Hi I am Ivy Hernandez, a Data science student at Southern Alberta Institute of Technology. I am discussing an article about a requirement on one of our course, which is Data043- Data Analysis Application. This article has been authored by a student as a part of a course requirement and does not necessarily reflect the views or opinions of the educational institution. The information presented in this article is based on the student's understanding and interpretation of the dataset, and should be used for educational purposes only. Readers are encouraged to independently verify the accuracy and reliability of the information provided herein before making any decisions or taking any actions based on it. The student and the educational institution shall not be held liable for any errors, omissions, or damages arising from the use of this information.

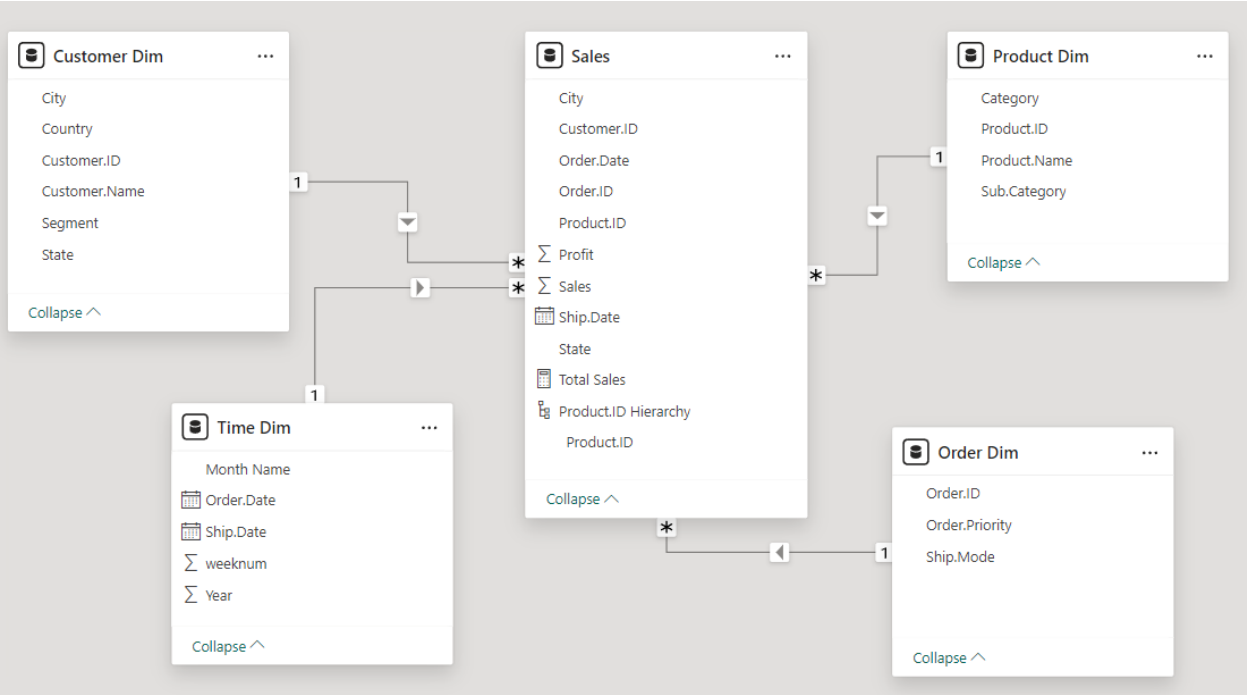
As a requirement for the final presentation for this course we need to find a open-source dataset and establish a data modelling tables for this data set as well as the visualization. So, what I do is I searched for some datasets on Kaggle, and after figuring out what I want. I ended up choosing the Superstore.csv file. Since retail is one of the biggest industries in Canada, using dataset related to the retail industry is a realistic and relevant scenario.

Real Canadian Superstore is a Canadian hypermarket chain owned by Loblaw Companies Limited, one of the largest food retailers in Canada. It offers a wide range of products, including groceries, clothing, electronics, home goods, and pharmacy items. The stores are known for their large size and competitive pricing, aiming to provide customers with a convenient one-stop shopping experience. Real Canadian Superstore operates across Canada, with locations in both urban and rural areas, serving millions of customers each year.

The objectives for choosing this company for data visualization is to analyze sales data to identify trends over time and gain insights into seasonal patterns or fluctuations in demand. We can also analyze profit margins across different product categories, customer segments, or geographic regions to optimize pricing strategies.Also it helps us to use this historical sales data to forecast future sales trends and make informed business decisions.And as a normal company we need to identify areas for operational improvement by analyzing order processing times, shipping costs, and customer service metrics.

So at first, I download this dataset on Kaggle, sort the data first trying to figure out the relationships between the given data after that I'm good to import this data csv file on power BI. And then I worked on doing the tables for data modelling, by duplicating all the datasets, and changing it as to their corresponding table names including the Fact Table which is my Sales Table, the Customer Dim, Time Dim, Product Dim and Order Dim as my dimensions table. I also hoose all the corresponding data needed by choosing and removing the unnecessary columns for the specific table. As I establish my tables, the next thing I do is to establish all the relationship between the tables. I remove the duplicates for each dimension tables. And I ended

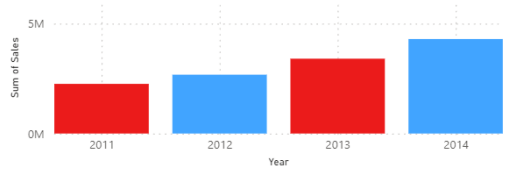
up having one to many relationship only, as the dimension tables have unique column which is linked on my fact table which has multiple transactions. And as I am exploring the tables, I have aggregation on my Time Dimension table which are the year and week number, it shows that it has the data type “number ‘ hence it is showing summation but there is no aggregation operations done in it.



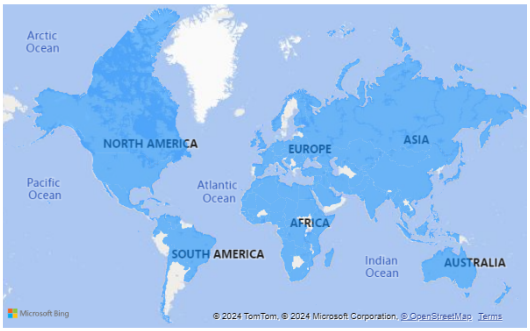
After that I immediately worked on my dashboard. I critically think what I want to visualize with the dataset I have. And I ended up having the **Total sales, Total Profit and Total customers** of the Real Canadian Superstore within the specific period. This helps in gaining a comprehensive understanding of the business's performance, customer behavior, and relationships between different metrics, ultimately supporting data-driven decision-making and strategic planning. Also I have the **Sales by Year** which can help you analyze the sales of the store yearly. It serves as a powerful analytical tool for understanding historical performance, guiding strategic decisions, setting realistic goals, monitoring progress, and communicating insights to stakeholders.**Top sales by Category** which allow you to identify which product categories are performing well or poorly over time. This information can be crucial for strategic decision-making, such as resource allocation, inventory management, and marketing efforts. I also did the **Sum of Sales by Sub Category** which allow you to visualize the specific product that are most sold on the store that can be also a great help on the inventory management. Lastly, I have the table for **Sales by Country** which provides a comprehensive view of your international sales operations, enabling data-driven decision-making and ultimately driving business success.

REAL CANADIAN Superstore

Sum of Sales by Year

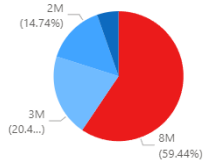


Sum of Sales by Country

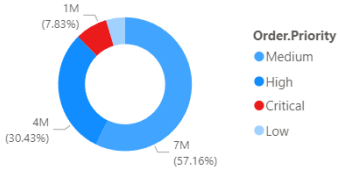


TOTAL PROFIT	TOTAL SALES	TOTAL CUSTOMERS
1.47M	13M	4873

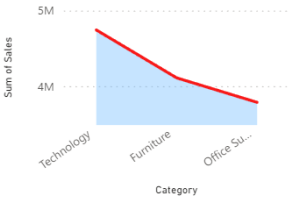
Sum of Sales by Ship.Mode



Sum of Sales by Order.Priority



Sum of Sales by Category



Sum of Sales by Sub.Category