CAPSTONE PROJECT -NORTHWIND TRADERS DATESET

By Fahmitha Fathima

**Overview of the project**

Northwind dataset is to analyse the sales data and provide valuable insights for business intelligence and decision making.

* Find out the total sales, average sales, profit margin, and growth rate of different products, categories, regions, and customers.
* Identify the best-selling and the least-selling products, categories, regions, and customers.
* Compare the sales performance across different dimensions and time periods to discover trends and patterns.
* Group customers based on their characteristics and purchase history.
* Find out what products are frequently bought together by different customers.
* Predict the demand and supply of different products and suppliers.
* Find out the optimal order quantity, reorder point, and safety stock level for each product and supplier.
* Evaluate the performance and satisfaction of different employees.
* Identify gaps and opportunities for improvement.

**Process**

1.**Data Acquisition from GitHub:**

Obtain the requisite dataset from a designated GitHub repository, containing essential information on university rankings, encompassing various countries and their performance across distinct ranking systems.

2. **Data Transformation and Enhancement:**

If necessary, execute data transformation procedures to ensure data quality and consistency. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.

3. **Connecting with Tools:**

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing.

4. **Problem Statement Solution in Power BI:**

Utilize Power BI to delve into the specified problem statements. Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions

5**. Exploratory Data Analysis (EDA):**

Perform exploratory data analysis using either Excel or SQL Workbench, depending on the complexity of the analysis. Extract meaningful patterns, relationships, and trends from the data to inform subsequent decision-making.

6. **Creation of Visual and Insightful PowerPoint**:

Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights.

7**. Detailed Documentation:**

Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations.

**Objective**

The Northwind dataset is a sample database that contains data about a fictional company called Northwind Traders, which imports and exports specialty foods from around the world. The dataset includes tables for customers, orders, order details, products, categories, suppliers, shippers, employees, and regions.

The objective of this project is to conduct an in-depth analysis of the Northwind dataset in order to explore patterns, trends, and factors influencing sales performance across different dimensions. The goal is to provide insights that can be used to enhance the business intelligence and decision making of the company. The project will involve the following tasks:

* Performing a comprehensive analysis of sales data, including variations across products, categories, regions, and customers.
* Identifying the best-selling and the least-selling products, categories, regions, and customers.
* Comparing the sales performance across different time periods to discover seasonal and temporal effects.
* Grouping customers based on their characteristics and purchase behavior.
* Predicting the demand and supply of different products and suppliers.
* Optimizing the order quantity, reorder point, and safety stock level for each product and supplier.
* Evaluating the employee performance and satisfaction of different roles and regions.
* Identifying gaps and opportunities for improvement.

The success of the project will be measured by the following metrics:

* The quality of the analysis
* The relevance of the insights
* The impact of the recommendations

This project is significant because it has the potential to improve the sales performance and profitability of the company. By understanding the factors that influence sales data, the company can better position itself to succeed in the competitive market.

**Significance**

The Northwind dataset can also be used for creating various projects and reports that demonstrate the ability to apply data analysis skills and techniques to solve real-world problems and provide valuable insights for business intelligence and decision making. For example, some possible projects and reports based on the Northwind dataset are:

* Analyzing the sales performance of different products, categories, regions, and customers.
* Identifying the customer behavior and preferences of different segments.
* Optimizing the supply chain and inventory management of different products and suppliers.
* Evaluating the employee performance and satisfaction of different roles and regions.
* Comparing the Northwind dataset with other similar datasets or real-world data sources.

**Conclusion**

 Harness the power of data visualization to drive better decision-making, streamline processes, and ultimately enhance your organization’s performance. With its customizable design, user-friendly interface, and advanced data analysis capabilities, the Beautiful Power BI Dashboard with Northwind Database is an invaluable tool for any data-driven organization.

**Data Dictionary :**

Table: Customer

* Fields:
  + customer: Unique identifier for each customer.
  + company name: Name of the company that the customer belongs to.
  + contact name: Name of the contact person for the customer.
  + contact title: Title of the contact person for the customer.
  + address: Address of the customer.
  + city: City of the customer.
  + region: Region of the customer.
  + postal code: Postal code of the customer.
  + country: Country of the customer.
  + phone: Phone number of the customer.
  + fax: Fax number of the customer.

Table: Order

* Fields:
  + order\_id: Unique identifier for each order.
  + customer\_id: Foreign key referencing the customer\_id field in the Customer table.
  + employee\_id: Foreign key referencing the employee\_id field in the Employee table.
  + order\_date: Date when the order was placed.
  + required\_date: Date when the order is required to be delivered.
  + shipped\_date: Date when the order was shipped.
  + ship\_via: Foreign key referencing the shipper\_id field in the Shipper table.
  + freight: Freight charge for the order.
  + ship\_name: Name of the person or company that received the order.
  + ship\_address: Address where the order was shipped to.
  + ship\_city: City where the order was shipped to.
  + ship\_region: Region where the order was shipped to.
  + ship\_postal\_code: Postal code where the order was shipped to.
  + ship\_country: Country where the order was shipped to.

Table: Order\_detail

* Fields:
  + order\_detail\_id: Unique identifier for each order detail.
  + order\_id: Foreign key referencing the order\_id field in the Order table.
  + product\_id: Foreign key referencing the product\_id field in the Product table.
  + unit\_price: Unit price of the product ordered.
  + quantity: Quantity of the product ordered.
  + discount: Discount applied to the product ordered.

Table: Product

* Fields:
  + product\_id: Unique identifier for each product.
  + product\_name: Name of the product.
  + supplier\_id: Foreign key referencing the supplier\_id field in the Supplier table.
  + category\_id: Foreign key referencing the category\_id field in the Category table.
  + quantity\_per\_unit: Quantity per unit of measure for the product.
  + unit\_price: Unit price of the product.
  + units\_in\_stock: Number of units in stock for the product.
  + units\_on\_order: Number of units on order for the product.
  + reorder\_level: Reorder level for the product.
  + discontinued: Boolean value indicating whether the product is discontinued or not.

Table: Category

* Fields:
  + category\_id: Unique identifier for each category.
  + category\_name: Name of the category.
  + description: Description of the category.

Table: Supplier

* Fields:
  + supplier\_id: Unique identifier for each supplier.
  + company\_name: Name of the supplier company.
  + contact\_name: Name of the contact person for the supplier.
  + contact\_title: Title of the contact person for the supplier.
  + address: Address of the supplier.
  + city: City of the supplier.
  + region: Region of

the supplier.

* postal\_code : Postal code of the supplier.
* country : Country of the supplier.
* phone : Phone number of the supplier.
* fax : Fax number of the supplier.
* homepage : Homepage URL of the supplier.

Table : Shipper

* Fields :
* shipper\_id : Unique identifier for each shipper.
* company\_name : Name of the shipper company.
* phone : Phone number of the shipper.

Table : Employee

* Fields :
* employee\_id : Unique identifier for each employee.
* last\_name : Last name of the employee.
* first\_name : First name of the employee.
* title : Title of the employee.
* title\_of\_courtesy : Courtesy title of the employee.
* birth\_date : Date of birth of the employee.
* hire\_date : Date when the employee was hired.
* address : Address of the employee.
* city : City of the employee.
* region : Region of the employee.
* postal\_code : Postal code of the employee.
* country : Country of the employee.
* home\_phone : Home phone number of the employee.
* extension : Extension number of the employee.
* photo : Photo of the employee.
* notes : Notes about the employee.
* reports\_to : Foreign key referencing the employee\_id field in the Employee table, indicating the manager of the employee.
* photo\_path : Path to the photo file of the employee.

Table : Region

* Fields :
* region\_id : Unique identifier for each region.
* region\_description : Description of the region.

**ER Diagram:**

****

**Problem Statements**

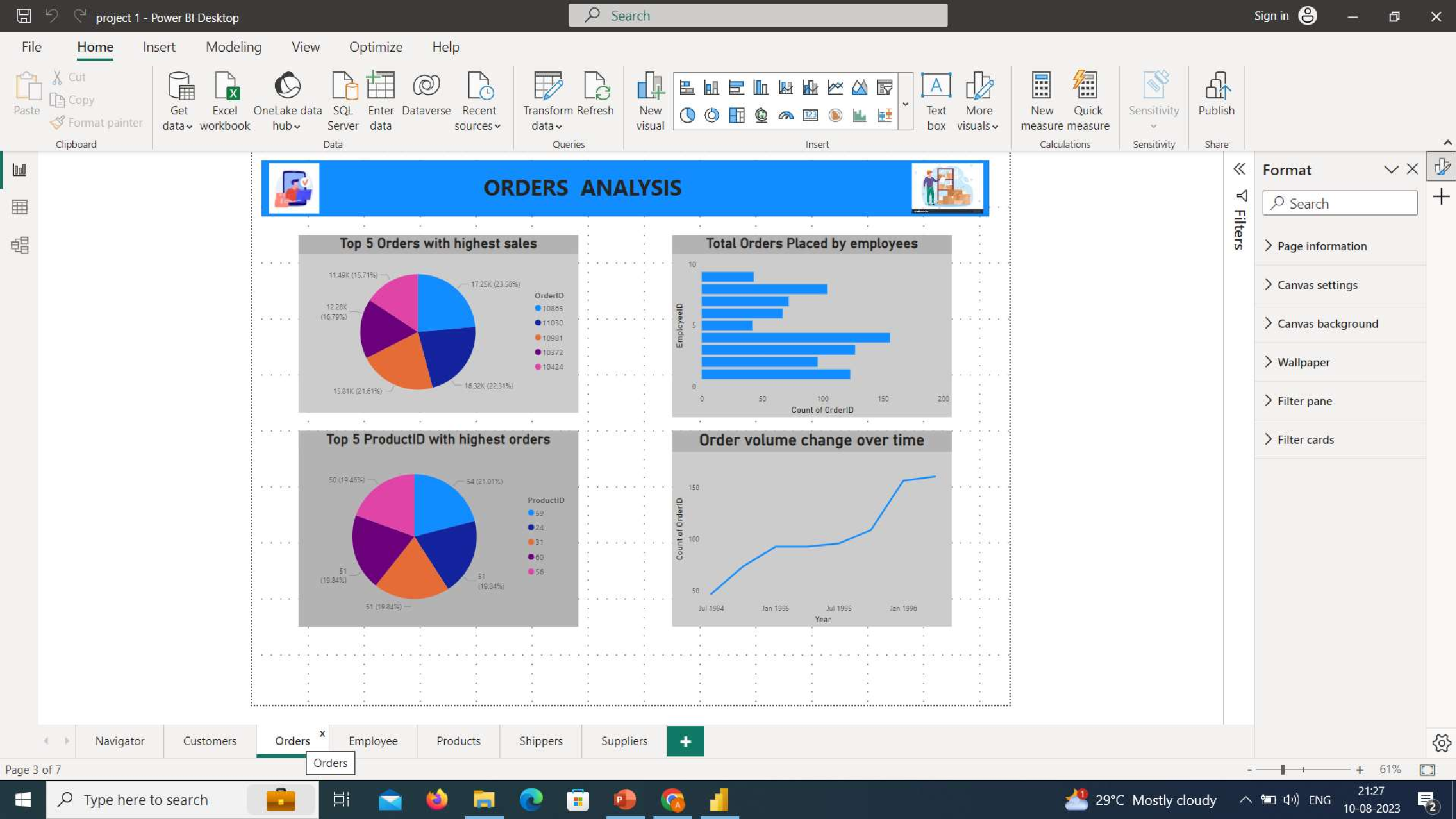
A screenshot of a computer

Description automatically generated

* + This dashboard primarily focuses on customer demographics and their

behavioral Analysis.

* First bar chart focuses on the customer segmentation by cities which gives us an insight to the different cities around the world which helps us to focus more on expansion of business over the world
* Second chart focuses on the behavioral analysis which is the process of examining and interpreting the actions, interactions, and patterns exhibited by customers when they interact with a product, service, website, app, or any other touchpoint associated with a business. Here we can see the Top 10 customers with highest number of order counts which clearly indicates the customers with a higher retention rate that indicates customer are happy with our products.
* Similar to first chart 3rd chart which is a pie chart which focuses on the customer demographics this time with top 5 countries which helps the business to focus more the countries and explore new opportunities in the countries in a view to expanding the business to more new cities.
* 4th chart is a pie chart which gives the distribution of customer by their Contact Title i.e., job role of different customers.
* 5th chart is a Area Chart which focuses on customer acquisition over time(i.e. month,year,quarter) which refers to the process of attracting and converting individuals into new customers for a business's products or services.



* This Dashboard Primarliy focuses on orders made by customers , revenue generated by the orders ,order

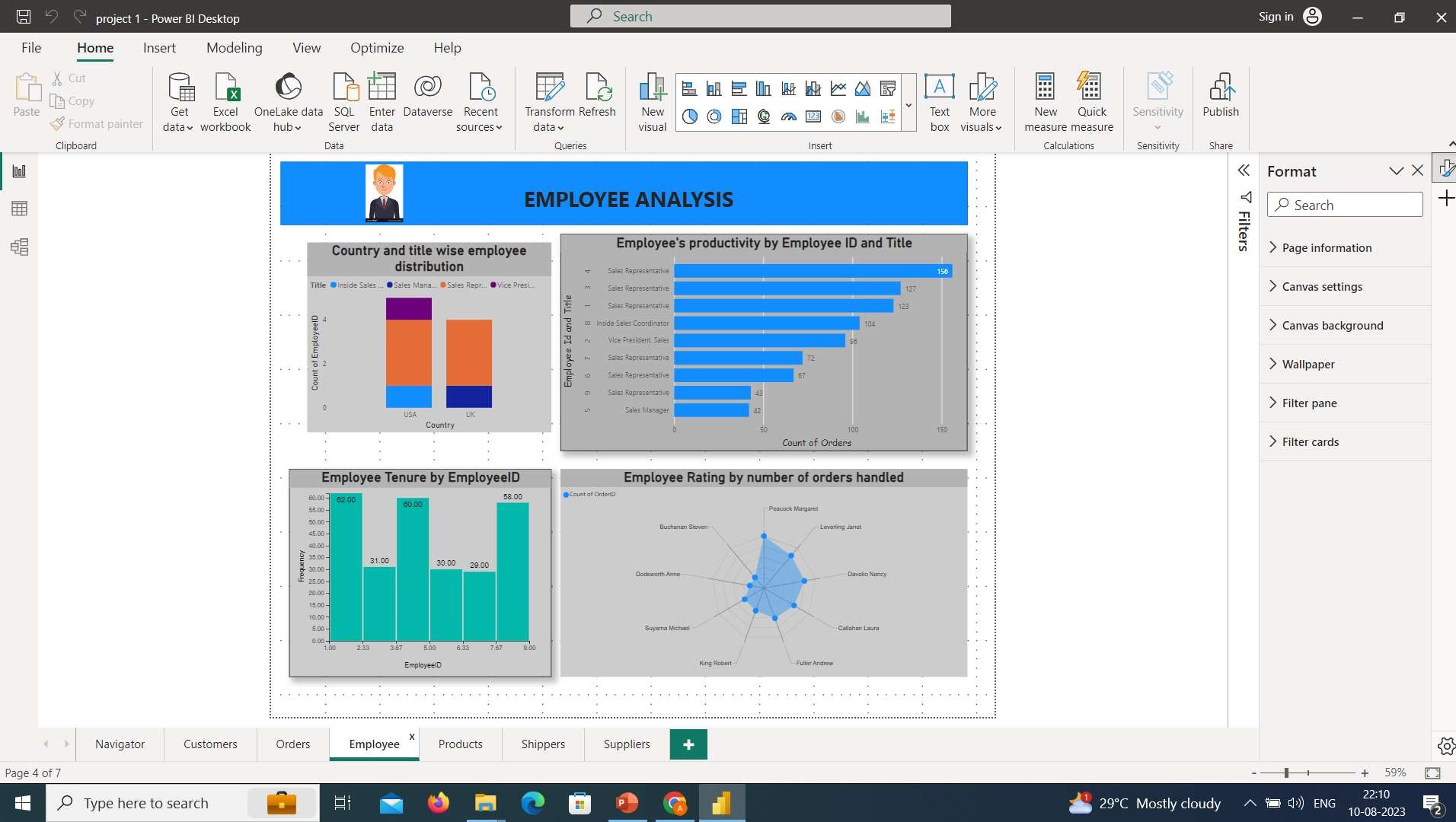
volume change ,Order placed,different types of products placed.

* 1st chart gives us a insight to the order values i.e. sales revenue generated by the orders.Here we can see

that this pie chart gives the top 5 orders with highest revenue which is a kpi for this business.

* 2nd pie chart gives us a insight of top 5 product placed with highest orders which signifies the demand of products amongst the customers .
* 3rd bar chart gives us a insight of Order placed by different employees over their period of service
* 4th Line chart gives us a insight of order volume change that is orders placed in each year.Here we can see

that in each year order count are rising which is a good sign for the business.



This dashboard focuses on employee demographics ,their

work tenure and also their ratings

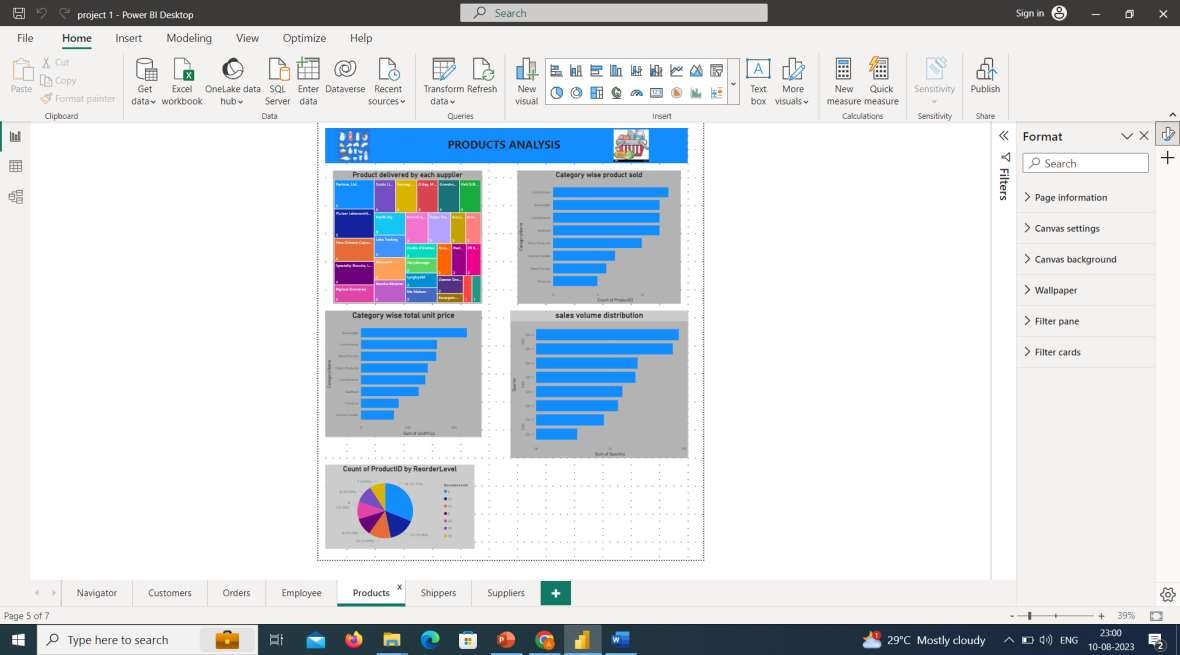
1st Stacked bar chart describe the employee demographics which includes country and title wise employee distribution

Second bar chart describes the employee productivity by employeeid and title which basically is their productivity count

Third chart which is a histogram describes the employee

tenure which shows their loyalty towards their company

Fourth chart which is a spider web tells about the employee rating given by different customers which is a key performance indicator in employee section.



* This dashboard comprises of Product analysis of different products sold across different categories
* 1st chart gives insight distribution of products sold by each supplier in a hierarchical manner
* 2nd chart gives a insight of category wise product sold across different categories
* 3rd chart gives a insight of total unit price category wise
* 4th and 5th chart gives a insight of sales volume that is order quantity over the years and product wise reorder level.