**Analysis on how the profit is affected in accordance with discount band**

**This project analyses data across various segments**

**based on discount from the year 2013 & 2014. This is a sample data source. I do not own this data.**

**Metadata of the dataset:**

1. Number of cells: 11216
2. Number of rows: 701
3. Number of columns: 13(before analysis) 16 after adding columns during analysis
4. Date of analysis: 2nd March 2022
5. Time of completion: 4pm GMT

Note: The sample data, which was downloaded for analysis, the dataset in which the analysis is done is attached along with this analysis report for reference.

**Summary:**

I structured my **analysis** to view the data from 3 perspectives: Profit, discount band and segment. Interesting part in the analysis was knowing which product made highest profits in segments and what was the effect of discount band on profits, focused more on region(country), segment, profit and discount band.

The following analysis was made and the best possible solutions including **visualisation** using excel **graphs** has been presented.

**Analysis was made on the below set of questions:**

1. What is the total profit in each segment?
2. In which region the highest profit was made and what was the discount band?
3. What is the manufacturing price of each product?
4. In which half of the year was the profit made?

My goal is to further delve into this dataset as my skills evolve. Below are some of the items I plan to add to this project.

1. Create visualizations excel graphs
2. Use Pivot tables to categorise the data, group the data and to perform advanced calculations.
3. Use appropriate functions for calculations*.*

**Phases of the Analysis**

*A screenshot of a computer

Description automatically generated with low confidence*

**Data Cleaning:**

Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset.

**Steps followed in data cleaning:**

1. Perform preliminary analysis: The data has been checked for its **accuracy, consistency, and completeness**.
2. Fixing misspellings/typos: The data has been checked and misspellings were corrected using the path **click Review > Proofing > Spelling**.
3. Removing Duplicates: Duplicate or repeated entries were deleted in the data set. **The path was data > Remove duplicates**
4. Checking for bias: The data was checked to make sure that the data obtained is not in the favour of a person or a company.

**Data Analysis:**

The process of systematically applying statistical and logical techniques to describe and illustrate, condense, and recap, and evaluate data.

This is the core phase of the project that involves finding meaningful insights after cleaning and organising the data properly.

**Steps followed in data Analysis:**

1. **Adjusting and formatting data:**
2. The currency value of Manufacturing price, sales price, gross sales, Discount, Sales, COGS and profit were converted into the same currency which is USD.
3. The Month and year were extracted from the date column in two separate columns using the following functions:

**Month: =Text(M2,”mmmm”)**

**Year: =YEAR(M2)**

1. If Sales > COGS(Cost of goods sold) then it is profit and if Sales < COGS then it is a loss. So the profit or loss of a particular record was identified using IF function

**=IF(sales>COGS,”profit”,”loss”) and the result obtained was dragged till the last record in the dataset using relative cell reference.**

1. Sum of all the unites sold, Total of manufacturing price, sum of gross sales, sum of net sales, sum of profit by deducting the loss were calculated at the bottom of the respective columns.
2. **The dataset has shown ##### error in COGS. It was rectified by widening the length of the cell.**

**2)Finding appropriate relationship and categorising data using PIVOT table:**

In the 1st sheet i.e; profit per segment, a **PIVOT table** was inserted, and a **Line Chart** was used to find the highest profit made by a segment in a particular year in the various ranges of discount band.

* Line graphs can also be used to compare changes over the same period for more than one group.

**For the year 2013:**

**1) When Discount band was high:**

The Government Segment made the highest profits, and the enterprise segment was facing huge losses.

**2)When Discount band was low:**

The Government segment made the highest profits, and the midmarket and enterprise segment made lowest profits. None of the segments were under loss.

**3)When Discount band was medium:**

The Government Segment made the highest profits, and the enterprise segment was facing losses.

**4)When there was no discount band:**

All the segments had neither profit nor loss, but the government segment was again benefitted with this type of discount band

**For the year 2014:**

**1)When discount band was high:**

The Government Segment as in boom, and the enterprise segment was facing huge losses.

**2)When Discount band was low:**

Small business and Government sector made good profits and none of the segments were in loss.

**3)When Discount band was medium:**

Small business and Government sector made good profits and enterprise segment was in loss.

**4)When there was no discount band:**

All the segments had neither profit nor loss, but the government segment was again benefitted with this type of discount band.

In the 2nd sheet i.e; manufacturing price per product, another **PIVOT table** was inserted, and a **bar graph** was used to find the highest manufacturing price of the product in 2013 and 2014.

* Bar graphs are used to compare things between different groups or to track changes over time. However, when trying to measure change over time, bar graphs are best when the changes are larger.

Through the bar graph, it was found that Amarilla and VTT were the products with the highest manufacturing price in the year 2013 as well as 2014.

In the 3rd sheet, quarterly and full year profits, PIVOT table was used again. A bar graph was used and was discovered that the government sector was the segment which made good profits in the 3rd quarterly of 2013 (September to December) and also in the entire year of 2014.

Data Visualisation:

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns.

The below represent the screenshot of dashboard depicting the analysis made:

*Graphical user interface

Description automatically generated*