

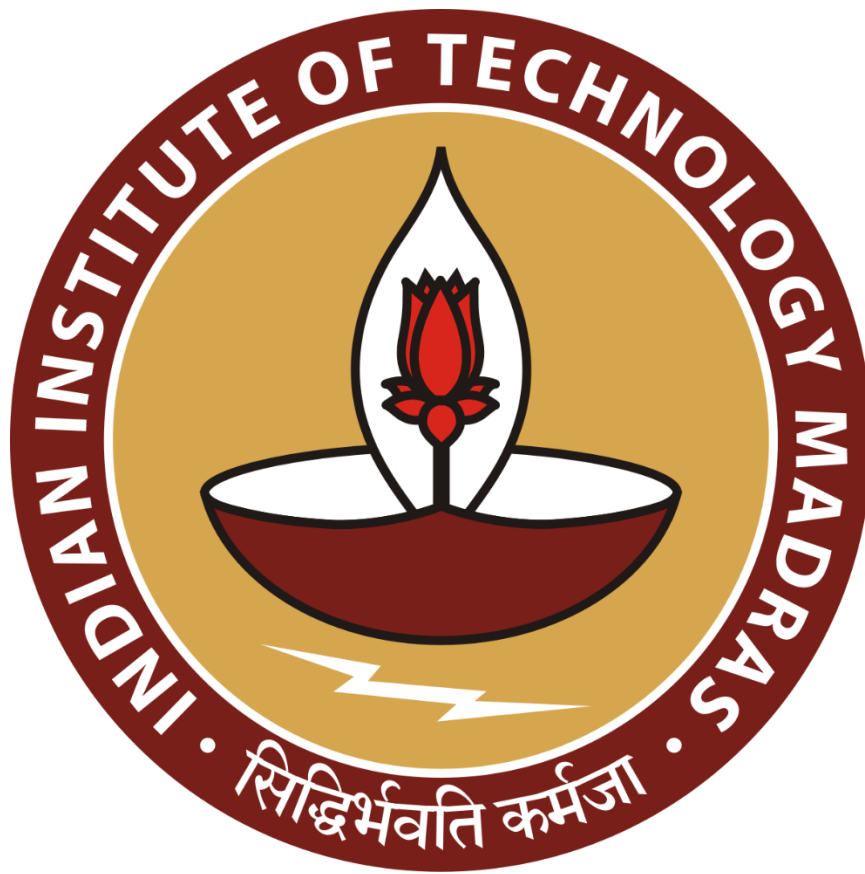
INVENTORY MANAGEMENT AND PARETO ANALYSIS OF A RETAIL LIQUOR STORE

Final report for the BDM Capstone Project

Submitted By:

Name: Harmanpreet Singh

Roll No: 21F3002026



IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

1. Executive Summary	3
2. Detailed Explanation of Analysis Process	4
2.1 Data Collection	4
2.2 Data Cleaning	6
2.3 Data Analysis	7
3. Results and Findings	10
3.1 Objective 1	10
3.2 Objective 2	15
3.3 Objective 3	16
4. Interpretation of Results and Recommendations	19

Executive Summary

This report focuses on a small retail liquor store chain, Legacy Liquor, a B2C business company, owning 4 stores all over Calgary, AB, Canada. Their stores specialize in vintage wines, premium scotch whiskeys and craft beers.

Some of the hiccups being faced by its latest store are, the store is stocking excess inventory, the products which have stopped serving as a commodity of choice and merely occupy the shelf area, thus holding in capital. Moreover, due to some poor organization, the store faces constant stock-outs of significant products and lastly, it is experiencing some decline in daily revenue.

The above stated issues are to be analyzed in-depth, more specifically, a proper system must be devised which could classify items serving as selection/choice from those acting as dead stock. In addition, in identifying most prominent products, a pareto analysis is the go-to tool. With key findings kept in mind, a possible solution addressed in the form of recommendations, which perhaps may ease the store's day to day functioning. Finally, the expected outcomes of this project, could potentially pinpoint store's management inaccuracies and the third issue, if tackled carefully, could not only identify the cause behind decline in sales, but also prevent it.

Detailed Explanation of Analysis Process

Data Collection

Every analysis, report, project, encompassing few steps or several stages, on a small business or a whole organization, starts with the process of data collection. It is a very crucial step in the chain of data analysis. In this project, primary data was collected from the company's newest addition, a retail liquor store by the name of Legacy Liquor, same as the company. The store, like all retail stores, has a point-of-sale (POS) system, integrated with debit machine and barcode scanner, for the cashier's ease and inventory management. Thus, all the inventory records are kept electronically, which helped a lot in the data collection for this project. The process began on March 25, 2023, through POSAL, a cloud-based POS service.

It collects various forms of data, like inventory records for keeping track of how many items the store has presently, when they were last purchased, and from which vendor and so on. Other forms, like sales trends, which items are top sellers, or fast movers. Furthermore, which items generate the highest revenue, which does not have to be same as the top sellers or fast mover items. It depicts all these results in the form of charts in a dashboard or, gives the user an option to generate custom reports based on date, beverage department and few other parameters.

For this project, a special feature which lets one see the detailed history of a specific item, from its purchase till its lifestyle trend throughout, was used. This was a tedious task, as it had to be done for each individual item which was selected for the analysis for all the issues.

POSAL

Legacy Liquor 17th Ave

Liquorcrossing

Log out

Account

Support

Inventory

Browse Inventory

Inactive Items

Purchase Orders

Physical Counts

Kits and Choices

Outgoing Transfers

Item History

Spplits and Combines

Bulk Inventory Change

Export

Daily History

Locate (Position) Date:

Search

?

Date	Regular Sales	Dollars	Cost	GM	Promo Sales	Dollars	Cost	GM
2022-11-05	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022-10-27	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022-07-09	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022-04-21	0	\$0.00	\$0.00	23.6%	0	\$0.00	\$0.00	0.0%
2021-12-20	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2021-10-16	0	\$0.00	\$0.00	0.0%	1	\$30.99	\$19.86	35.9%
2021-08-18	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%

Export

Customers

Management

Reports

TIP: You can sort most lists on any column with an underlined header. Click the column header to sort it on that column. Click the same header again to reverse the sort order.

Weekly History

Week	Regular Sales	Dollars	Cost	GM	Promo Sales	Dollars	Cost	GM	
2022	45	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022	44	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022	28	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2022	17	0	\$0.00	\$0.00	23.6%	0	\$0.00	\$0.00	0.0%
2021	52	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%
2021	42	0	\$0.00	\$0.00	0.0%	1	\$30.99	\$19.86	35.9%
2021	34	1	\$25.99	\$19.86	23.6%	0	\$0.00	\$0.00	0.0%

Export

Monthly History

(Image 1 – POSAL Item History)

The above image shows how the item history page of the Posal cloud service looks. As seen, it has tabs for daily, weekly and monthly sales history of the item whose SKU we enter. The “export” option under each tab, downloads an XML file, containing information for the respective sales trends. Luckily, these XML files could be converted to MS Excel workbooks with a click of a button.

The store sells a total of 9 beverage departments, but only 6 were chosen for the analysis, due to time constraints. These six beverage departments are:

1. Whisky
2. Rum
3. Vodka
4. Gin
5. Tequila
6. Wines, classified further based on country of origin –
 - 6.1 Canada
 - 6.2 United States of America
 - 6.3 Australia
 - 6.4 Argentina
 - 6.5 France
 - 6.6 Spain

For each beverage department, at least 15 SKUs were monitored, with Wines as an exception. For 6 countries of origin listed here, 10 items were analyzed thus pushing total SKUs for wines to 60, which is a good representation of the overall inventory.

For the sake of simplicity, data for objective I and II, spans for the whole year 2022. On the other hand, for objective III, daily sales records, from the store's physical ledger, were stored in a spreadsheet, from March 2022 to April 2023, for a total of 14 months.

In addition, there is a column called "Popularity" in the SKU Master sheet of each beverage department, which describes the popularity of that item in Canada. These popularity ratings were collected from 2 websites. Vivino, a common database for all the wines sold in Canada (<https://www.vivino.com/ca/en>) and Distiller, a general website for spirits. (<https://www.distiller.com>)

Data Cleaning

Most of the times, raw data collected for various types of analysis, is unstructured and contains a lot of unnecessary information, sometimes, it even has missing information. All these issues demand a thorough data cleaning process, which could take a lot of time. In this analysis, data cleaning took 5 times the amount of time it took for collection.

The above-mentioned XML files in the data collection process, unfortunately, are not human readable. Therefore, they were stored as Excel workbooks, and the task did not end there. There were a lot of columns which were irrelevant to the analysis here. As the store, sometimes, sells items at discounted prices, 4 columns of sales, which essentially points to the same thing, were converted into one, after accounting considerations. In addition, there was no filter for

downloading XML files for a particular time period, no separate column for total sales numbers and total revenue. Thus, “Record a Macro” feature of excel, for these repeated tasks was used. Furthermore, there was no use of original SKUs, and they were masked with custom ones. For instance, the original SKU 7880, is named W01 (W for whisky).

For objective III, as mentioned earlier, only physical records in a ledger were present, describing the daily sales for each month. Daily sales, were not noted in a single column, but they were recorded in several, each justifying the mode it was collected in. Essentially, there were 2 basic forms of sales, those which happened in the store itself, and the rest, online. Moreover, to save time, entries were recorded to nearest 100 dollars, as it did not impact much in overall analysis.

Data Analysis

Through extensive use of “Record Macro” feature, all separate files of each item pertaining to a department were combined into a single workbook with that department as the name, with data for the time span mentioned. Now, the next step after cleaning was to analyze it.

In a separate sheet, apart from sales numbers, item’s popularity and total revenue generated from an SKU, 2 more columns, “Probability” and “Score” were added and below is the explanation for each.

“Probability” – For objective I, identifying dead stock from selection, a scale had to be devised which could separate the two. And in the search for it, the probability of an item being picked up by a customer, had to be considered. It is nothing more than total quantity sold in year divided by the maximum quantity sold in that beverage department and then scaled to 100.

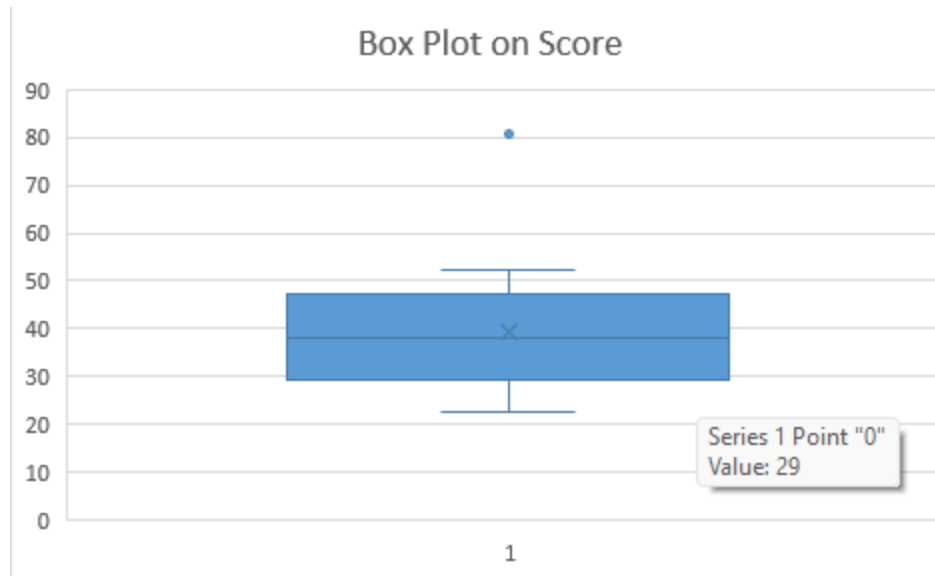
$$Probability = \frac{Item's\ total\ quantity\ sold}{Max\ of\ quantity\ sold\ in\ the\ department} \times 100$$

“Score” – The final item for the scale, is giving a score (out of 100) to each item in the department, which considers both “Popularity” of the item (looked-up from web) and its “Probability.”

$$Score = \frac{Probability + Popularity}{2}$$

Now that a scoring system was developed, further calculation and the binary classification of SKUs into supplements (choice) and dead inventory was possible. The key feature of the scale was a single variable, “Score”, which made visualizing results a simple task, in a form of “Box and Whisker” plot. To further clarify, box and whisker plots are used, wherever, depicting the spread and skewness of data points is a requirement. Thus, for each department, a separate plot was interpreted, leading to the first quartile (Q1), a value under which 25% of the data points fall. Hence, the tail portion of the revenue pareto could be further classified into, “Choice” (acting as supplements to top selling products) and “Dead stock” (items occupying shelf real-estate). In a nutshell, the Q1 value acted as a decision stump for the binary classification.

For instance, below is a box and whisker plot on the “Score” column of the department “Whisky”, and it is clear from just by hovering over the lower end of the box that the value of first quartile of Score is 29 in this case.



(Image 2 - Box and Whisker plot for finding Q1)

Moreover, for objective II, to identify top selling SKUs, the pre-existing “Sales Data” spreadsheet was used. The pareto analysis, which states that top 20% of the products contribute to 80% in total revenue generated, was used for identifying top selling items in each category. However, there was one thing to be considered that there were not enough wines belonging to each country for this analysis, as there were not significant items to begin with.

Finally, objective III, to determine the reason behind declining sales, 14 months data was analyzed, first to pin-point when the drop happened and later, factors which contributed to it. In addition, the owners made aware of the high competition existing in the area, hence, geo analysis was done, in order to find a potential competitor in the proximity. Indeed, it was a new entrant, which caused a significant decline in sales.

Results and Findings

Several aspects were studied, as laid out in the “data analysis” paragraph, which lead to important results under each objective. Some of these findings are stated below.

Objective I

The main goal of objective I was to identify items only occupying shelf area, holding capital and thereby, not allowing for the purchase of more important top sellers or even supplements, to those from items acting as a choice to top sellers. This was achieved through formulating a scoring system which assigns a score (out of 100) to each item in its beverage category. Once again, the score takes into account, both the item’s probability being sold (normalized) and its overall popularity, noted from various web resources. It is an average of both the aspects and was thought to be decent measure to solve the problem of identifying.

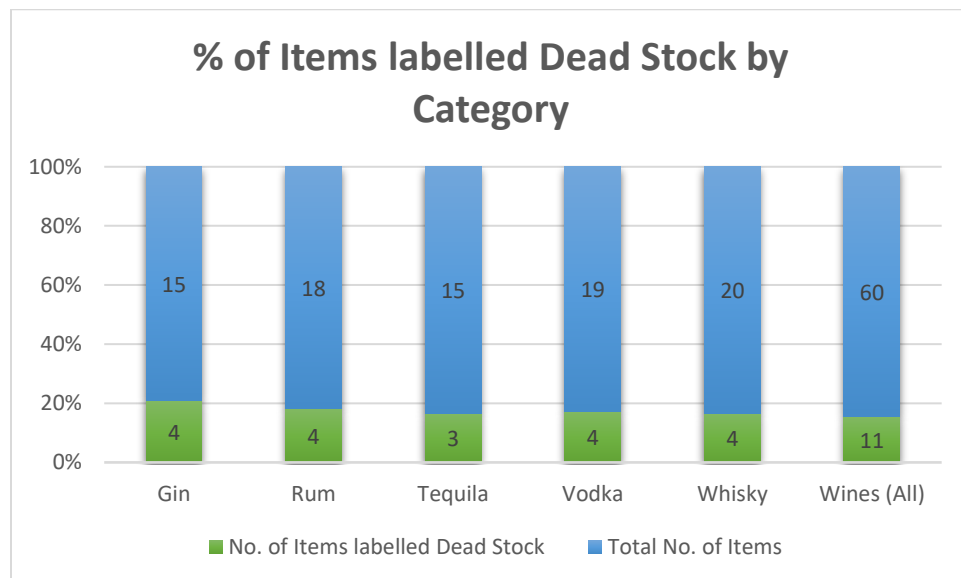
Department	Q1 Value	No. of Items labelled Dead Stock	Total No. of Items
Gin	31	4	15
Rum	34	4	18
Tequila	39	3	15
Vodka	29	4	19
Whisky	29	4	20
Wines (All)	43	11	60

(Table 1 – Distribution of Q1 values and Dead-Stock per Department)

Table 1 describes the distribution of the first quartile and the total number of items which were labelled as result of falling below the Q1 value, in each beverage category. There are some points to note here, for instance, ignoring wines, the number of items labelled dead stock remain around

4, despite increasing the set of all SKUs for department. Furthermore, value for the department whisky is the lowest, thus indicating that almost all the whisky brands are performing well in terms of sales and brand recognition. On the other hand, except wines, department tequila has the highest Q1 with 39, and hence, explaining two-fifths of the stock in this category, usually go unnoticed. Finally, to do the above analysis for each country a wine originates from, was painstakingly tedious and data was insufficient too. Hence, all the subcategories of wines, were merged into one, creating a larger pool of items to produce conclusive results.

The same results could be viewed in the form of a stacked column chart as added below.

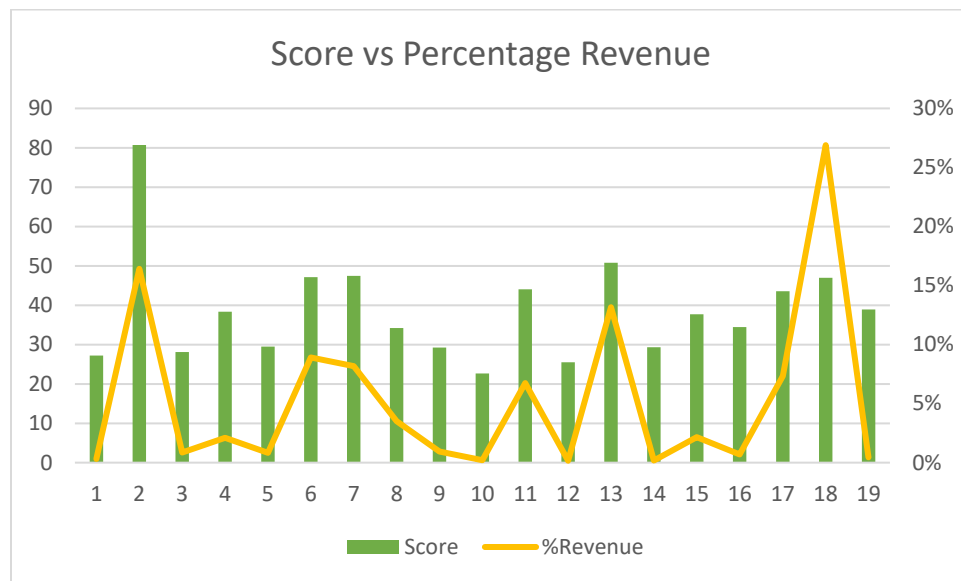


(Image 3 – Percentage of items termed Dead-Stock in each category)

The above image represents a 100% stacked column chart, with 6 columns representing 6 different beverage departments on horizontal axis. The vertical axis, on the other hand, represents the percentage of items labelled as dead inventory from the total number of SKUs present. It is clear from the chart that, roughly one-fifth of the items just occupy shelf space and do not contribute much to the sales. The values present inside the columns represent total number of items and number of dead-stock items, respectively. Stacked column chart felt like a right

choice here because, it helps us visualising 2 variables in the vertical axis, with changing entities/values on x-axis.

Another important characteristic, the variation of “Score” among the items of a department. Below are few such observations for the category “Whisky” and “Tequila”.



(Image 4 – Clustered Column for Score vs Percentage Revenue)

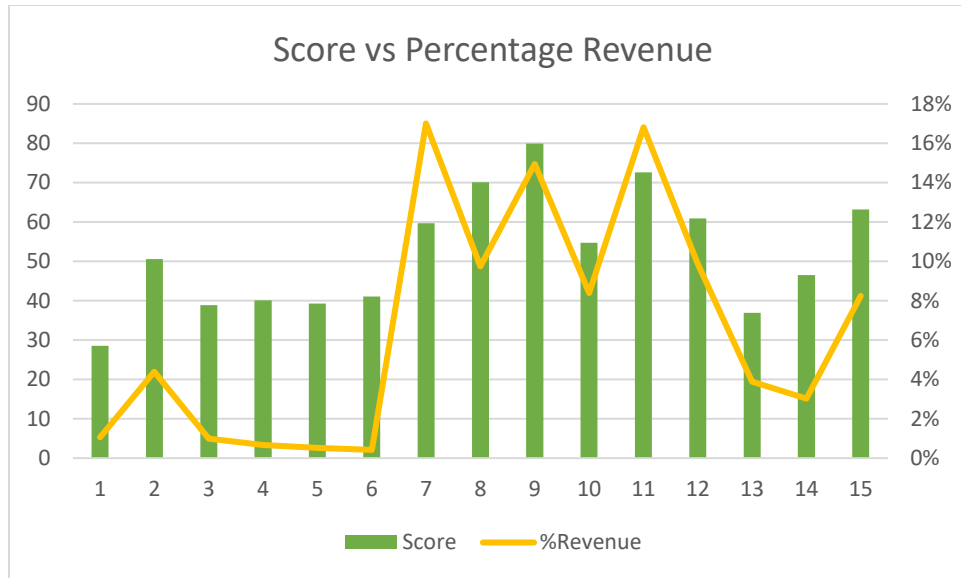
Image 4, describing the variation of score, an item receives in contrast to the revenue it generates, thus, showing the correlation between the two. (SKU 7880 was omitted here, as it was an outlier). Another important tool for visualizing correlations between variables is a heatmap generated from the underlying correlation matrix. The below image (Image 5) demonstrates a strong correlation of 0.95 of “Score” with the variables “Sales” and “Probability”. It was plotted using Seaborn library in Google Colab.



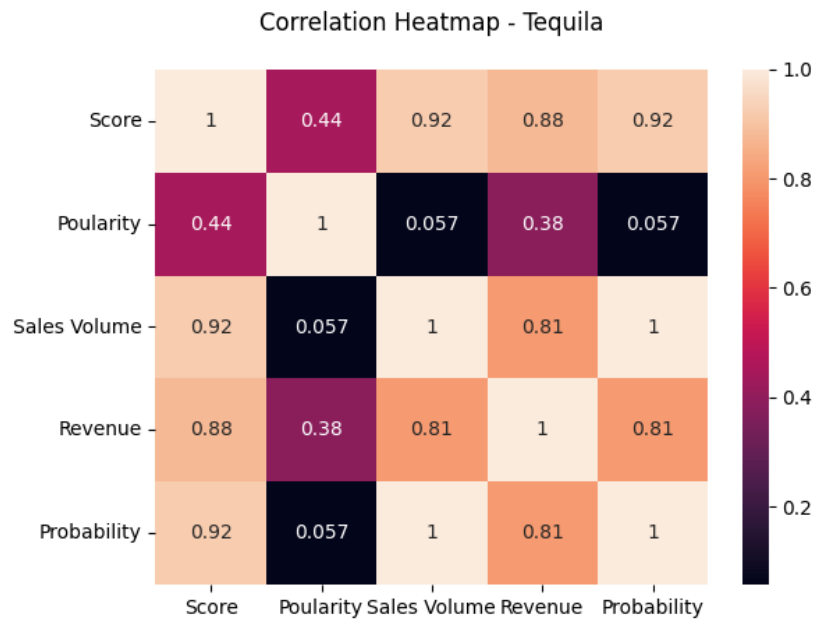
(Image 5 – Correlation Heatmap between variables for “Whisky”)

For department “Tequila” –

This department was chosen for depiction, as it sits on the opposing end of Whisky category, with a higher Q1 value, and it has less variation in its “Score” variable.



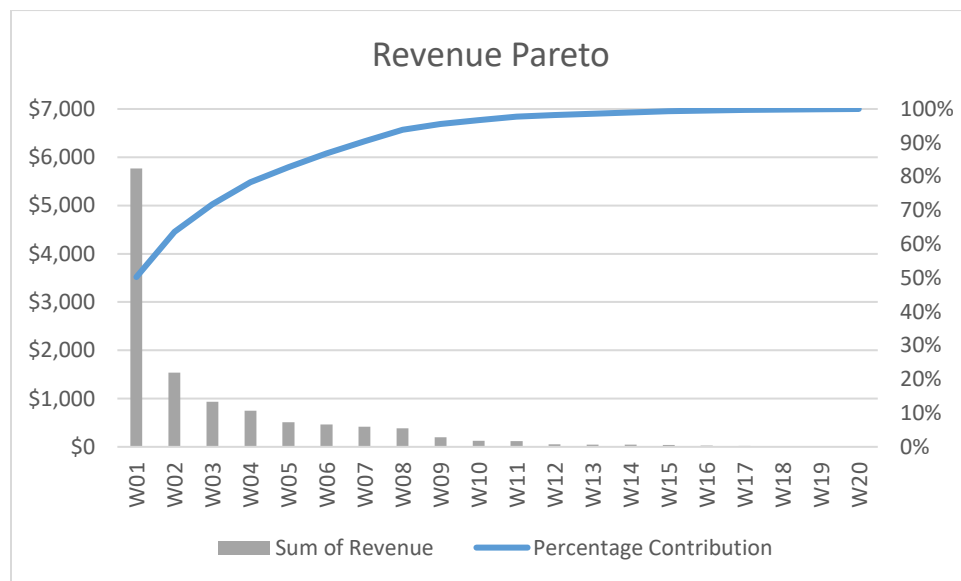
(Image 6 – Correlation between Score and the Revenue an item generates in the “Tequila” category)



(Image 7 – Correlation Heatmap between variables for “Tequila”. Here, the map shows that there is a strong correlation of 0.92 between the Score and the Probability variables.)

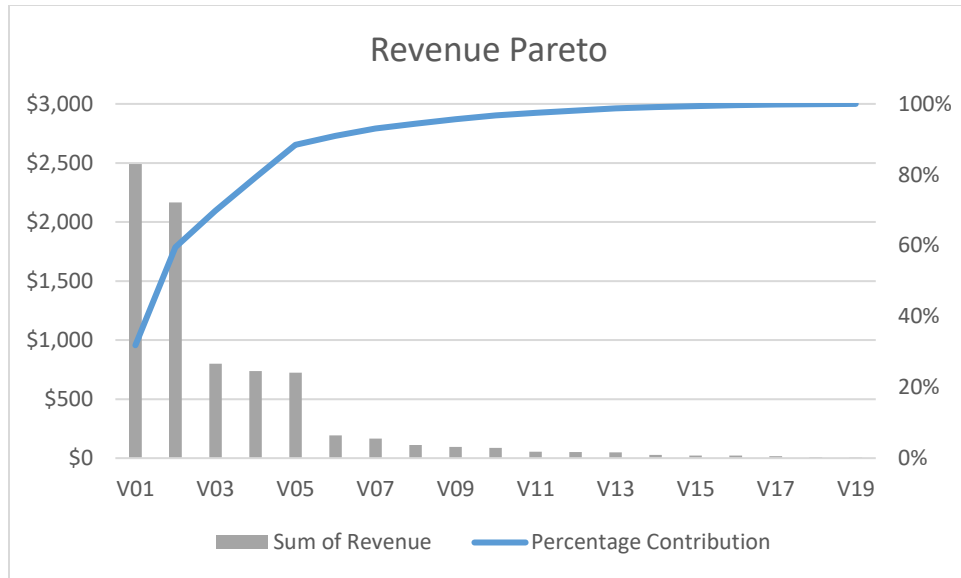
Objective 2

The pareto principle, or the law of the vital few, which states that 80% of the consequences come from the 20% of the causes, was used to identify top selling SKUs, thus solving the second objective. For all categories, except wines, this principle came in handy to identify such items, in terms of the revenue they generate. For instance, below are some of the graphs showing that 20 percent of the items being sold encompass 80 percent of the revenue generated, pertaining to that department.



(Image 8 – Chart showing revenue pareto analysis for department ‘Whisky’)

It is evident from the above chart that, 4 items (20%) ranging from W01 to W04, contribute to 80 percent of the 20 items tracked for the analysis.

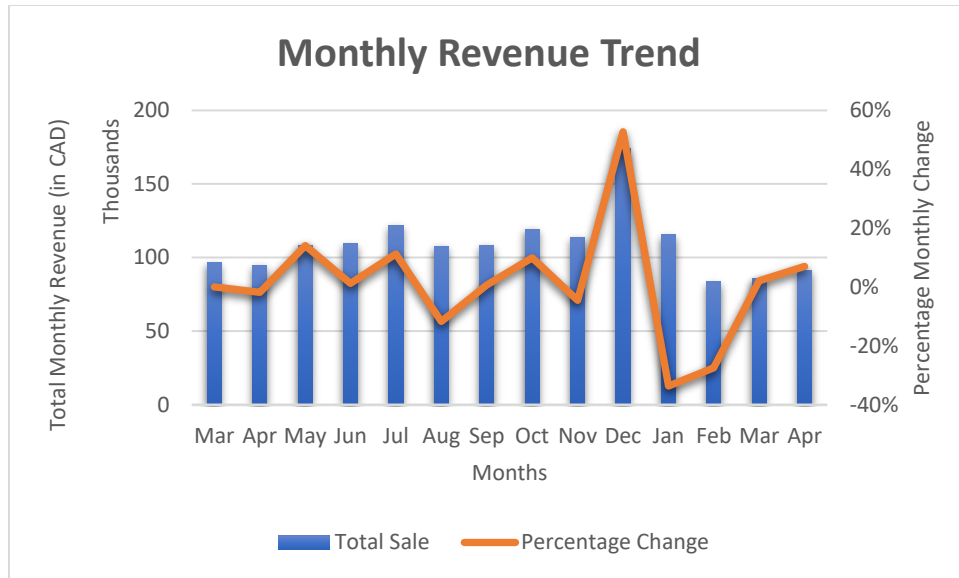


(Image 9 – Chart showing revenue pareto analysis for department ‘Vodka’)

Although not smooth, the blue trend line in the Image 9, shows that the revenue pareto holds, with again, 4 items (20%) contributing to 4/5th of the total profits.

Objective 3

In order to find the cause behind plummeting daily sales, first, the trend had to be visualized. Once again, a clustered column chart marking “Months” on horizontal axis and variables “Total Monthly Revenue” and “Percentage change in monthly revenue” on vertical axis.



(Image 10 – Chart showing the decline in monthly sales along with percentage change)

From the above chart, it can be interpreted that there is a drop in sales in the beginning of year 2023, more particularly, for the months of January and February. Upon some research, it came to surface that, after the privatization of retail liquor stores in Calgary, the number of stores in the city, jumped from 25 to 340 (at present). Therefore, cut-throat competition and low profit margins were inevitable. Not only that, the existing stores starve for survival, but a threat of new entrant worsened circumstances. Indeed, in this case, there was a new player in the market, which led to the dip in sales for Legacy Liquor.

Interpretation of Results and Recommendation

1. The results obtained after the thorough examination in identifying dead inventory, states that almost one-fifth of the inventory merely occupies shelf real estate and contributes a negligible amount to sales. These are the items that do not act as a supplement to top contributors and prevent further expansion of product variety. To prevent this, staff responsible for ordering inventory should, at minimum, consider the product's popularity from the websites stated earlier.

In addition, a customer request form should be kept, such that, if a case exists that the product is not popular or has less chances of being sold, then at least, there is one fixed customer who is certain to buy it, and therefore, a return on investment could be guaranteed over a steady period.

2. A stock-out not only impacts the steady flow of revenue, but also, the store's reputation. Hence, key SKUs should always exist on the shelf. Revenue pareto(s) hold for every beverage department, and the top sellers are already highlighted in this project, though, a subset of the total inventory, but generalizes a pattern. In most cases, it is recommended to order the top movers in bulk, whenever they are listed on monthly or quarterly sales, thereby, maximising the profits.
3. Made clear from the aerial image of Google Maps, that these retail stores depend on the number of customers visits per day, and not on the profit margins, as all stores are price setters, indicating a perfect competition type of market. Moreover, it states that there is low barrier to entry, same type of products being sold, hence lowering the profits.

In such a condition, a new entrant like, “Metro Wines” could greatly impact sales and therefore, a decline of 33 percent in monthly sales was seen. Scenarios like this should be dealt with precision and a considerable amount of support from the government could help for the long run. More precisely, a petition could be signed by the pre-existing stores, that there should be a restriction on the new entry, where the conditions are already worse. A circle of fixed radius should be devised, for the ease of operations of stores existing inside it. On the other hand, the store could try to stock in different products from its competitors, therefore, separating from the general market.

In a nutshell, all the points listed here are in the form of recommendations to the business owner. Therefore, necessary actions, if feasible, should only be taken in accordance with their business practices and ethics.