Figure 8 below evaluates the forecasting performance of these models for one- to twelve-month-ahead forecasts. The plot shows that the forecast error increases as the forecast horizon extends, which is expected. It can be seen that the Seasonal Naive method fails this test by producing the largest errors. In contrast, the Theta method with additive seasonality performs very well, maintaining the lowest error across all forecast horizons. Therefore, this test reveals the superiority of the Theta method in forecasting this time series.

Figure 8: Cross-Validation of the Top Models



4.4 Residual Diagnostic of the Best Forecast Method

The “residuals” in a time series model are the differences between the actual observations and the corresponding forecasted values after fitting the model.

Residual = Actual value−Forecast value

A good forecasting method should produce residuals with specific characteristics: they should be uncorrelated and have a mean of zero. If there are correlations between these residuals, it indicates untapped information that should have been used to enhance forecast accuracy. Similarly, if the residuals have a mean different from zero, the forecasts are biased. If either property is not met, modifying the forecasting method can lead to improved forecasts.

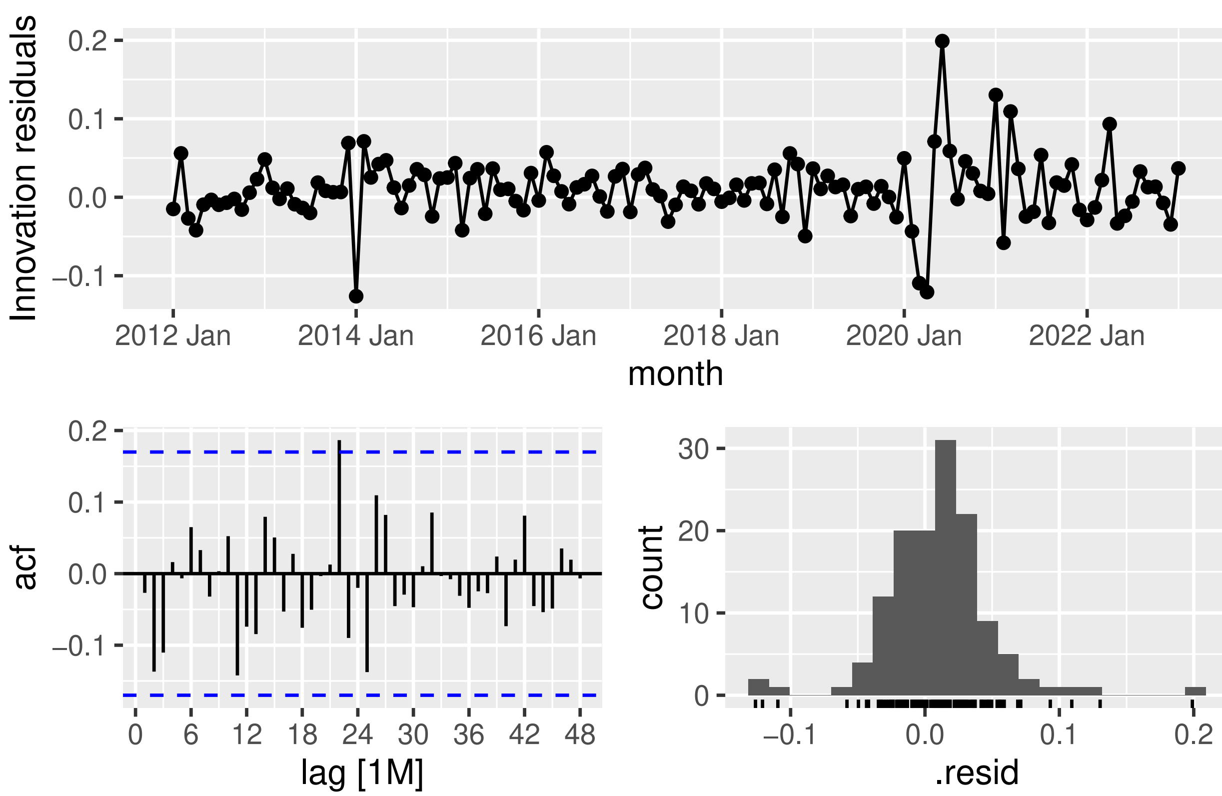
Figure 9 produces residual diagnostic tests for the Theta method. On average, the mean of the residuals is close to zero, indicating that, on average, our model forecasts the actual values fairly accurately. A time plot of the residuals shows that their variation remains relatively constant over the historical data, except for one outlier. Therefore, the residual variance can be treated as constant. The autocorrelation function (ACF) graph shows no significant correlation in the residuals. However, the histogram suggests that the residuals may not follow a normal distribution, as the right tail appears somewhat long even when the outlier is ignored. This implies that while the forecasts from this method are likely to be quite accurate, the prediction intervals, which assume a normal distribution, may be unreliable. To check for autocorrelation further, we can use the Ljung-Box test. This test returns a relatively large p-value, indicating that the residuals are not significantly different from a white noise series. This suggests that the method captures all the available information when producing forecasts.

Overall, these residual diagnostics suggest that the Theta method effectively captures the information in the data and is likely to produce accurate forecasts.

Table 3: The Ljung-Box Test for the Residuals

Ljung-Box Stat	Ljung-Box p-value
29.888	0.754

Figure 9: Residual Diagnostics of the Top Model



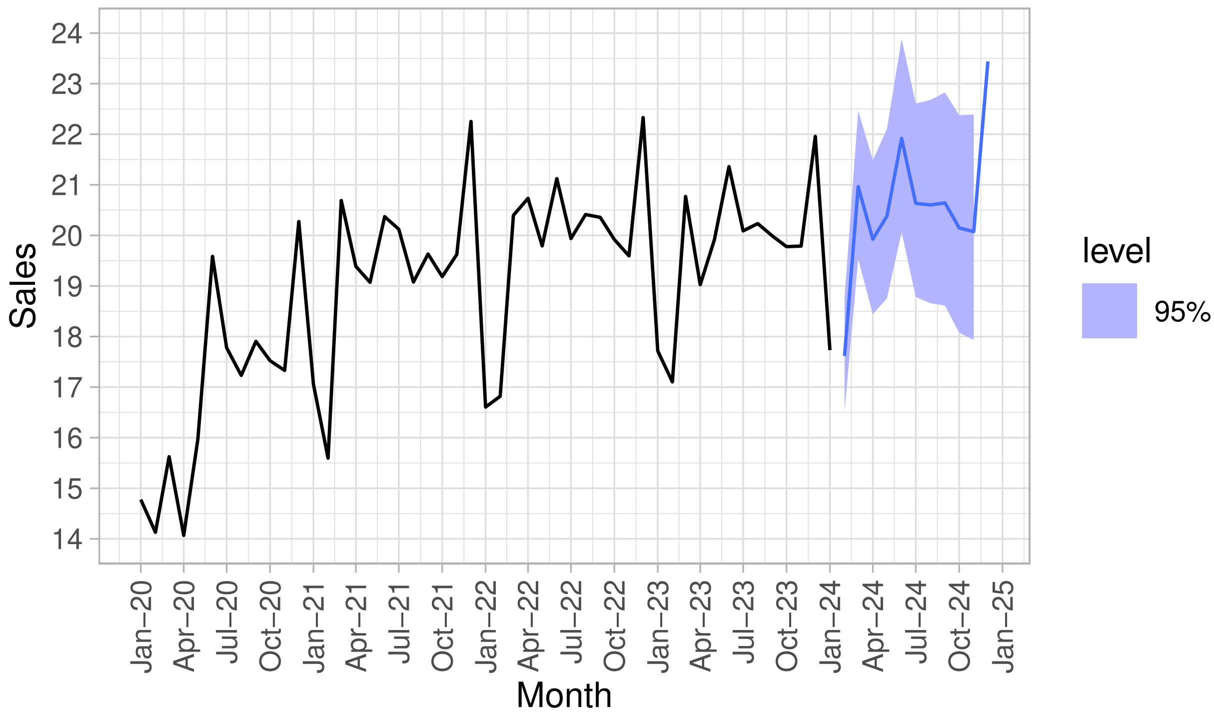
4.5 Forecasts

As a final step in forecasting, this report employs the Theta method with additive seasonality on the data from January 2012 to January 2024 (the latest available data month) to predict retail sales for the next eleven months, spanning February to December 2024. The forecasted values are presented in Figure 10.

Similarly, Table 4 presents the back-to-school season 2024 retail sales predictions, comparing the predicted values with the inflation-adjusted values from the same months in 2023. The back-to-school season in 2024 looks promising for the retail sector, with total sales for the season expected to reach approximately $84 billion. This marks an average monthly increase of 2.6 percent compared to the previous year.

- **June 2024:** Sales are predicted to reach $21.92 billion, $0.56 billion more than in June 2023, reflecting a 2.6 percent growth.

Figure 10: Retail Sales Forecasts (February 2024-December 2024)



- **July 2024:** Sales are expected to decrease slightly to $20.63 billion compared to June 2024, but they will still show a $0.54 billion increase and 2.7 percent growth over July 2023.
- **August 2024:** Sales are anticipated to be $20.60 billion, marking a $0.37 billion increase and a 1.8 percent growth from August 2023.
- **September 2024:** Sales are expected to rebound to $20.64 billion, with a $0.65 billion increase and a 3.3 percent growth compared to September 2023.
- **Back-to-School Season 2024:** Total sales for the season (June-September) are expected to be $83.80 billion, up from $81.67 billion last year, indicating a $2.13 billion increase and a 2.6 percent growth.

Overall, these predictions highlight a growing sales volume during the back-to-school season 2024 compared to last year.

Table 4: Retail Sales Forecasts for June 2024 - September 2024

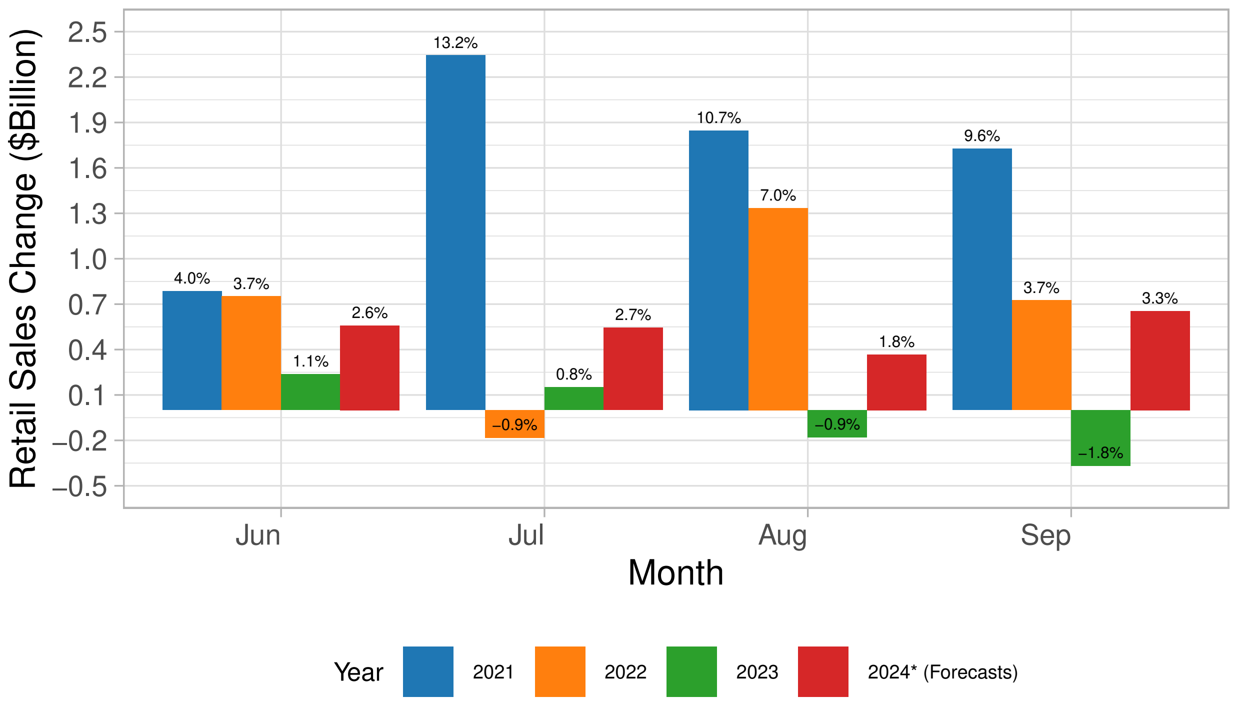
Month	Predicted Sales	Change ($Billion)	Growth (%)
2024 Jun	21.92	0.56	2.62
2024 Jul	20.63	0.54	2.71
2024 Aug	20.60	0.37	1.82
2024 Sep	20.64	0.65	3.27

To further understand these forecasts, Figure 11 provides a detailed month-by-month comparison of these predictions with inflation-adjusted retail sales from 2021 to 2023. The data indicate mostly positive growth in back-to-school season retail sales over the past several years, although the growth rates have fluctuated and slowed down recently.

- For June 2024, the report projects retail sales to reach $21.92 billion, up from about $21.36 billion in June 2023, continuing the historical upward trend. However, the growth rates for June have varied: sales increased by 4.0 percent in 2021, dropped to 3.7 percent in 2022, sharply declined to 1.1 percent in 2023, and are projected to rise to 2.6 percent in 2024.
- For July 2024, projected retail sales are expected to reach $20.63 billion, up from $20.09 billion in July 2023. The growth rates for July started at 13.2 percent in 2021, dropped to -0.9 percent in 2022, increased to 0.8 percent in 2023, and are anticipated to be 2.7 percent in 2024.
- For August 2024, the report projects retail sales to reach $20.60 billion, up from about $20.23 billion in August 2023. The growth rates for August followed a similar declining pattern, starting at 10.7 percent in 2021, dipping slightly to 7.0 percent in 2022, dropping sharply to -0.9 percent in 2023, and are forecasted to increase to 1.8 percent in 2024.
- For September 2024, projected retail sales are expected to reach $20.64 billion, up from $19.99 billion in September 2023. The growth rates for September have also shown a consistent decline: they started at 9.6 percent in 2021, decreased to 3.7 percent in 2022, sharply dropped to -1.8 percent in 2023, but are forecasted to rise to 3.3 percent in 2024.

Overall, while retail sales volume continues to trend upward, the growth rates indicate a decelerating pace of increase. This suggests a potential slowdown in market dynamics, likely influenced by shifts in consumer behavior and purchasing power, potentially due to high inflation rates in recent years.

Figure 11: Retail Sales Change and % Change from Same Month Last Year



5 Summary

The success of retailers during the back-to-school season hinges on accurate forecasting. By anticipating shopping trends, retailers can tailor their strategies to optimize inventory, pricing, and marketing efforts. This report examines and predicts retail sales in North Carolina for the 2024 back-to-school season (June-September).

The analysis begins by examining historical sales data to uncover patterns and trends. The data reveals consistent long-term growth in sales volume, although there are fluctuations from year to year. Additionally, the data shows two seasonal peaks: the first and largest peak during the holiday season and a second, relatively smaller peak in June.

Predicted sales for the back-to-school season from June to September 2024 show a consistent increase compared to the previous year, reaching about $84 billion. The month-by-month projections show that July 2024 will experience the highest sales growth of 2.7 percent from July 2023, followed by June 2024, with a growth of 2.6 percent from June 2023. On the other hand, August 2024 will see the lowest growth for the season, with a 1.8 percent increase compared to August 2023. In September 2024, sales are expected to rebound, growing by 3.3 percent compared to September 2023. These projections indicate growing sales values during the back-to-school season, with some months seeing higher growth rates than others.

The forecasts are based on sales tax data from various retail sectors across all regions but may not capture specific geographical variations or sector-specific sales patterns. Some geographical areas and sectors may experience growth trends that are different from those projected in these aggregated values. Therefore, it's important to consider these limitations when using these forecasts for decision-making purposes.