Table of Contents

[Abstract/Overview 2](#_Toc157466235)

[Assignment Summary 2](#_Toc157466236)

[Assignment Background 3](#_Toc157466237)

[Assignment Tasks 3](#_Toc157466238)

[Part 1: Creating dashboards to address business questions 3](#_Toc157466239)

[Part 2: Further Analytics 3](#_Toc157466240)

[Conclusion 3](#_Toc157466241)

[Data Cleaning 3](#_Toc157466242)

[Creating dashboard to answer business question (for Part 1) 7](#_Toc157466243)

[1. Period with Most Sales 7](#_Toc157466244)

[1.1 Dashboard Components 7](#_Toc157466245)

[1.2 Insights Derived: 8](#_Toc157466246)

[2. Top 5 Most Popular Products Across Sales Channels 9](#_Toc157466247)

[2.1 Dashboard Components 9](#_Toc157466248)

[2.2 Insights Derived: 10](#_Toc157466249)

[3. Top & Bottom 5 Branches in Terms of Profit 11](#_Toc157466250)

[3.1 Dashboard Components 11](#_Toc157466251)

[3.2 Insights Derived: 12](#_Toc157466252)

[4. Top 3 Manufacturers by Sales 13](#_Toc157466253)

[4.1 Dashboard Components 13](#_Toc157466254)

[4.2 Insights Derived: 13](#_Toc157466255)

[5. Conclusion 15](#_Toc157466256)

[6. Dax Formula Used 16](#_Toc157466257)

[Further analysis (for Part 2) 18](#_Toc157466258)

[Univariate Analysis: 18](#_Toc157466259)

[Dashboard 1: Sales Distribution, Monthly Trends, and Promotional Impact 18](#_Toc157466260)

[Dashboard 2: Product Performance, Subcategory Analysis, and Profit Metrics 19](#_Toc157466261)

[Dashboard 3: Store Analysis and Regional Performance 20](#_Toc157466262)

[Conclusion: 20](#_Toc157466263)

[Bivariate Analysis: 21](#_Toc157466264)

[Dashboard 1: Probing Sales and Basket Size Variations by Branch 21](#_Toc157466265)

[Dashboard 2: Displaying Geographic and Transactional Insights 22](#_Toc157466266)

[Dashboard 3: Analysing Product and Regional Performance 23](#_Toc157466267)

[Conclusion: 23](#_Toc157466268)

[Multivariate Analysis: 24](#_Toc157466269)

[Branch Region Analysis: 24](#_Toc157466270)

[Channel Type Analysis: 25](#_Toc157466271)

[Business Perspectives 25](#_Toc157466272)

[Data Mining 26](#_Toc157466273)

[Data mining with K-means Clustering: 26](#_Toc157466274)

[Methodology: 26](#_Toc157466275)

[Data Scaling and Model Fitting: 26](#_Toc157466276)

[Interpreting Results: 26](#_Toc157466277)

[Business Insights: 27](#_Toc157466278)

[Conclusion: 27](#_Toc157466279)

[Summary 28](#_Toc157466280)

[Introduction and Project Overview 28](#_Toc157466281)

[Part 1: Creating Dashboards to Address Business Questions 28](#_Toc157466282)

[Part 2: Further Analytics 28](#_Toc157466283)

[Data Cleaning 28](#_Toc157466284)

[Conclusion and Business Implications 28](#_Toc157466285)

[Overall Implications 28](#_Toc157466286)

[Conclusion 29](#_Toc157466287)

# Abstract/Overview

## Assignment Summary

I'll be putting my acquired knowledge and skills from the Data Exploration and Analysis (DEA) module to the test through a thorough case study in this assignment. The goal is to use analytics methods on actual data and offer useful insights for business decision-making. I'll have the chance to improve TechScape Singapore's sales strategies and overall business success by digging into the offered datasets.

## Assignment Background

Future-focused technology company TechScape Singapore was founded in 2020 and focuses on cutting-edge electronics and intelligent solutions. I'll be entrusted with using data to assist the company's changing demands as it embraces the digital age. TechScape has demonstrated its dedication to data-driven decision-making by taking steps to form a Business Analytics team and consider an Enterprise Data Warehouse (EDW). The task relies on merging various datasets and employing data modeling approaches given the variety of CSV files containing crucial company data and the initiative to develop a central repository.

## Assignment Tasks

### Part 1: Creating dashboards to address business questions

My responsibility in this portion of the assignment is to develop dashboards that, using the supplied datasets, respond to particular business queries. I'll have to respond to inquiries like determining the peak sales period, listing the top 5 products across all sales channels, and rating the profitability of branches. I also have to create a dashboard that displays the Key Performance Indicators (KPIs) regarding the top 3 manufacturers in order to provide depth.

### Part 2: Further Analytics

I'll go a step further in Part 2 and analyze the data using advanced analytics methods. This requires the use of univariate, bivariate, and multivariate analysis. My objective is to find hidden patterns, paying close attention to whether some products show clustering based on geographic regions or consumer purchase trends. For enhancing business success, these insights are essential. I will go into detail about the methods I chose, outlining how they reveal additional layers of knowledge about the industry. My findings will be successfully illustrated through the use of visualizations.

## Conclusion

By bridging the gap between theory and practice, this assignment enables me to put the lessons I learned from the DEA module into practice. I have the opportunity to contribute to insightful decision-making by immersing myself in TechScape Singapore's data landscape. Building useful dashboards and discovering cutting-edge insights are completely in line with the changing world of corporate analytics, where data is the compass for success. I'm developing my analytical skills as I work through this assignment, and I'm also getting a firsthand understanding of how important data is in forming contemporary company strategy.

# Data Cleaning

A screenshot of a computer

Description automatically generated

While looking at the dataset, I found out that in the promotion file, both Holidays have the same date and discount percentage, however the name is slightly different.

A screenshot of a computer

Description automatically generatedMaking both PromotionName the same

A screenshot of a computer

Description automatically generated

Replace “\_” with “/” so that we can change this column datatype from string into Date

A screenshot of a computer

Description automatically generated

Change datatype in sales table to int

A screenshot of a computer

Description automatically generated

Change the properties of all data  
All the Key (Unique identifier of each Category): Don’t summarise

Longitude & Latitude: Don’t summarise

A white surface with black dots

Description automatically generated with medium confidence

Creating Calendar Table

# Creating dashboard to answer business question (for Part 1)

## 1. Period with Most Sales

The dashboard created to answer the query "Which period has the most sales?" is insightfully analyzed in this report. This dashboard's main goal is to identify the time frame, such as a certain year or quarter, that has the biggest sales. The dashboard enables TechScape Singapore to identify seasonal trends and better efficiently deploy resources by visualizing this information.

A screenshot of a computer

Description automatically generated

### 1.1 Dashboard Components

#### 1.1.1 Line and Clustered Chart - Sales, Profit, and Profit Margin:

This visual component combines a line chart showing the profit margin with clustered bar charts showing sales and profit. The chart illustrates the overall performance of various time periods by classifying the data into years and quarters.

#### 1.1.2 Waterfall Chart - Yearly Sales Growth Rate:

The waterfall chart demonstrates the yearly sales growth rate in a dynamic way, highlighting the numerous aspects that contribute to the rise or decline of sales over the course of a year.

#### 1.1.3 Line Chart - Sales by Year:

This line chart gives a visual depiction of the sales trend across time, making it possible to see general growth trends.

#### 1.1.4 Donut Chart - Sales by Seasons:

The donut chart divides sales data into seasons and provides information on how sales volume varies throughout the year.

#### 1.1.5 Heatmap - Sales by Year and Month:

The heatmap displays sales data over various years and months to make it easier to spot peak times.

#### 1.1.6 Slicer Section - Individual Promotion Details:

The slicer gives users the ability to filter data depending on specific promotions. Users can examine each promotion's unique performance metrics, such as sales volume, sales volume, transactions, and discount percentage.

### 1.2 Insights Derived:

Through the dashboard, valuable insights were uncovered:

#### 1.2.1 Highest Sales Period:

The heatmap, line, and clustered charts all provide a clear visual picture of the highest sales times. This made it possible to recognize peak sales periods right away.

#### 1.2.2 Annual Growth Trends:

The waterfall chart showed the factors that contributed to yearly sales growth or decline, providing information about the forces that affected sales performance.

#### 1.2.3 Seasonal Analysis:

The donut chart showed how sales were distributed over the year's seasons, revealing any recurring trends or consumer preferences.

#### 1.2.4 Sales Trend:

To help identify long-term growth patterns, the line chart showing sales by year provided a thorough overview of sales trends throughout time.

#### 1.2.5 User-Driven Exploration:

The slicer section gave users the ability to look at specific promotion details and gain a deeper understanding of how each promotion performed.

## 2. Top 5 Most Popular Products Across Sales Channels

The dashboard is designed to answer the query "What are the Top 5 Most Popular Products for the Different Sales Channels?". The dashboard seeks to offer insights into the most well-liked items across various sales channels, helping TechScape Singapore to improve its marketing approaches and successfully meet customer preferences.

A screenshot of a computer

Description automatically generated

### 2.1 Dashboard Components

#### 2.1.1 PIE Chart - Top 5 Products by Sales:

The sales distribution among the top 5 goods is clearly shown visually in the PIE chart. This quick overview makes it easier to find the products that contribute the most to sales.

#### 2.1.2 Heatmap - Product Performance Over Years:

The performance of the leading products throughout several years is shown on the heatmap. We can see patterns and changes in popularity over time thanks to this heatmap.

#### 2.1.3 Line Chart - Monthly Revenue, Profit, Profit Margin and Sales Quantity:

The line graphs depict the top 5 items' monthly sales, profit, profit margin, and sales volume trends. This is to identify patterns and potential seasonality.

#### 2.1.4 Treemap - Sales Amount by Product Category and Subcategory:

The treemap shows how sales amounts are distributed among various product categories and subcategories. Understanding which categories generate the most revenue is made easier by this information.

#### 2.1.5 Interactive Section - Specific Product Information:

Users can filter data using the interactive function of the slicer depending on specific items. This section offers a deeper look into the products ranking, amount sold and its total sales.

### 2.2 Insights Derived:

Through the dashboard, valuable insights were uncovered:

#### 2.2.1 Product Distribution:

The PIE chart made apparent which products made up the top 5, with Camcorders accounting for the largest share of total sales.

#### 2.2.2 Seasonal Performance:

The heatmap showed Camcorders and Projectors & Screens have a decreasing trend with its sales decreasing by year. Product E, on the other hand, is increasing steadily indicating its increase in demand from over the years.

#### 2.2.3 Sales, Profit, Profit Margin and Sales Volume Trends:

Sporadic surges for particular items during particular months were visible on the line charts. These increases can be attributed to focused marketing activities or outside factors.

#### 2.3.4 Category Insights:

The treemap visualization effectively explained how sales amounts were distributed throughout different categories and subcategories. This knowledge could be used to inform strategic choices for concentrating resources on high-revenue categories.

## 3. Top & Bottom 5 Branches in Terms of Profit

The dashboard designed to answer the query "What are the Top & Bottom 5 branches in terms of profit?" is thoroughly analyzed in this report. This dashboard's goal is to determine which divisions are most and least profitable for the business. The dashboard helps TechScape Singapore make data-driven decisions by visualizing this data.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

### 3.1 Dashboard Components

#### 3.1.1 Top & Bottom 5 Branches on the BAR Chart by Profit:

The top five most profitable and the bottom five least profitable branches are effectively showed using the BAR chart. The branches with the highest and lowest profitability are highlighted.

#### 3.1.2 Scatter Plot - Profit vs. Sales Per Basket:

The scatter plot shows the relationship between profit and sales per basket for each branch. When profitability and average sales per basket are compared, this helps reveal any patterns or outliers.

#### 3.1.3 Spark Line Chart - Trend of Profit and Profit Margin:

For the chosen branches, the sparkline chart shows the trend of profit and profit margin over time. The identification of any peaks or valleys in profitability.

#### 3.1.4 Matrix Table - Branch Details:

The matrix table presents a structured overview of branch-related data, including branch location, employee count, area measurement, and sales per basket. This comprehensive viewpoint lends perspective to the statistics on profitability.

#### 3.1.5 Treemap - Sales Amount by Branch Region and Location:

The distribution of sales volumes across various branch region and locations is efficiently communicated by the treemap. This aids in comprehending regional performance.

#### 3.1.6 Slicer Section - Individual Branch Details:

The section enables users to filter data for specific branches. Users of this interactive feature can examine in-depth data about the profitability, discounts and region of each branch.

### 3.2 Insights Derived:

Through the dashboard, valuable insights were uncovered:

#### 3.2.1 Branch Profitability:

The BAR chart made it evident which branches were the top five most and least profitable, focusing attention on the branches that made up the highest and lowest percentages of the business's profitability.

#### 3.2.2 Profit vs. Sales Per Basket:

The scatter plot demonstrated whether there is a relationship between greater sales per basket and greater profitability, enabling the identification of trends in branch performance. In this case, the greater the average sales per basket, the greater the profitability.

#### 3.2.3 Profitability & Profit Margin patterns:

For the selected branches, the sparkline charts showed the profitability patterns over time. This understanding makes it easier to spot branches that could need tactical interventions. The sparkline profit margin chart showed how profit margins remained stable throughout time. Any anomalies can be a cause for concern.

#### 3.2.4 Branch Information:

The matrix table provided an ordered overview of the main branch information and contextualized the profitability data by including other pertinent data.

#### 3.2.5 Regional Insights:

The treemap helped to comprehend the overall performance of the region by seeing the distribution of sales quantities across various branch regions and locations.

## 4. Top 3 Manufacturers by Sales

In order to answer the query, "Who are the Top 3 manufacturers in terms of sales?" this research goes deep into the dashboard analysis. The dashboard's goal is to pinpoint the vendors driving the majority of sales so that TechScape Singapore may make educated choices about its product lineup and business alliances.

A screenshot of a computer

Description automatically generated

### 4.1 Dashboard Components

#### 4.1.1 Stacked Bar Chart - Profit & Total Cost:

The stacked bar chart contrasts profit and total cost for the goods produced by each manufacturer. This comparison clarifies how profitable goods from various manufacturers are.

#### 4.1.2 Line Charts - Monthly Revenue, Profit, and Profit Margin:

The line charts indicate the trends in monthly revenue, profit, profit margin, and sales quantity for the top-selling goods of the top 3 manufacturers. This gives information on how well their products have performed over time.

#### 4.1.3 Matrix Table - Manufacturer Details:

The matrix table offers a comprehensive picture of data on manufacturers, including the manufacturer's top product, produced sales amount, sales volume, and profit. The sales data are contextualized by this matrix.

#### 4.1.4 Slicer Section - Individual Manufacturer Details:

The section enables users to filter data for specific manufacturers. Users can look for specific details on the best-selling products and related data from each manufacturer.

### 4.2 Insights Derived:

Through the dashboard, valuable insights were uncovered:

#### 4.2.1 Profitability and Costs:

The stacked bar chart made it easy to quickly compare the profitability and overall cost of each manufacturer's products.

#### 4.2.2 Trends Analysis:

The manufacturers' top-selling products were emphasized in the line charts for revenue, profit, profit margin, and sales volume. Understanding product performance requires this knowledge.

#### 4.2.3 Sales Quantity Patterns:

The line chart showing sales quantity patterns can be used to spot spikes in demand and changes in a product's level of popularity.

#### 4.2.4 Manufacturer Insights:

The matrix table gave a consolidated picture of important manufacturer information, making it easier to comprehend how manufacturers affect sales and profitability in general.

#### 4.2.5 Detailed Exploration:

The slicer functionality allowed users to delve further into the performance of a single company and their products.

## 5. Conclusion

In conclusion, the four meticulously constructed dashboards help to provide a comprehensive overview of the business environment in TechScape Singapore. These visual representations, which were created to respond to specific business questions, are an excellent example of how data-driven insights can be used to inspire strategic choices and improve overall business performance.

The "Which Period has the Most Sales?" dashboard makes it easier to identify busy shopping seasons and seasonal patterns. Users may thoroughly investigate promotions using the interactive aspects, which improve promotional methods.

Actionable insights into product distribution, trends, and profitability are provided by the dashboard examining the "Top 5 Most Popular Products Across Sales Channels". TechScape can efficiently optimize marketing strategy and cater to client preferences by combining visualizations and interactive elements.

A thorough overview of branch profitability, trends, and related information is provided by the dashboard that focuses on the "Top & Bottom 5 Branches in Terms of Profit". TechScape can enhance branch performance and overall business profitability through interactive features and visualization.

The "Top 3 Manufacturers in Terms of Sales" dashboard provides key insights into manufacturer profitability, sales trends, and related information. This gives TechScape the ability to choose alliances and product offerings strategically.

These dashboards demonstrate the value of data visualization in converting complicated datasets into understandable insights. The interactive elements enable users to investigate, examine, and draw conclusions that may be put into practice, encouraging an environment conducive to informed decision-making. These dashboards are essential tools for forming future plans, maximizing profitability, and providing superior value to consumers and stakeholders as TechScape Singapore embraces the changing world of business analytics.

## 6. Dax Formula Used

*A screenshot of a computer program

Description automatically generated*

A screenshot of a computer program

Description automatically generated

# Further analysis (for Part 2)

## Univariate Analysis:

This section delves into the method used to carry out the univariate analysis using three different dashboards. Each dashboard was carefully created using a variety of visual analytics tools to study particular areas of TechScape Singapore's business operations. The selection of the tasks and visualizations is explained, as well as how they could lead to more insightful conclusions.

### Dashboard 1: Sales Distribution, Monthly Trends, and Promotional Impact

A screenshot of a computer screen

Description automatically generated

#### Step 1: Sales Channel Distribution (PIE Chart)

Understanding how various sales channels affect the company's overall revenue was the goal of this visual. A PIE chart was selected because it efficiently depicts proportions. With that, the most profitable channels are immediately clear, making it possible to allocate resources strategically.

#### Step 2: Monthly Sales Trends (Bar Chart)

The goal was to find any seasonal patterns or trends by choosing a bar chart that showed sales each month. This option gives decision-makers the ability to spot changes throughout the year, assisting in the optimization of inventory and marketing initiatives.

#### Step 3: Geographical Sales Distribution (Map Chart with Bubble Size)

To visually depict sales in various places, a map chart with bubble sizes to illustrate sales quantity was chosen. The justification for this representation was that it not only offers geographical information, but also enables stakeholders to estimate the size of sales in each location.

#### Step 4: Promotional Sales (Bar Chart)

A bar chart that displays the total sales amount by promotion aids in assessing the efficacy of various promotions. This duty enables management to identify the promotions that result in significant revenue increases and modify their promotional tactics as necessary.

### Dashboard 2: Product Performance, Subcategory Analysis, and Profit Metrics

A screenshot of a computer screen

Description automatically generated

#### Step 1: Sum of Sales by Product Category (Bar Chart)

The goal of choosing a bar chart to show the total sales by product category was to pinpoint the most lucrative product categories. Making educated decisions about product offers and marketing tactics is made easier with the aid of this representation.

#### Step 2: Sales Distribution by Product Subcategory (Histogram)

To identify distribution trends within each subcategory, a histogram was chosen to represent sales by product subcategory. This process gives TechScape information about the popularity of particular subcategories, which helps them focus on those areas while developing new products.

#### Step 3: Profit Distribution by Product (Histogram)

To assess the profitability of all products, a histogram displaying the distribution of profit by product was selected. This knowledge is essential for locating goods that considerably boost the bottom line of the business. And the optimal profit range that TechScape may consider pricing their products at.

#### Step 4: Unit Cost Distribution by Product (Histogram)

The goal was to understand cost structures across various products, hence a histogram was chosen to show the distribution of unit costs by product. Pricing methods and cost optimization benefit from this knowledge.

### Dashboard 3: Store Analysis and Regional Performance

A screenshot of a graph

Description automatically generated

#### Step 1: Number of Stores in Each Region (Bar Chart)

Comparing regional presence is made possible by using a bar chart to show the number of stores in each region. This information aids in determining regional coverage and possible growth locations.

#### Step 2: Sales Distribution in Each Region (Bar Chart)

Using a bar chart to visualize the distribution of sales by region makes it easier to see whether regions have larger or lower sales volumes. This knowledge directs the use of resources and focused marketing initiatives.

#### Step 3: Number of Employees Distribution by Store (Histogram)

To comprehend the distribution of employee counts, a histogram was used to depict the number of employees by store. This data is essential for assessing the potential effects of number of employees on sales and customer satisfaction.

#### Step 4: Area Size Distribution by Store (Histogram)

To comprehend the distribution of store sizes, a histogram was used to depict area sizes by store. This data is essential for assessing the potential effects of store size on sales and customer satisfaction.

### Conclusion:

The methodical methodology used for each univariate analysis dashboard was motivated by the need to offer TechScape Singapore's decision-makers insightful and useful information. Stakeholders obtain a greater understanding of sales trends, product performance, store efficiency, and geographical discrepancies by carefully choosing visuals that match the specific objectives of each dashboard. This method not only improves the readability of the data, but also equips TechScape to take data-driven decisions that improve company processes and spur growth. These dashboards provide TechScape with useful tools for navigating the intricacies of the business landscape by fusing data, visualization, and strategic insights.

## Bivariate Analysis:

The systematic method used to conduct bivariate analysis across three different dashboards is thoroughly described in this section. Each task's purpose and selected visualization are described in detail, highlighting any insights that these analyses may be able to add to TechScape Singapore's understanding of how their business is run.

### Dashboard 1: Probing Sales and Basket Size Variations by Branch

A screenshot of a computer

Description automatically generated

#### Step 1: Sales and SalesPerBasket by Branch (Scatter Plot)

The rationale behind this visual was to identify any potential relationships between overall sales and the average sales per basket in various branches. The visualization can identify patterns and outliers using a scatter plot that may show branches with synchronized sales and basket size growth. This insightful information can guide targeted efforts like improved store layouts or custom marketing campaigns.

#### Step 2: SalesAmount by ProductCategory and ProductSubcategory (Treemap)

To create a hierarchical depiction, I decided to use a treemap to illustrate sales distribution within product categories and subcategories. This method quickly identifies important revenue generators, assisting TechScape in identifying profitable product categories. As a result, this supports inventory management and focused marketing initiatives to increase sales where they are most needed.

#### Step 3: SalesAmount by ProductSubcategory and ProductCategory (Box Plot)

To highlight sales concentration and potential outliers, a box plot was chosen to depict the distribution of sales quantities within product categories and subcategories. This research reveals the distribution within high-performing categories in addition to identifying them. TechScape can proactively manage inventory and enhance product offers by taking these precisions into account.

### Dashboard 2: Displaying Geographic and Transactional Insights

A screenshot of a computer

Description automatically generated

#### Step 1: SalesAmount by Branch Region and Location (Treemap)

A treemap was used to visualized to present a hierarchical view of geographical dynamics, sales distribution across branch regions and locations. This gives stakeholders the ability to pinpoint the key revenue drivers in each region, ultimately influencing decisions about resource allocation and tactical expansion.

#### Step 2: Number of Transactions & AreaSize by Branch (Scatterplot)

It is intended to shed light on trends in store efficiency and customer engagement by using a scatter plot to examine the link between transaction count and retail area size. When deciding on store architecture and management, it can be determined whether greater store areas result in more customer activity by visualizing this correlation.

#### Step 3: Sales & SalesPerBasket by Branch (Scatterplot)

To determine whether branches with greater sales also have higher average sales per basket, a scatter plot will be used to investigate the relationship between total sales and sales per basket across various branches. These data point up potential cross-sell or upsell opportunities, increasing overall revenue.

### Dashboard 3: Analysing Product and Regional Performance

A screenshot of a computer

Description automatically generated

#### Step 1: SalesAmount by BranchRegion and ChannelType (Stacked Bar Chart)

The aim is to compare distribution in a stacked bar chart to show sales amounts across branch regions and channel types. By identifying regional preferences for different buying channels, this data enables TechScape to adjust marketing campaigns as necessary.

#### Phase 2: SalesAmount by BranchRegion and ProductCategory (Heatmap)

To discover regions and product categories that perform well, a heatmap is chosen to show sales distribution across branch regions and product categories. TechScape can identify relationships between particular product categories and regional preferences thanks to this combined analysis, which helps with inventory management and targeted advertising.

#### Step 3: SalesAmount by BranchRegion and PromotionName (Clustered Bar Chart)

Finding the effectiveness of a promotion can be done by choosing a clustered bar chart to compare sales amounts across branch regions and promotion names. This analysis informs the effectiveness of advertising in various geographic contexts, enabling regionally specific promotional methods.

### Conclusion:

These three dashboards' bivariate analyses serve as an example of how visual exploration can reveal complex correlations in data. Whether it be sales trends, geographic dynamics, or promotional efficacy, each visualization and analysis that is chosen is supported by a reasoning that is concise and in line with the relevant business questions. With the use of this organized method, TechScape Singapore is able to acquire detailed insights into their business operations, enabling them to make well-informed decisions for resource optimization, marketing strategy improvement, and ultimately improved business performance.

## Multivariate Analysis:

By taking into account numerous factors at once, multivariate analysis is essential for gaining deeper insights from complex datasets. Through this method of research, we can spot trends, correlations, and patterns that may not be obvious through univariate or bivariate analysis. In this section of the study, we'll go through two different multivariate analysis scenarios that each use Principal Component Analysis (PCA) to reduce dimensionality and visualize data. To better comprehend the business dynamics, we'll go into more detail about the step-by-step procedure and the reasoning behind the decisions made.

### Branch Region Analysis:

A graph of a graph

Description automatically generated with medium confidence

In the first case, our goal is to comprehend the relationships between the variables "TotalDiscountAmount," "TotalCost," "SalesAmount," "EmployeeCount," and "AreaSize" in the context of several branch regions. The purpose of this analysis is to see whether certain regions show comparable patterns of these variables and perhaps to find unobserved trends that could have an impact on the business strategy.

To ensure that all variables are on comparable scales, we first standardize the data using the Standard Scaler. The dimensionality is then reduced to two main components using PCA. The greatest substantial variation in the data is captured by these components. To displaying the findings, we construct a scatter plot in which each point represents a branch location, and its placement is dictated by the values of 'col1' and 'col2'. To identify clusters or patterns, we employ colour to distinguish between various branch sections.

The second plot shows the proportion of variation that each primary component contributes to. We may determine the ideal number of components to preserve by using this plot to better understand how much information each component retains.

### Channel Type Analysis:

A graph of a graph of a graph

Description automatically generated with medium confidence

In the second case, the emphasis is shifted to examine how the same set of variables connect to various channel types. Channel types indicate many ways to make purchases and could exhibit unique patterns depending on the factors taken into account. The company can modify tactics and promotions based on the clients' preferred channel types by knowing these tendencies.

We normalize the data and perform PCA for dimension reduction, which is similar to the first situation. This time, however, we base our research on "ChannelType" rather than "BranchRegion." The scatter plot now shows how various channel types align in the reduced-dimensional space defined by 'col1' and 'col2'. Once more, colour is employed to distinguish between different channel types, making it easier to spot clusters or patterns.

The variance explained plot continues to play a crucial part in our research, helping us choose how many components to keep.

### Business Perspectives

Finding patterns and linkages among several variables that may influence business decisions is the justification for these multivariate analyses. We can find branches or channel types that behave similarly by displaying clusters or patterns, which can direct focused marketing campaigns or operational improvements. Additionally, the variance explained plot enables us to determine the amount of data that the reduced-dimensional representation captures and aids in choosing the right number of components.

In conclusion, multivariate analysis offers a powerful way to convert highly dimensional data into insights that guide corporate initiatives, especially when using methods like PCA. By considering several variables at once, we are able to develop a more thorough understanding of the underlying dynamics and relationships, allowing us to make data-driven decisions that will eventually improve business performance. This report section's step-by-step methodology illustrates how PCA can be used to analyze complex datasets and get useful insights.

## Data Mining

### Data mining with K-means Clustering:

A graph with colored dots

Description automatically generated

K-means clustering is a fundamental data mining approach that is essential for spotting innate patterns and structures in datasets. Data points are organized into clusters based on their commonalities, simplifying the identification of hidden insights and assisting in decision-making. We used K-means clustering to find significant connections between product sales amounts and quantities in our investigation of TechScape Singapore's business data, revealing light on possibly different client preferences and purchase behaviours.

### Methodology:

We concentrated on the 'SalesAmount' and 'SalesQuantity' properties of various items in order to fully utilize the potential of K-means clustering. These characteristics offer insightful data about each product's demand and sales performance. We sought to discover unique categories by grouping products with comparable sales patterns in order to inform marketing tactics, inventory control, and product promotion.

### Data Scaling and Model Fitting:

To make sure that all attributes had a uniform scale, we normalized the data using the Standard Scaler before employing the K-means algorithm. This phase is essential to ensuring that the clustering process is not dominated by qualities with vastly different values. After scaling the data, we used the K-means algorithm with four clusters as the first parameter. The 'k-means++' initialization method was chosen to improve the algorithm's stability and convergence.

### Interpreting Results:

A scatter plot was used to display the resulting clusters identified using K-means clustering, where each data point represented a product and was color-coded in accordance with its cluster assignment. We were able to recognize distinct product categories with comparable sales behaviour thanks to this representation. For instance, similar sales figures and quantities for products in the same cluster may indicate shared consumer preferences or shopping patterns.

### Business Insights:

The use of K-means clustering in this situation provides insightful information that will help TechScape Singapore's business plan. The business may decide wisely about product marketing, supply management, and client engagement by examining the various clusters. For instance, items in a high-sales cluster might point to a certain class of goods that are well-liked by consumers, which would cause the business to devote more resources to promoting them. Similar to this, products in a cluster with low sales may be a sign that they could use some improvement or that they require clever marketing efforts to increase their sales.

### Conclusion:

With K-means clustering, we now have a more comprehensive view of TechScape Singapore's product sales environment. Using this method, we may group things according to how well they sell, exposing hidden trends and patterns that might help company decision-makers make wise choices. TechScape Singapore can now optimize its product offerings, improve customer experiences, and effectively manage resources with the findings after performing K-means clustering, all of which support the expansion and success of the business.

# Summary

The Report explores data analytics in great detail, using a variety of methods to glean insightful information that can guide strategic choices. The report is broken up into numerous sections, each of which uses data-driven analysis and visualizations to discuss a particular facet of the company's operations. The full research is summarized below, emphasizing the most important conclusions and the results for TechScape Singapore.

## Introduction and Project Overview

The report begins by describing the backdrop and goals of the assignment. It presents TechScape Singapore, a tech business with a keen eye for the future that places an emphasis on cutting-edge electronics and clever solutions. In line with TechScape's dedication to data-driven solutions, the assignment's goal is to provide insights for business decision-making using analytics approaches.

## Part 1: Creating Dashboards to Address Business Questions

Four dashboards are created in this section, each of which addresses a different business question. The dashboards are made to offer information on peak selling times, top-selling items across all modes of distribution, branch profitability, and top-selling manufacturers. The organization acquires a better grasp of its sales trends, product performance, branch profitability, and manufacturer partnerships through these representations.

## Part 2: Further Analytics

Advanced analytics methods are covered in more detail in the second section. To find hidden patterns and correlations in the data, univariate, bivariate, and multivariate analysis are used. Bivariate and multivariate analyses use Principal Component Analysis (PCA) to show intricate interactions among several variables, while univariate analysis focuses on understanding the distributions and patterns of individual variables.

## Data Cleaning

To guarantee the correctness and dependability of the study, the report recognises the need of data cleaning. It describes how to find and fix errors, including duplicate promotions, changes to the data structure, and the establishment of a calendar table for time-based analysis.

## Conclusion and Business Implications

The importance of bridging the theory and practice gaps in data analytics is emphasized in the report's conclusion. It emphasizes how important data visualization is in converting complicated datasets into clear insights. TechScape Singapore is able to make well-informed decisions, increase profitability, and benefit stakeholders because to the dashboards and analysis built.

## Overall Implications

The significance of data-driven decision-making in contemporary business environments is highlighted in this Report. The study gives TechScape valuable insights on sales trends, product performance, branch profitability, and regional preferences through the use of numerous analytical techniques. The ability of the business to strategize effectively, spend resources wisely, and ultimately strengthen its competitive advantage in the market is improved by the display of data.

## Conclusion

The TechScape Singapore Business Insights Report is a thorough demonstration of how data analytics may influence strategic decision-making, to sum up. The study produces insightful data using dashboards, advanced analytics, and multivariate analysis, which is used to guide TechScape's business strategy. This study demonstrates the transformative power of data analytics, highlighting how data is the compass pointing modern businesses toward success in a dynamic business environment.