Project:
Executive Dashboards
The Coffee Cup

Data Clean and Transformation Project solely done on Tableau

Tarun Sharma

Contents

Document Overview

Following document is a report highlighting detailed analysis of company - The Coffee Cup on the basis of data collected over a period of 2 years from 2019 to 2020.

This report will help us to know what data should be captured in 2021 and solution to fill the gaps in the data in upcoming years. Analysis of the data will be including the list of questions the data answers as well as the list of questions that data does not answer but should.

Report majorly consists of analysis, results and conclusions derived from executive dashboard made on Tableau for The Coffee Cup based on the data provided by the company executives. Report will also be consisting of comments on data quality (good, bad, adequate) and recommendations to improve the quality of the data. All the steps taken to clean the data for the attached dashboard will be noted and logged for future reference.

The Coffee Cup provides the world's ultimate coffee shop experience with-

- **The highest-quality products**
- **Most inviting stores**
- **Friendliest staff**
- **Best value**

Report consists of an executive dashboard to support the strategy map to communicate the answers to the questions the data answers. A few performance metrics to consider are: Cost of Goods, Coffee Store Performance, Customer Loyalty, Profits and Sales & Product Inventory.

Data Analysis

Data Dictionary

Profit	Number		
Margin	Number		(Selling price of the item sold)-(Cost price of the item sold)
Sales	Number		Selling price of the item sold
Cost of Goods Sold	Number		Cost price of the item sold
Total Expenses	Number		Total expenses is extra cost to sell the item
Marketing	Number		Part of Total Expenses (guess)
Inventory	Number		Stock of Products in numbers
Budget Profit	Number		Predicted Profits
Budget Margin	Number		Predicted Margins
Budget Sales	Number		Predicted Sales
Budget COGS	Number		Predicted Cost of Goods Sold
Date	Date		Date of transaction (Jan2019-Dec2020)
Market	String		Central, West, North, Atlantic, Prairie
Province	Location		Provinces of Canada
City	Location		City names
Store#	String		Store Name (multiple stores with same Store#)
Market Size	String		Small, Major
Product Type	String		Category
Product	String		Sub-category
Type	String		Category
Margin Calculated	Number	Calculated Field	(Sales)-(Cost of Goods Sold)
Profit Calculated	Number	Calculated Field	(Margin Calculated)-(Total Expenses)
Other Expenses	Number	Calculated Field	(Total Expenses)-(Marketing)
Budget Expenses	Number	Calculated Field	(Budget Margin)-(Budget Profit)

Table 2.1 Data dictionary for Coffee Cup Data

Data Quality

1.1.1. Date field

•Dates in the dataset ranges from 31st January 2019 to 31st December 2020

•Dataset contains 177 entries from last day of every month in 2019 and 2022 i.e. dates are limited to 28th, 30th and 31st in the dataset. (177*2*12=4248= Number of Rows in Dataset)

Year of D..	Month of D..	Day of D..	
2019	January	31	177
	February	28	177
	March	31	177
	April	30	177
	May	31	177
	June	30	177
	July	31	177
	August	31	177
	September	30	177
	October	31	177
	November	30	177
	December	31	177
2020	January	31	177
	February	28	177
	March	31	177
	April	30	177
	May	31	177
	June	30	177
	July	31	177
	August	31	177
	September	30	177
	October	31	177
	November	30	177
	December	31	177

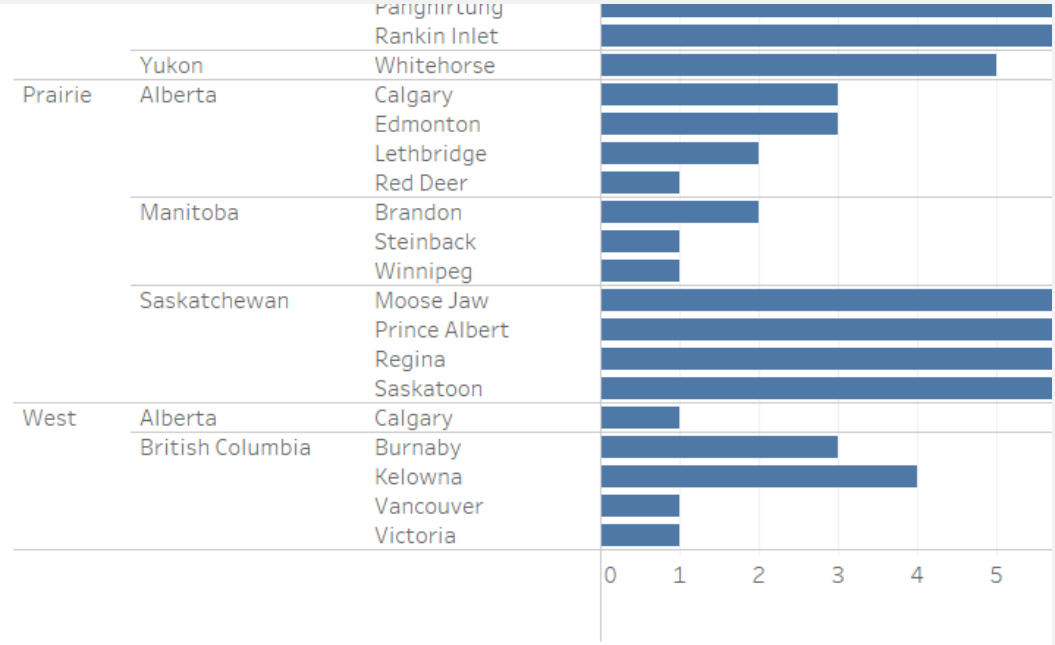
•<u>Two(2) entries in the dataset are annotated as 2012 (shown in Figure below), and hence been changed to 2019 in both the instances</u> for further analysis of data.

1946	152	94	2101	370	500	840	340	31/01/2020	West
1947	152	94	2101	260	380	640	260	31/01/2020	Atlantic
1948	58	23	558	360	400	460	60	31/01/2020	Atlantic
1949	127	93	2449	-240	-160	0	160	31/03/2012	Central
1950	146	113	1118	-190	-60	60	120	31/03/2012	Atlantic
1951	94	50	1288	-120	-40	150	190	31/03/2019	West

1.1.2. Market/Province/City consistency of stores

•In the given dataset, Market is assigned as a 'String' instead of 'Location' as it represents clustering of the Canadian Provinces into Central, West, North, Atlantic, Prairie

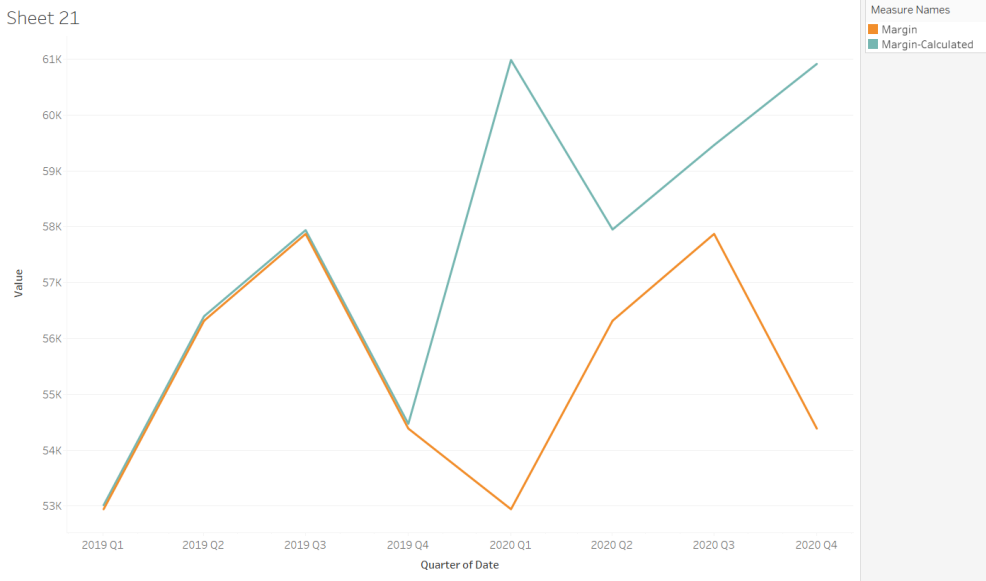
•One(1) Stores# from Calgary, Alberta has been incorrectly assigned to 'West'. A new calculated field has been made by the name of 'Province (group)' to assign each province into right 'Market'.



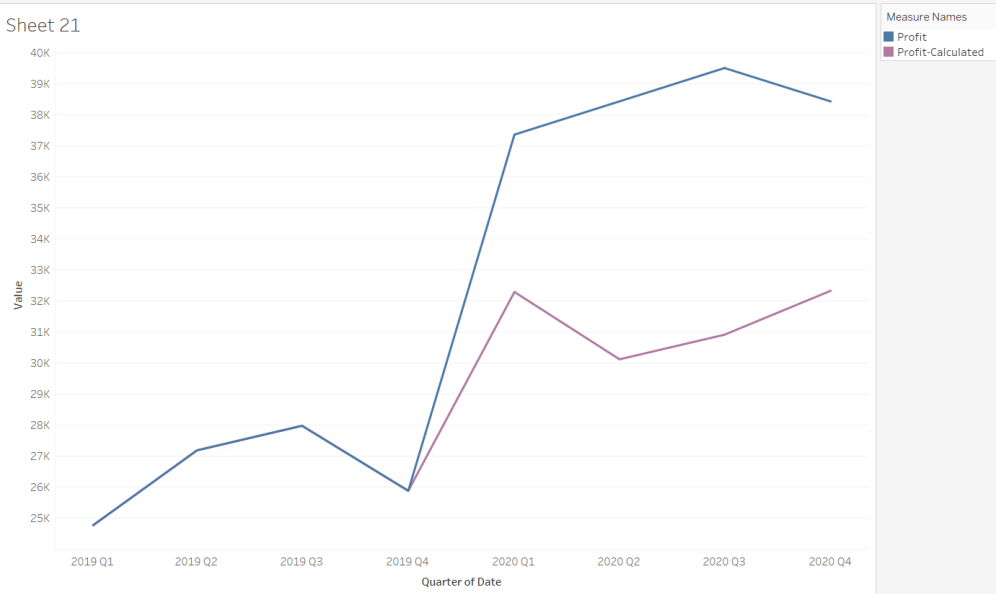
•<u>There are multiple stores in many cities within a province with multiple Store ID. This hinders the analysis of the individual store performance.</u>

1.1.3. Calculated field consistencies

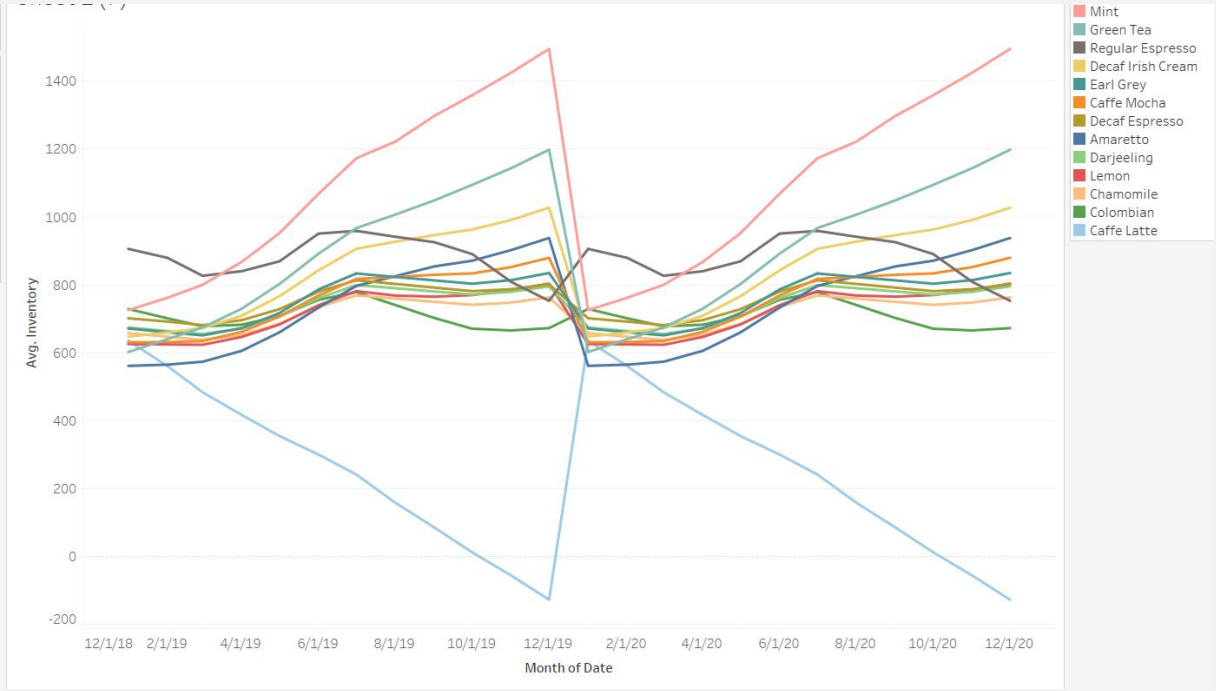
•Margin is the result after subtracting COGS from Sales. Provided 'Margin' column doesn't match with the calculation. As shown in the figure below, provided 'Margin' column diverges from 'Margin-Calculated' after 2019.



•Profit is the result after subtracting Total Expenses from Margin. Provided 'Profit' column doesn't match with the calculation. As shown in the figure below, provided 'Profit' column diverges from 'Profit-Calculated' after 2019.



•Cyclical nature of inventory with exact repeating pattern. Inventory of Caffe Latte goes to negetive on each end of the year which is odd.



Data Scope

1.1.4. Key Performance Questions (KPQs) dataset can answer *(Answered in Dashboard attached)*

- **What are the Products/Product-Types with most and least Profits?**
- **What are the Provinces/Regions/City with most and least Profits?**
- **What is the relation between Marketing Expenses and Profits?**
- **Are Profits or any other parameter cyclical in nature?**
- **Is there any Year-on-Year increase in profits?**
- **Any difference between projected & realized profits and how to improve Profits & lower down Total Expenses**

1.1.5. Key Performance Questions (KPQs) dataset **cannot** answer

- **What Stores are generating least and most Profits?**
  Dataset contains Store names that have multiple locations, so individual performance of the store cannot be calculated
- **What is the trend throughout the week?**
  As we have only 177 entries from last day of every month in 2019 and 2022, we can not analyze daily and weekly performance of Products and Stores
- **What is the breakdown of Total expenses?**
  Further analysis on Total Expenses can help find Cost Saving methods but we only have 'Marketing' and lack other constituents.
- **Was Customer service reason for the losses?**
  There is no Data regarding Customer experience and Staff performance for this kind of analysis

Conclusions & Recommendations

1.1.6. Conclusion

•What are the Products/Product-Types with most and least Profits?

Most Profits- Colombian, Darjeeling, Decaf Espresso

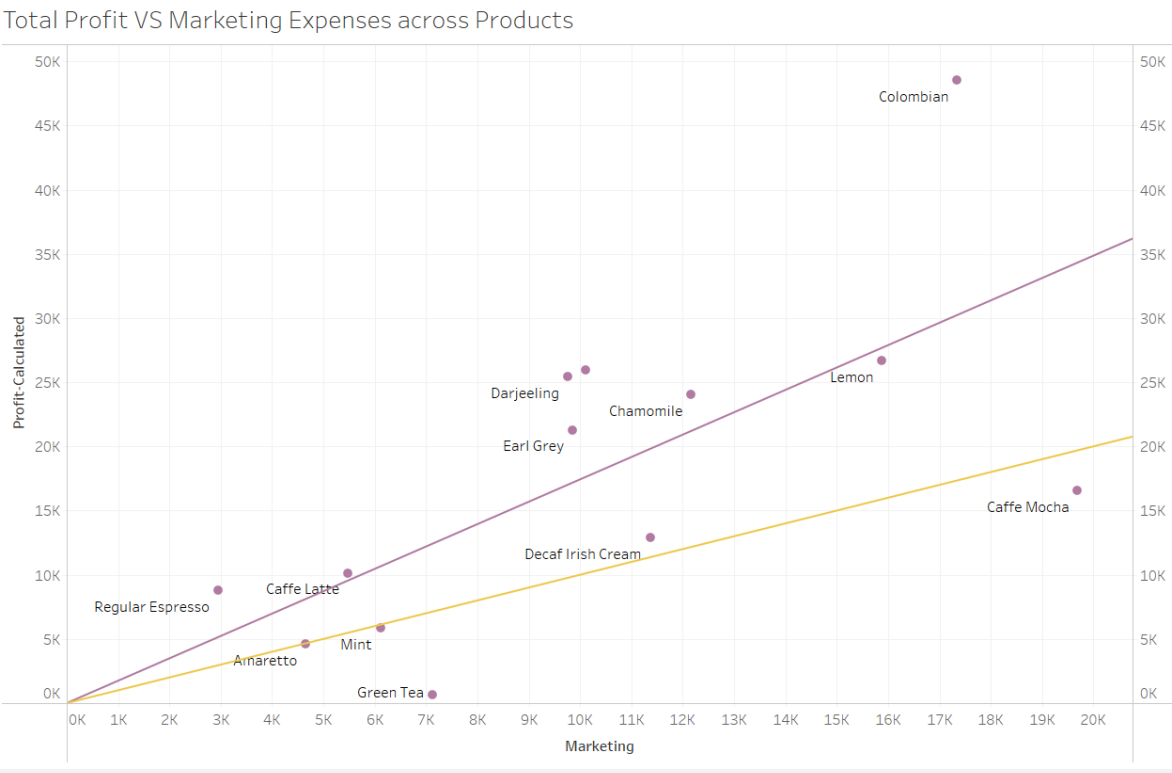
Least Profits- Green Tea, Café Mocha, Mint tea

•What are the Provinces/Regions/City with most and least Profits?

Most Profits- Nova Scotia, Alberta, British Columbia

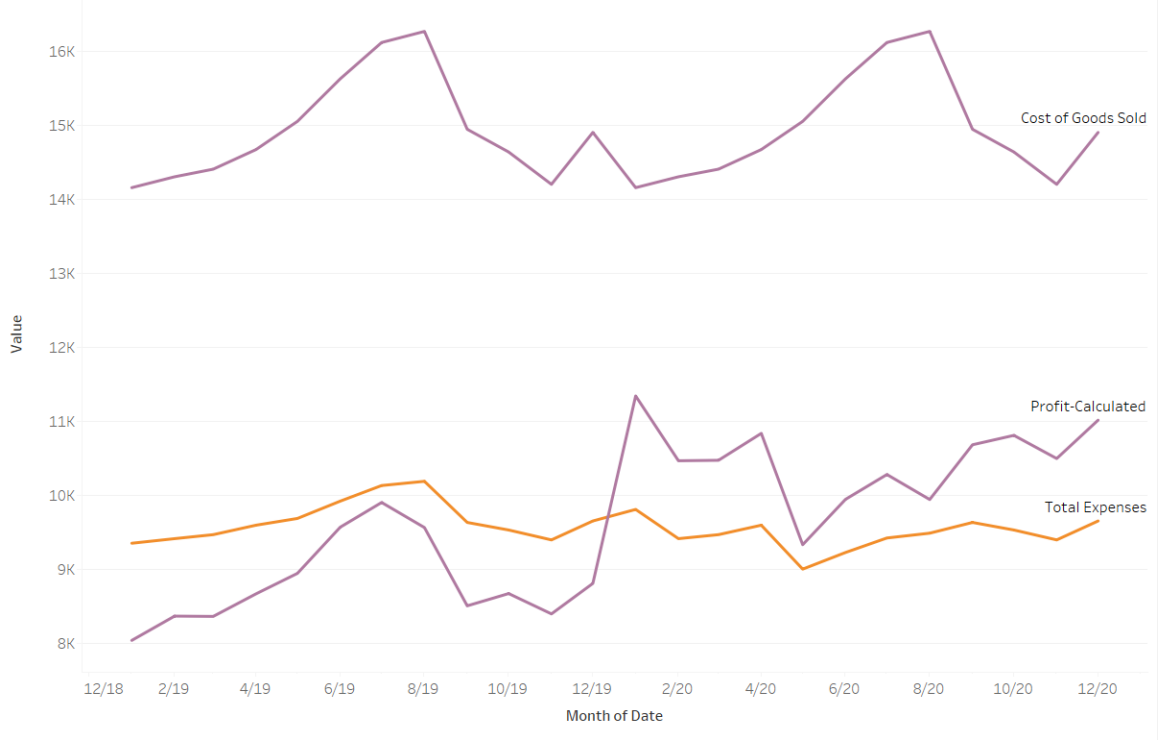
Least Profits- North Provinces, Saskatchewan, PEI

•What is the relation between Marketing Expenses and Profits across Products?



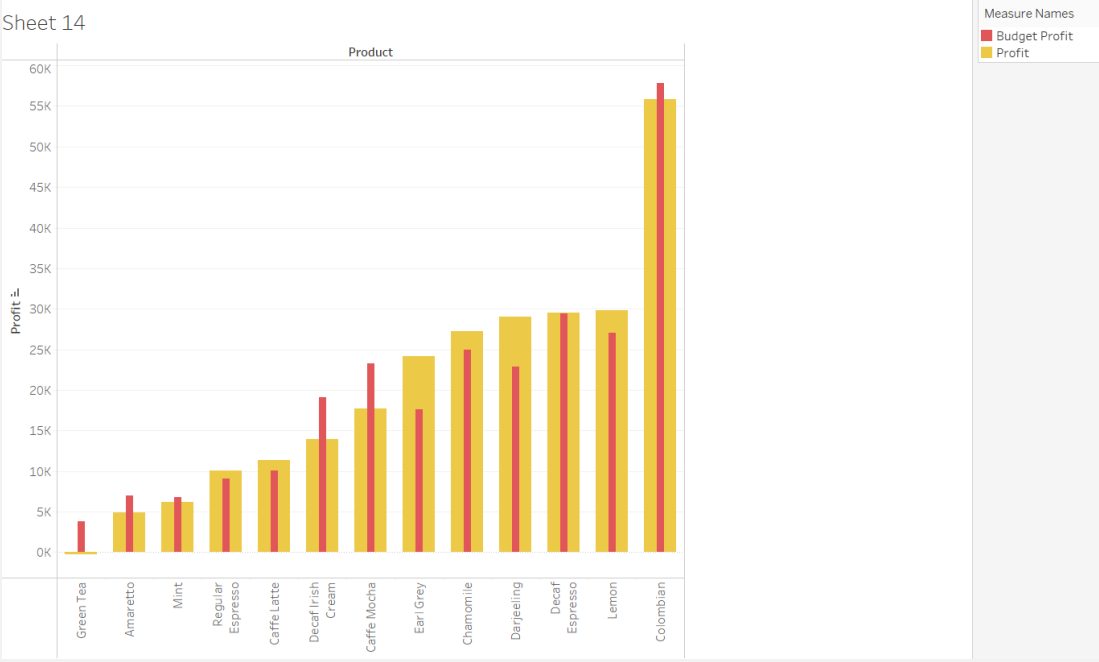
•Is there any Year-on-Year increase in profits?

Yes, around 40%



•What are the Products/Product-Types with most and least Profits?

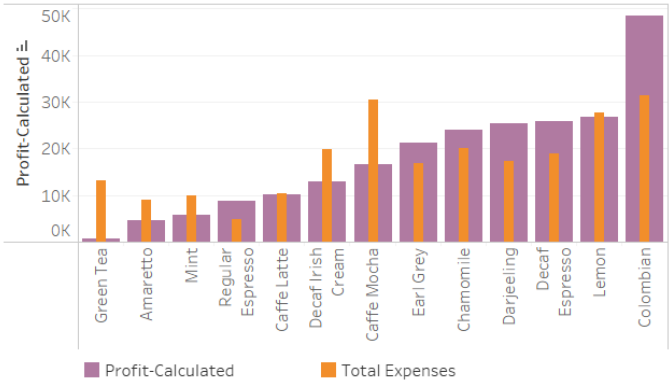
Most Profits- Colombian, Darjeeling, Decaf Espresso



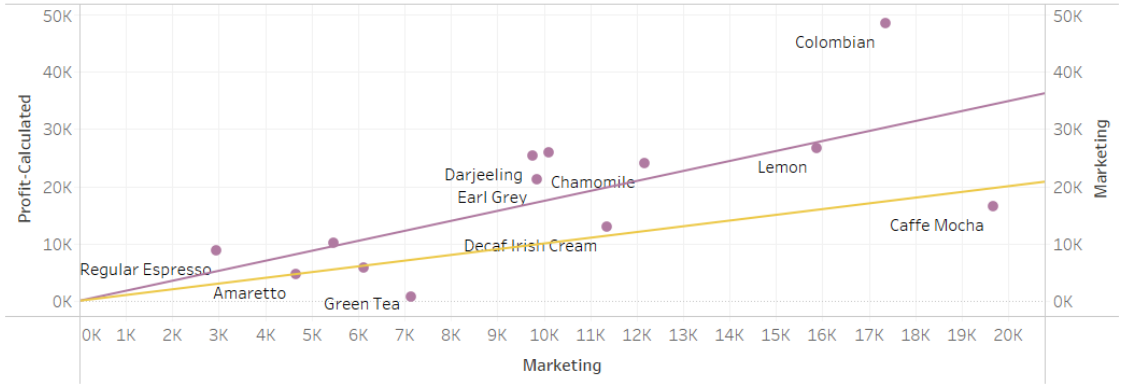
Executive Dashboard

The Coffee Cup Executive Dashboard

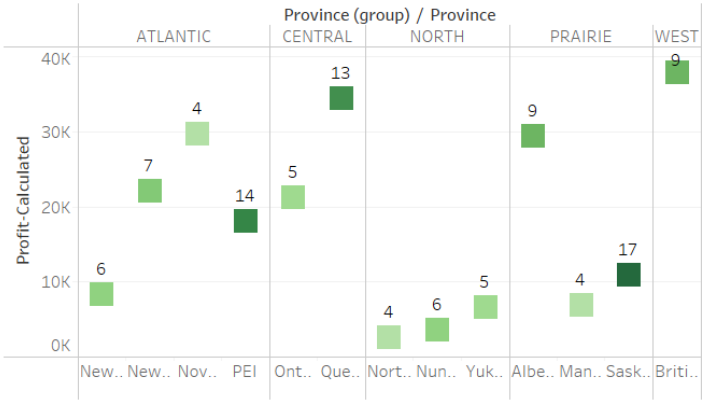
Total Profit VS Total Expenses across Products

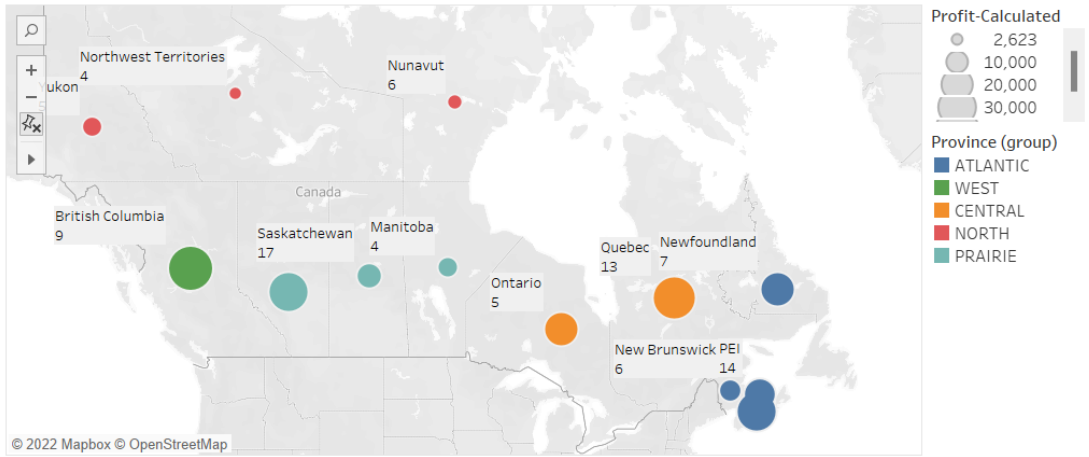


Total Profit VS Marketing Expenses across Products



Total Profit per Distinct Store across provinces





Sheet 21 | Sheet 21 (2) | Sheet 1 | Sheet 11 | Sheet 1 (2) | Sheet 19 | Sheet 1 (7) | Sheet 1 (3) | Sheet 14 | Sheet 17 | Sheet 17 (2) | Sheet 20 | Sheet 22 | Sheet 23 | **Dashboard 1** | Dashboard 2