**Robustness Diagnostics**

Phionah Nakakeeto

Grand Canyon University

MIS-620: Descriptive and Diagnostic Analytics

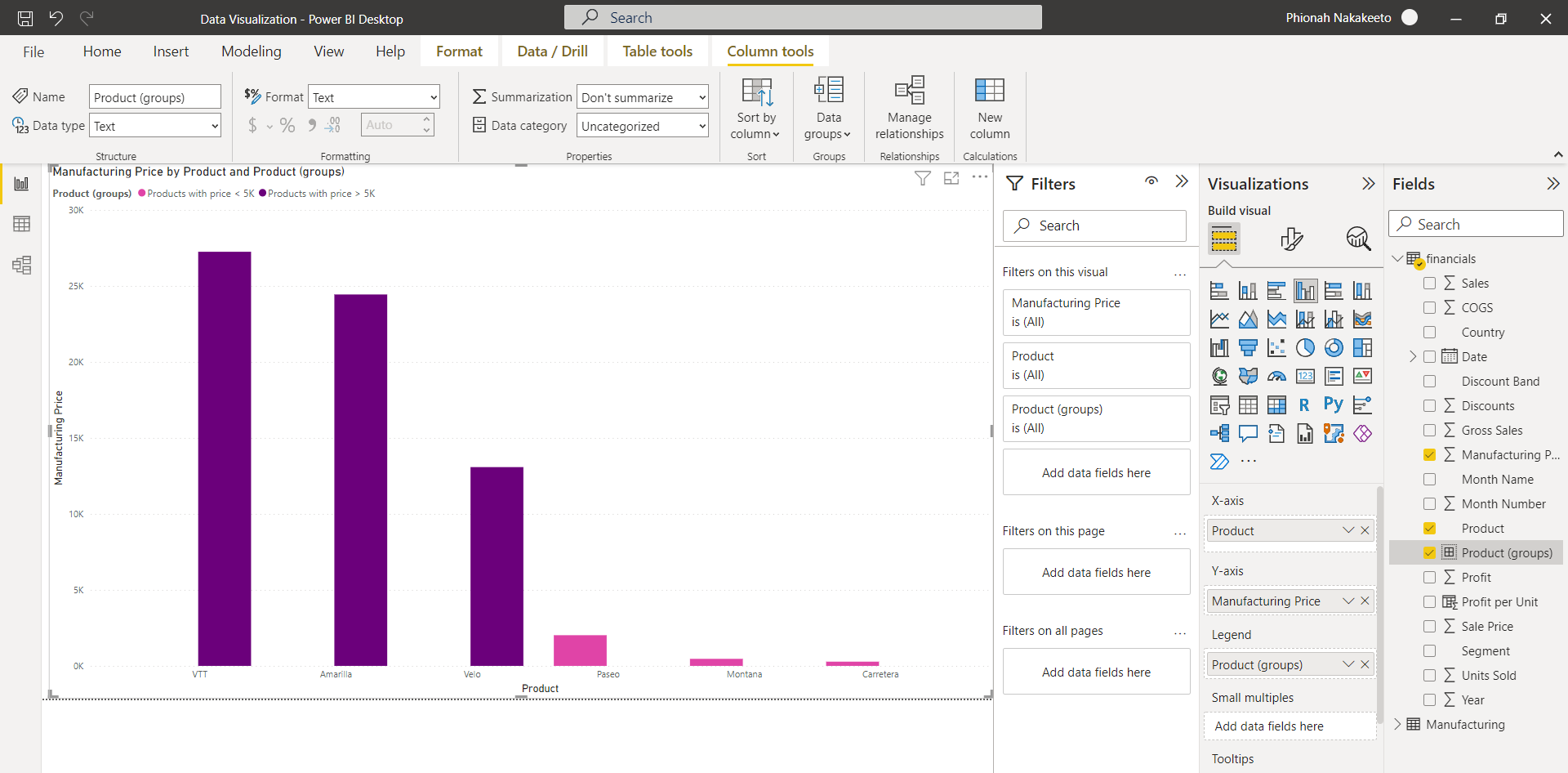
Dr. Omari Williams

August 3, 2022

**Using Power BI, group data in the following scenarios and use the groups to provide an analysis of the data.**

**Using the product variable, group the data into two groups based on manufacturing price.**

I identified and created 2 groups; Manufacturing price greater than 5000, and manufacturing price lower than 5000.



**Which products did you place into each group and why?**

Products VTT, Amarilla, and Velo have a price greater than 5000 whereas products Paseo, Montana, and Carretera have a price lower than 5000.

**What would you call your two groups?**

My 2 groups are named Products with price > 5K AND Products with price < 5K.

**What is an alternative way you could have grouped the data?**

I could have also grouped the data based on the 2 Products with the most significant and lowest manufacturing prices i.e., Montana and Carretera.

**Why is your choice of groups superior, in your view?**

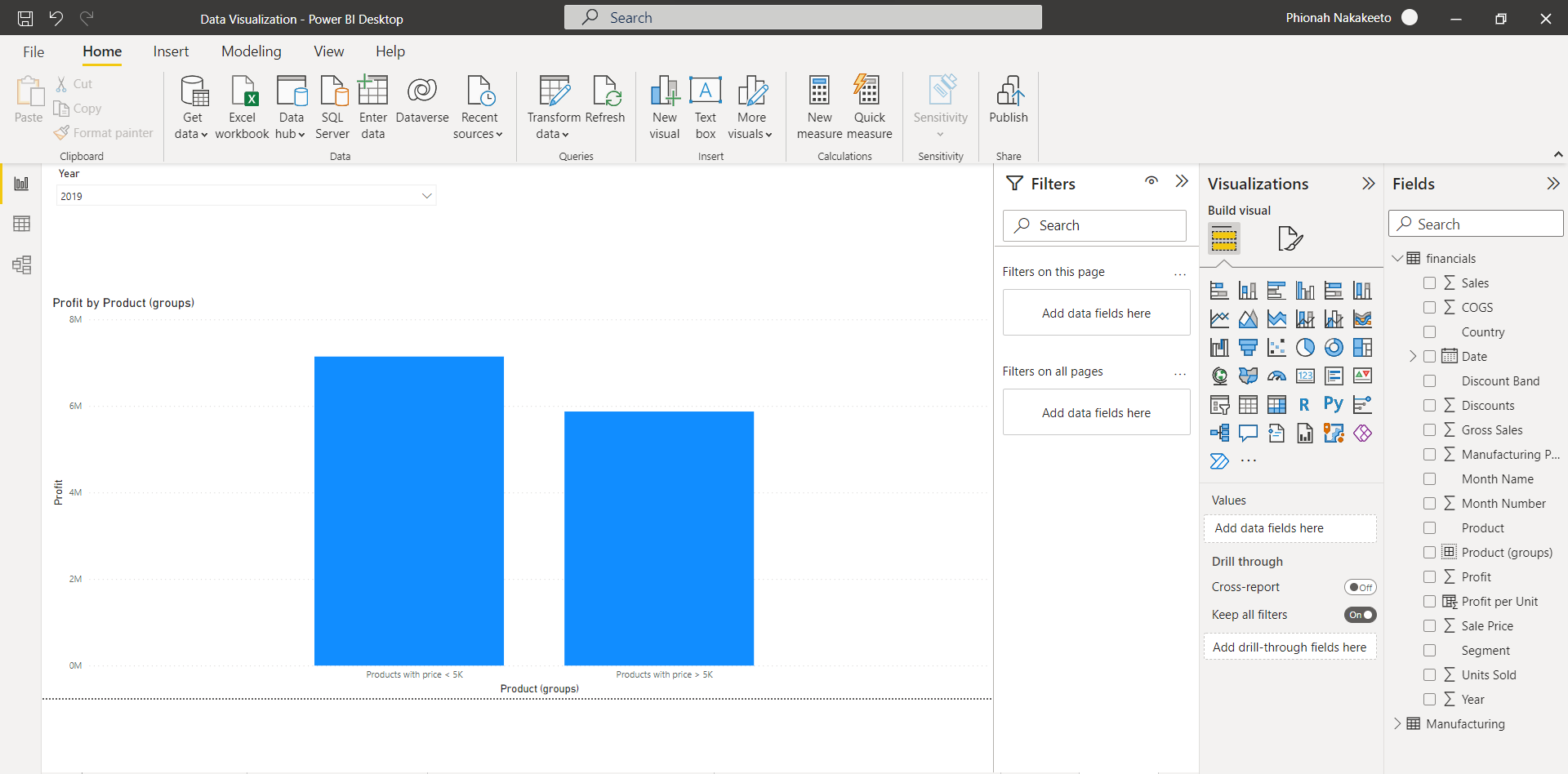
The 2 products have prices that are too different from the rest of the products i.e., outliers and this requires further analysis to investigate the reasons for the significant details.

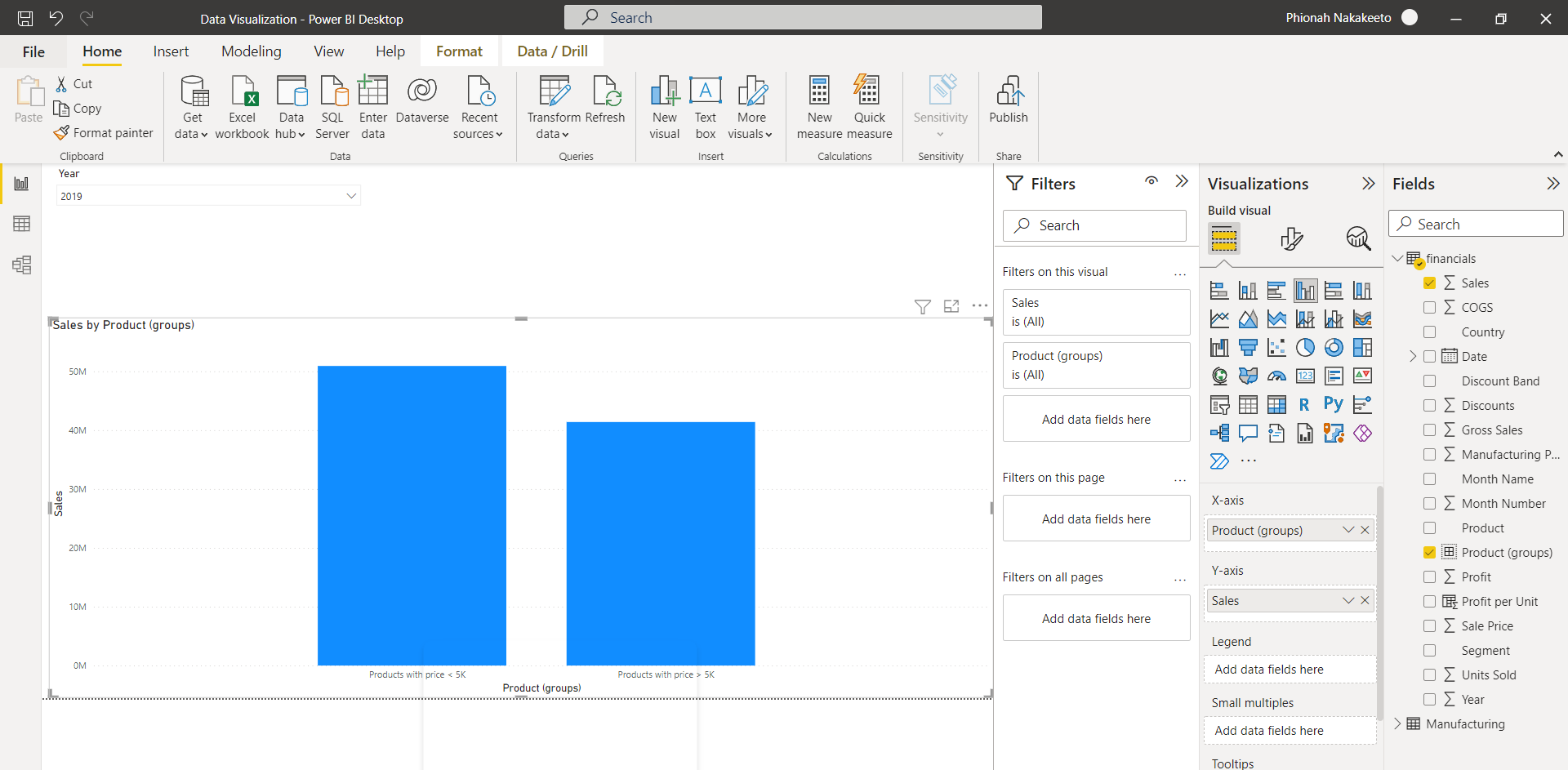
**Using these groupings, assess overall profitability, unit sales, profit per unit, and manufacturing defect rate for 2019. What patterns, if any, do you notice about the groups?**

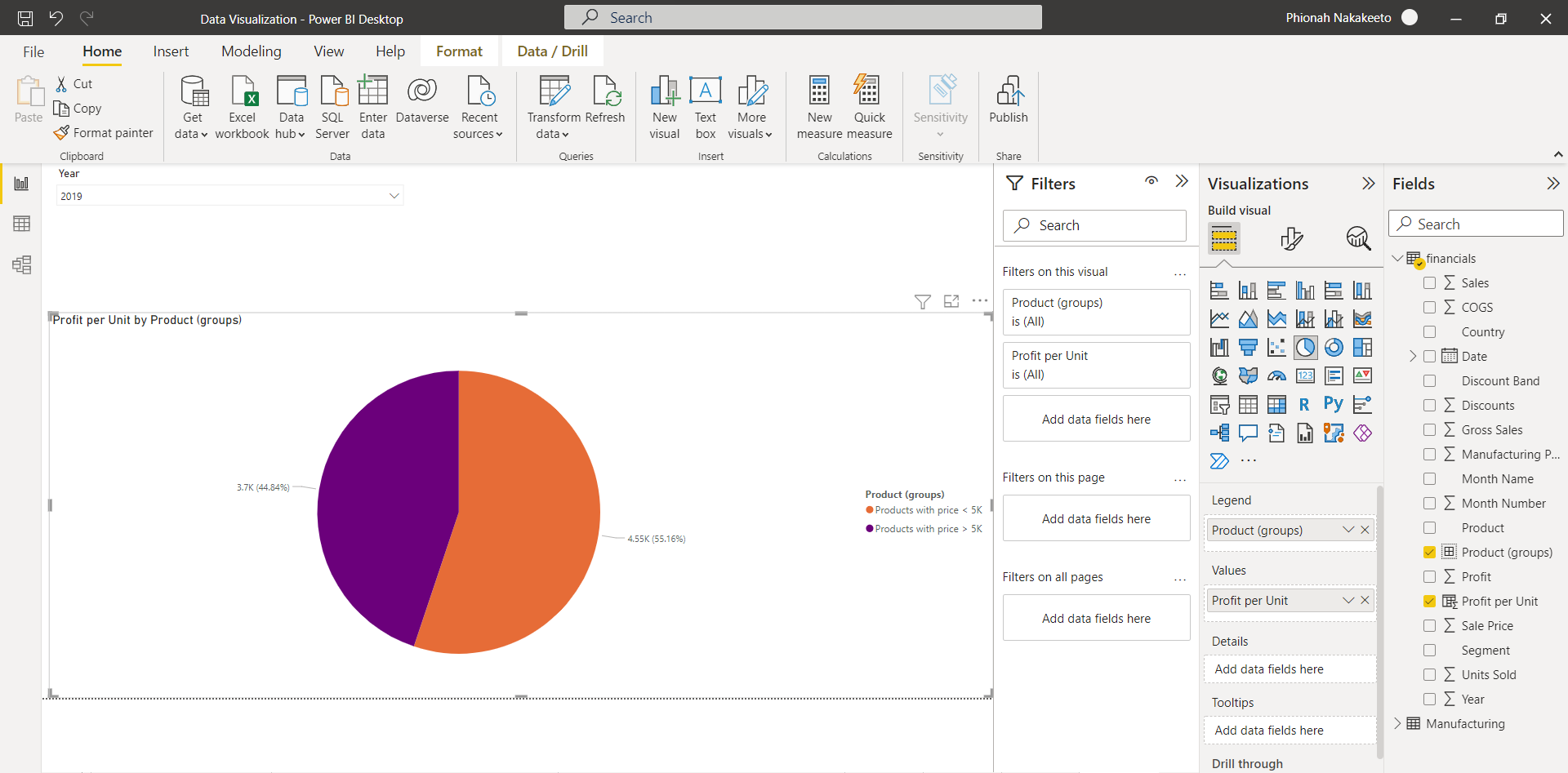
I have noticed that the products with price < 5K are more profitable than those with price > 5K.

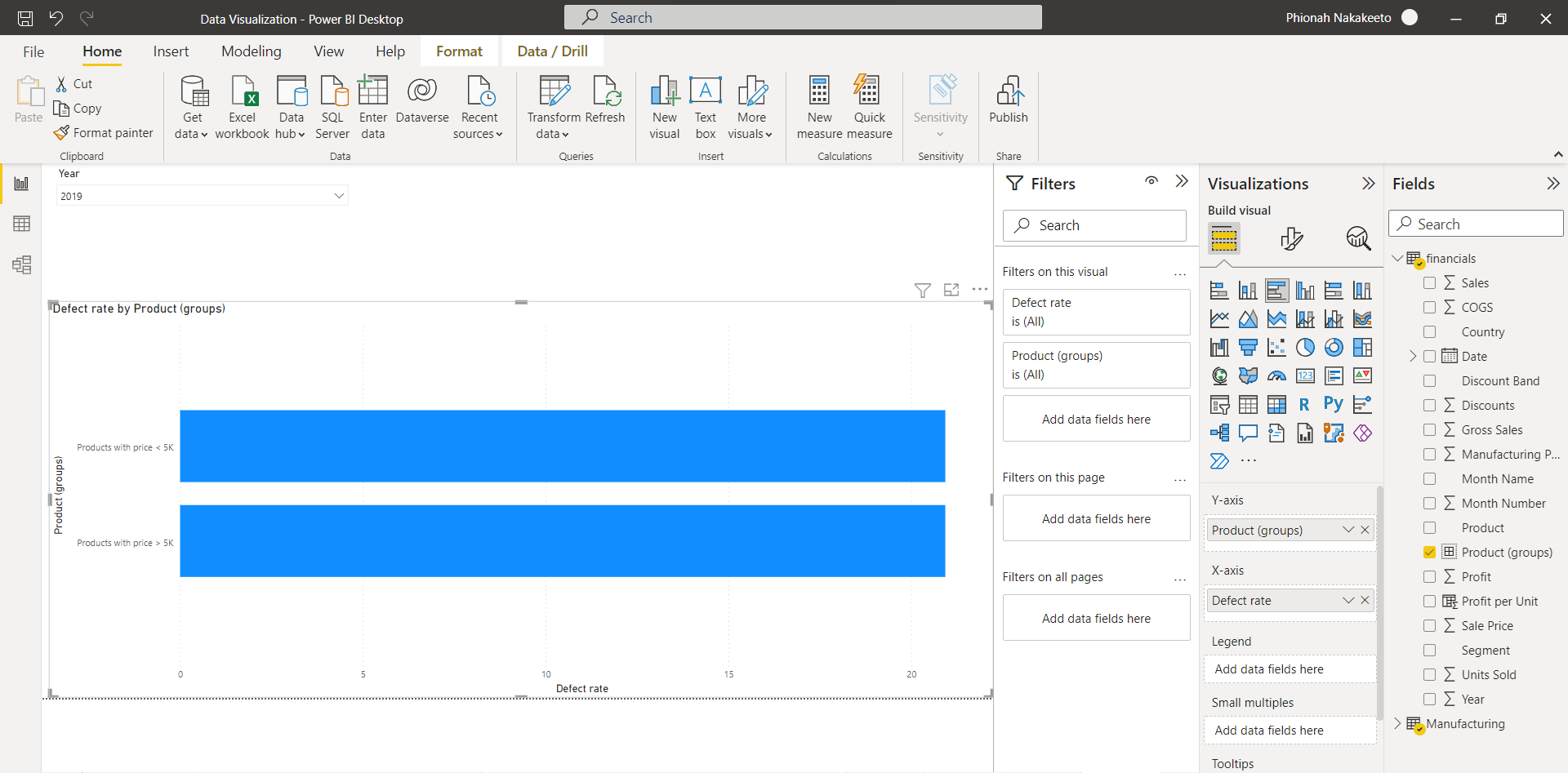
The Products with price < 5K also have greater sales and higher profit per unit than those with prices > 5K.

However, the manufacturing defect rate for the two product groups is the same.









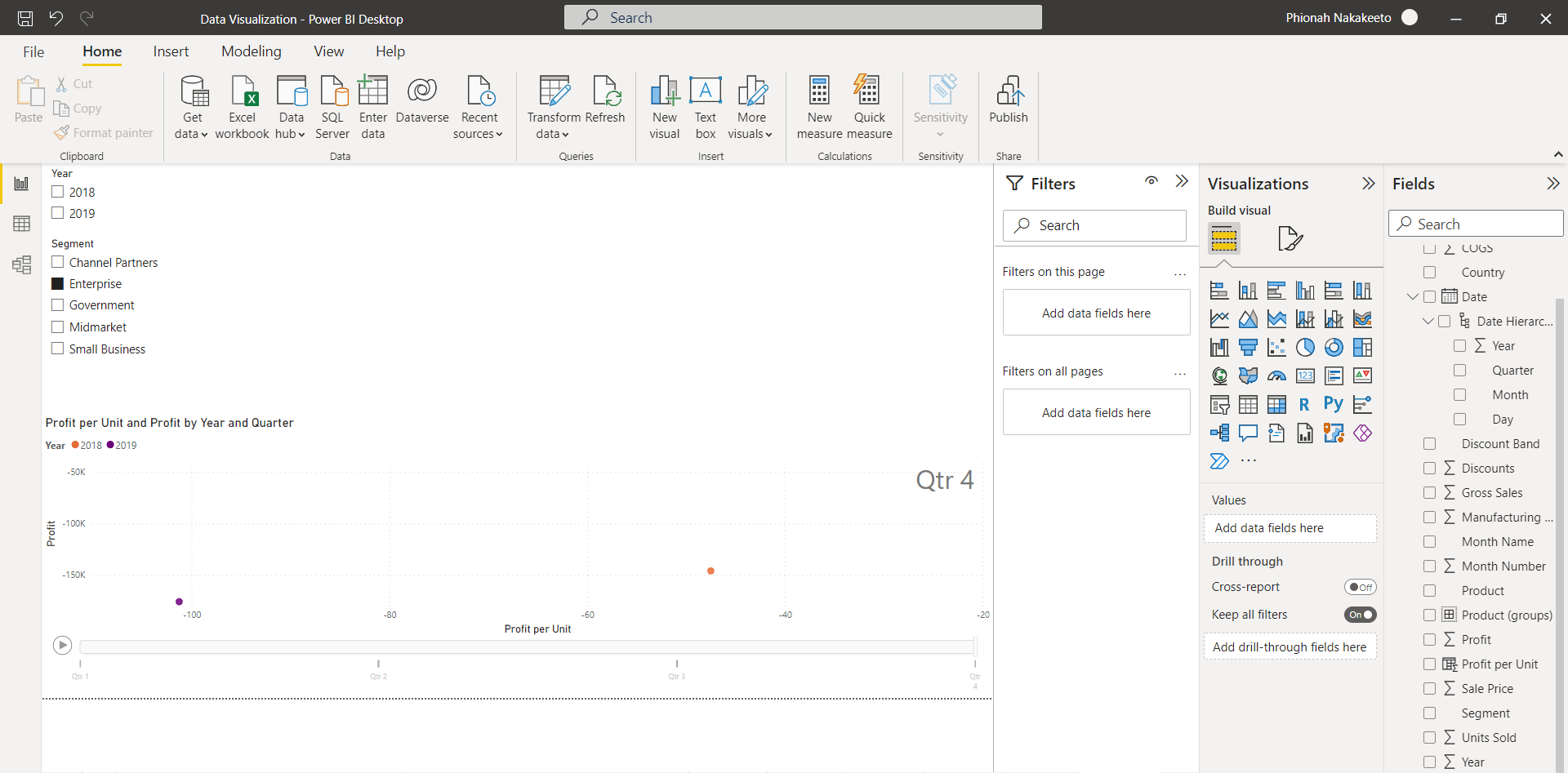
**Which group do you think performed more strongly?**

The Products with price < 5K performed more strongly.

**Compare the profitability of the Enterprise segment for the last quarter of 2018 and 2019. Calculate the profit per unit sold.**

**Did profitability improve in 2019 compared to 2018?**

No, profitability did not improve in 2019. Instead, it reduced.



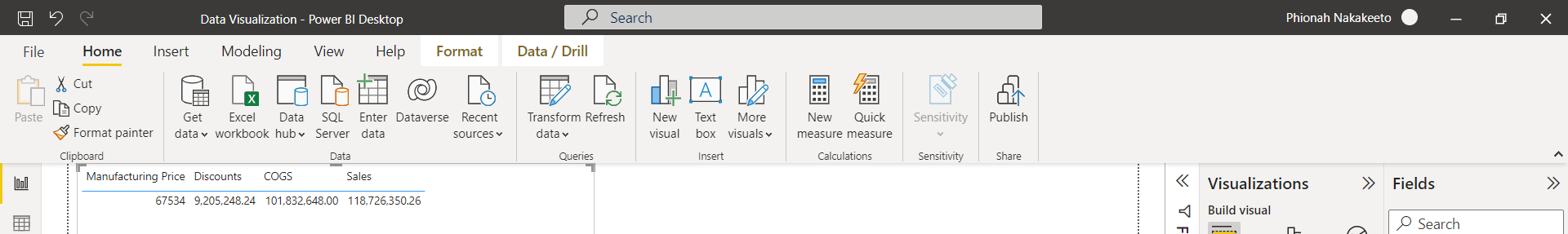
**Why do you think sales are unprofitable?**

Sales are unprofitable because some products' prices are high, leading to reduced sales and overall reduced profit.

**Look at the variables associated with the costs of production and sales, such as manufacturing cost, discount, and COGS. Is there anything unusual in the data that warrants further investigation? If so, what additional information would you need to make recommendations about how to improve profitability of this segment?**

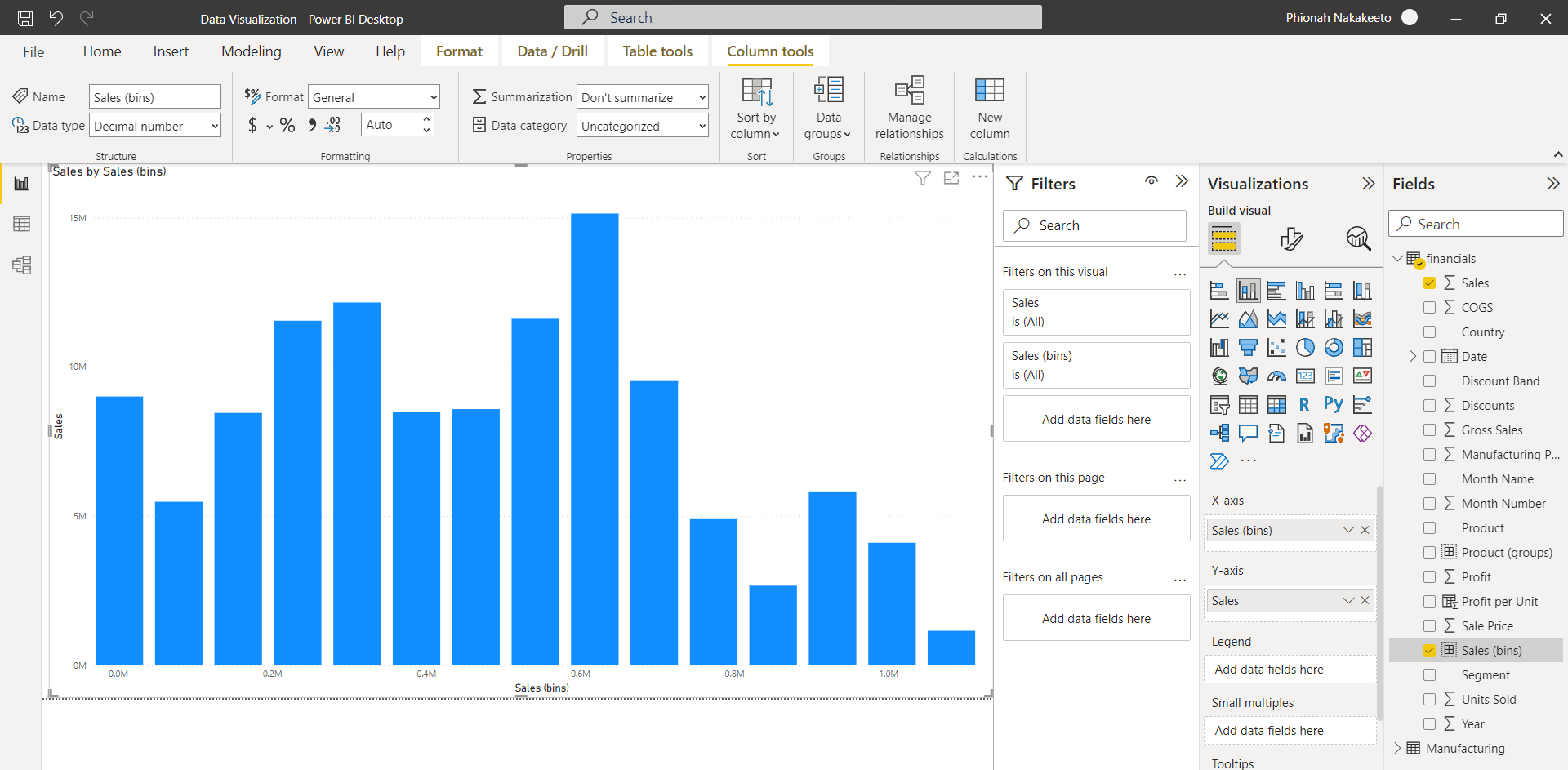
The total COGS is very high. The total sales are only slightly higher than total COGS which means that there are lower profit margins realized. The total COGS is also much higher than the total manufacturing price which is unusual. This would require further analysis.

The company should find ways of reducing the total COGS for the Enterprise segment in order to maximize profit.



**Use Power BI to create the following graphical representations of the data set and answer the following analysis questions.**

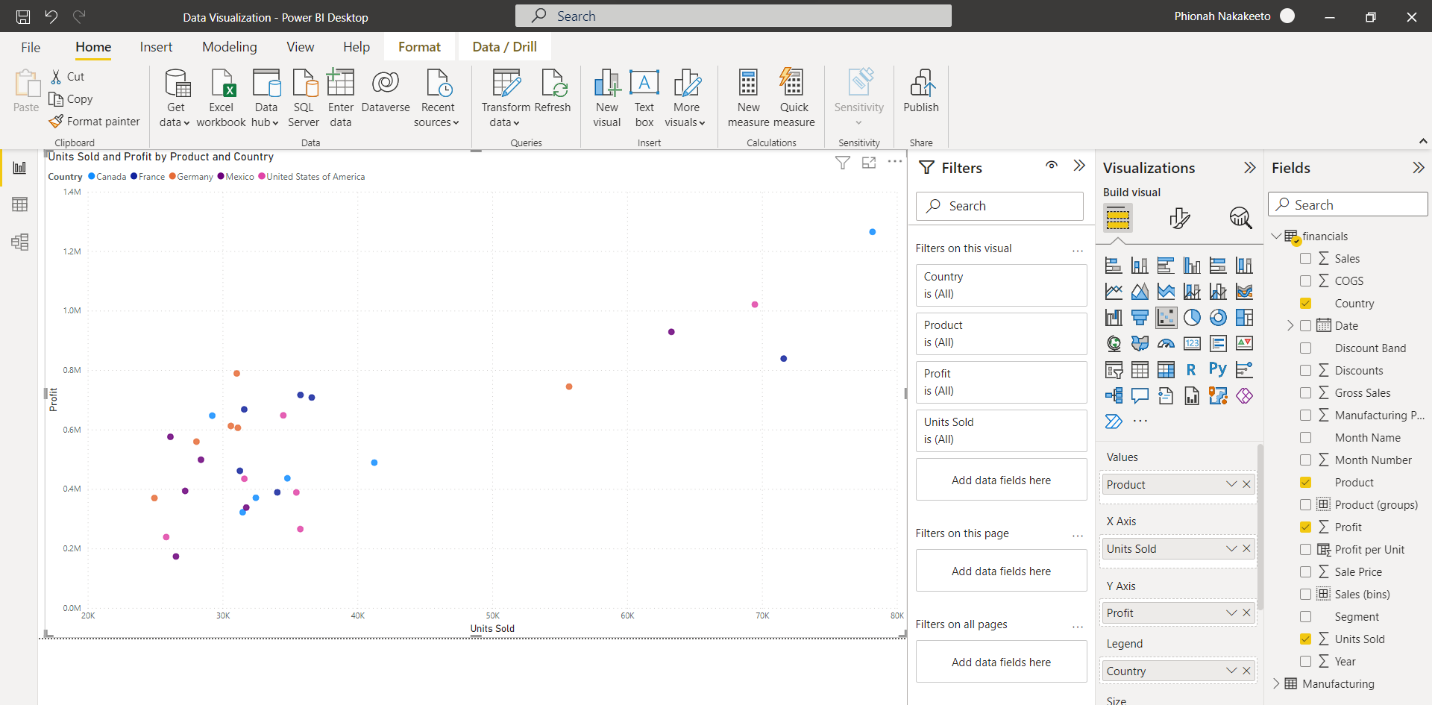
**Create a histogram using the sales variable, showing the sales data distribution. Use 15 bins for the histogram.**



**Create a scatterplot showing units sold and profit by product and by country. What patterns do you notice related to unit sales and profit as it relates to the six products? Is there a product that stands out from the others? If so, how, and what does it mean for business planning?**

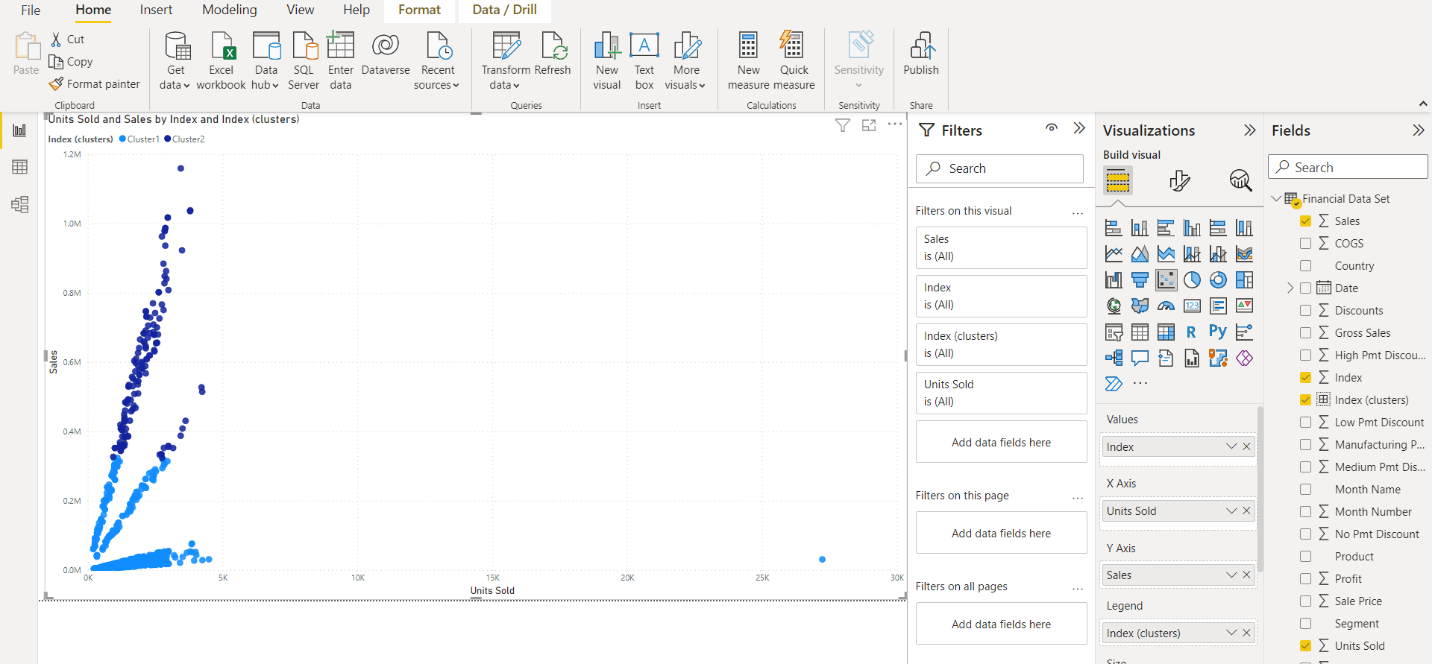
The scatter plot shows that most of the products are not making as much profit i.e., the amount of units sold is more than the realized profit. However, there are some outliers of the product Paseo.

Further investigation should be conducted to find out why one product appears to have relatively more profit and sell more than the rest.



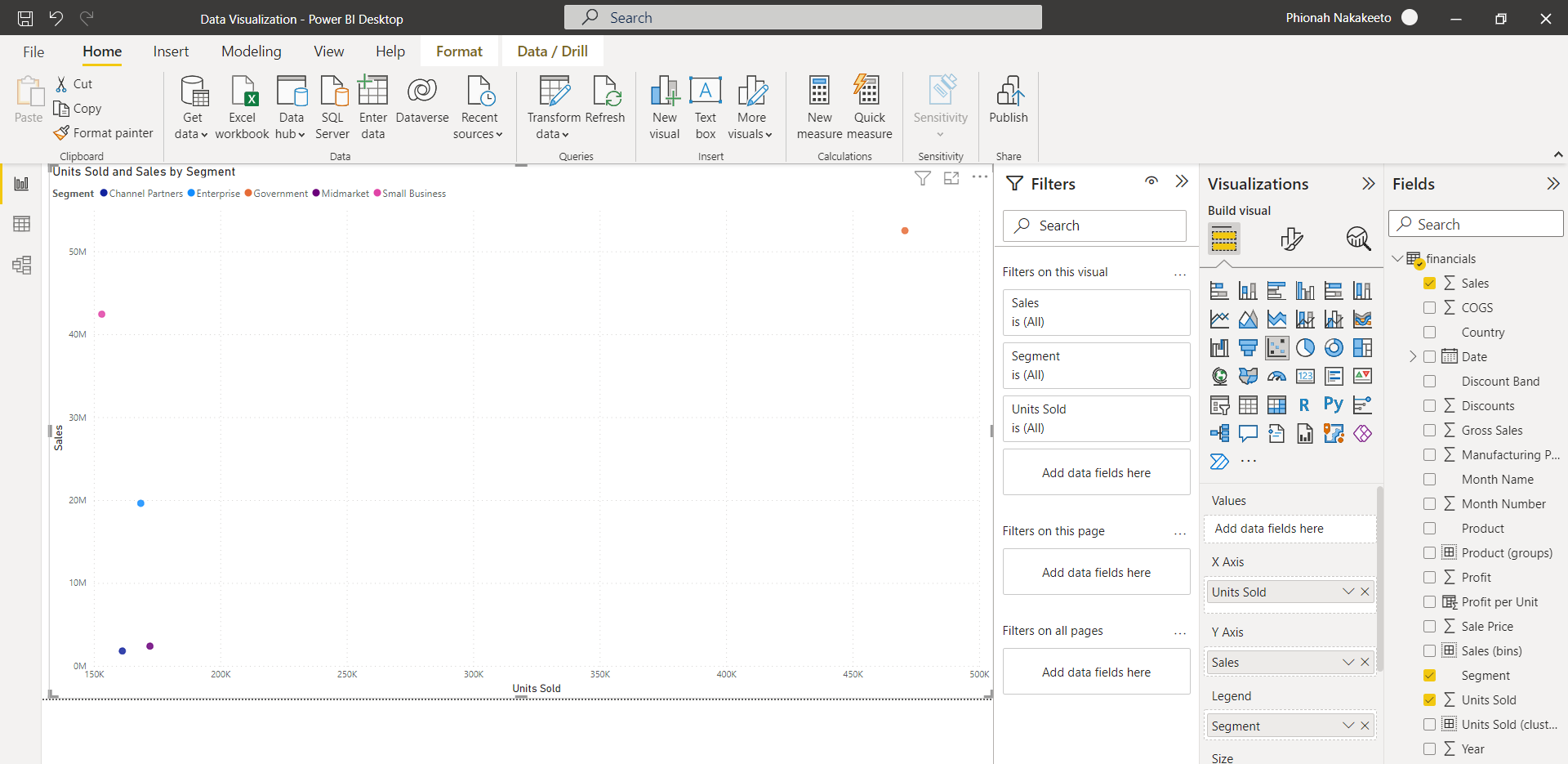
**Create a scatterplot showing units sold and sales by index (transaction) number. Use Power BI's cluster functionality to automatically find clusters in this scatterplot. How many clusters did Power BI identify? Explain whether these clusters make sense to you and whether they would be useful for business purposes.**

Power BI could identify 2 clusters in the scatter plot. The clusters don’t make sense because they show that there are higher sales with fewer units sold. There’s one outlier that shows a significantly higher number of units sold and very low sales.



**Using Power BI, create a scatterplot showing units sold and sales by segment. Do the clusters Power BI identified in Step 3 rely on segment for clustering purposes? In your view, does it make more sense to rely on the clusters identified by Power BI or to rely on segment as the clustering characteristic? Explain your answer.**

Yes, the clusters Power BI identified in Step 3 rely on Segment for clustering purposes and yes, it would make more sense to rely on Segment as the clustering characteristic because it gives us a better idea of why the shape of the scatter plot with Power BI clusters is the way it is. We can now tell which Segment influences the most sales and which one doesn’t.



**Using Power BI, create a box and whisker plot showing each country's profit per transaction. (Note: Power BI does not have built-in functionality for box and whisker plots. You should use a box and whisker chart that you can download and import into your Power BI interface.) In your plot, set the whisker type to less than 1.5 the interquartile range and ensure that each country's transactions are plotted in a separate box. What insights can you learn from the box and whisker plot? Does the visual support that there are substantial differences between countries in average profit per transaction?**

The whisker plot shows that Germany and France have a relatively higher Profit per Unit than Mexico, the USA, and Canada which supports the claim that the Sales department has made about some countries having a higher profit per transaction on average. Therefore, the company should proceed to expand operations in Europe where profitability is higher than in the North American market.

