# **Problem Statement**

The dataset consists of information about customers, employees, shipping details, and warehouse data. The data seems to be the records and database of a retail or e-commerce company that is involved in obtaining goods from suppliers and shipping them to customers with the help of employees.

A company’s net profit should be the primary goal of an e-commerce company. Without sufficient profit, financial resources would deplete without other investments and business failure is imminent. Hence, I am keen to explore ways to improve profits of the company, by looking at it from two angles. Firstly, the overall profit made by the employees and their offices. Secondly, the overall profits from the items and the companies which supplies these items.

# **Dimensional Model Building and Data Exploration**

## Dimensional Model Building

In order to ensure that I have as much information as possible for data exploration, and then providing recommendation for the problem statement, I wanted to ingest all the information available into PowerBI.

Each of the tables were first opened on excel to check for the features of each table, potential links and keys were then identified. The dimensional model was then subsequently created and joined using PowerBI.

The process of loading, merging, appending, and linking tables to understand the relationships between each of the tables clearer were done step-by-step as described below.

Loading in Employees Table and New Employees Table

* Append new employees to employees table, naming it “Employees\_Combined”.
* Removed “Employment Type” feature as only two of the employees (The new ones) have them.
* Hide the New Employees table.

Loading in Employee Team Table (Employee Team and Team sheets)

* Merge the two sheets – Employee Team and Team using Employee ID – so that we can see each the team name of each employee ID.
* Hide the team sheet.
* Merge Employee Team Table into “Employees\_Combined”.

Loading in Customer Details Table and Divisions Table

* Merge Customer Details Table with Division Table using division ID.
* Hide the division table.
* Rename to “Customers\_Dimension”.

Loading in Offices Table

* Link with the “Employees\_Combined” Table using office ID.
* Rename to “Offices\_Dimension”.

Loading in Orders Table and Shipping Info

* Inner merge for Orders Table and Shipping Info Table (Based on Order ID, Line No., and Product ID) creating one “Order\_Shipping\_Combined”.
* Remove duplicated columns.
* Hide individual shipment and order table.
* This will be the Measure/Fact table and will link to the four main dimension tables.
* Renamed to “Order\_Shipping\_Combined\_Measure”.

Loading in Transaction Details Tables

* Link with “Order\_Shipping\_Combined\_Measure” by Order ID.

Loading in Shippers Table and Shippers Info Table

* Merge both to transaction details.
* Note that not all shipper ID has company name (I.e 4 and 5 is missing).
* Hide both tables.

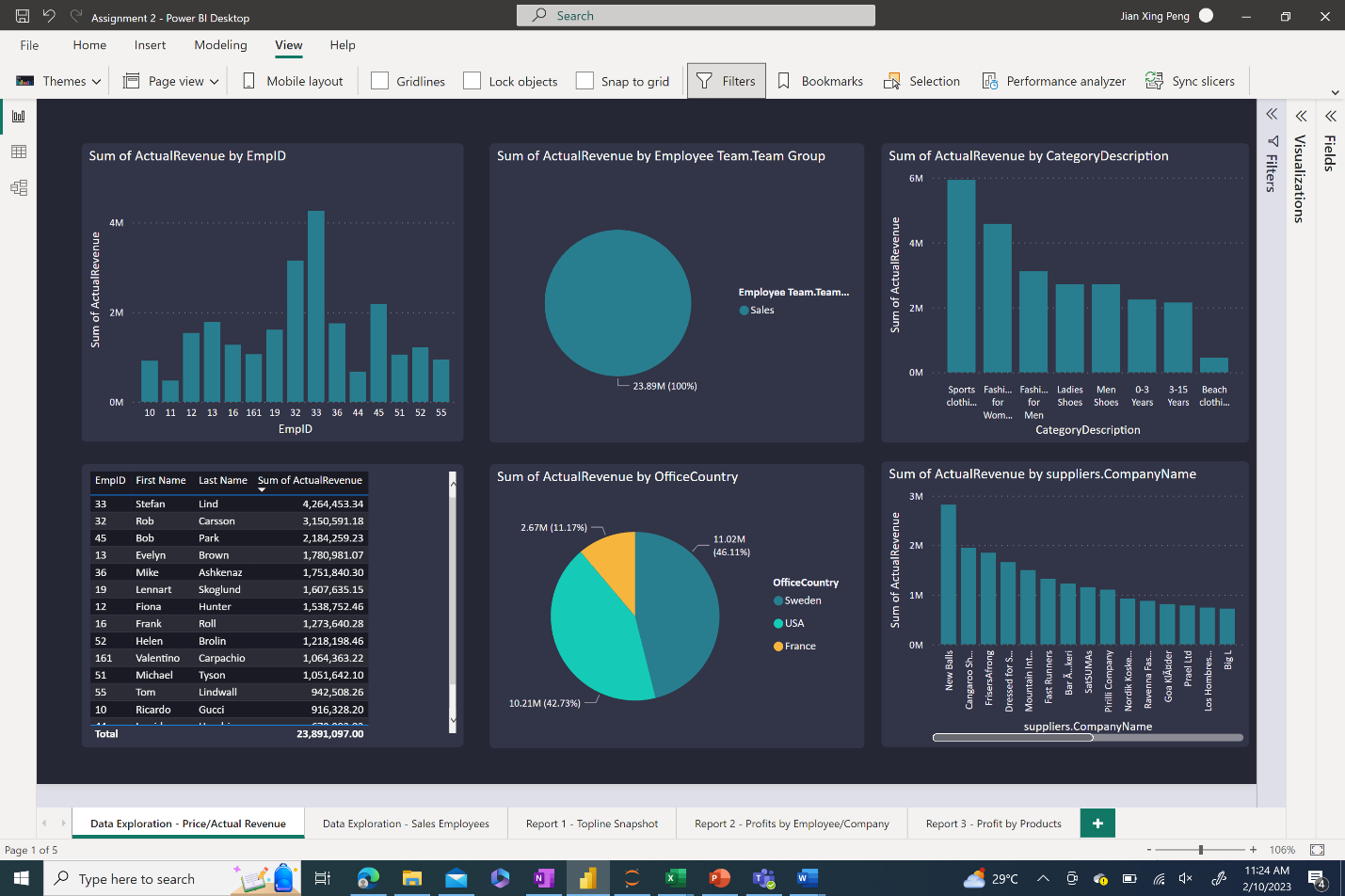
Loading in Product Table, Category Table, and Supplier Table

* Link Product Table with “Order\_Shipping\_Combined\_Measure” by Product ID.
* Merge Product Table with Category Table and Supplier Table.
* Note that one supplier ID is missing in Supplier Table that is present in product table.
* Hide Category Table and Supplier Table.
* Rename as “Product\_Dimension”.

## Data Exploration and Creation of New Measures

Creation of new measures in “Order\_Shipping\_Combined\_Measure” for exploration.

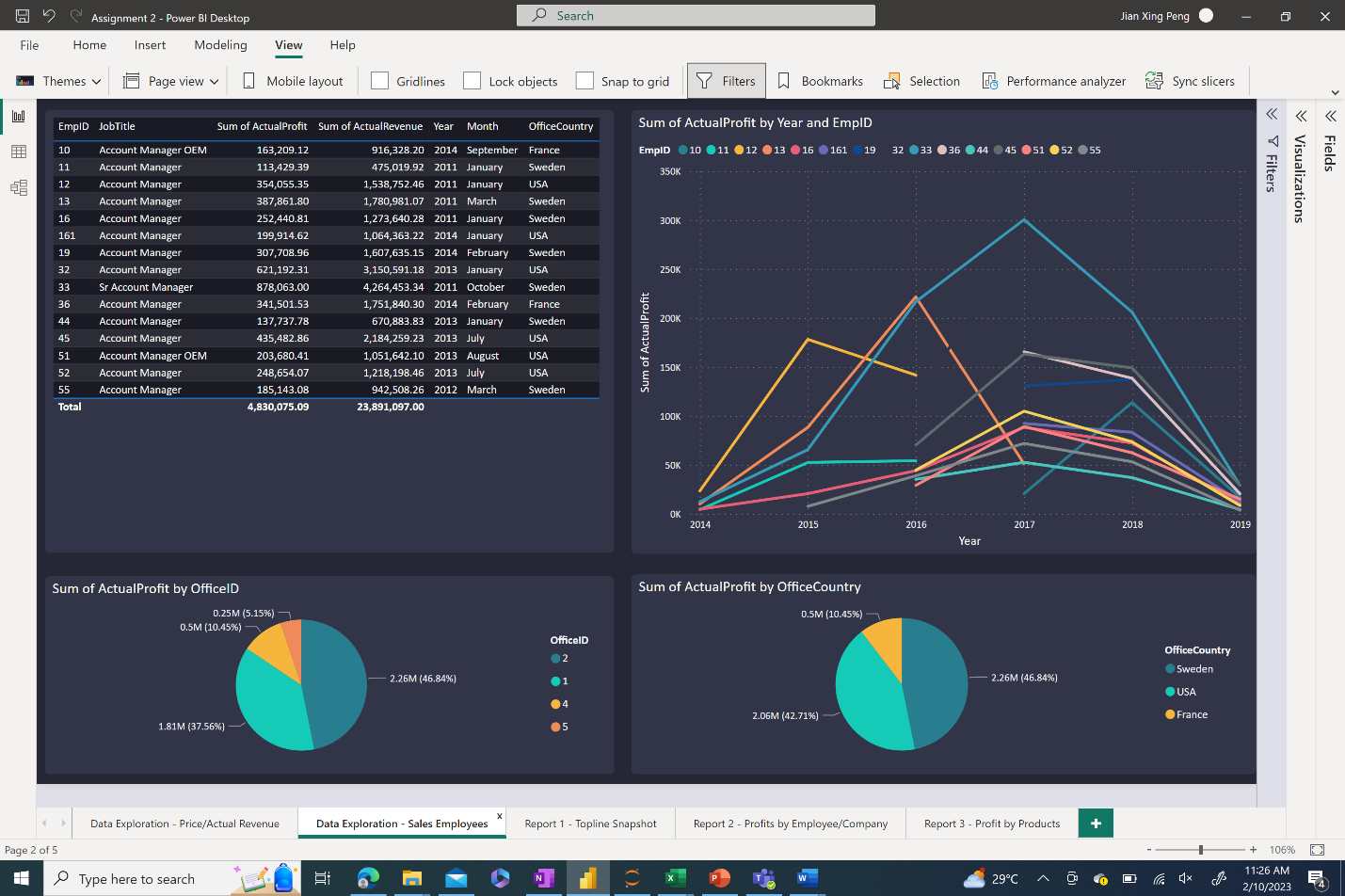
* Created new measure of UsualRevenue (Quantity X UnitPrice)
* Created new measure of ActualRevenue (UsualRevenue X (1-Discount))
* Created new measure of ActualCost (Products\_Dimension[UnitCost] \* Quantity)
* Created new measure of ActualProfit (ActualRevenue – ActualCost)
* Renamed “Quantity” to “Units” to more appropriately refer to the unit of items and reduce confusion on the different “Quantity” variable across tables



*Table 1. Data Exploration – Prices/Actual Revenue*

From these overall charts, we can tell that

1. Only employees from the sales team have sales revenue attached to them in the database.
2. Stefan Lind has made the most revenue out of all the employees over the years.
3. The largest revenue is from the Sweden branch, followed by USA, then France.
4. Sports Clothing as a category has the largest revenue.
5. New Balls is the supplier supplying the goods that made up the largest revenue.



*Table 2. Data Exploration – Sales Employees*

From these charts, we can

1. Identify the best and worst employees, for further action.
2. Employee 33 is one of the best employees around, with the title of senior account manager, having both the highest revenue and profit overall, and having the highest revenue and profit for quite a few of the years.

## Final Dimensional Model

The overall dimensional model is a snowflake schema. The table and its attributes are described below and can be seen in Table 3 after creation of the features through data exploration.

* Fact/Measure Table – Order\_Shipping\_Combined\_Measure
  + OrderID
  + Line No
  + ProductID
  + UnitPrice
  + Quantity
  + Discount
  + CustomerID
  + EmployeeID
  + ShipperID
  + ShipmentDate
  + ShipperName
  + UsualRevenue
  + ActualRevenue
  + ActualCost
  + ActualProfit
* Dimension Tables
  + Employees\_Combined\_Dimension – Contains information of employees and their team
    - EmpID
    - First Name
    - Last Name
    - DateOfBirth
    - StartDate
    - EndDate
    - OfficeID
    - JobTitle
    - TeamID
    - TeamGroup
  + Office\_Dimension – Contains information of the offices
    - OfficeID
    - OfficeAddress
    - OfficeCity
    - OfficeCountry
    - OfficeFax
    - OfficePhone
    - OfficePostalCode
    - OfficeStateProvince
  + Customer\_Dimension – Contains information of customers and their division
    - CustomerID
    - CompanyName
    - ContactName
    - City
    - Country
    - DivisionID
    - Address
    - Fax
    - Phone
    - PostalCode
    - StateProvince
    - Division Name
  + Product\_Dimension – Contains information of the products and its suppliers
    - CategoryID
    - ProductID
    - ProductName
    - QuantityPerUnit
    - SupplierID
    - ItemCost
    - ItemPrice
    - UnitsInStock
    - UnitsOnOrder
    - UnitCost
    - UnitPrice
    - CategoryName
    - CategoryDescription
    - Supplier.Address
    - Supplier.City
    - Supplier.CompanyName
    - Supplier.ContactName
    - Supplier.Country
    - Supplier.Fax
    - Supplier.Phone
    - Supplier.PostalCode
    - Supplier.ID
    - Supplier.BandID
    - Supplier.Band
  + Transaction\_Details\_Dimension – Contains information of the transaction order
    - OrderID
    - CustomerID
    - EmployeeID
    - ShipperID
    - FreightWeight
    - OrderDate
    - Shipper Name

A picture containing schematic

Description automatically generated

*Table 3. Final Dimension Model – Highlighted in yellow is the Measure table while in cyan are the Dimension tables. Those not highlighted are the original tables which have already been combined into the highlighted tables.*

# Data Visualization Building

“Simply Modern Dark” theme by Metricalist was used for the base theme and colour. Screenshots from the dashboard will be uploaded in the tables below.

Do note that all tables are interactive and can be filtered simply by clicking on figures or slicers. Some DAX language is used during the creation of measures and mentioned in the earlier section.

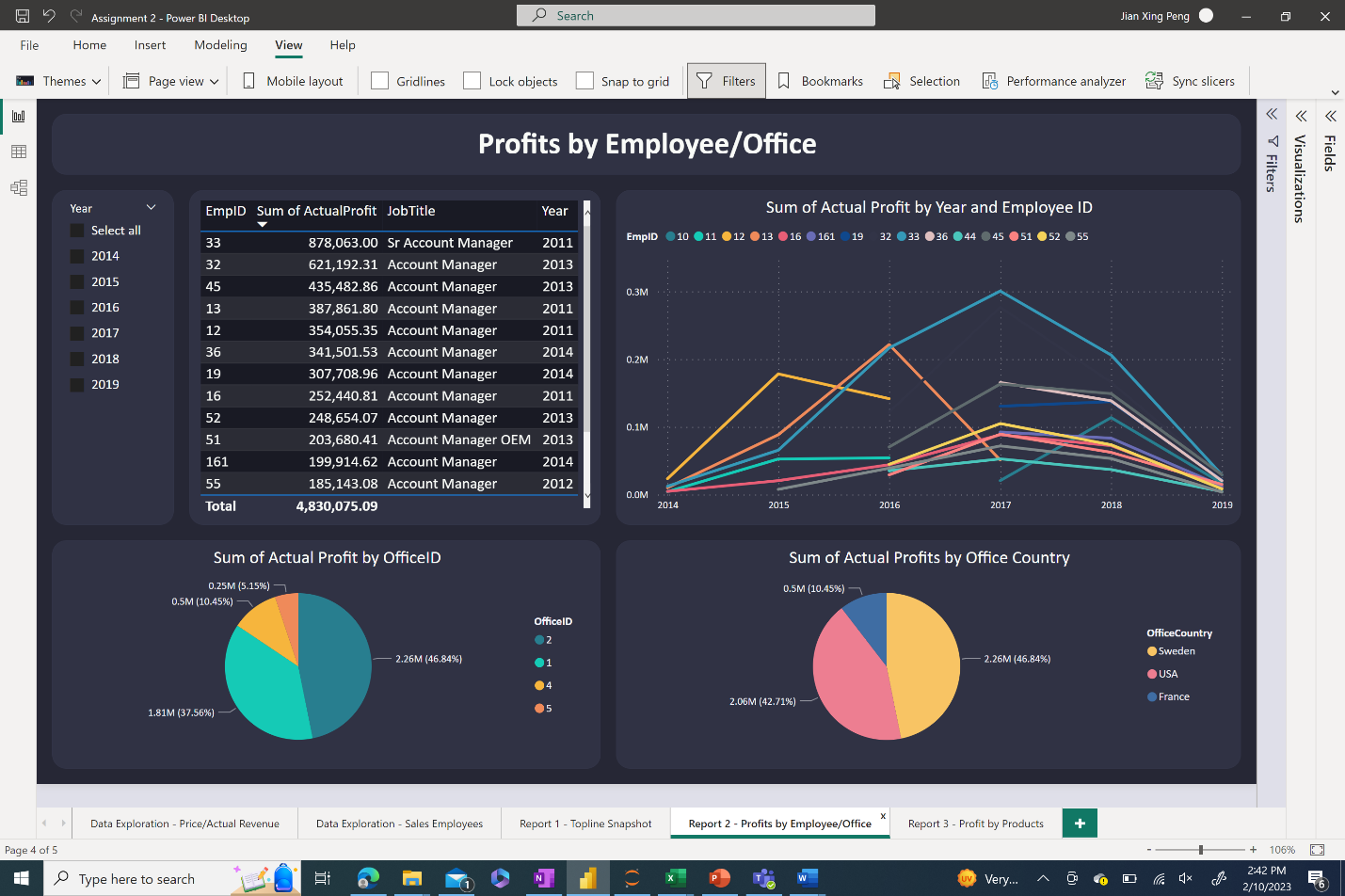
## Report 1 – Topline Snapshot

Graphical user interface, chart

Description automatically generated

*Table 4. Topline Snapshot – Overview of Data*

## Report 2 – Profits by Employee/Company



*Table 5. Profits by Employee/Office*

## Report 3 – Profits by Products

Graphical user interface

Description automatically generated

*Table 6. Profits by Products*

# Business Insight and Recommendation

Do note that in 2019, only the first quarter of the shipping information has been included in the dataset. Hence, we will not be including 2019 when deliberating some of the analyses related to the year.

## Report 1 – Topline Snapshot

The aim of the topline snapshot report was to give an overview of the data regarding the problem statement we have created.

In this report, here are some of the findings and recommendations:

* The overall profit that the company has earned was 4.83 million dollars over the years 2014 to 2019, with its cost of 19.1 million dollars, and revenue of 23.9 million dollars.
* Sweden has the largest profits at 2.26 million dollars, compared to USA at 2.06 million dollars, and France at 0.5 million dollars.
* Women’s clothes has the highest profit in terms of the categories of items, at 1.1 million dollars. On the other hand, swimwear has the lowest profit at around 100 thousand dollars.
* The supplier that has helped to make the most profit was New Balls at around 474 thousand dollars. On the other hand, Executive Clothing GMBH has only around 3 thousand dollars.
* The year with the highest profit was 2017 at around 1.6 million dollars. This was then followed by a decrease in profit the following year.

## Report 2 – Profits by Employee/Office

The aim of the report on profits by employee/office was to take a deeper dive into the employees and their offices, to figure out who are the employees and offices that are raking in the profits. Identifying them can help identify the factors that have helped them achieve such high profits. A good idea would be to follow up with a more qualitative interview with these employees and offices to understand how they have managed to achieve this.

Identifying employees and offices that are performing poorly is crucial as well. Management can then step in to review their processes and identify pain points that might be stumbling them.

Here are some of the findings and recommendations:

* Employee ID 33 is one of the best sales employees, having the highest overall profit over the years at 878 thousand dollars, and also topping the sales for three consecutive years.
* Profits for all employees dropped from 2017 onwards. This likely points to the fact that it was not the fault of the employees that caused a drop in the profit. It could be due to systemic issues, or other issues that might not be in the control of the sales employees themselves.
* Office 2 has the highest profit overall, at 2.26 million dollars, 46.8% of the overall profits.
* Office 4 is rather resilient. Despite only starting up in 2017, it was able to increase its profit in 2018, although all the other offices had a decrease in profits. It would be good to investigate the reasons why this office specifically was able to do this.
* Although the overall profit across the years was highest for Sweden, profits in USA surpassed Sweden from 2017 onwards.

## Report 3 – Profits by Products

The aim of the report on profits by product was to take a deeper dive into the products and the companies that help supply these products. Understanding the profits made by the products, the quantity shipped, and their company can help determine demand for the type of products and the potential profits.

Here are some of the findings and recommendations:

* Sportswear has the highest units sold when compared to the different categories, although it comes in second to women’s clothes when comparing profits.
* Women’s clothes has the highest profit when compared to the other categories, at 1.1 million dollars.
* Specifically, Halter Dresses are making a huge proportion of the overall profits, at 267 thousand dollars. It would be good to work with Nordik Koskenkorva to import in more Halter Dresses to make an even larger profit as there is proof that the demand is there. Both profits and units sold are high.
* On the other hand, swimwear has the lowest profits out of all the categories, even though the units sold are not the lowest. It might be good to perhaps focus efforts on the more profitable categories if the profit margin for swimwear is determined to be too little over the years. Another recommendation would be to only bring in Riviera Swimsuits, which has the highest profit out of all the swimwear.
* The US-Master Jeans should be discontinued as it brought in a loss of about 4.5 thousand dollars across the years.