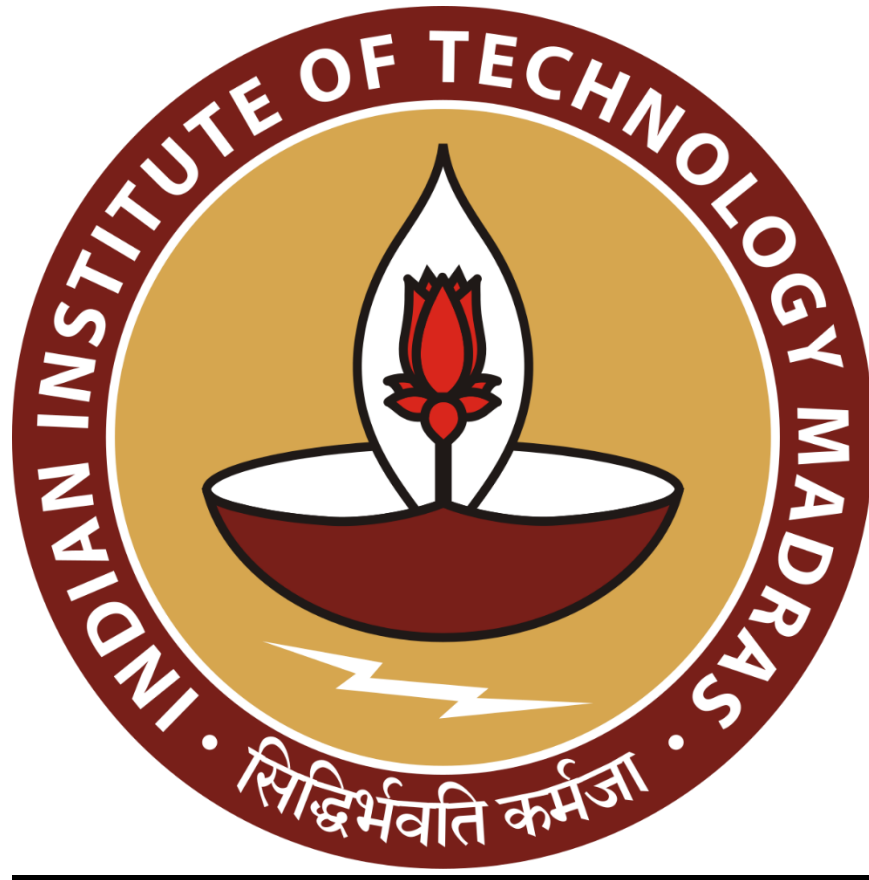


**INDIAN INSTITUTE OF TECHNOLOGY, MADRAS**



---

**Mid-Term Report**

**INVENTORY MANAGEMENT AND PARETO ANALYSIS**  
**OF A RETAIL LIQUOR STORE**

**Submitted By:**

**Name:** Harmanpreet Singh

**Roll No:** 21F3002026

## **Executive Summary**

This report aims to solve 3 business problems faced by a retail liquor store, Legacy Liquor, situated in Calgary, AB, Canada. The store, 1 out of 4 owned by the company with the same name is a prominent one in a posh area. Due to high competition and ever-increasing inventory list, the owner has encountered some hiccups along the way.

After a careful consideration, to identify the dead stock, that the store is holding, and to point out key SKUs for stock-out prevention, it was justified to capture sales trend for most products sold for each beverage department (Wines, Whisky, Rum, Vodka, Gin, Tequila). To solve the latter problem, pareto analysis was sufficient, but to identify dead inventory, a scoring system was devised which considers both the product's popularity and its probability of being sold. Thus, the tail part of revenue pareto could now be further classified into products providing selection (choice) to customers, and those occupying shelf real estate and holding capital.

Moreover, to solve the final objective, geo analysis was done, in which all the nearby competitors were studied and as it turned out, the cause of decline in daily sales was indeed a new entrant.

## **Proof of Originality**

Link to google drive folder, containing all the files –

<https://drive.google.com/drive/folders/1DtHYVS8s5vFly9g4TxlCd-QtJ9HY0I1r?usp=sharing>

Link to primary data, collected on several spreadsheets

[https://drive.google.com/drive/folders/1v3db8-8d0z8m3M5YiURDJPkxzRWcgj\\_H?usp=sharing](https://drive.google.com/drive/folders/1v3db8-8d0z8m3M5YiURDJPkxzRWcgj_H?usp=sharing)

Link to a short video of interaction with the business owner –

[https://drive.google.com/drive/folders/1Ly08tQSg-\\_CDdH9\\_XIAppwIsP3jU\\_cB?usp=sharing](https://drive.google.com/drive/folders/1Ly08tQSg-_CDdH9_XIAppwIsP3jU_cB?usp=sharing)

Link to the letter authorized by company, Legacy Liquor –

<https://drive.google.com/file/d/1kCXyxka8UBHl-WZy8HzJ0mHDAJ3BjqAZ/view?usp=sharing>

Link to few photos describing each beverage department –

<https://drive.google.com/drive/folders/1zPT->

[JtFHHDHUuMqUf7VbyYXDD5VhCeIZ?usp=sharing](https://drive.google.com/drive/folders/1zPT-JtFHHDHUuMqUf7VbyYXDD5VhCeIZ?usp=sharing)

## Metadata

The data collection process started on March 25, 2023, after a letter authorizing consent was received. Majority of the data was digital, in a point-of-sale system called POSAL (<https://cloud.posal.com/>), inventory and sale records of all the SKUs were obtained with the help of this cloud service. On the other hand, sales record of the store, maintained in the form a hard copy ledger, was digitized in spreadsheet.

For this project, inventory and sales record, for problem objective I and II, span for the whole year 2022 and daily sale records for objective III, were collected from March 2022 to April 2023, hence 14 months.

There are 3 sheets for each beverage department (stored as beverage\_type.xlsx) and wines (stored as country\_of\_origin.xlsx). The first sheet is mainly used as analysis space (named “Pareto”), but 2<sup>nd</sup> Sales Data, and 3<sup>rd</sup> SKU Master are described in detail below. (Sheets named “Box Plot” and “Dead Stock” were added later for analysis and results.)

Sales Data (Columns and Tuples explained) –

Column Name	Data Type	Description
SKU	General	Stock Keeping Unit
Date	Date	Date on which sold
Regular Sales	Integer	Qty sold at regular price
Regular Sales Price	Integer (in CAD)	Regular Price
Regular Sales Cost	Integer (in CAD)	Cost of the item
Promo Sales	Integer	Qty sold at promotion price
Promo Sales Price	Integer (in CAD)	Promotion Price
Promo Sales Cost	Integer (in CAD)	Cost of the item
Sale Price	Integer (in CAD)	Net sale price
Cost Price	Integer (in CAD)	Net cost price
Sale Volumes	Integer	Total Qty sold on a particular date

Revenue	Integer (in CAD)	Total revenue generated on a particular date
---------	------------------	--

SKU Master (Columns and Tuples Explained) –

Column Name	Data Type	Description
SKU	General	Stock Keeping Unit
Name	Text	Name of the item
Popularity	Integer	Popularity of the item (out of 100)
Sales	Integer	Qty sold in year 2022
Revenue	Integer (in CAD)	Revenue generated from the item in year 2022
Probability	Integer	Probability of the item to be sold (scaled to 100)
Score	Integer	Combination of Popularity and Probability (out of 100)
Dead Stock	Categorical (True, False)	Item categorized as Dead Stock or Not

DailySales (Objective 3, Columns and Tuples Explained) –

Column Name	Data Type	Description
Date	Date	Date of record entry
Total Sale	Integer (in CAD)	Total sale in the day
Cash	Integer (in CAD)	Income in form of cash
Debit	Integer (in CAD)	Income collected on debit machine
SkipTheDishes	Integer (in CAD)	Income on online food delivery service #1

Uber Eats	Integer (in CAD)	Income on online food delivery service #2
Doordash	Integer (in CAD)	Income on online food delivery service #3
In-store	Integer (in CAD)	Total sale happened in store
Online	Integer (in CAD)	Total sale through online food delivery services

## **Descriptive Statistics**

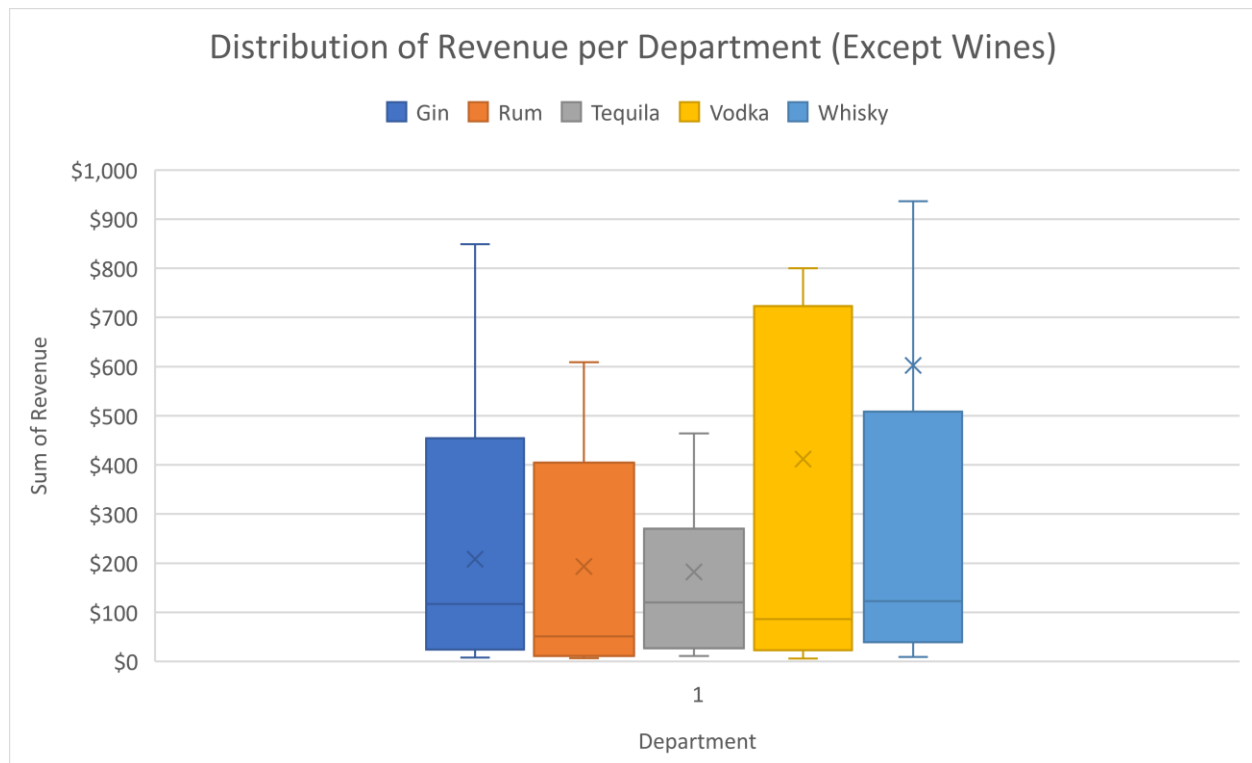
For Objective 1 and 2 –

### **Measures of Central Tendency**

Table shown below mentions measures of central tendency for the variable, sum of revenue, put forth by individual departments.

<b>Department</b>	<b>Mean (in CAD)</b>	<b>Median (in CAD)</b>
Gin	208	117
Rum	193	51
Tequila	182	120
Vodka	412	86
Whisky	603	122

### **Measures of Variability**

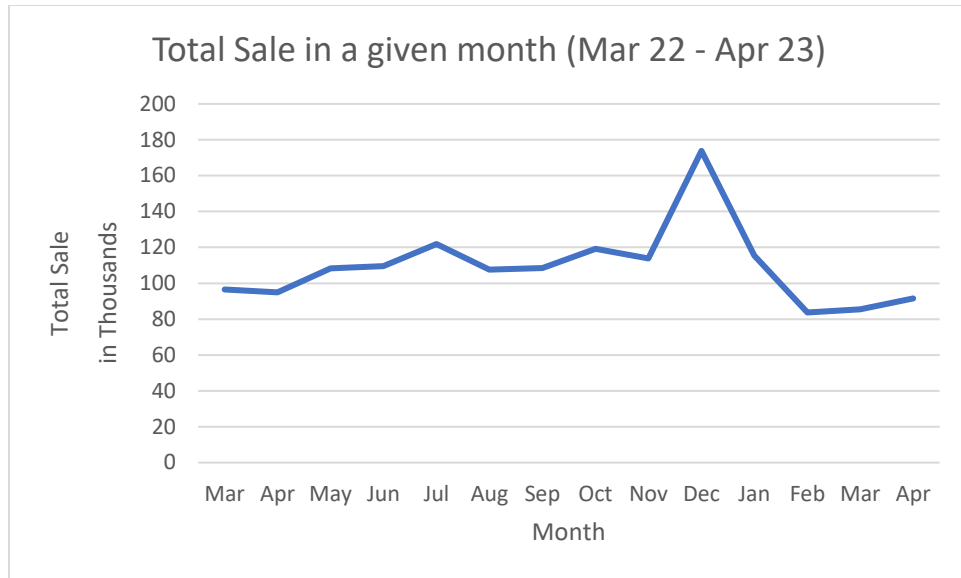


(The chart above describes the distribution of revenue (contributed by various products) in a particular beverage department. In this case, box and whisker was chosen because the spread of same variable (sum of revenue by an item) is visualized in different departments.)

It can be inferred from the chart that average of revenue generated by individual items in the year 2022, is highest for department Whisky and lowest for Tequila. On the other hand, variation of revenue generated by individual items in year 2022 is highest for Vodka, whereas lowest for Tequila.

### 3) Objective 3 – Finding the cause of plummeting daily sales –

For this objective, the monthly sale data for 14 months was recorded and here are some key findings:



(Bar chart describing the trend of monthly sale of the store from March 2022 to April 2023. Because there was only one variable, choice of bar chart was optimal.)

Total Sale in 14 months	1,530,200 CAD
Total Sale for Financial Year 2022 (Apr 22 – Mar 23)	1,342,100 CAD
Highest Sale	December 2022 (173,800 CAD)
Lowest Sale	February 2023 (83,700 CAD)

Measures of central tendency	Measures of variability
Mean: 109,300 CAD	Standard Deviation: 22,164 CAD
Median: 108,350 CAD	Variance: 491,273,846
Mode: 1	

## **Detailed Explanation of Analysis Process**

### **Collection**

The data process began on March 25, 2023, and was mostly collected through a cloud-based service, which the store uses for most of its ordering, inventory management and other office use. To begin with, 17-18 SKUs for each of the 6 beverage departments (mentioned above), were noted, though a subset of all the items in that department but a good representation of its performance. Wines were an exception with 10 SKUs for 6 different countries, 60 in total, and hence a decent quantity. Then, XML file (by default) of sale record of each SKU was opened with MS Excel and was formatted to desired outcome. Moreover, there is a column called “Popularity” in SKU Master sheet of each beverage department, which describes the popularity in terms of brand recognition, and was noted down from websites like Vivino (<https://www.vivino.com/CA/en>) for wines and Distiller (<https://www.distiller.com/>) for the rest.

## **Cleaning**

As mentioned above, the XML files from POSAL (cloud-based POS system), were not in desired format, and there were numerous columns which were redundant and irrelevant. Furthermore, it contained records from year 2020. Hence, were filtered for the year 2022. All this work was quite repetitive, therefore, macro feature of Excel, became quite handy in this scenario.

For the final objective, daily sale records of the past year had to be looked at, which, unfortunately, were not digital. Thus, to save time, all the entries were rounded off to nearest hundred, as the impact due to this in the sales trend was minimal.

## **Analysis**

After storing files for all the SKUs, a common excel file for a particular department (say Gin) was made, with one sheet (sales data) spanning the sale records for the whole year, and other (sku master) for identifying the different SKUs with their name and popularity. Then, creating a pivot table for measuring total sales volume and total revenue generated for each SKU in that department, was the optimal method. Therefore, columns like Sales and Revenue were added to SKU Master, beside Popularity of each SKU. Since, a product’s popularity was not the only deciding factor, a different column called Probability (scaled to 100), which is just the normalized version of Sales Volume (qty sold of a given SKU/maximum of qty sold among all SKUs), which in simple terms is a proxy of the probability of a given product being picked up by a customer. Finally, a column called “Score” which is the average of Popularity and Probability,



was designed, as a result, now each SKU had its score out of 100, which could separate dead inventory from supplements (selection/choice) in a hope to solve objective 1.

Furthermore, the resultant Score column now served as a base for further calculations. Since it was a combination of both Popularity and Probability, a single variable, box and whisker plot was perhaps the most suitable option to depict the spread and skewness of the data points, thus focusing on first quartile (value under 25% of the data points fall) and using the value of Q1 as a further classifier on separating the tail part of the revenue pareto. **If the score of an item is less than Q1, then it is labelled as dead stock.** In a nutshell, Q1 was used as a binary classifier to separate “Dead Stock” items from those which serve as “Selection”.

In addition, the same Sales Data sheet and pivot table were used to perform pareto analysis for each beverage department, thus, identifying top ~20% items which contributed to 80% of the sales. An exception for wines, because only 10 SKUs (approximately 25 in total) were chosen per country of origin, due to time constraints, and hence not achieving the 20% aspect.

The analysis phase for the last objective was merely subjective, and hence, some key points and observation were only made on common facts. Like, the opening of a new store nearby, the distance between these stores and number of stores in a 1-kilometer radius.

## **Results and Findings**

- 1) For identifying the items which solely holds in capital and occupy shelf space, thereby, not allowing purchase of current in-demand and popular products, the overall idea of devising a score system was beneficial. Thus, leading to a **decision stump** (obtained from the value of first quartile in box plot), which classified items into “Dead Stock” and “Not” in a following distribution:

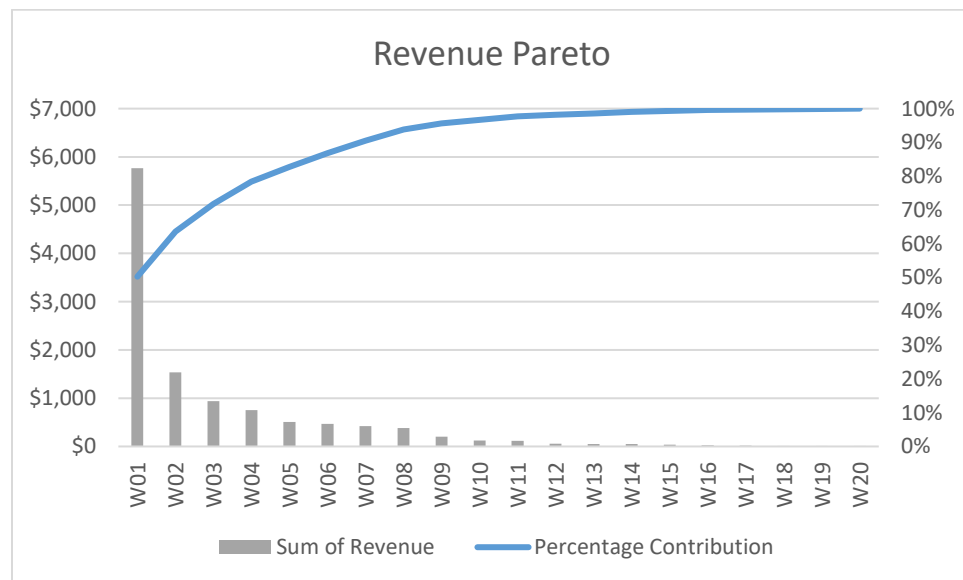
<b>Department</b>	<b>Value of First Quartile</b>	<b>No. of Items labelled Dead Stock</b>	<b>Total No. of Items</b>
Gin	31	4	15
Rum	34	4	18
Tequila	39	3	15

Vodka	29	4	19
Whisky	29	4	20
Wines (Argentina)	52	2	10
Wines (Australia)	46	2	10
Wines (Canada)	41	2	10
Wines (France)	44	1	10
Wines (Spain)	31	2	10
Wines (USA)	41	2	10

(Table depicting the results of Objective 1, showing number items labelled “dead stock’ in each beverage department)

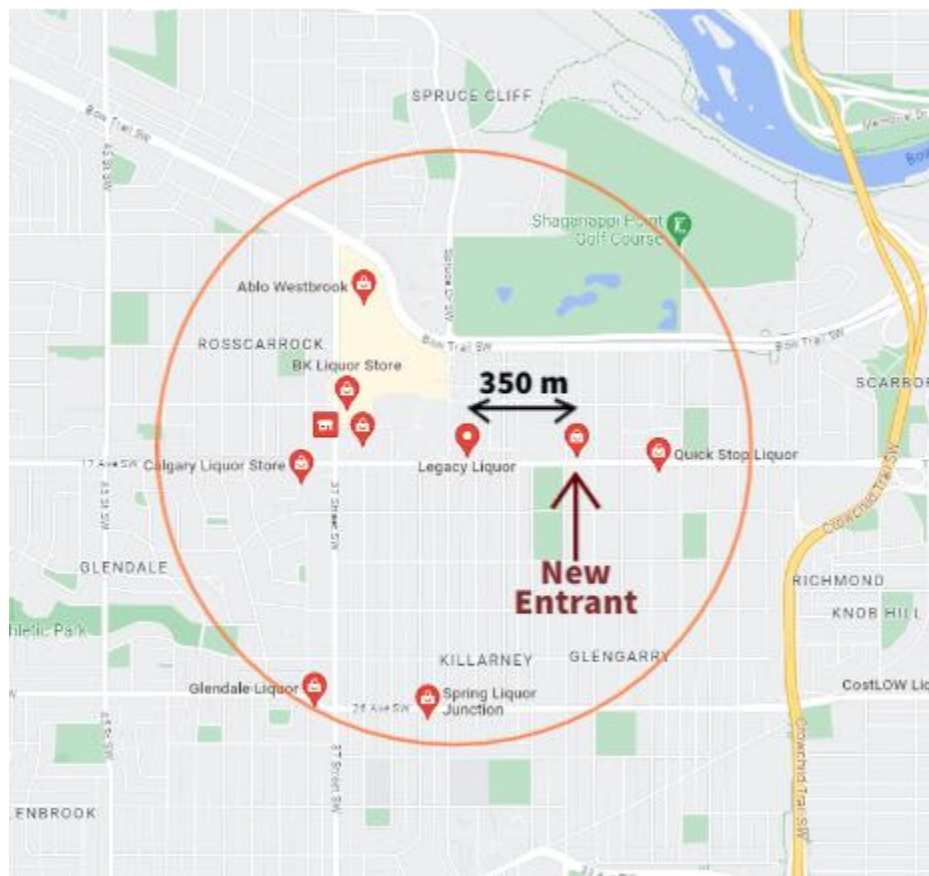
Q1 turned out to be a good classifier for this separation as items labelled **True** were not only at the bottom part of revenue pareto but also less popular.

- 2) Due to a step further into classification in “Wines” department, based on country of origin, favourable results in revenue pareto graph were not achieved but for the remainder, top SKUs which contribute to 80% of revenue were determined. Graph below is one such instance.



(Graph showing revenue pareto analysis for department ‘Whisky’)

- 3) Before 1990, there were only 25 liquor stores in Calgary, Alberta. However, after privatisation of sales, the number boomed to 340 (at present). Hence, cut-throat competition and low profit margins is inevitable. In a circle of radius 1 km with Legacy Liquor at centre, there were 8 competitors till the end of year 2022. But during the mid of January, a new entrant, Metro liquor store, came into picture. As a result, the store started to observe a decline in daily income (27% decrease in total monthly sale for Feb 23, and steady rate for following months).



(Image depicting closeness of competitors to Legacy Liquor (centre) and its new rival Metro Wines, just at 350 m. Source: Google Maps)

In addition, this new entrant is not the only factor, but it is common knowledge that Canada, currently is in recession and people are losing jobs at a rapid rate, resulting in less buying power and overall shortage in demand.