**Design Document**



TOPIC : Northwind Traders a gourmet food supplier company

Student Name: Yogendran Suraskumar

Student ID: 000736199

Instructor : Ms Monireh Rezai Rad

Email : [yogen24hrk@gmail.com](mailto:yogen24hrk@gmail.com)



**Table of Contents**

1. Introduction ---- Page 3

* Purpose
* Scope

2. System Overview ---- Page 3

* 2.1 System Architecture
* 2.2 System Components

3. Data Management ---- Page 4

* 3.1 Data Sourcing
* 3.2 ETL Process

4. Data Modeling ---- Page 4/5

* 4.1 Entity-Relationship Diagram
* 4.2 Schema Design

5. Visualization & User Interface ---- Page 5/6

* 5.1 Layout

- Sales Trends

-Product Performance

* 5.2 Features & Functionalities

6. Technical Requirements ---- Page 7

* 6.1 Software & Tools
* 6.2 Hardware Requirements

7. Milestones & Timeline ---- Page 7

8. Conclusion ---- Page 7

**Introduction**

1.1 Purpose

The purpose of this document is to outline the design of a data science project to study and analyze the sales data of Northwind Traders. The goal is to identify the best and worst-performing product categories and determine which countries have the highest sales so it will help in strategic decision making.

1.2 Scope

This document covers the overall system architecture, data management strategies, data modeling, visualization approaches, technical requirements, and project timelines. It does not cover the implementation details or the deployment of the system.

**System Overview**

2.1 System Architecture

- Data Sourcing: Data will be sourced from Maven Analytics.

- Data Storage: Data will be stored on Google drive.

- Microsoft Azure data studios: The ETL (Extract, Transform, Load) process will be used to prepare the data.

- Data Modeling: The data will be modeled using an Entity Relationship Diagram (ERD) using a star schema model.

- Visualization: Power BI will be used for data visualization.

2.2 System Components

- User Interface: A Power BI dashboard providing interactive visualizations.

- Backend Processing: ETL processes implemented using DAX calculations.

- Database: Google drive for storing cleaned and transformed data.

**Data Management**

3.1 Data Sourcing

- Maven Analytics: Provides quality and clean datasets suitable for business intelligence projects.

3.2 ETL Process

- Extract: Data will be extracted from CSV files provided by Maven Analytics.

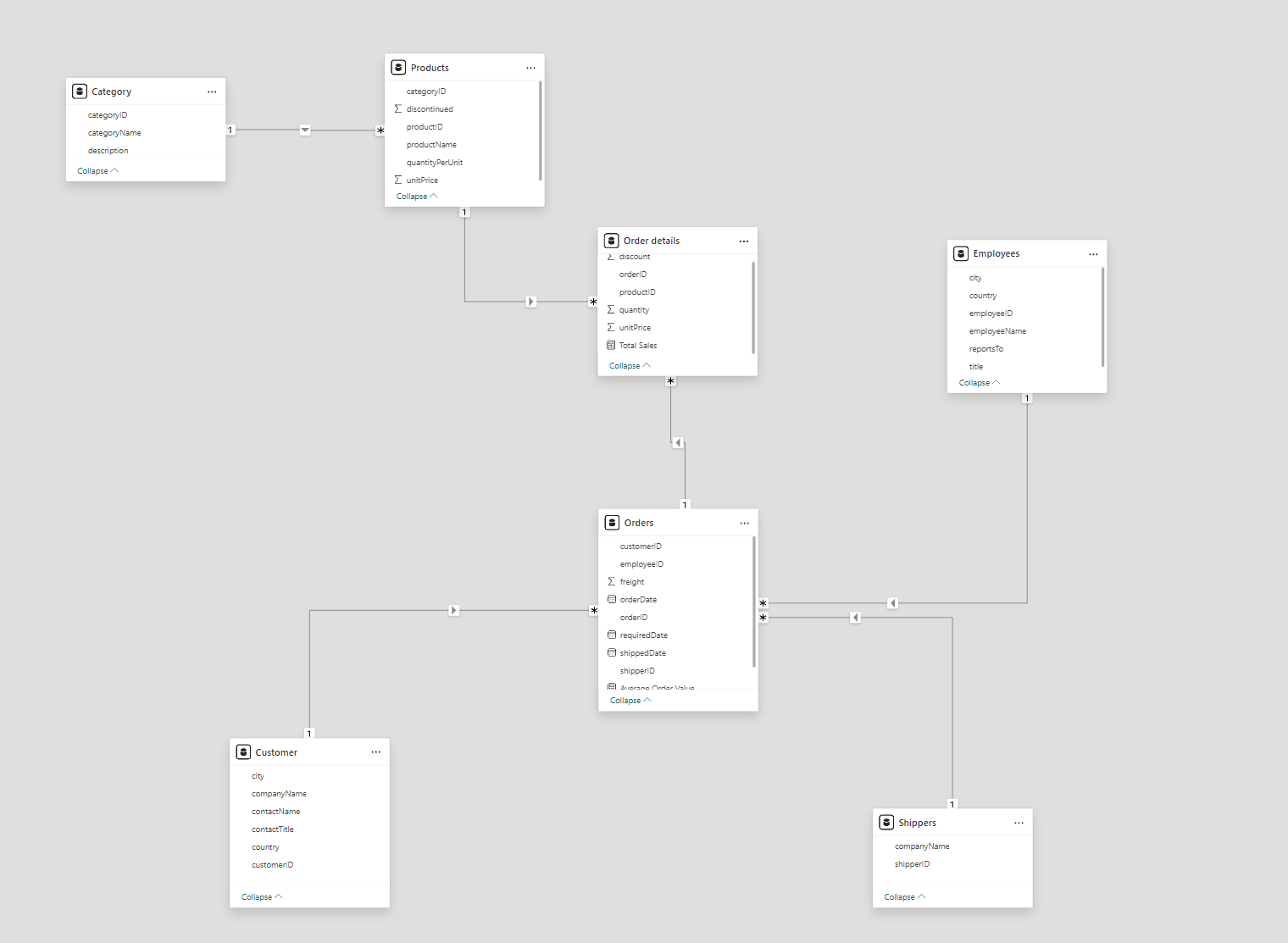
- Transform: Data cleaning is done by handling missing values and correcting data types. Calculated fields will be created for analysis.

- Load: Transformed data will be stored into Google Drive for further use.

**Data Modeling**

4.1 Entity-Relationship Diagram

The ERD will include entities such as Products, Categories, Orders, and Customers, showing their relationships.



4.2 Schema Design

- Products Table: Contains columns like ProductID, ProductName, CategoryID, UnitPrice.

- Categories Table: Contains columns like CategoryID, CategoryName.

- Orders Table: Contains columns like OrderID, CustomerID, OrderDate.

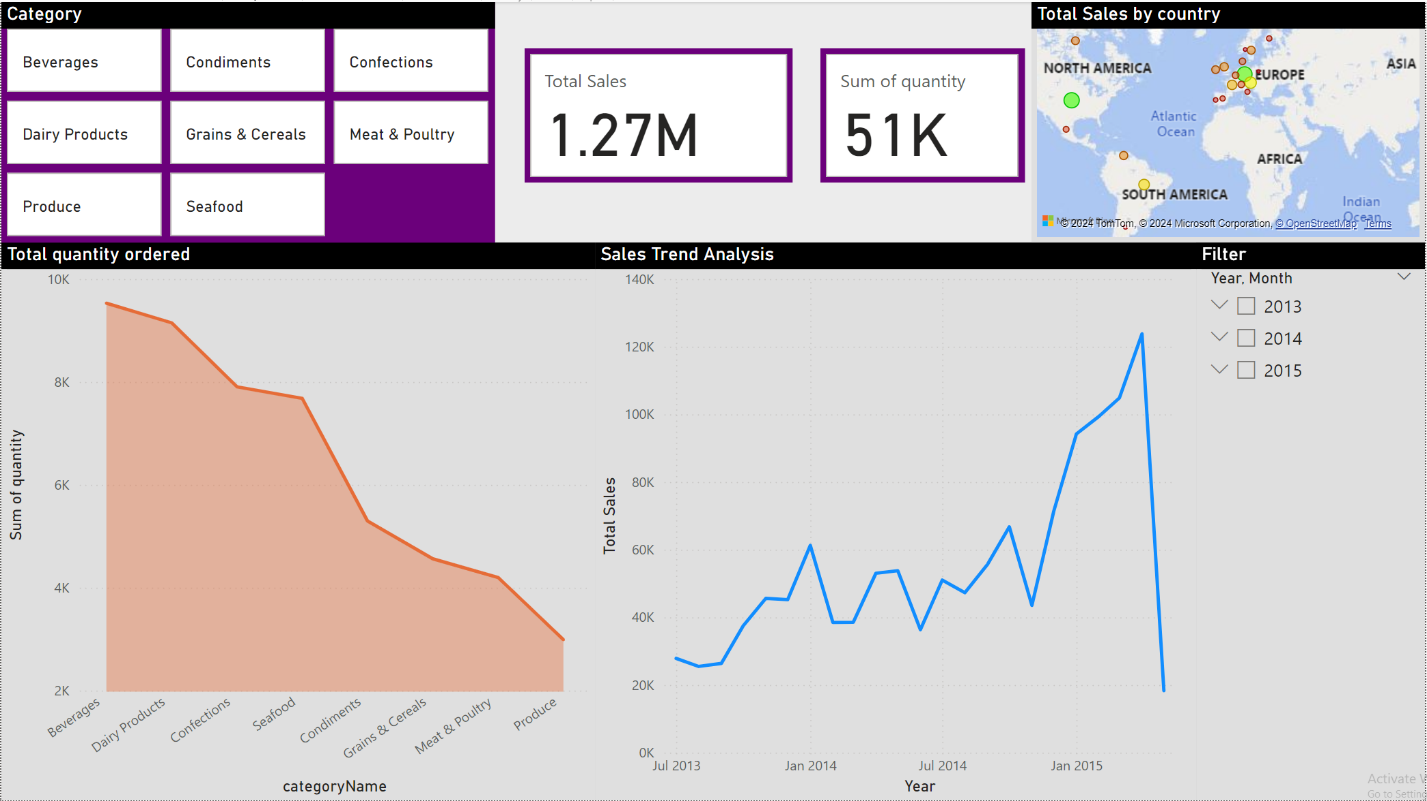
- Customers Table: Contains columns like CustomerID, CustomerName, Country.

**Visualization & User Interface**

5.1 Layout

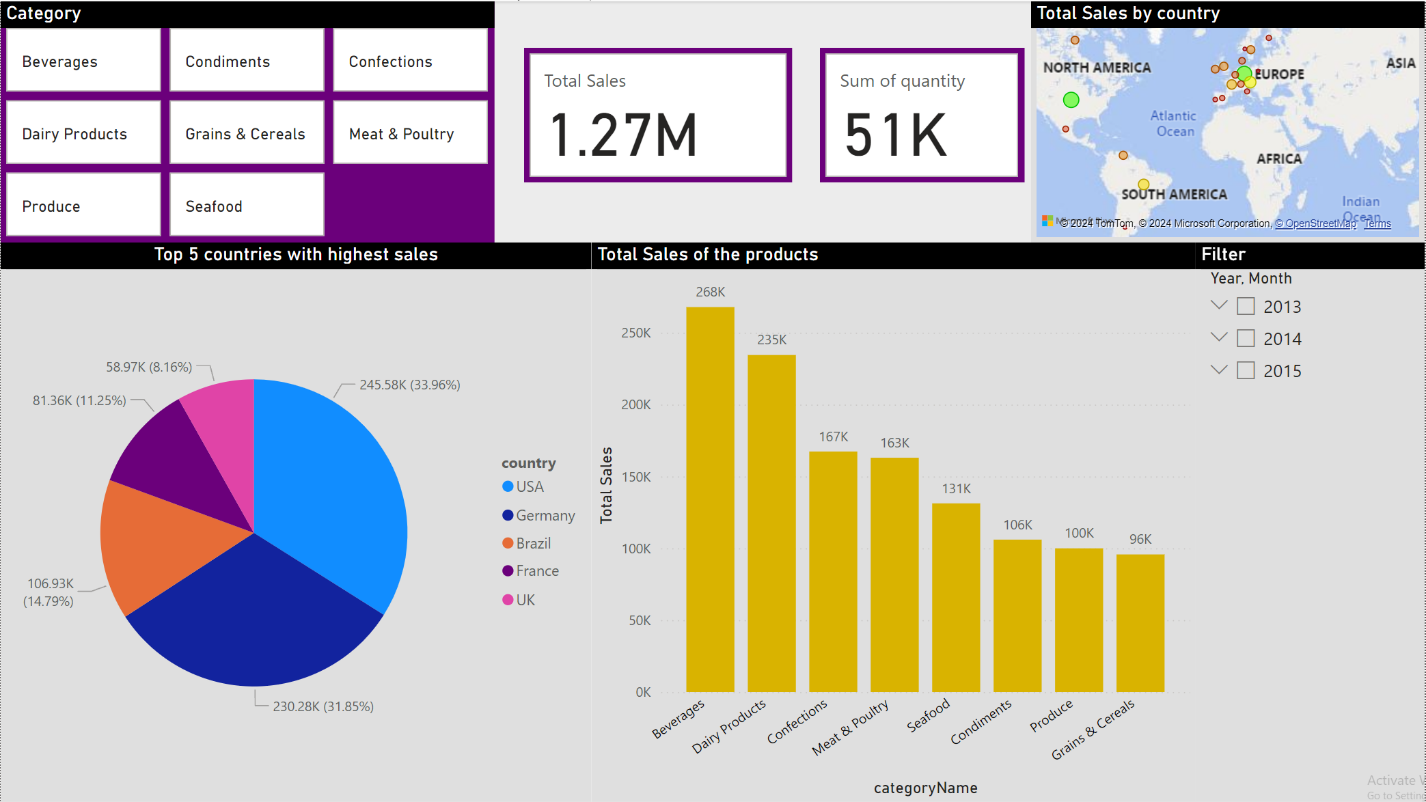
The dashboard will be divided into sections such as:

Sales Trends



This dashboard focuses on the total quantity ordered and sales trend analysis over time. The top section has tiles for various product categories and displays the same total sales and quantity figures. The map shows sales by country. The left graph shows the total quantity ordered for each product category, where Beverages lead, followed by Dairy Products, Confections, and so on down to Produce. On the right, the sales trend analysis graph shows the total sales over time from July 2013 to early 2015, revealing an upward trend with periodic fluctuations and a peak in early 2015 before a sudden drop. This dashboard helps in understanding the quantity distribution and the temporal sales performance.

Product Performance



This dashboard provides an overview of sales data categorized by different product categories and countries. At the top there are tiles representing different product categories such as Beverages, Condiments, Confections etc. The total sales amount to 1.27 million, with a total quantity sold of 51,000 units. A map on the right displays the total sales by country, with markers indicating the sales distribution geographically. Below, a pie chart highlights the top 5 countries with the highest sales: USA, Germany, Brazil, France, and the UK. The bar chart next to it shows the total sales of each product category, with Beverages and Dairy Products leading, followed by Confections, Meat & Poultry, Seafood, Condiments, Produce, and Grains & Cereals.

5.2 Features & Functionalities

-Interactive Elements: Charts and graphs for dynamic data exploration

-Filter & Search Option: Options to filter by date range, product category, and country.

-Navigation: Interactive navigation between different sections of the dashboard.

**Technical Requirements**

6.1 Software & Tools

- Power BI: For data visualization.

-Google drive: For storing data

-Microsoft Azure Data studios : For data preprocessing and ETL

- DAX: For advanced data calculations.

6.2 Hardware Requirements

- Google account: For Google drive storing and ensuring high availability.

-Microsoft account: For Power Bi full access and Azure Data Studios

- Computational Requirements: Sufficient CPU and RAM for data processing and visualization.

**Milestones & Timeline**

- Project Design Completion: (10/06/2024)

- Data Collection and Preparation: (10/06/2024)

- Data Modeling and ETL Implementation: (11/06/2024)

- Dashboard Development: (12/06/2024)

- Final Review and Presentation: (14/06/2024)

**Conclusion**

This design lined out in this document is a systematic approach to analyzing Northwind Traders sales data. The results gained from this project will help in making data driven decisions for Northwind Traders while optimizing product offerings and improving sales strategies. More enhancement can be done with this project by integrating more data sources and implementing advanced predictive analytics.