Analytical Solution

SAFFRON RETAILERS LTD.

PRESENTED TO:
EXECUTIVE DIRECTORS

CONTACT

Andrew Da Costa
Binal Patel
www.saffronretailors.ca

Addressing the diminishing sales trend spanning the past four years necessitates a multifaceted approach. This encompasses a synthesis of techniques including Exploratory Data Analysis (EDA) to uncover patterns and anomalies, trend analysis to understand the trajectory of the decline, and the potential utilization of predictive modeling to forecast and decipher contributing variables. This amalgamation of methodologies aims to holistically dissect the complex factors influencing the observed decline in sales and offer practical business solutions that would increase the market performance of Saffron Retailers if they were to be fully implemented.

Data Cleaning and Preprocessing:

Initially, a CSV file was imported into Visual Studio Code, yielding a dataset containing 15 columns and 9993 rows of historical transactional data. It was evident that certain rows contained missing data, prompting the elimination of those problematic entries. Additionally, to enhance data quality and integrity of the analytical solution, the rows with inadequate or nonsensical quantities were eliminated, recognizing that values for sales, quantities of items sold, or individual unit prices, cannot logically be zero, negative, or string values. Any instances of duplicated data were scrupulously identified and removed from the dataset. As a result, the dataset underwent a reduction in size to 9986 rows.

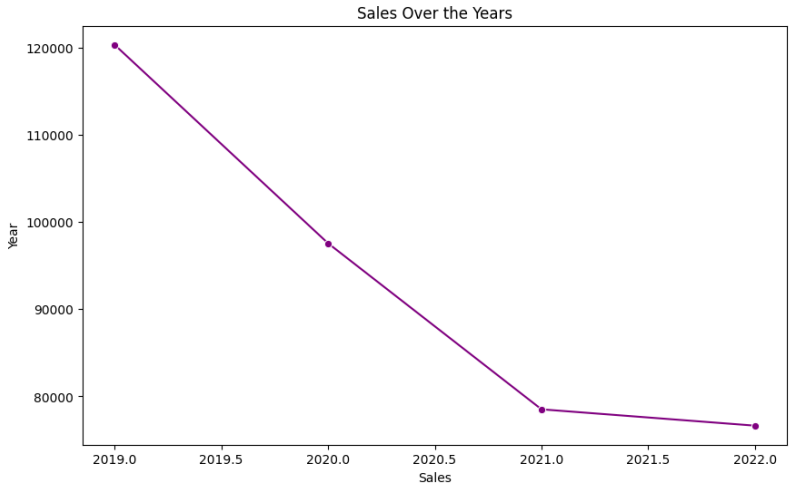
Following this, the number format for units sold, price of individual units, and sales, were all converted into a standardized format. The currency symbols were removed, and these values were converted into the cost format, allowing further and easier manipulation of these columns during the analysis. The process of cleansing categorical attributes was the next action taken. Any extraneous white spaces that might have been present at the beginning or end of the entries in categorical fields were removed. By undertaking this action, the consistency and accuracy of these descriptive fields is ensured, specifically when analyzing sales in various regions, or when analyzing the distribution of types of products sold at Saffron stores.

Although originally included in the dataset, some columns that are not relevant to this specific analysis were dropped from the dataframe. Since this study is not focusing on the individual customers themselves, columns such as OrderID, Customer Name, and Payment type were all deemed to be irrelevant for this analysis.

Lastly, after meticulously cleaning and transforming the data, the modified dataset was archived into a new CSV file, that could later be used in PowerBI to construct visualizations. This refined data set will also be further pushed through Python's predictive modelling capabilities. The process of cleaning and refining the data is now concluded, allowing further utilization and exploration of the data to understand Saffron's decline in sales, and the suggestion of solutions to increase profit margins. This meticulous approach underscores Saffron's commitment to working with a dataset that is both reliable and primed for meaningful analysis.

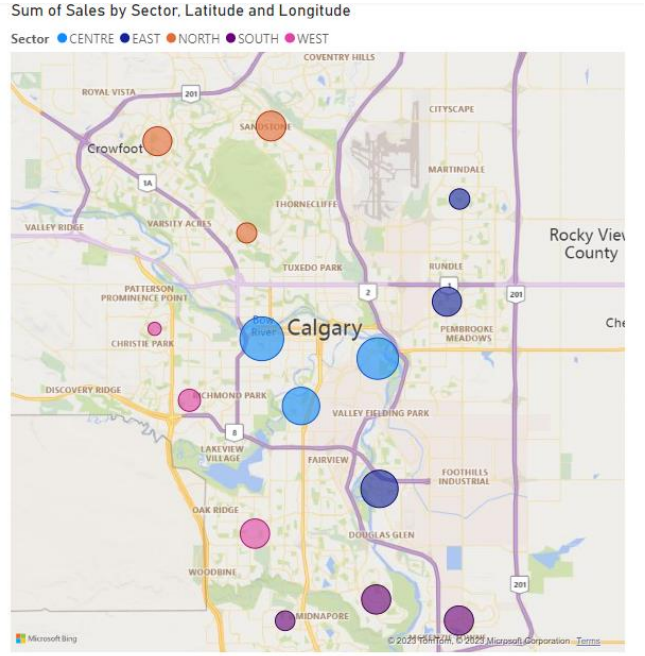
Exploratory Data Analytics (EDA):

It is important to extensively explore the historical transactional data, to better understand, as well as prove to the board of directors that there is in fact a decline in sales. Exploratory analysis will also help identify factors that may be the root causes of this decline. To prove that sales are declining, the most logical step would be to plot the summation of sales by year. The following plot represents the decline in sales.



The decline in sales over the past four years is evident based on the above plot. The sales significantly decreased from 2019 to 2022, so there is no doubt COVID played a major factor in reducing the foot traffic within the stores. The slope of the plot between 2021 and 2022 is small, therefore suggesting that the decline is starting to plateau, and the business is recovering from the effects of the pandemic. Saffron can take advantage of this knowledge, and with practical thinking, they could rebound from this decline in sales.

It is important to identify the sectors of Calgary that are not performing well. The following map, which shows the locations of all stores within Calgary, indicates the total sales of the store based on bubble size, for the year 2022. It is obvious that West Calgary has not performed as well as the other sectors. The potential solutions that could be implemented to bolster sales based on regional analysis will be explored later in this report.



It is important to understand how different product offerings impact net profits. The products that are positively and consistently contributing to the net profit should be recognized such that their sales can be maximized to generate increased profits. Monte Carlo simulations were run for each product category to uncover which products influence profits the most, the results are as follows.

Product Category: Bakery
Mean Simulated Net Profit: 44.748422530657606
Standard Deviation of Simulated Net Profit: 4.47945365063647

Product Category: Beverages
Mean Simulated Net Profit: 16.68231283475467
Standard Deviation of Simulated Net Profit: 1.6532045014168262

Product Category: Eggs, Meat and Fish

Mean Simulated Net Profit: 70.39025098454152
Standard Deviation of Simulated Net Profit: 7.30772065581987

Product Category: Food Grains
Mean Simulated Net Profit: 42.86459189814885
Standard Deviation of Simulated Net Profit: 4.11867806340860

Product Category: Fruits and Veggies
Mean Simulated Net Profit: 36.39395100750741
Standard Deviation of Simulated Net Profit: 3.727068507595947

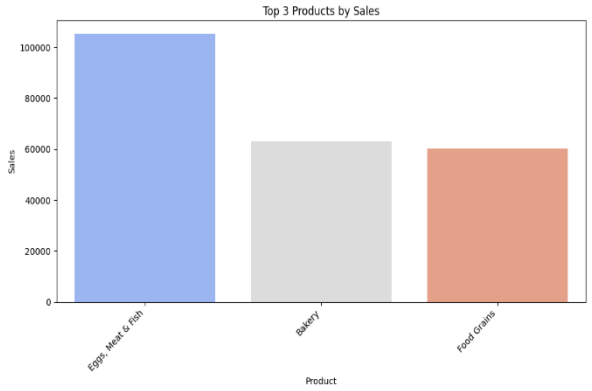
Product Category: Oils and Masala
Mean Simulated Net Profit: 26.83033318794355
Standard Deviation of Simulated Net Profit: 2.733162858152448

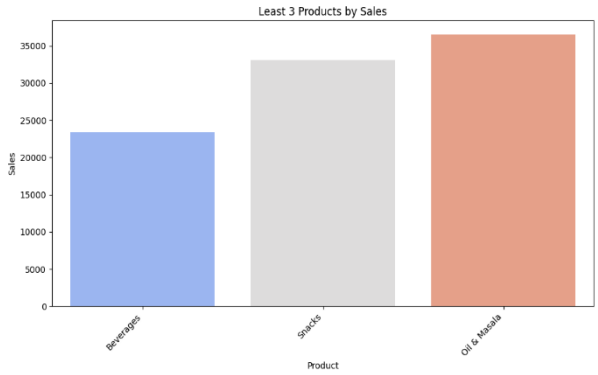
Product Category: Snacks
Mean Simulated Net Profit: 21.801511870461155
Standard Deviation of Simulated Net Profit: 2.3181334445268176

Based on the insights gained above, the focus should be on promoting and selling product categories that have higher mean simulated net profits and lower standard deviations. For instance, "Eggs, Meat and Fish" and "Fruits & Veggies" have higher mean net profits, indicating strong potential for driving net profits. These categories could be prioritized in marketing strategies.

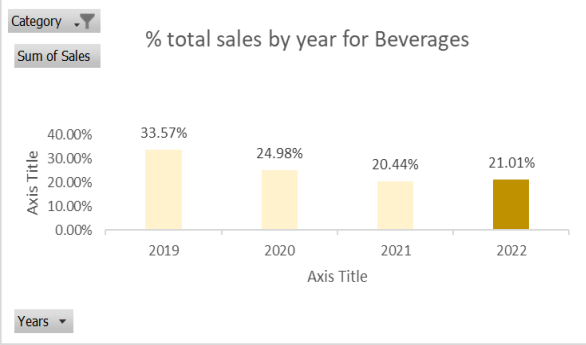
Analysis on Products:

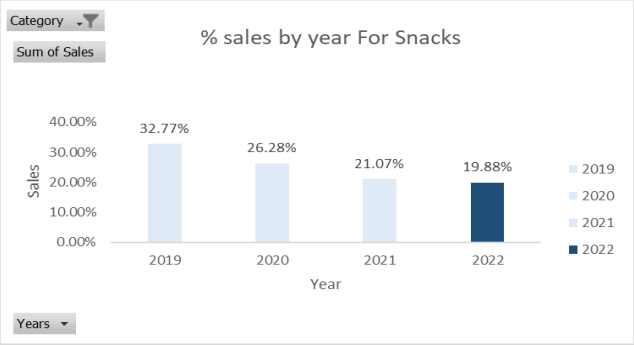
Through investigation the best-performing product category overall within Saffron's offerings emerged as eggs, meat and fish. Subsequently, the weakest performing product category overall is beverages.





Moving forward, a deeper exploration was undertaken, focusing specifically on the category of beverages. The objective was to gain insights into the yearly sales performance of beverages. Notably, the chart provided below shows a declining trend in sales of beverages over the years. However, an exception was noted in the year 2022, where sales figures surpassed those of 2021, indicating a slight upward trajectory during that period. With this analysis in mind, attention was then shifted to the penultimate category, namely snacks, to glean analogous insights.



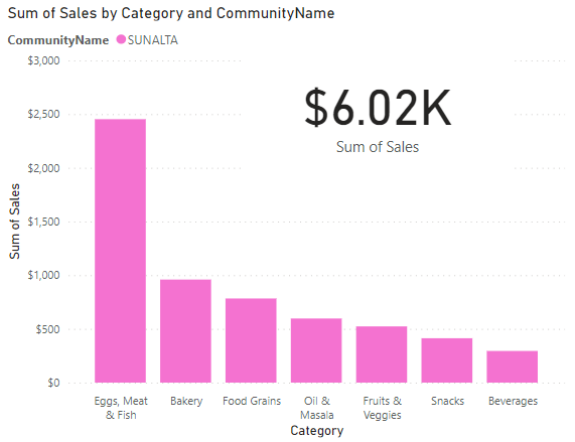


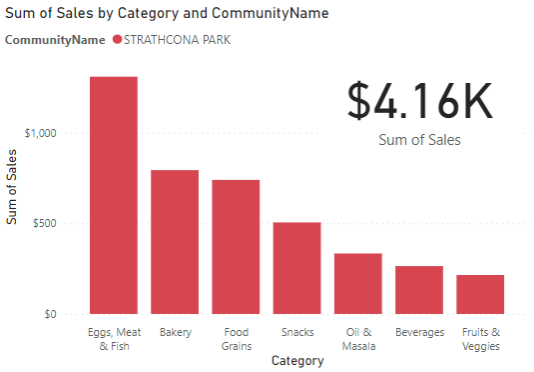
The evident observation is that sales within the snacks category are experiencing a consistent decline, signifying the necessity for a comprehensive and in-depth analysis to uncover the underlying factors contributing to this downward trend.

Based on the analysis above, it may be beneficial to simulate a scenario in which a product with declining sales, such as snacks, is replaced by a better selling product to bolster sales at a specific store. This is an important consideration as products that aren't selling take up important shelf space, and there is higher potential for low selling products to spoil or expire before leaving the store, effectively wasting the company's money.

One of the major objectives for this study is to determine how to increase the sales at stores such as Strathcona Park. Before implementing the previously mentioned scenario, it is important to note

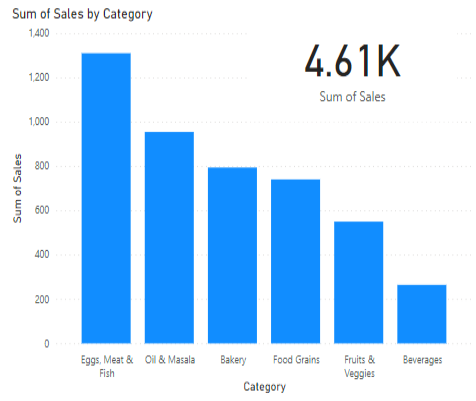
the sales distribution of products within specific stores. The following two bar charts, show the highest and lowest performing stores across Calgary in 2022.

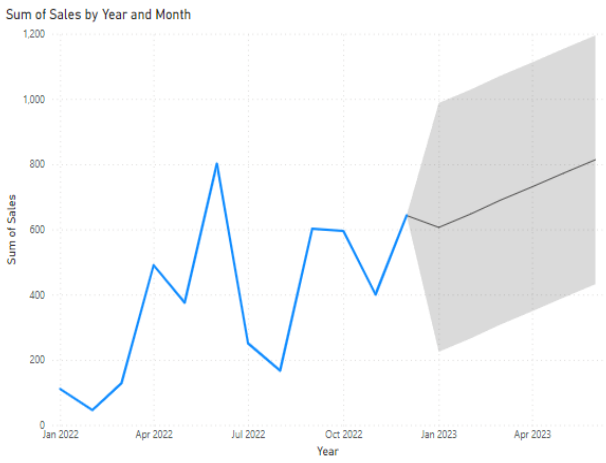




It is obvious that the Sunalta store significantly outsold the Strathcona Park store, across all products. However, Oil and Masala is selling relatively lower than expected at both stores, considering Saffron Retailers caters more to those in the Indian community. At the Strathcona Park store, Fruits and Vegetables are also unexpectedly the lowest selling product.

A few assumptions must be made so that the scenario is relatively practical and considerations as to its implementation can be made. Using data from 2022, the transactions associated with Strathcona Park can be altered, such that those with a declining product such as snacks are replaced by a higher selling product or a practical product such as Masala. This follows the assumption that the stores are still selling equivalent quantities of product that is physically on the shelves, regardless of what that product may be. However, the sales will be affected due to the difference in product prices. A six-month forecast into 2023 can also be developed. The following two plots are a representation of this simulated scenario with respect to the Strathcona Park Store.

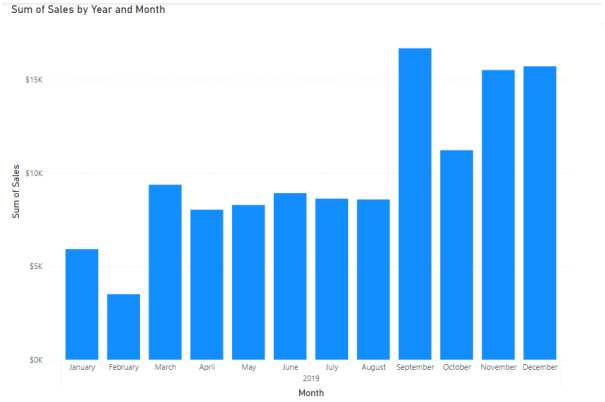


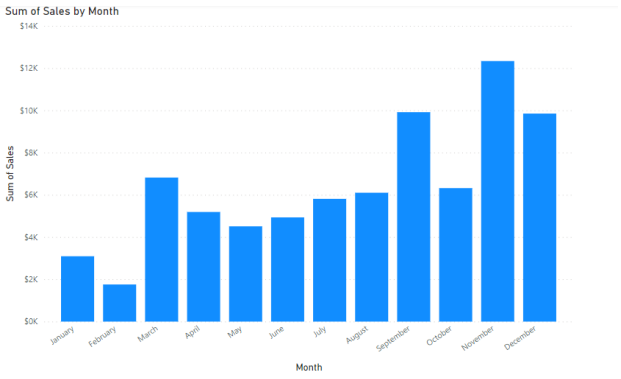


The transactions with snacks at the Strathcona Park store, were replaced by a combination of Masala, Fresh Fruits and Fresh Vegetables. As seen from the plot above, Masala is now the second highest selling product at this store, and the annual sales in 2022 increased when compared to the original scenario. This is important when considering that Saffron caters towards the Indian community, and Masala is a product that should generate high sales due its use in many Indian dishes. The forecast of six months into 2023 also shows that the store will generate an increase in profit. This scenario is just one of many that could be presented to the Board of Directors to convince them of changes that will increase profits based on an analysis of the products themselves.

Seasonality Analysis:

The following two charts display the sales of Saffron stores for 2019 and 2022. The trend for the years in between is also similar. It is obvious that around the time of March, and between September and December sales are higher than the rest of the year. This is important to note, as important Indian holidays or festivals, such as Holi and Diwali occur around these months. A seasonality analysis is import because the business should understand that certain products need to be consistently fully stocked, specifically products such as Masala. This will satisfy the product demand around these important celebrations for the Indian community living in Calgary.





An important assumption to make based on these plots, is that the same seasonality trend will continue beyond 2022. This becomes apparent when comparing the four-year gap between 2019 and 2022. This is a business solution in its own right, because the Board of Directors should be aware of these seasonality trends and appropriately compensate by ordering an excess of products that are known to sell higher during these months.

Association Rule Mining:

Another angle of analysis performed on the historical data set was determining what the most common combinations of products were that customers tend to purchase together. Association Rule mining is an important part of Machine Learning, it can be used to create targeted product bundles or promotions to increase sales of products by understanding how customers strategically shop for products. The results of Association Rule mining are below.

1. Most Frequent Product purchased with other products: 'Snacks' with a Support value of 0.6124595469255664 (61.2%).

This is particularly interesting, because the support value of Snacks is high, but based on previous analysis, the sales of snacks has been declining over the years.

2. Second Most Frequent Product: Beverages with a Support value of 0.6067961165048543 (60.7%)

This is also interesting as the overall sales for Beverages is the lowest of all products.

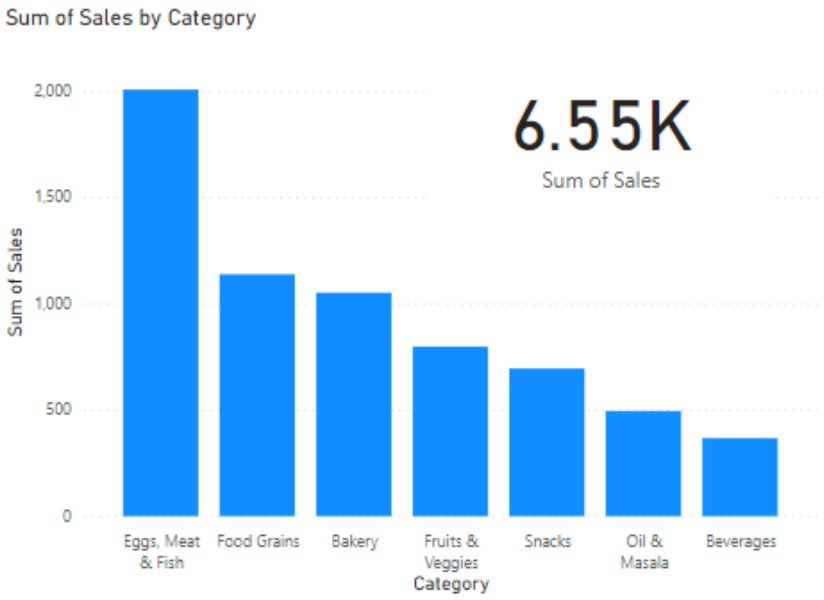
3. Third Frequent Itemset: Fruits and Veggies with a Support value of 0.6059870550161812 (60.6%)

Regional Analysis:

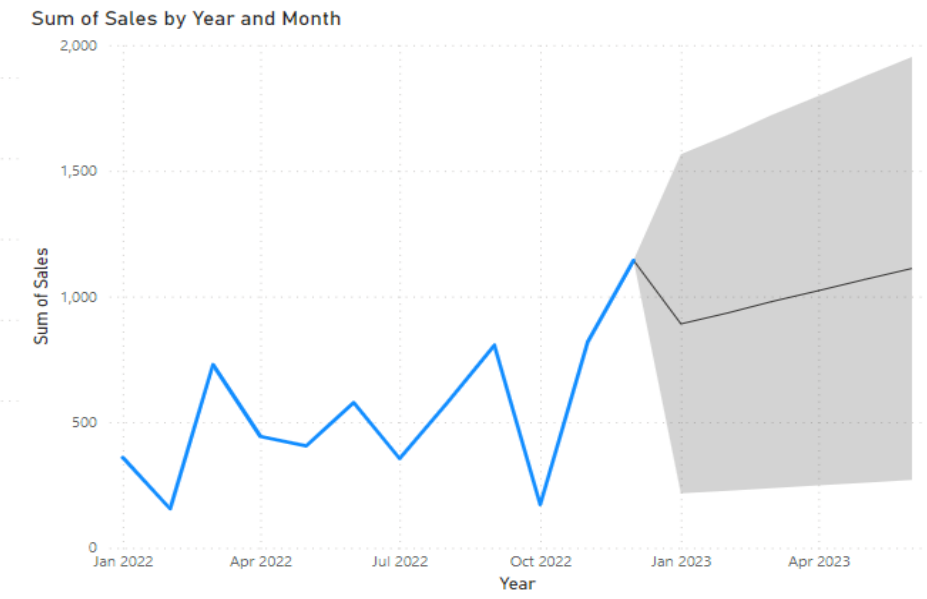
As previously stated, the store in Strathcona Park, is significantly underperforming when compared to other stores in Calgary. In fact, there is one store in each of the north, south, east and west sectors that is underperforming. The number of brick-and-mortar stores may be too significant. Therefore, it is important to simulate a scenario in which these underperforming stores are sold off, and the redistribution of resources is made across the company to potentially bolster sales at other locations. Having fewer physical locations may boost sales as there will be an increase in customer foot traffic from those that live further away, and those customers may purchase larger orders to reduce the frequency of the long commute. Although this angle is not considered in this analysis, there is the obvious advantage of gaining money from the sale of a store and reducing the distribution/transportation costs by sending product to fewer stores.

Another major objective of this study is to determine how to increase the sales while looking from a regional sales perspective. This report will offer one scenario, that could be considered among many others, that simulates changes made on a higher, regional level, rather than simply making changes at individual locations or with products sold on shelves. If Saffron Retailers were to sell their location at Strathcona Park, rerouting 50% of their customers to the location in Glamorgan and the other 50% to the location in Charleswood. A comprehensive study could be performed on the raw sales data, and the distribution of product types sold after the influx in customers to these locations.

Using the data from 2022, the transactions from Strathcona Park were randomly split between Charleswood and Glamorgan, assuming customers would have no real preference as to what store they go to, as they are both approximately equidistant from the Strathcona Park area. An assumption here would be that the customers would have the same order regardless of which store they go to, and therefore the transactions can simply be added to the sales data of a particular store. What is important to note here, is that predictive modelling or forecasting can be performed on this alternative scenario to understand how the sales would increase six months into 2023 based on these business changes. The following plot describes the total sales, as well as the product distribution, at the Charleswood store. It is safe to assume that the Glamorgan store will have similar results.



As expected, the sum of sales increased for the Charleswood store as it gained new customers, and each product sold in greater quantities. A predictive or forecast model can be used to understand how this new distribution of product sales will impact the yearly profits.



The forecast of six months into 2023 at the Charleswood store, shows that the store will eventually generate profit as there is a slight trend upwards. This slow rise is most likely a result of less frequent trips to Saffron stores from those customers that live further away. The customers are still coming to the store, but the generation of profits is slower. This business solution, based on regional analysis, could be presented to the Board of Directors to convince them of changes that will potentially increase profits.

Risks vs. Rewards

Theare are a few risks associated with this project. Rather than implementing a solution without further studying its implications, Saffron Retailers chose to risk approximately four months of declining business operations, to conduct an extensive and thorough analysis on how they can not only recover from the decline in sales, but also take advantage and prosper. The longer the analysis takes, the more money the business will effectively lose as there would be no practical solutions put into place. Another risk taken with this sort of study is that the implementation costs of certain solutions may be more than the business can afford. There are risks in deciding which approach to use for increasing sales, some may take longer to complete, cost more money, require a higher skill level and therefore increases hiring costs, and some may lead to unachievable solutions.

The rewards associated with the project potentially outweigh the risks with respect to this project. Of course, one of the most valuable rewards would be uncovering a solution to the declining sales. After the analysis has been completed, the Board of Directors will be presented with suggestions

to improve on the profitability of the business, as well as on customer satisfaction. The project will reward the business with the tools and knowledge to implement inherent solutions at their stores, ultimately driving the success and growth of Saffron Retailers Ltd. There is also potential to improve on the data fluency and culture of the entire organization. This would not only be considered a reward for this project, but for all future projects as well, as the ability to conduct these studies will be greatly improved. After implementing solutions, there is potential to increase the brand power of Saffron Retailers within Calgary.

Conclusion

After thorough analysis of the transactional data, there were significant insights uncovered that will both explain the decline in sales and suggest advantageous solutions to the board of directors. Saffron's highest earning products emerged as eggs, meat and fish, whereas the weakest earner overall is beverages. Through analysis however, it was determined that snacks were significantly declining in sales on a yearly basis, and a potential candidate to be replaced at stores with a more consistent product. An important discovery was that the yearly sales show seasonality trends. The times of the year that have higher sales are generally around the months of Indian holidays, and the company can use this to their advantage when considering stocking and selling specific products. Association Rule mining was also performed, and it was deemed that the products that were most often purchased together with another product are Snacks with a support value of 61.2%, Beverages at 60.7% and Fruits and Vegetables at 60.6%.

It was uncovered that at least one store in each sector represented by the cardinal directions was not performing well, and ultimately leading to the decline in sales for the company. The weakest performing store, Strathcona Park, was a perfect candidate to conduct some analyses on, to better understand how to improve business performance. After developing and modelling a scenario in which snacks were replaced by masala and fresh fruits and vegetables at the Strathcona Park store, a significantly higher profit was forecasted to occur six months into 2023. The other scenario that was developed involved splitting the customer base that shopped at Strathcona Park and redirecting them to the Charleswood and Glamorgan stores. Although the potential increase in sales was not as significant in this scenario, there was a slight increase in profits noted by the forecast into 2023. These two scenarios may be the solutions the board of directors needs to consider when deciding how to increase Saffron Retailers market presence and grow into a prominent player in the grocery industry.

Business Questions

After analysing and predicting we are ready to answer business questions

1. What are the overall sales trends over time, and are there any seasonal patterns or fluctuations?
   In general, the sales are declining over time, and therefore solutions must be put into place to increase profits. There are however seasonal trends when looking closer at each year. These seasonal patterns tend to occur around Indian Holidays, where the increase of sales occurs.

2. What are the most common combinations of products (itemset) that customers tend to purchase together? How can this information be used to create targeted product bundles or promotions to increase sales?
   Using **Association Rule Mining** in Machine Learning We able to extract some information about our data set

   Most Frequent Itemset: 'Snacks'
   Support: 0.6124595469255664 (62%)
   The output is showing that the item 'Snacks' is the most frequent item in our dataset, appearing in around 61.2% of the transactions. This information is valuable because it indicates that 'Snacks' is a popular product that customers tend to purchase frequently.
   Second Frequent Itemset: Beverages
   Support: 0.6067961165048543

   third Frequent Itemset: Fruits Veggies
   Support: 0.6059870550161812

Items	Support
Snacks	62%
Beverages	60%
Eggs Fruits	43%
Snacks Vegetables	42%

3. How do different product offerings impact net profits? Can we identify specific products that consistently contribute positively to net profit? Should we focus on promoting and selling these products more?
   To address this problem, we ran **monte Carlo stimulation** for each category and found interesting results which are as follows
   Product Category: Bakery
   Mean Simulated Net Profit: 44.748422530657606
   Standard Deviation of Simulated Net Profit: 4.47945365063647

   Product Category: Beverages
   Mean Simulated Net Profit: 16.68231283475467
   Standard Deviation of Simulated Net Profit: 1.6532045014168262

   Product Category: Eggs Meat Fish
   Mean Simulated Net Profit: 70.39025098454152
   Standard Deviation of Simulated Net Profit: 7.30772065581987

   Product Category: Food Grains
   Mean Simulated Net Profit: 42.86459189814885
   Standard Deviation of Simulated Net Profit: 4.11867806340860

   Product Category: Fruits Veggies
   Mean Simulated Net Profit: 36.39395100750741
   Standard Deviation of Simulated Net Profit: 3.727068507595947

   Product Category: Oil Masala
   Mean Simulated Net Profit: 26.83033318794355

Standard Deviation of Simulated Net Profit: 2.733162858152448

Product Category: Snacks
Mean Simulated Net Profit: 21.801511870461155
Standard Deviation of Simulated Net Profit: 2.3181334445268176

Based on the insights gained, we can consider focusing on promoting and selling categories that have higher mean simulated net profits and lower standard deviations. For instance, "Eggs Meat Fish" and "Fruits & Veggies" have higher mean net profits, indicating strong potential for driving net profits. These categories could be prioritized in marketing strategies.

4. What strategy was employed to address the issue of declining sales?
n response to declining sales, our approach was to concentrate efforts on two specific store locations, namely Strathcona Park and Charleswood stores. By analyzing sales data for the year 2022, we identified the underperforming category and strategically transformed it into the top-performing category. This adjustment was made with the aim of revitalizing sales. Subsequently, utilizing this revamped strategy, we forecasted sales for the upcoming six years. As evident in the provided data, this proactive initiative resulted in a modest upward trend in sales, signifying a positive trajectory.

	2022	Predicted 2023
Strathcona Park	6030.6	6909.3
Charleswood Park	4610.3	5649.7