|  |  |
| --- | --- |
| Submitted By: | Damian (224106Y) |
|  | Jade (224495K) |
|  | Wei Jun (221814H) |
|  | Jonathan (224548E) |
| PEM Group: | BA2203 |
| Tutor: | Ms Teo Miow Ting |
| Deadline: | 18 Dec 2022 |

ITB131 Data visualisation

Assignment

**Table of Contents**

[**Task 1: Design a Data Model in Power BI** 2](#_Toc122298159)

[Creation of Date Table 2](#_Toc122298160)

[Creation of Calculated Columns/Quick Measures 3](#_Toc122298161)

[Renaming and Removing of Unnecessary Columns 3](#_Toc122298162)

[Arranging Columns into Display folders 5](#_Toc122298163)

[Creation of Hierarchies 7](#_Toc122298164)

[Configuring Table and Column Properties 7](#_Toc122298165)

[Model Relationships 9](#_Toc122298166)

[**Task 2: Design Reports in Power BI** 10](#_Toc122298167)

[Sales Analysis by Products 10](#_Toc122298168)

[Sales Analysis by Employees 10](#_Toc122298169)

[Sales Analysis by Suppliers 11](#_Toc122298170)

[Sales Analysis by Customers 11](#_Toc122298171)

[Sales Analysis by Shippers 12](#_Toc122298172)

[**Task 3: Data Analysis in Power BI** 13](#_Toc122298173)

[Analysis Report of Customers by Damian 13](#_Toc122298174)

[Analysis Report of Suppliers by Jade 15](#_Toc122298175)

[Analysis Report of Employees by Wei Jun 16](#_Toc122298176)

[Analysis Report of Products by Jonathan 19](#_Toc122298177)

# **Task 1: Design a Data Model in Power BI**

Designing the data model for the Northwind Database, involves establishing relationships between tables, configuring table and column properties to enhance the user-friendliness and usability of the data model.

We will be performing the following task for each table in Northwind Database:

* Configuring table and column properties
* Creating hierarchies
* Arranging columns into display folders for readability
* Creating calculated columns/quick measures/measures

## Creation of Date Table

We created a date table, as it allows date and time-based analysis which enables the use of DAX time intelligence functions that would not normally work without a date table.

Graphical user interface, application

Description automatically generatedOur date table **‘Master Date’** contains the following columns with its respective datatype:

|  |  |
| --- | --- |
| Column Name | Datatype |
| Date | Date |
| Month | Text |
| Month Name | Text |
| MonthKey | Whole number |
| Quarter | Text |
| Year | Whole number |

The hidden **MonthKey** column allows us to sort the months **chronologically**. An additional Hierarchy, “Calendar” is also created.

Text

Description automatically generated with medium confidence

The Master Date table has a **one-to-many relationship** with the Orders table, connected through the Date and Order Date columns.

Graphical user interface, application

Description automatically generated

## Creation of Calculated Columns/Quick Measures

**Order Table**

* Created a calculated column “Days to Ship” to calculate number of days to shipping.



* Created a calculated column “Days to Ship Bin” to group days to ship in bins of 5 days.

Text, letter

Description automatically generated

**Order Details Table**

* Created measure named *“*Sales*”* to calculate total sales.



* Created measure named *“*Sales Bin*”* to calculate group sales into bins of $500.

Text

Description automatically generated

* Created a measure named “Total Discount” to calculate the discount given.



**Products Table**

* Created “Stock Buffer” to calculate the stock buffer for each product. 
* Created “Replenish Indicator” to determine if product needs restocking.

## Renaming and Removing of Unnecessary Columns

**Customer Table**

To ensure readability, the following columns were **renamed**:

* *“CompanyName”* to *“Company Name”*
* *“ContactName”* to *“Contact Name”*
* *“ContactTitle”* to *“Contact Title”*

The following columns are deemed unnecessary towards the analysis were **removed**:

* Address, PostalCode, Fax, Phone, Region

**Order Table**

To ensure readability, the following columns were **renamed**:

* *“OrderDate”* to “*Order Date*”
* “*RequiredDate*” to “*Required Date*”
* “*ShippedDate*” to “*Shipped Date*”
* “*ShipVia*” to “*Ship Via*”
* “*ShipName*” to “*Ship Name*”
* “*ShipCity*” to “*Ship City*”
* “*ShipCountry*” to “*Ship Country*”

The following columns are deemed unnecessary towards the analysis were **removed**:

* ShipAddress, ShipRegion, ShipPostalCode

**Order Details Table**

To ensure readability, the following columns were **renamed**:

* “*UnitPrice*” to “*Unit Price”*

**Territories Table**

To ensure readability, the following columns were **renamed**:

* “TerritoryDescription” to “Territory Description”

**Region Table**

To ensure readability, the following columns were **renamed**:

* “RegionDescription” to “Region Description”

**Suppliers Table**

To ensure readability, the following columns were **renamed**:

* “CompanyName” to “Company Name”
* “ContactName” to “Contact Name”
* “ContactTitle to” “Contact Title”

The following columns are deemed unnecessary towards the analysis were **removed**:

* “Address”, “PostalCode”, “Fax”, “Region”, “Homepage”

**Employees Table**

The following columns are deemed unnecessary towards the analysis were **removed**:

* “TitleOfCourtesy”, “Address”, “Region”, “PostalCode”, “Notes”, “PhotoPath”, “HomePhone”, “Photo”, ”BirthDate”, “ContactTitle”, “HireDate”

Additional task performed:

* Merged “FirstName” and “LastName” Column into “Employee Name”

**Shippers Table**

To ensure readability, the following columns were **renamed**:

* “CompanyName” to “Company Name”

**Categories Table**

To ensure readability, the following columns were **renamed**:

* “CategoryName” to “Category Name”

The following columns are deemed unnecessary towards the analysis were **removed**:

* “Picture”

**Products Table**

To ensure readability, the following columns were **renamed**:

* “ProductName” to “Product Name”
* “UnitPrice” to “Unit Price”
* “UnitsInStock” to “Units In Stock”
* “UnitsOnOrder” to “Units On Order”
* “ReorderLevel to “Reorder Level”

The following columns are deemed unnecessary towards the analysis were **removed**:

* “QuantityPerUnits”

## Arranging Columns into Display folders

**Customer Table**

Created Display Folders:

* “Location Details”

Text

Description automatically generated with medium confidence

**Orders Table**

Created Display Folders:

* “Order Details”

**A picture containing text

Description automatically generated**

* “Shipping Details”

A picture containing graphical user interface

Description automatically generated

**Order Details Table**

Created Display Folders:

* “Sales Info”

**A picture containing table

Description automatically generated**

**Supplier Table**

Created Display Folders:

* “Supplier Contact”

A picture containing graphical user interface

Description automatically generated

* “Supplier Location.”

Graphical user interface, text, application

Description automatically generated

**Employees Table**

Created Display Folders:

* “Employment Details”

A picture containing text

Description automatically generated

* “Geography”

Text

Description automatically generated

**Products Table**

Created Display Folders:

* “Inventory Details”

Table

Description automatically generated with medium confidence

* “Product Details”

Graphical user interface

Description automatically generated with medium confidence

## Creation of Hierarchies

**Customer Table**

Created ‘Customer location’ Hierarchy

Text

Description automatically generated with low confidence

**Orders Table**

Created ‘Ship Destination’ Hierarchy

A picture containing diagram

Description automatically generated

**Supplier Table**

Created ‘Supplier Location’ Hierarchy

**Diagram

Description automatically generated**

**Employees Table**

Created ‘Employee Location’ Hierarchy

Text

Description automatically generated with low confidence

## Configuring Table and Column Properties

**Customer Table**

The following columns were set to the appropriate data categories respectively:

* “City” > City
* “Country” >Country/Region

Graphical user interface, application

Description automatically generated with medium confidenceA picture containing chart

Description automatically generated

**Order Table**

The following columns are set to their appropriate data categories:

* “Ship City” > City
* “Ship Country” > Country/Region

Graphical user interface, application

Description automatically generated with medium confidenceA picture containing chart

Description automatically generated

**Order Details Table**

The following columns are set to their appropriate datatypes:

* “Unit Price” > Currency, 2 Decimal point
* “Total Discount” > Currency, 2 Decimal point

Application

Description automatically generated with low confidence Application

Description automatically generated with low confidence

**Territories Table**

The following columns are set to their appropriate data categories:

* “Territory Description” > State or Province

Graphical user interface, application

Description automatically generated with medium confidence

**Region Table**

The following columns are set to their appropriate data categories:

* “Region Description” > Country/Region

Graphical user interface, application

Description automatically generated with medium confidence

**Suppliers Table**

The following columns are set to their appropriate data categories:

* “City” > City
* “Country” > Country/Region

Graphical user interface, application

Description automatically generated with medium confidenceA picture containing chart

Description automatically generated

**Employees Table**

The following columns are set to their appropriate data categories:

* “City” > City
* “Country” > Country/Region

Graphical user interface, application

Description automatically generated with medium confidenceA picture containing chart

Description automatically generated

**Products Table**

The following columns are set to their appropriate datatypes:

* “Unit Price” > Currency, 2 Decimal point

Application

Description automatically generated with low confidence

## Model Relationships

This is an overview of the relationships within the Northwind database.Graphical user interface, application, Teams

Description automatically generated

In addition, we have set the **cross-filter direction** of the following relationships to **'both'** to enable filtering in both directions.

* Products and Categories

Graphical user interface, application

Description automatically generated

* Suppliers and Products

Graphical user interface, text, application

Description automatically generated

* EmployeeTerritories and Territories

Graphical user interface, application

Description automatically generated

# **Task 2: Design Reports in Power BI**

## Sales Analysis by Products

Chart, treemap chart

Description automatically generated

## Sales Analysis by Employees

Graphical user interface, application

Description automatically generatedGraphical user interface, chart, application

Description automatically generated

## Sales Analysis by Suppliers

Graphical user interface

Description automatically generated

Graphical user interface, application

Description automatically generated

## Sales Analysis by Customers

Graphical user interface

Description automatically generated

Graphical user interface, application

Description automatically generated

## Sales Analysis by Shippers

Graphical user interface, application, Teams

Description automatically generated

# **Task 3: Data Analysis in Power BI**

## Analysis Report of Customers by Damian

To begin the analysis, a Scatter Chart is used to identify the relationship between sales and the quantity of products ordered by each customer. The scatter chart allows us to see the strength of the relationship between these variables, which in this case appears to be strong.

Chart

Description automatically generated

This suggests that the sales generated by each customer is closely tied to the quantity of products they order. Additionally, the scatter chart allows us to identify outliers among the customers. By examining the chart, we can see that there are three customers - **'QUICK-Stop'**, **'Save-a-lot Markets'**, and **'Ernst Handel'** - that stand out as being significantly different from the rest of the group.

Further analysis suggests that they have **significantly higher** quantities of products ordered and higher sales compared to the rest of the customer group. This may be due to a variety of factors, such as the size of the business, the type of products they order, or their location.

Chart, bar chart

Description automatically generated

By understanding the factors that contribute to these customers' higher sales and order quantities, we may be able to identify opportunities for improving our sales and customer relationships. Using a decomposition tree, we can do a deeper dive into one of the customers by doing a sales breakdown.

Teams

Description automatically generated with medium confidence

In this deeper dive we will focus on this customer, QUICK-Stop. As shown in the decomposition tree, Majority of the sales generated from QUICK-Stop falls under the **Beverage ($36,216)** category followed by **Confections ($18,530)** and **Dairy Products ($13,800)**. With the top-selling product being **‘Cote de Blaye’** under the Beverage category**,** which generated a total of **$22,924.** This suggests that QUICK-Stop may be a mid-large sized specialty store or a restaurant that primarily sells beverages, confections, and dairy products.

According to the forecast, it is expected that QUICK-Stop will generate approximately **$90,349** in sales on 2023, a significant increase from their current sales. In addition, the number of orders for from their beverages category is expected to significantly increase from 404 in 2020, to an estimated **634** in 2023.

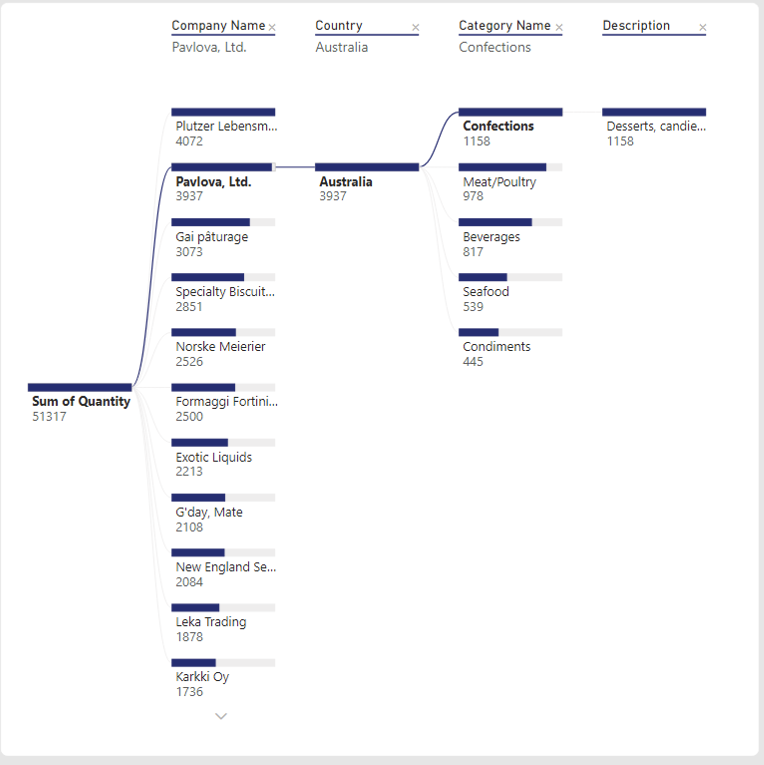
|  |  |
| --- | --- |
| Chart, line chart  Description automatically generated | Chart, line chart  Description automatically generated |

This indicates that QUICK-Stop is expected to grow significantly in the coming years, resulting in more sales and higher orders from them, particularly in the Beverage category.

In conclusion, as the earlier scatter chart suggests, sales generated by each customer is closely tied to the quantity of products they order, therefore, it is crucial to maintain a sufficient inventory to maintain and drive-up sales figures.

## Analysis Report of Suppliers by Jade

I use a tree composition to start off my report because it can show data in various dimensions, which will make data searching accessible and easily.



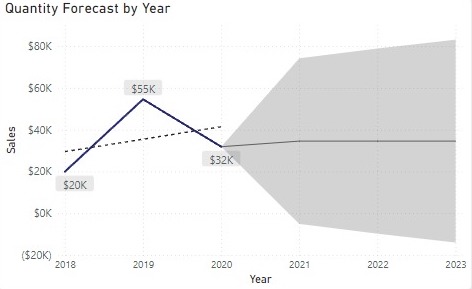
I use data from the sales of an Australian company called Pavlova Ltd.

I decided to analyse the sales of this company, based on categories and the description data. Users who want to see the breakdown for just one breakdown instead of all the supplier's breakdown should use this visualization.

As in my report, users can view the supplier categories and descriptions are given. You can command it to find the dimension to drill down based on specific criteria when you want to concentrate on a specific tree branch on my tree composition.

I am focusing on the total sales to the breakdowns of Pavlova to analyse data. The user can view a breakdown of sales, including the total amount and the growth rates of each category.

I included the country in the analysis field because this is where Pavlova is based. As there are various providers supplying, hence there must be different categories; I believe that stating category names is crucial. The description comes next. Taking in view that not everyone has all the products knowledge, important information about the products category is included for clearer clarification for users.



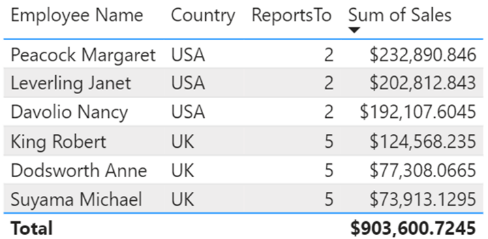
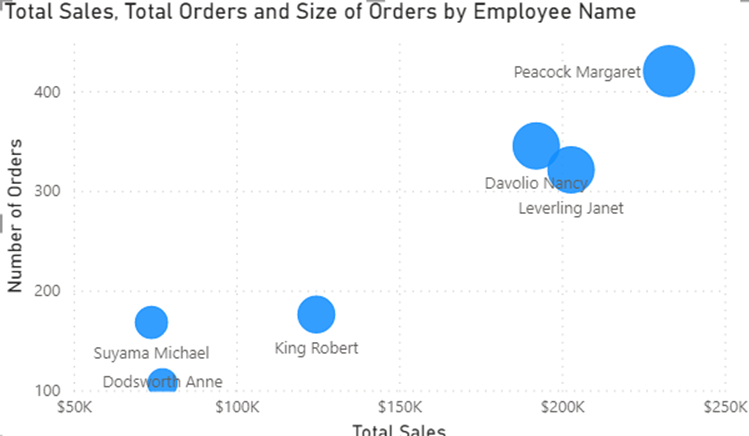
In order to understand more about Pavlova's profitability in the future, forecasts can be used to detect seasonality in the data and present findings from several data sets.

The model predicts that the profits will remain steady over time at a profit of roughly $1259. As can be seen from their total sales for each category, meat/poultry generates the highest profit at $32,698.38, while beverages come in at $10,672.65. Therefore, of all the categories, the meat and poultry category are the most profitable.

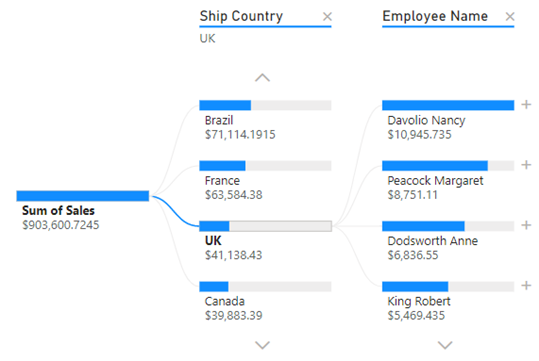
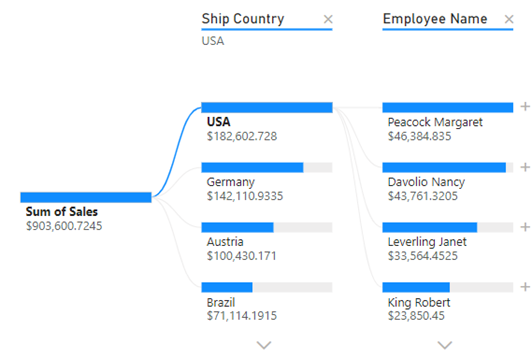
Finally based on my research, I’d also recommend additionally that they can expand by offering vegetables and fruits because currently they (only) offer meat/poultry; seafood; confections; sauces and beverages. This will help them grow their sales significantly. Vegetables and fruits are statistically profitable in all nations, so if Pavlova supplies them, they should anticipate a rise in sales. Secondly, from my tree composition, it indicates that they are making more money from their meat/poultry category. Therefore, I suggest a variety of meat preparations. Since all their meats are processed, they might, for instance, include pan-frying, stir-frying, roasting, etc

## Analysis Report of Employees by Wei Jun

Using a scatter chart to analyse the total sales, number of orders and size of orders.

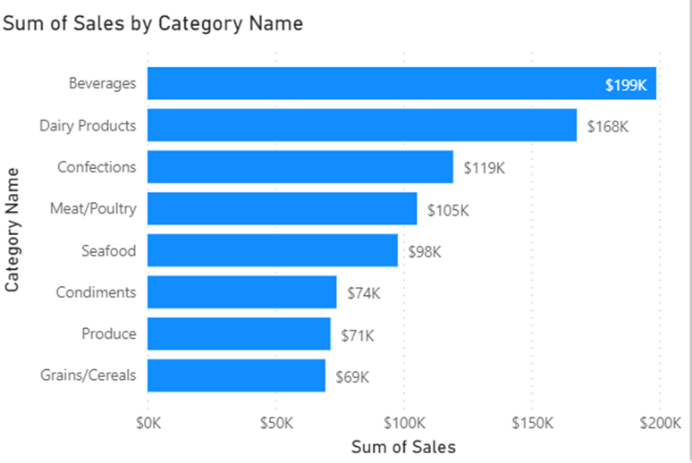
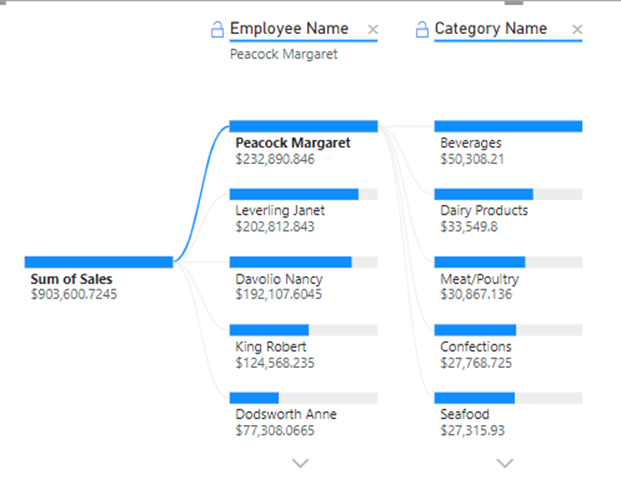


The three sales representatives, “Peacock Margaret”, “Davolio Nancy” and “Leverling Janet” have far surpassed the other Sales representatives. Of which, the sales representatives who report to Employee ID 2 ( Fuller Andrew, Vice President, Sales ) and based in USA has 227.64% more sales when compared to the sales representatives who report to Employee ID 5 ( Buchanan Steven, Sales Manager) and are based in London. There is the same number of employees reporting to Fuller and Buchanan (3). Sales representative in USA made $627,811 whereas sales representatives in UK only made $275,789 over the span of the FY2019 and FY2020 (FY = Financial Year).



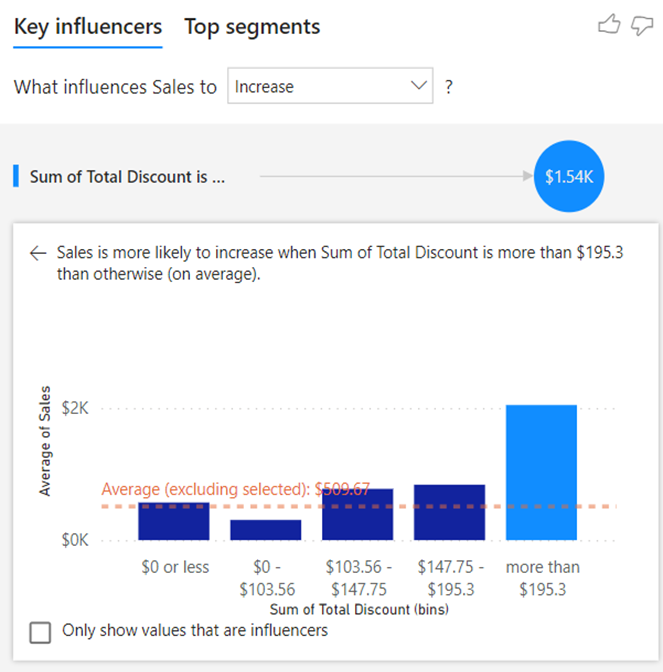
USA is the largest market comprising of $182,602 (20.21%) out of $903,600. The top 3 sales representatives have sold the most in USA market are all based in the USA. However, the same cannot be said about UK market. Even though there are 3 sales representatives in the UK, but they are not able to sell as well to their domestic market when compared to their US colleagues.

From the above I would recommend placing a Vice President of Sales in the UK to lead and guide the 3 sales representatives to grow the UK market.



Beverages has the highest market share out of all categories. Each sales representative has 21.60% (Peacock Margaret), 22.07% (Leverling Janet), 24.26% (Davolio Nancy) , 22.45% (King Robert), 25.41% (Dodsworth Anne) and 12.79%( and Suyama Michael) from selling beverages. (Arranged from highest total sales to the lowest). The 2 sales representatives with lowest overall sales are Dodsworth Anne and Suyama Michael, of which beverages were not their highest sale. For Suyama (Lowest overall sales), beverages only made up 12.79% of his total sales which is especially low compared to all the other sales representatives.

As beverages is the company’s top selling product, therefore, I would recommend all sales representatives especially those who have lower sales, to focus on selling more of beverages. Not only that, as Grains/Cereals, Produce and Condiments are not performing as well as the other categories, the company should do a market survey on what customers prefer, to improve the products and increase sales.



Using Key influencers, sales are more likely to increase when the discount amounts to more than $195.30. The average of Sales more than doubles when discounts are more than $195.30. Hence when trying to sell products, Sales representative should try to push products with more discounts.

In conclusion, to help the company expand, they should try to expand into the UK market as there is already personnel there, and they would understand their local market well. Not only that, they should also focus on improving the lower performing products to increase sales and push for products which have discounts.

## Analysis Report of Products by Jonathan

Timeline, Teams

Description automatically generated

Through using the decomposition tree for a deeper analysis, beverages can be identified as the top contributor among the 8 categories as it has gained a revenue of $267, 868.18, as we go further down the decomposition tree, Cote de Blaye was found to be the product that has played a role in gaining 52.79% of the category under beverages.

Chart

Description automatically generated with low confidence

As we move further down to using the scatter chart itself, we can identify that the relationship between the Sum of Sales and Quantity is relatively strong if we were to look at the products that have been the most quantities sold through the trend line, Camembert Pierrot was the most consistent source of revenue, whilst in comparison to the highest sold item which is Corte de Blaye, the amount of quantity sold was below average. From this graph, we can determine that the top 3 products that has stable number of sales and quantity sold are Camembert Pierrot, Raclette Coudavault and Gorgonzola Telino.

Chart, bubble chart

Description automatically generated

Using key influencers visual allows us to determine what actions to take to gain more revenue, and in this case increasing the Sum of Quantity sold by from 49 or more will yield us the average sales increase of 1,16k, units in stock will have to be 22 or less to gain $277.90 and Selling the product ID of 55 – 64 ( [55] Pate Chinois, [56] Gnocchi di nonna alice, [57] Ravioli Angelo, [58] Escargots do Bourgogne, [59] Raclette Courdavault, [60] Camembert Pierrot, [61] Sirop derable, [62] Tarte ai sucre, [63] Veggie – spread, [62] Wimmers gute semmelknodel ) would increase the average sales by $332.20.

In conclusion, based on the scatter chart and key influencers, we have narrowed down to a few factors that can be taken into consideration for increasing the revenue by focusing on the categories under grains/cereal, and among the products that we can consider focusing on increasing within grains/cereal are Gnocchi di nonna Alice, Ravioli Angelo and Wimmers gute Semmelknodel. Outside the categories of grains/cereal there are Escargots do Bourgogne, Raclette Courdavault, Camembert Pierrot, Sirop derable, Tarte ai sucre and Veggie – spread. This will allow us to consistently increase revenue in comparison in attempting to focus more on sales that fetch high value but in lower demand products such as the beverage that we have mentioned previously ‘Cote de Blaye’.