2024 Forecasting Report

“ABC Store”



**TEAM:**

Padmini Poojari

Ravi Teja Lakkakula

Jamie Hardy

Under Guidance of:

Edward Williams

DS630-Applied Forecasting Term Project Report

08 December 2023

Contents

[1.](#_heading=h.35nkun2) Brief background of enterprise: 2

[2.](#_heading=h.1ksv4uv) Description of the business problem: 2

[3.](#_heading=h.44sinio) Data collection and analysis: 3

[4.](#_heading=h.2jxsxqh) Forecasting approaches used: 3

[4.1. Exploring forecasting techniques which better fit our data: 3](#_heading=h.z337ya)

[5. Analysis of output results & recommendations 3](#_heading=h.3j2qqm3)

6[. Future work: 3](#_heading=h.4i7ojhp)

## Background of Enterprise:

Our Academic Project: ABC Retail Store in Plymouth, MI

Discover the thriving retail store that has been serving the community for the past 30 years. ABC Retail Store, located in Plymouth, MI, offers a wide range of products to cater to the needs of its dedicated local customer base. This established store has undergone significant changes over the years, one of which includes the addition of a laundromat to complement the overall business growth.

Proudly serving customers in Plymouth, Redford, and Livonia, ABC Retail Store provides a welcoming ambiance and an exciting atmosphere. With regular promotional events every two weeks, customers can enjoy tastings and be the first to try new products on the market.

Recognizing the importance of digitalization, ABC Retail Store has embraced online presence, making it effortless for customers to connect and shop via various social media channels. Additionally, partnering with well-known delivery companies like DoorDash, Grubhub, and UberEATS ensures that customers can conveniently order items from the comfort of their homes.

The outstanding reputation of ABC Retail Store is evident in their impressive Google rating of 4.7 and a perfect 5 out of 5 on Facebook. Experience the excellence of ABC Retail Store and discover why they continue to be a trusted and beloved part of the community.

## Description of the Business Problem:

The current challenge faced by the store is uncertainty regarding the demand for items and anticipated sales in the upcoming months. A successful business is expected to have consistent growth, but if there is significant volatility with frequent fluctuations, it becomes concerning to understand the factors driving this roller coaster effect on sales. Additionally, the store made a minor renovation change in February and now wants to determine if it had any impact on sales. This information will guide their decision-making for an upcoming major renovation, taking into account sales results and future growth prospects. December is a crucial month for this retail store, as it involves holiday shopping for Christmas and New Year’s. Therefore, each year it remains unpredictable how much the holiday season will affect sales. By analyzing sales data, the business management can better gauge the level of aggressiveness required to meet the demand and maintain appropriate stock levels.

## Data Collection and Analysis:

Our data consists of net sales, which are the sales after subtracting customer returns, discounts, and other deductions. We collected this data from the point-of-sale system, covering the period from November 2019 to October 2023. In order to assess our business growth, we also obtained retail sales data from the state of Michigan for comparison purposes.

Data collection is never perfect, so we applied data cleaning techniques to address missing values. It is also important to remove outliers for consistency, although in our case, outliers during the holiday season, such as New Year and Christmas sales, are significant as the business consistently reached new highs.

To gain a better understanding of the characteristics of our gathered data, we plotted it as a time series using Python libraries such as pandas and matplotlib. Through the seasonal decompose method, we determined the percentage of the trend, seasonal, and random error components in our data. The results showed that our data exhibits a strong trend component at 42%, a significant seasonal component at 24%, and a random error component at 34%.

Summary statistics of our data include:

A screenshot of a computer

Description automatically generated

A graph showing sales and sales

Description automatically generated

Fig 1: Sales data Histogram

## Forecasting Approaches Used:

A diagram of a forecasting method

Description automatically generated

Fig 2: (Snippet from Textbook)

The operational phase of the forecasting process involves four key steps:

* **Collect Data:** Gather historical sales data and relevant information like marketing spend, economic indicators, and competitive trends.
* **Examine Data Patterns:** Look for trends and seasonality in the data to inform the choice of a suitable forecasting method.
* **Choose Forecasting Method:** Select from various forecasting methods based on the business's characteristics and data, considering strengths and weaknesses.
* **Determine Accuracy:** Evaluate forecasting method performance on a held-out test set to ensure generalizability and effectiveness on future data.

Forecasting is an iterative process, which must be regularly repeated due to rapidly changing business environments. For example, a retail company may forecast sales for a quarter by collecting historical data; identifying trends such as higher holiday season sales; selecting an appropriate method to account for seasonality and trends; assessing accuracy; and proceeding with forecasting. To stay competitive, businesses must continually re-evaluate and modify their forecasts. Forecasting is essential for making informed decisions regarding inventory, staffing, and marketing, making it a valuable tool for businesses of all sizes.

### 4.1. Exploring forecasting techniques that better fit our data:

A graph with blue lines and numbers

Description automatically generated

Fig 3: Basic Time series plot for 1-year sales data

Upon visualizing the time series plot of the sales data, we have observed Seasonality, Trend, and Variability. Seasonal pattern of recurring peaks and troughs throughout the months however prominent peaks are in December and July with an upward trend in overall sales over time suggesting a potential autocorrelation. There is a sharp increase in sales starting from October 2022, peaking in December 2022, and a continued strong trend in subsequent months. Variability, fluctuations in sales observed in certain months.

Understanding these patterns can be valuable for:

* Strategic planning
* Inventory management
* Anticipating future sales trends
* Conducting further statistical analysis and modeling for deeper insights.

This forecasting helps in strategizing inventory management, the company can leverage the graph to forecast monthly product sales, facilitating a balance where sufficient inventory meets demand without excess that might tie up funds or risk obsolescence. Also helps to plan marketing campaigns. The company can use the graph to identify the months when sales are typically the lowest. This is a good time to run marketing campaigns to boost sales and train employees when it is slower to better their performance in the busiest months.

#### 4.1.1. Autocorrelation for sales data

A graph with lines and numbers

Description automatically generated

Fig 4: Checking Autocorrelation

Echoes in the Data: Autocorrelation Signals Sales Trends

The rhythmic rise and fall of sales is evident in the time series plot. Here, we see the “ghost of past performance”. The autocorrelation analysis reveals this, confirming its presence through positive ACF values and a robust LBQ statistic. This means that past sales figures cast a long shadow, shaping the landscape of future sales. This realization ushers in a new era for sales forecasting, where time series models that embrace the past are imperative to predicting the future with greater accuracy.

#### 4.1.2. Naïve forecasting Model

Naive forecasting is an approach that relies solely on recent information for prediction. The naive forecast assumes the current value of the series is 100% accurate. In short-term weather forecasting, this method is commonly referred to as the “no change” forecast. The naive forecast rapidly tracks changes but also can track random fluctuations mistakenly as fundamental changes.

A screenshot of a table

Description automatically generated

The calculated MAPE is 0.0879. From this result, the following conclusions can be drawn:

The Naive technique demonstrates accurate predictions, evident from the small mean absolute percentage error (MAPE) of 8.79%. A smaller MAPE is desirable for accurate forecasting.

Given the consistent and predictable nature of the data, it is probable that the Naive method performs better than more sophisticated forecasting approaches.

#### 4.1.3. Single Exponential smoothing Forecasting Technique:

A screenshot of a computer

Description automatically generated

Fig 5: Single Exponential Smoothing

For single exponential smoothing or Holt’s Method, MAPE is 0.08, or 8%. Ideally, any forecasting technique is considered to be good if MAPE is < 5%

A graph of a graph of sales

Description automatically generated with medium confidence

Fig 6: Smoothing plot for Single Exponential Smoothing

The chart indicates that the company's sales exhibit a reasonably robust and foreseeable pattern. The company can strategically anticipate the seasonal fluctuations and benefit from the encouraging upward trajectory. Nevertheless, it is crucial for the company to acknowledge the fluctuations in the data and remain adaptable, ready to modify plans as necessary.

Several graphs and charts

Description automatically generated

Fig 7: Residual plot for Single Exponential Smoothing

A couple of observations from the residual plots, Forecasts tend to be higher than actual sales due to the method's emphasis on recent upward trends. Accuracy is better for recent months as this method is more suitable for short-term forecasting.

It is less accurate for earlier months since older sales data is given less weight and is less relevant to current trends.

The residual plots indicate that single exponential smoothing is a fairly suitable match for the sales data, but there's room for enhancement. Considering a more advanced method, like double exponential smoothing, may improve forecast accuracy, especially for earlier months.

**4.1.4. Holt’s Model:**

A graph with blue and red lines

Description automatically generated

**Holt’s method of Forecasting:**

**A screenshot of a computer

Description automatically generated**

Fig 8: Holt’s method of SmoothingA graph with numbers and lines

Description automatically generated

Fig 9: Holt’s method of Smoothing plot

The forecasting approach does not account for the data's seasonal pattern. As a result, forecasts made during the peak and trough months might be less accurate. The data's slight upward trend is not considered by the forecasting method. Future sales could be underestimated because of this.

There is some volatility in the residual plot, particularly for the earlier months. This implies that there may be occasional errors in the forecast values.

Even MAPE value suggests that the method doesn’t accurately forecast future sales.

#### 4.1.5. Winter’s Forecasting Technique:

A screenshot of a computer

Description automatically generated

Fig 10: Winter’s Method for sales

Forecasts for 2024 using Winter’s method as follows:

A screenshot of a graph

Description automatically generated

Fig 11: Forecasts for 2024 using Winter’s Method

A graph with numbers and lines

Description automatically generated with medium confidence

Fig 12: Winter’s Method plot for sales

From the plot, Winter’s forecasting method appears generally effective, closely aligning forecasted and actual sales values. However, it overlooks the data's seasonal pattern, potentially impacting accuracy during peak and trough months. Additionally, the method doesn't account for a slight upward trend, risking underestimation of future sales. Residual plot volatility, especially in earlier months, indicates occasional inaccuracies. While the method is reasonable, there's room for enhancement; employing a more advanced model like seasonal ARIMA could refine forecast accuracy.

Several graphs and charts

Description automatically generated

Fig 13: Residuals plot using Winter’s Method

Based on the residual plot, there is still room for improvement and hence the following suggestions can be made to enhance the forecasting model:

1. To identify the seasonal pattern in the data, use a seasonal ARIMA model.
2. To consider the data's slight upward trend, apply exponential smoothing with a trend component.
3. Monitor the forecast's accuracy and tweak the forecasting model as necessary.

#### 4.1.6. Trend analysis for our data:

A close up of a number

Description automatically generated

A graph with blue and red lines

Description automatically generated

Fig 14: Trend analysis plot

The trend analysis graph shows a dynamic picture of sales, showcasing a well-managed business with rhythmic ups and downs. The smooth curve suggests steady customer flow and revenue, while subtle fluctuations hint at underlying factors like seasons and marketing dollars spent or not spent. Notably, the consistent upward trend reflects business growth and adaptability. The cyclical nature indicates seasonal patterns, offering insights for inventory, staffing, and marketing planning. Overall, the graph portrays a healthy, growing business, but a comprehensive analysis should consider factors like market share and profitability.

This depicts a harmony of seasons telling the story of success and resilience in business.

#### 4.1.7. ARIMA model:

A graph with lines and numbers

Description automatically generated

Fig 13: Trend analysis plot

It is clear from the above picture that the autocorrelation coefficients slowly reduce as we move to the right though the data remains quite high for the given periods which indicates a significant autocorrelation.

A graph with lines and numbers

Description automatically generated

Fig 14: Partial Autocorrelation function for sales

From the above figure, there is a noticeable partial autocorrelation that is evident, with a peak at lag 1. Therefore, it can be said that the provided data requires the first difference. Then we take the log e for the retail store sales values to smoothen the data fluctuations. The significant lag at 12 period in the autocorrelation function of the first difference gives us a hint that the fit seasonal model period should be 12. The autocorrelation function and partial autocorrelation function for the logarithmic sales values reflect the same plots with significant autocorrelation indicated by slowly reduced values as we move to the right and there is significant peak at lag 1 for the partial autocorrelation indicating significant autocorrelation. Hence the ARIMA (0,1,0) (0,1,1) 12 should be used for the given data.

A graph with lines and numbers

Description automatically generated

Fig 15: Autocorrelation function for Ln sales

A graph with lines and numbers

Description automatically generated

Fig 16: Partial Autocorrelation function for Ln sales

A screenshot of a computer

Description automatically generated

A screenshot of a graph

Description automatically generated

Differencing: 1 regular, 1 seasonal of order 12

Number of observations: Original series 48, after differencing 35

A close-up of a number

Description automatically generated

A screenshot of a graph

Description automatically generated

A graph with lines and numbers

Description automatically generated

A graph with lines and numbers

Description automatically generated

In the above figures we do not see significant autocorrelation with which we can say that the model is best fit for given data as the series is stationary. For further confirmation we will test the level of significance as from the output chi-square value at lag 12 is 16.28 and the p-value is 0.092 which is greater than the level of significance of 0.05, concluding that model is good. The forecast values for the next three months are 11.71, 11.83, and 11.65 for November, December and January respectively.

A screenshot of a data table

Description automatically generated

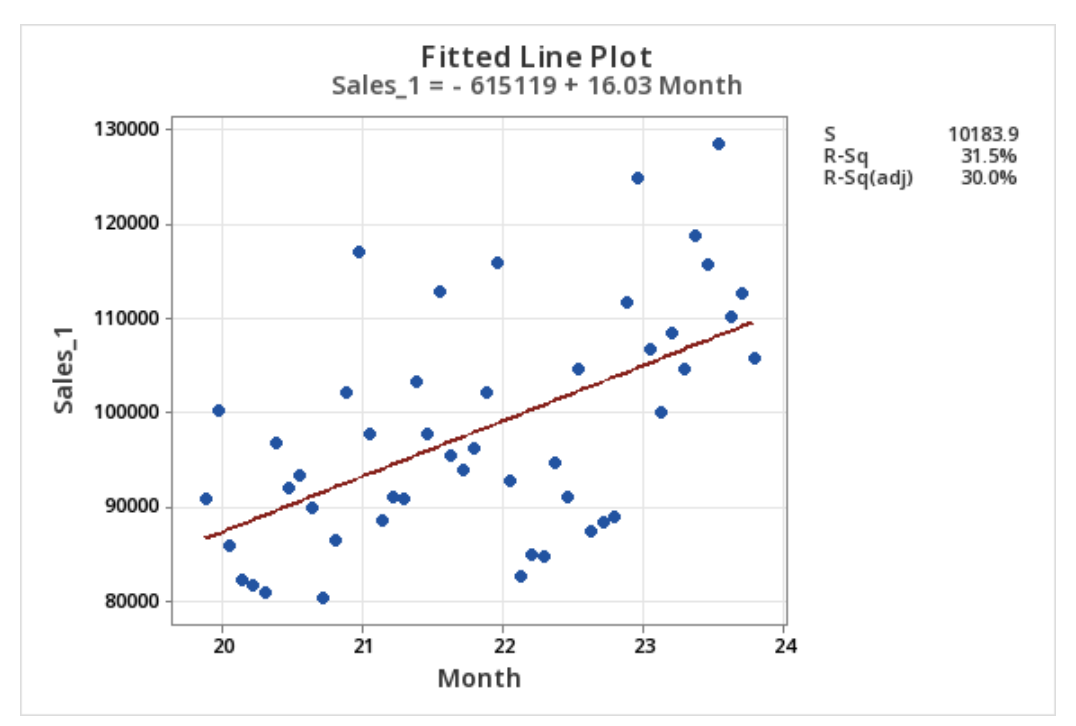
A graph of residual plots

Description automatically generated with medium confidence

#### 

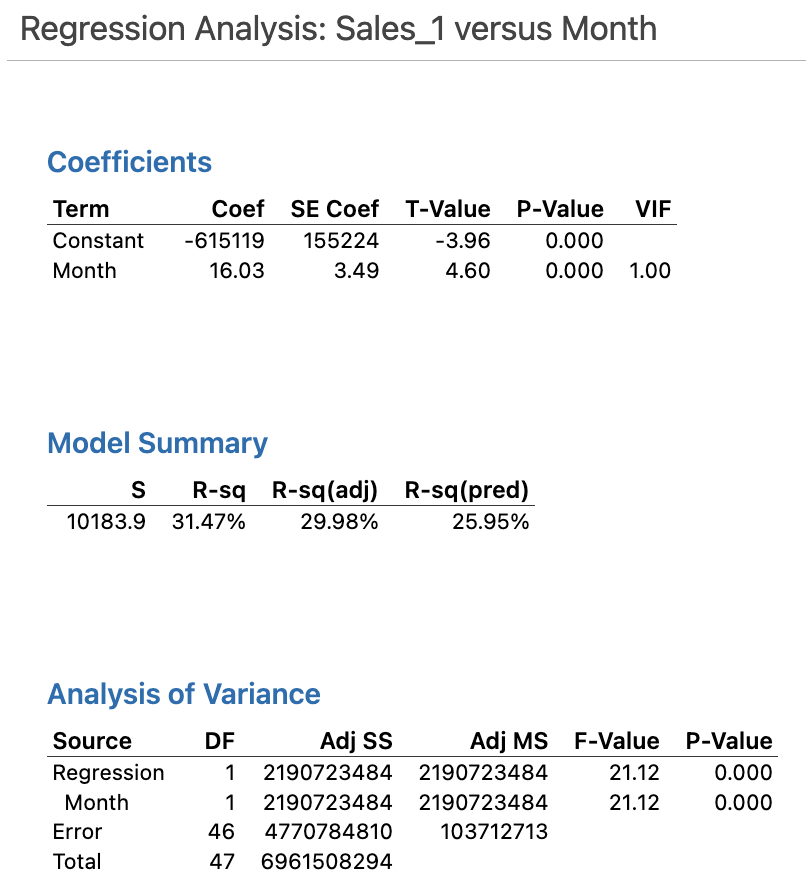
#### 4.1.7. Regression Analysis:

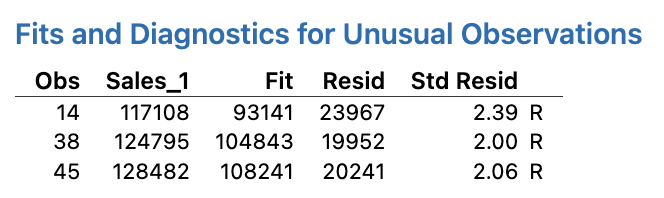
#### 



Above, the intercept term is -615,119, which represents the value of y when our month is 0. However, when considering months, month “0” is not meaningful. Here, it is more valuable to interpret the slope coefficient.

The slope of 16.03 tells us for each additional month, the sales data is expected to increase by 16.03 units. Which assumes a linear relationship between the variable month and sales data.

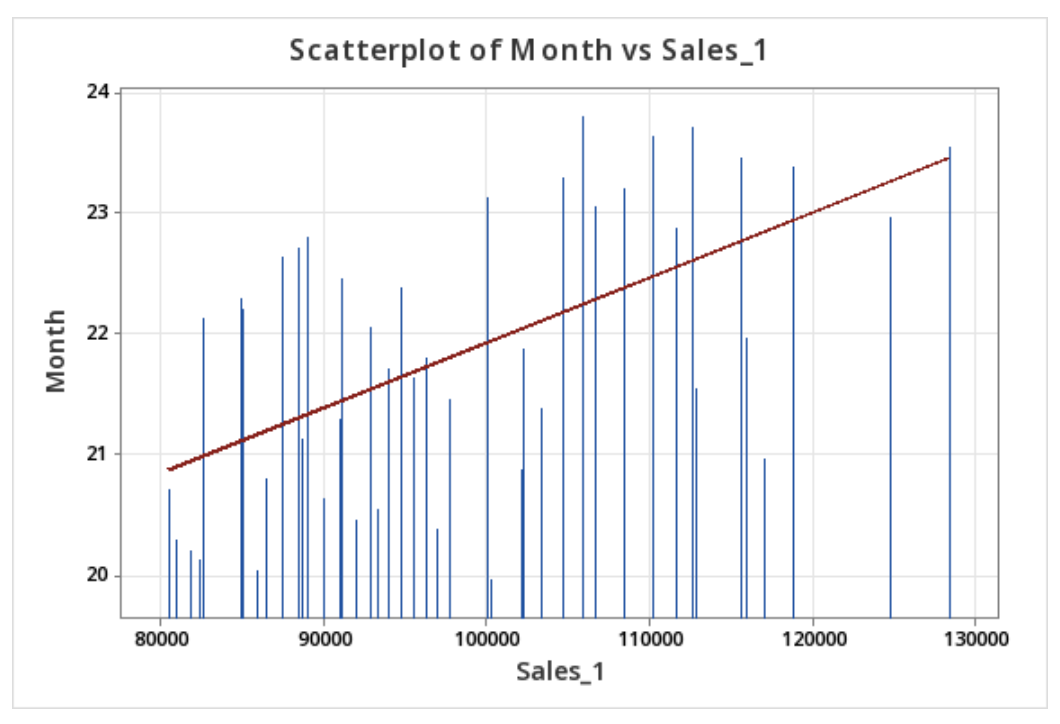






The Durbin-Watson Statistic of around 1.21 indicates that there is autocorrelation.

#### 4.1.8. Scatterplot of Month vs Sales



In the above scatter plot, we can see the red trend line which shows there is a positive correlation between months and sales, as months increase (or go on) sales increase as well. The blue bars show the variation in sales from month to month. You can also see that as we approach months 23 and 24, AKA months November and December (of year 2) the sales are higher, indicating that sales are increasing before/around holiday months.

## Analysis of Output Results & Recommendations:

Domain knowledge serves as the foundation for judgmental forecasts and modifications. It refers to any information, apart from time series data, that is relevant to the forecasting process. In simpler terms, non-time series information is considered as domain knowledge. In theory, incorporating more data should lead to more accurate forecasts. Hence, projections that combine domain expertise and historical data are expected to outperform those that rely on only one type of information. This is the case with the ARIMA model. If we have validated the forecasted sales, we can analyze the patterns of previous years around the same time. This gives us some understanding of how the sales are likely to be projected.

After reviewing the results of the Time Series Plot, there are many thoughts that ABC Store should consider and implement, if not already doing so, moving forward. One main consideration being the seasonality of their business. Things that ABC Store can do to prepare better knowing they see more foot traffic and produce a higher sales volume is to prepare early. ABC Store can do things strategically such as staff their business accordingly for these peak times. They could also put an emphasis and focus on training their employees in the times proven less busy. These things can ensure that their customers are happy and being helped while shopping, hopefully creating a sense of brand loyalty for future sales, making ABC Store the first store that comes to their minds’. Furthermore, ABC Store can ensure that their inventory is more properly planned moving forward, ensuring that their top-selling items are in stock and available during these busier times. Additionally, implementing marketing campaigns and special promotions, catering toward those top-selling products can help get even more customers in the door.

When analyzing the autocorrelation, seeing that the data shows a positive ACF and robust Ljung-Box (LBQ) statistic, this suggests the presence of serial autocorrelation in the time series data. The student group recommends that ABC Store ask for these insights on a more regular basis. The reason for this is because they could plan marketing efforts and quickly see results post-campaigns monitoring performance on what worked and what did not.

Further recommendations for ABC Store’s management include converting from storing the data manually for the laundromat to converting to digital which helps the management to analyze how well the laundromat business is doing in real time. Additionally, they can use that to see how much the laundromat influences the sales of their store. Also, it is suggested for the business to focus on people from Plymouth, Livonia, and Redford as these three areas constitute the majority of the customer base. By doing that, the advertising revenue will be less as they are targeting a smaller radius. Social media platforms provide an opportunity to target a specific audience such as men or women, and be age specific. From an objective standpoint, the minor renovation that they did in February turned out to be a good step which resulted in an increase in sales over the following periods.

The recommendations to ABC Store’s management includes converting from storing the data manually for the laundromat to converting to digital which helps the management to analyze how well the laundromat business is doing in real time. Additionally, they can use that to see how much the laundromat influences the sales of their store. With just the online presence people may not be able to recognize the business until they advertise and attempt to grow to reach new customers. Hence, we recommend the business to start advertising which will bring in new leads. Also, it is suggested for the business to focus on people from Plymouth, Livonia, and Redford as these three areas constitute the majority of the customer base. By doing that, the advertising revenue will be less as we are only targeting a small radius. Additionally, social media platforms provide an opportunity to target a specific audience such as men or women, and be age specific. From this project’s objective standpoint, the minor renovation that they did in February turned out to be a good step which resulted in an increase in sales over the following periods.

## Future work:

In later projects with ABC Store, we will analyze the impact of renovating stores based on sales volume. This analysis will help determine the extent to which the addition of a laundromat contributes to an increase in sales. The renovation includes upgrading the washers and dryers in the laundromat, which will attract customers and benefit the retail store.

To assess the success of the laundromat and our advertisement spending, it is important for management to track these factors. This information is crucial for various aspects of the business, including forecasting future sales. Without tracking this data, we cannot assess the company's financial health, make strategic decisions, or evaluate overall growth or decline. Understanding and analyzing this data is essential for making informed decisions that have the potential to grow the business. Our future work depends on it, so we strongly encourage ABC Store to button-up on their current practices.