

Bharati Vidyapeeth
(DEEMED TO BE UNIVERSITY), College of
Engineering, PUNE- 43
COMPUTER ENGINEERING

Report for the fulfilment of Vocational Course

Topic – Data Modelling In Power BI

REPORT PREPARED BY –

Roll Number	Name	PRN
01	Arnab Adhikary	2114110064

SUBMITTED TO –

Dr. Sandeep Vanjale

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to Dr. Sandeep Vinjale for his invaluable guidance and support throughout the preparation of this report on Data Modelling in Power BI. His expertise and insights have been instrumental in shaping this work. I am truly thankful for his encouragement and feedback, which have significantly enriched the content of this report.

Thanks again to all who helped us

Teacher's Signature.....

CONTENTS

S.NO	TITLE	PAGE NO.
1.	Importance of Data Modelling	05-06
2.	Data Modelling in Power BI	07-08
3.	Process of Data Modelling in Power BI 2.1) Data Preparation 2.2) Creating Relationships 2.3) Defining Calculated Columns and Measures 2.4) Optimizing Data Model	09-10
4.	Best Practices for Data Modelling in Power BI	11-13
5.	Projects	14-23
6.	Conclusion	24
7.	Certificate	25

Importance of Data Modelling

Data modelling is a fundamental aspect of database design and plays a crucial role in ensuring that data is organized, structured, and represented accurately. In the context of business intelligence and analytics, effective data modelling is essential for deriving meaningful insights and making informed decisions. Here's why data modelling is important:

1. Organizing and Structuring Data:

Data modelling helps in organizing and structuring data in a way that reflects the real-world entities and their relationships.

By defining the structure of the database, data modelling ensures that data is stored in a logical and efficient manner.

2. Data Consistency and Integrity:

Properly designed data models enforce data consistency and integrity by defining rules and constraints.

Data modelling ensures that data is accurate, complete, and free from errors or inconsistencies.

3. Facilitating Analysis and Reporting:

A well-designed data model makes it easier to retrieve and analyze data for reporting and decision-making purposes.

By organizing data into meaningful entities and relationships, data modeling enables users to perform complex analysis and generate insightful reports.

4. Supporting Business Processes:

Data modelling aligns with the business processes and requirements, ensuring that the database structure supports the needs of the organization.

It allows for the customization of data structures to meet specific business objectives and goals.

5. Scalability and Performance:

An optimized data model improves the performance and scalability of the database system.

By eliminating redundancy and optimizing queries, data modeling enhances the efficiency of data retrieval and processing.

6. Adaptability to Changes:

Data modeling provides a flexible framework that can adapt to changes in the business environment, data sources, or requirements.

It allows for the seamless integration of new data sources and the modification of existing structures as needed.

7. Enhancing Data Governance and Security:

Data modeling supports data governance by defining access controls, security policies, and data ownership.

It helps in maintaining data quality and security by implementing measures to protect sensitive information.

In summary, data modeling is essential for creating a robust and efficient database structure that supports the organization's data management, analysis, and reporting needs. By investing in effective data modeling, organizations can derive maximum value from their data assets and gain a competitive advantage in today's data-driven business landscape

Data Modelling in Power BI

Data modelling in Power BI is the process of structuring and organizing data to create a meaningful and efficient data model that supports analysis and reporting. Power BI provides powerful tools and features to facilitate the data modelling process, allowing users to transform raw data into valuable insights. Here's an overview of data modelling in Power BI:

1. Data Preparation:

Power BI allows users to connect to various data sources such as Excel, SQL Server, CSV files, web sources, and cloud-based services.

The first step in data modelling is data preparation, where users can clean, transform, and shape the data using the Power Query Editor.

Power Query Editor provides a user-friendly interface to perform tasks such as removing duplicates, filtering rows, renaming columns, and applying transformations to the data.

2. Creating Relationships:

After preparing the data, the next step is to create relationships between different tables in the data model.

Power BI uses a relational model where data is organized into tables, and relationships are established between these tables.

Users can define relationships based on common fields (keys) between tables, such as primary keys and foreign keys.

Power BI automatically detects and suggests potential relationships based on column names and data types.

3. Defining Calculated Columns and Measures:

Once relationships are established, users can define calculated columns and measures using DAX (Data Analysis Expressions).

Calculated columns are virtual columns that derive values based on expressions or formulas applied to other columns in the table.

Measures are calculations performed on the data, such as sum, average, count, etc., and are used in visualizations to analyse the data.

4. Optimizing Data Model:

To optimize the data model for better performance and efficiency, users can take the following steps:

Remove unnecessary columns and tables that are not required for analysis.

Use calculated columns and measures instead of loading redundant data into the model.

Optimize relationships by defining appropriate cardinality (one-to-many, many-to-one) and cross-filter direction.

Use hierarchies to organize data and enable drill-down analysis.

Apply data compression and storage mode settings to minimize the size of the data model.

5. Best Practices for Data Modelling in Power BI:

Use meaningful names for tables, columns, and measures to improve readability and understanding.

Follow a star schema or snowflake schema to organize data into fact and dimension tables.

Use surrogate keys to establish relationships between tables instead of using natural keys.

Regularly update and refresh the data model to reflect changes in the underlying data sources.

6. Visualization and Analysis:

Once the data model is created, users can use Power BI's rich visualization capabilities to analyse and visualize the data.

Users can create interactive reports, dashboards, and visualizations to gain insights and make data-driven decisions.

In summary, data modelling in Power BI is a critical step in the process of transforming raw data into valuable insights. By following best practices and utilizing Power BI's features and tools, users can create efficient and optimized data models that support analysis, reporting, and decision-making.

Process of Data Modelling in Power BI

Data modelling in Power BI is the process of structuring your data to optimize analytical performance and create clear, insightful reports. Here's a breakdown of the key steps involved:

1. Understanding The Data:

Data source identification: Start by identifying the various data sources you'll be using, such as databases, spreadsheets, or cloud storage.

Data exploration: Explore the structure and content of each data source. Familiarize yourself with the tables, columns, and data types involved.

Data quality assessment: Evaluate the quality of your data. Look for inconsistencies, missing values, or errors that might need cleaning or transformation before modelling.

2. Data Ingestion and Transformation:

Import data: Import your data sources into Power BI Desktop. You can connect directly to databases, load files, or use Power Query Editor for more advanced data manipulation.

Data transformation (optional): In Power Query Editor, you can clean, shape, and transform your data as needed. This might involve removing duplicates, formatting data types, splitting or merging columns, or creating calculated columns.

3. Defining Relationships:

Identify relationships: Determine the relationships between your tables based on common fields. For example, a Sales table might have a foreign key referencing a Customer ID in a Customers table.

Establish relationships: Create relationships between tables in the Power BI relationship's view. Specify the matching columns and the cardinality (one-to-one, many-to-one, or many-to-many). Many-to-many relationships are not generally recommended due to potential data redundancy and performance issues. Consider creating bridge tables if necessary.

4. Defining Dimension and Fact Tables:

Dimension tables: Identify the tables that contain descriptive attributes used to categorize and contextualize your data. These tables typically have a unique identifier column and might have hierarchical structures. Examples include Customers, Products, Dates, or Locations.

Fact tables: Identify the tables that store the quantitative data you'll be analysing for insights. These tables often summarize business events or transactions and might have foreign keys referencing dimension tables. Examples include Sales figures, Inventory levels, or Customer interactions.

5. Defining Measures:

Measures: Create calculations to summarize or analyse your data. Measures are typically based on data points in the fact table and can involve aggregations (e.g., SUM, AVERAGE), calculations, or comparisons.

6. Validation and Iteration:

Validation: Test your data model to ensure it produces the expected results. Use visualizations and explore data to identify any inconsistencies or errors.

Iteration: Refine your data model as needed. You might need to adjust relationships, create calculated columns, or modify measures based on your analysis and exploration.

Additional Considerations:

Star schema vs Snowflake schema: Decide whether a Star schema (simpler structure with central fact table) or a Snowflake schema (more normalized with separate dimension attribute tables) is more suitable for your data complexity.

Data security: Implement appropriate data security measures to restrict access to sensitive information based on user roles and permissions.

By following these steps and best practices, you can create a well-structured data model in Power BI that facilitates efficient data analysis and generates valuable insights for your business.

Best Practices in Data Modelling in Power BI

Here's a breakdown of best practices for data modeling in Power BI to ensure your data model is optimized for performance, clarity, and maintainability:

Data Source and Transformation:

Understanding your data: Before diving into modeling, thoroughly explore your data sources. Identify tables, columns, data types, and potential quality issues that might require cleaning or transformation.

Data transformation best practices: When transforming data in Power Query Editor, prioritize efficiency and maintainability. Use clear and descriptive names for columns and steps. Consider reusability by creating reusable functions for common transformations.

Data filtering: If possible, filter data at the source (database) to reduce the amount of data loaded into Power BI and improve performance.

Table Structure and Relationships:

Star vs. Snowflake schema: Choose the schema that best suits your data complexity. A Star schema is simpler and faster for basic analysis, while a Snowflake schema is more normalized and efficient for complex data hierarchies.

Dimension tables: Dimension tables should be normalized to reduce redundancy and improve data integrity. Use descriptive names for columns and attributes. Consider implementing hierarchical structures within dimensions if your data has natural hierarchies (e.g., product categories with subcategories).

Fact tables: Fact tables should be de-normalized to avoid complex joins during analysis. Include foreign keys referencing dimension tables. Store pre-aggregated data in fact tables for faster querying, especially for large datasets.

Relationship management: Clearly define relationships between tables based on matching columns. Set the appropriate cardinality (one-to-one, one-to-many, many-to-one). Avoid many-to-many relationships directly between tables. If necessary, create bridge tables to resolve these relationships.

Data Modelling Best Practices:

Data types: Ensure data types are consistent and appropriate for the data they represent. Using the correct data type can improve performance and avoid errors.

Minimize cardinality changes: Limit the use of functions that change the cardinality of a column within a table. This can impact how filters and aggregations work in your model.

Calculated columns vs. measures: Use calculated columns for transformations that are needed throughout your model, while measures are best suited for calculations specific to a particular analysis.

Date tables: Create a dedicated date table with various date hierarchies (year, month, day, etc.) for flexible time-based analysis.

Data partitioning: For very large tables, consider data partitioning to improve query performance. Partitioning allows Power BI to focus on specific data subsets based on your filters.

Model documentation: Document your data model, including data sources, transformations, relationships, and measure definitions. This improves clarity and maintainability for yourself and others who might work with the model in the future.

Additional Tips:

Model validation: Test your data model thoroughly. Use visualizations and explore data to identify any inconsistencies or errors in calculations or aggregations.

Performance optimization: Monitor model performance and identify bottlenecks. Consider techniques like data partitioning, filtering at the source, or optimizing complex calculations for better query speed.

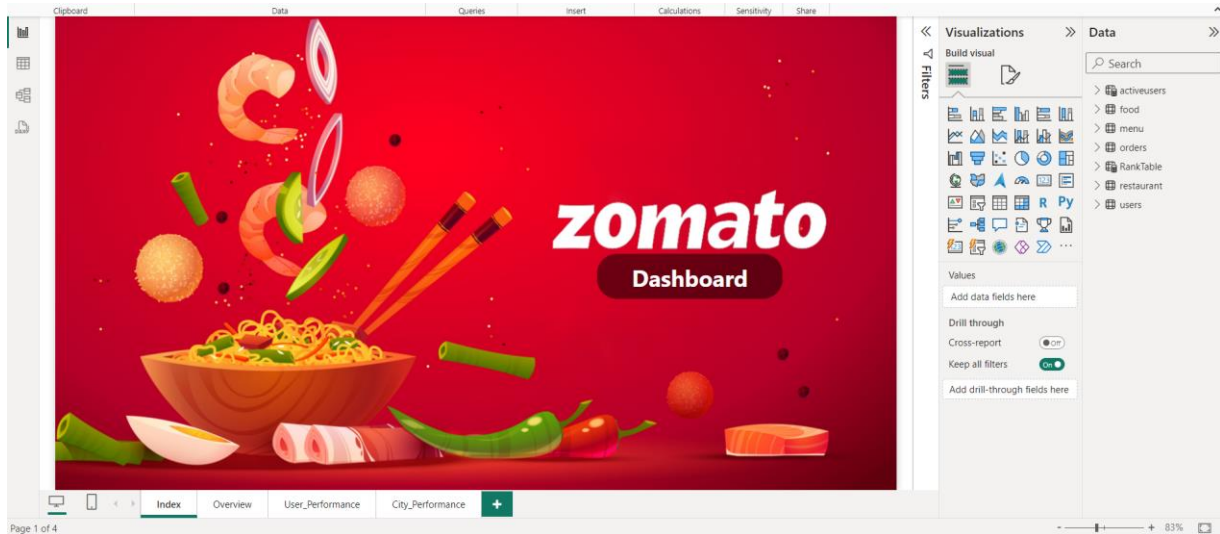
Security considerations: Implement data security measures to restrict access to sensitive information based on user roles and permissions.

By following these best practices, you can create robust and efficient data models in Power BI that empower insightful data analysis and informed decision-making within your organization.

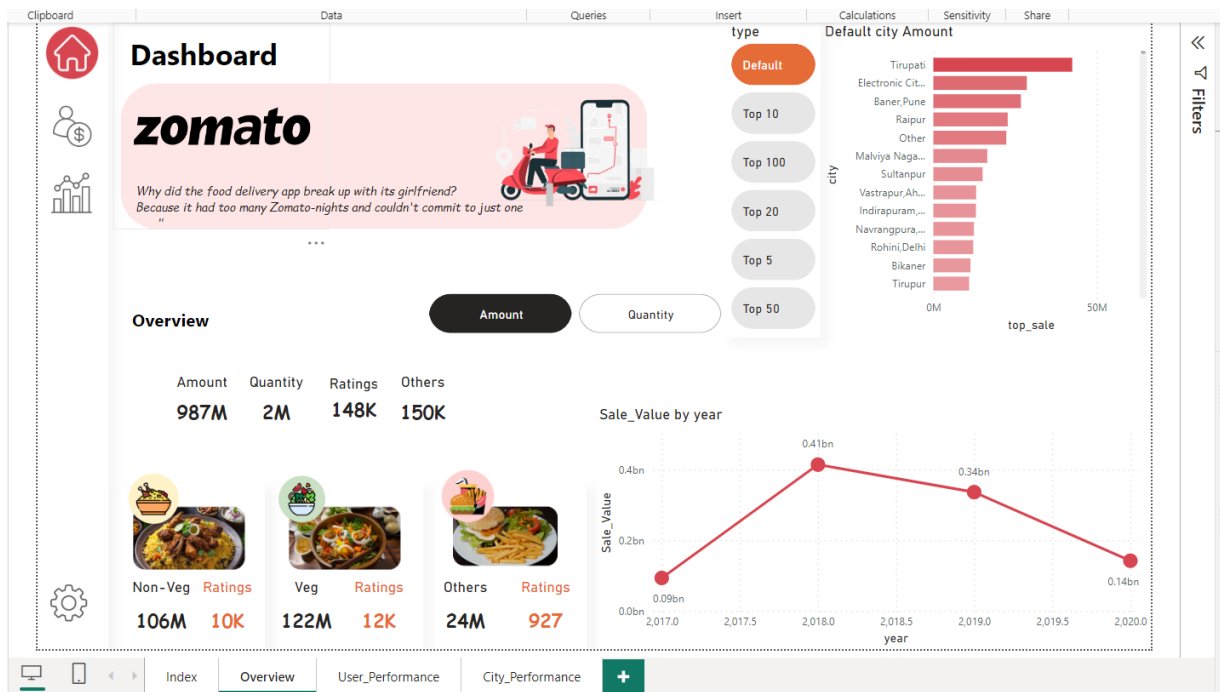
Projects:

1.) Zomato Dashboard:

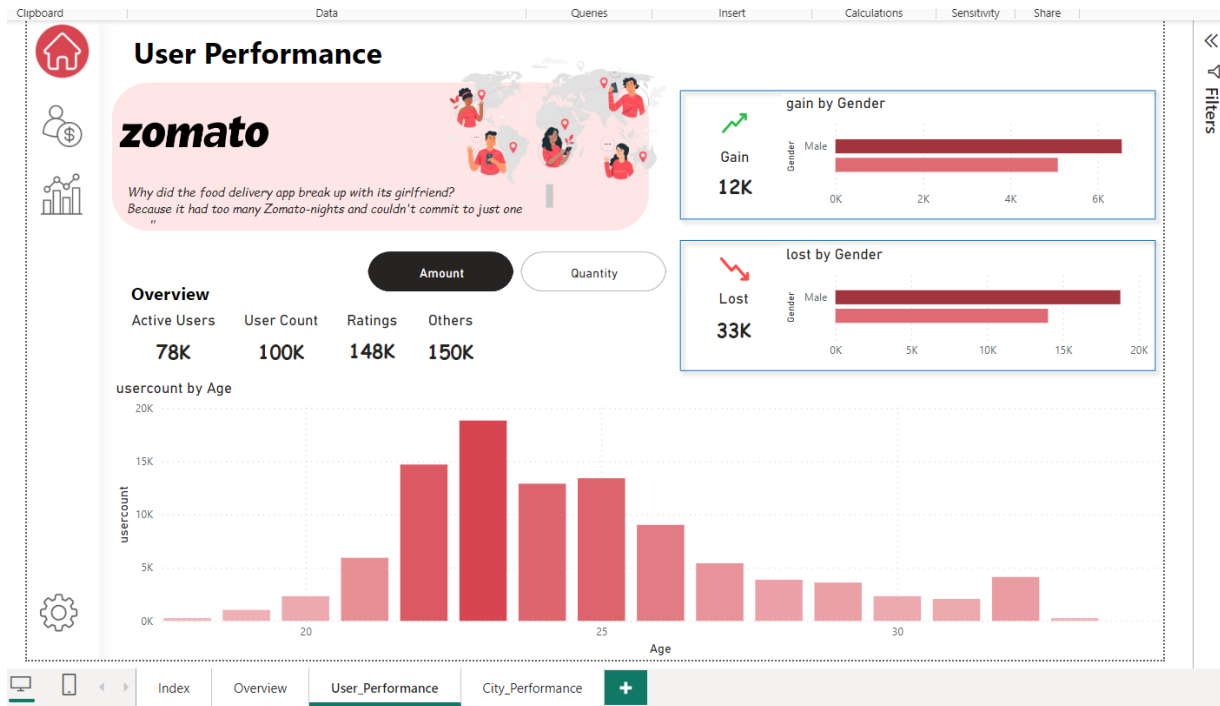
→ Index page



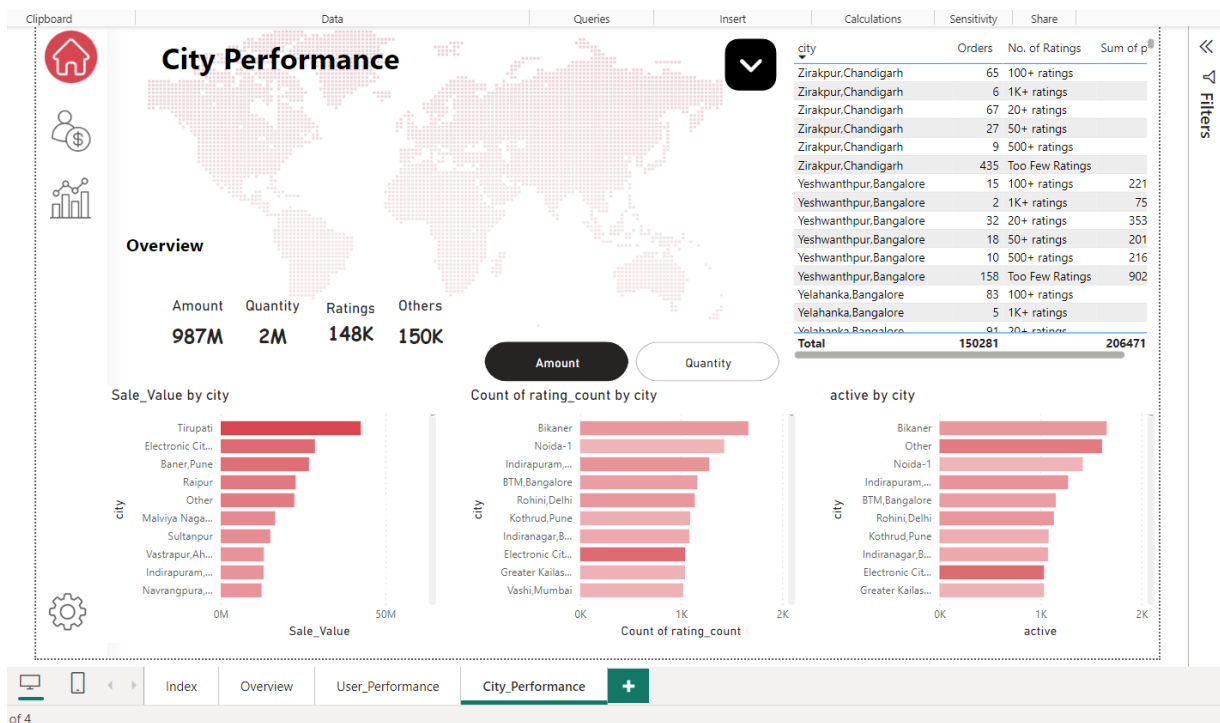
→ Overview



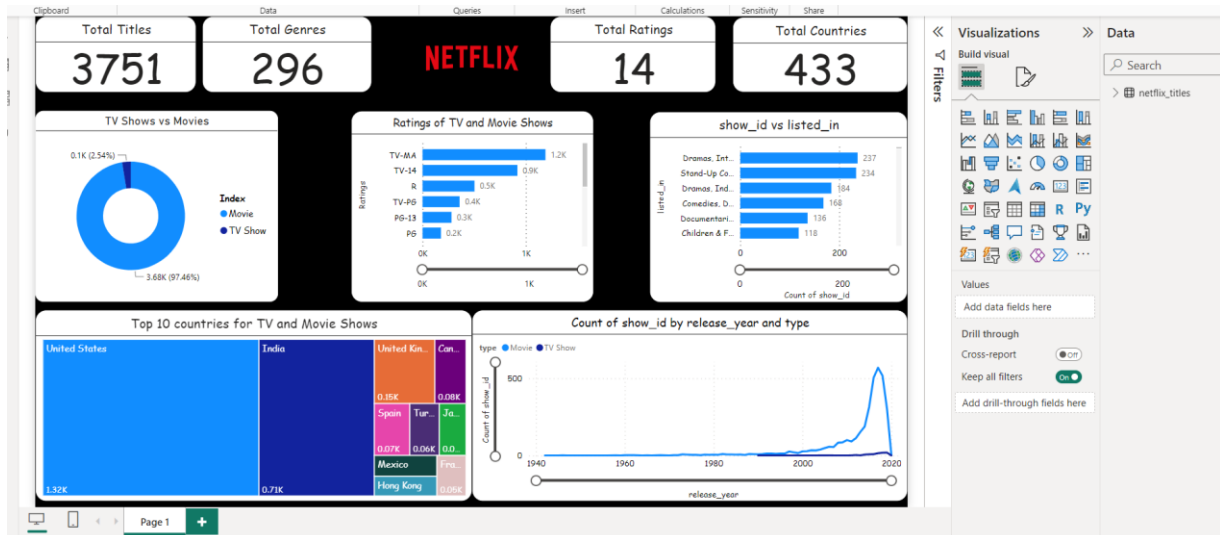
→ User Performance Page:



→ City Performance:



2.) Netflix Dashboard:



3.) Empowering Business with Effective Insights:

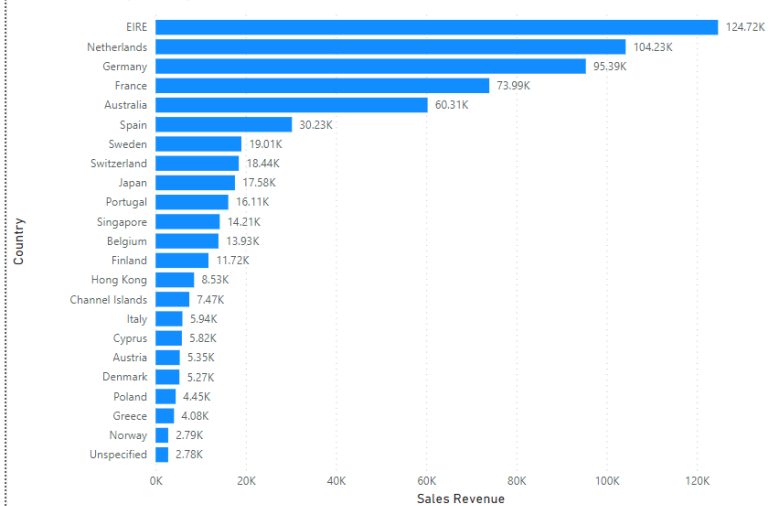
→Country wise Performance For sales



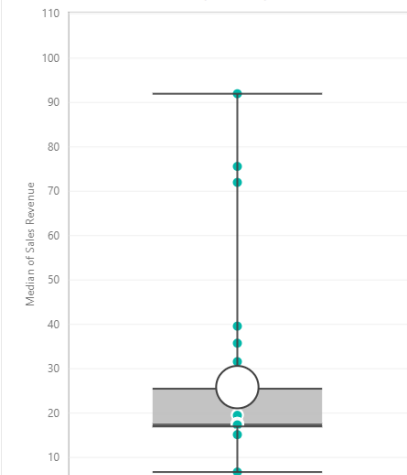
→Sale Revenue Performance by Countries

Sales Revenue Performance by Countries

Sales Revenue by Country



Median of Sales Revenue by Country

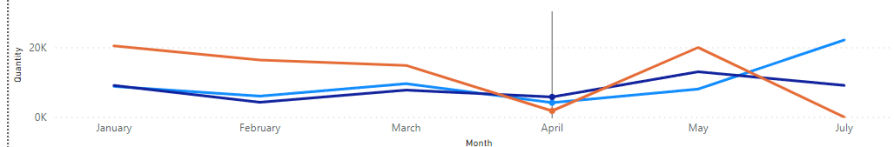


→Sales Trends

Sales Trend

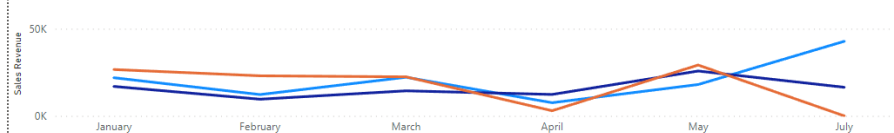
Quantity by Month and Country

Country ● EIRE ● Germany ● Netherlands



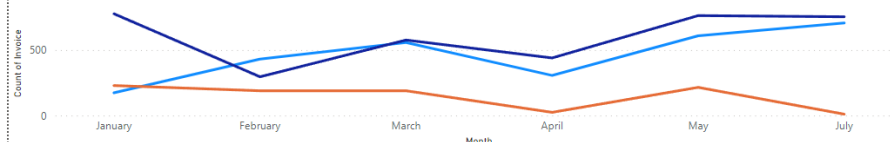
Sales Revenue by Month and Country

Country ● EIRE ● Germany ● Netherlands

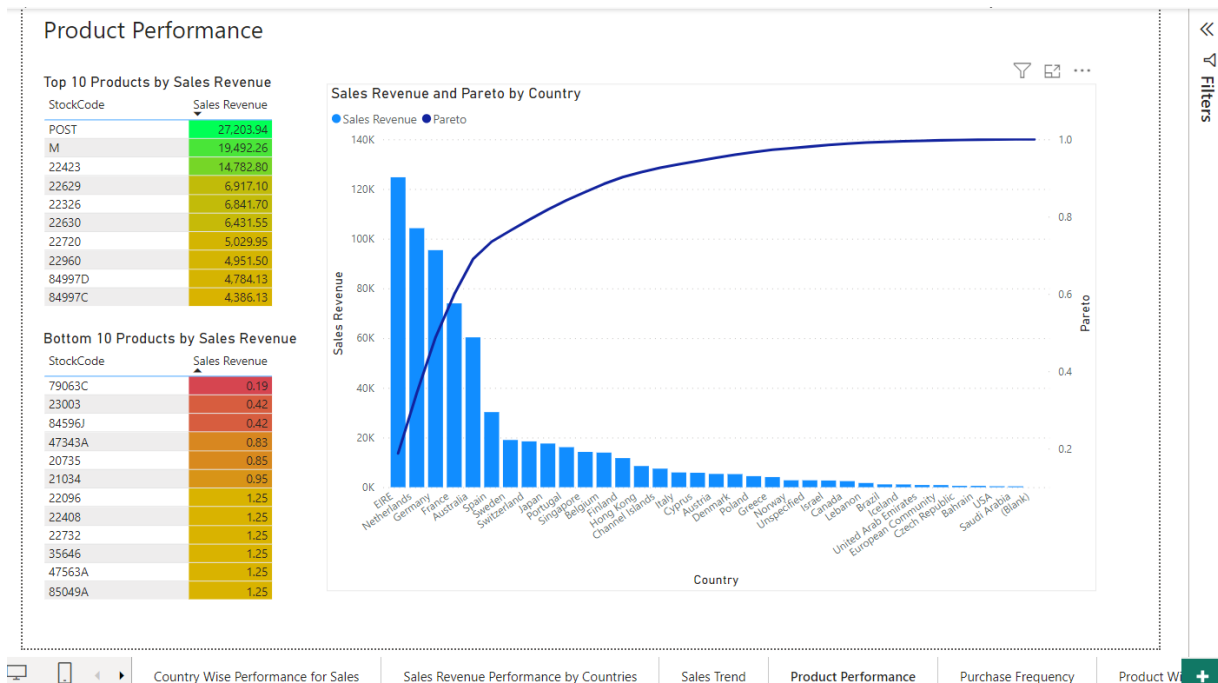


Count of Invoice by Month and Country

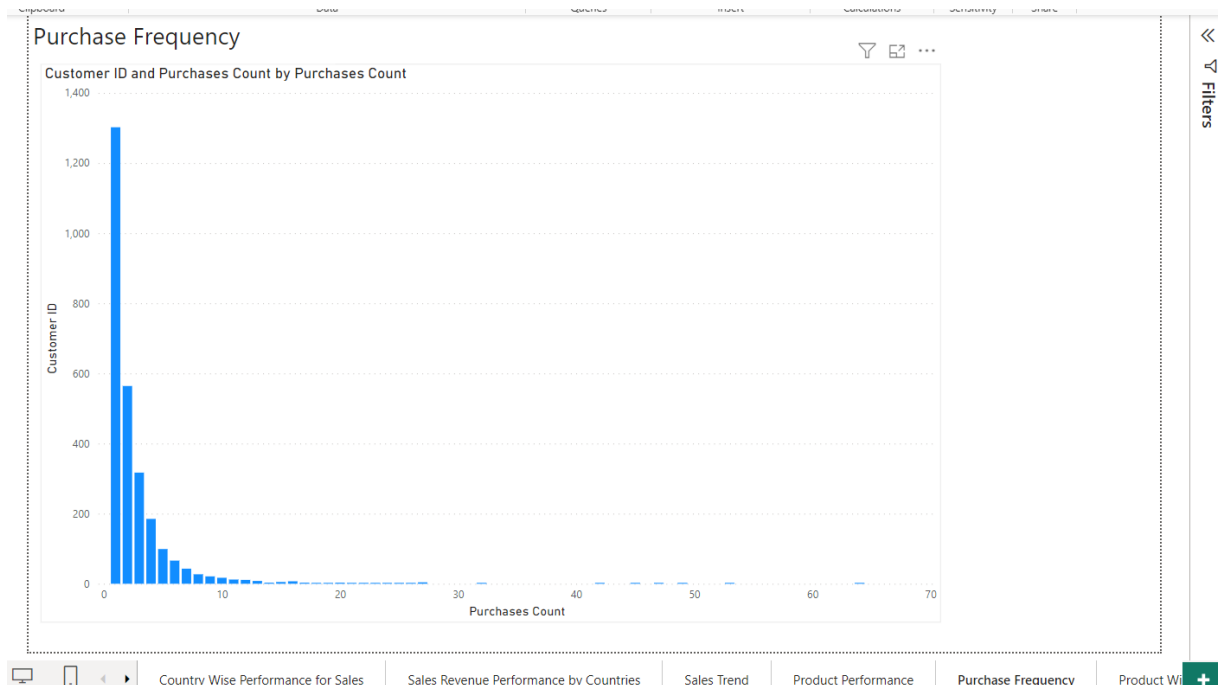
Country ● EIRE ● Germany ● Netherlands



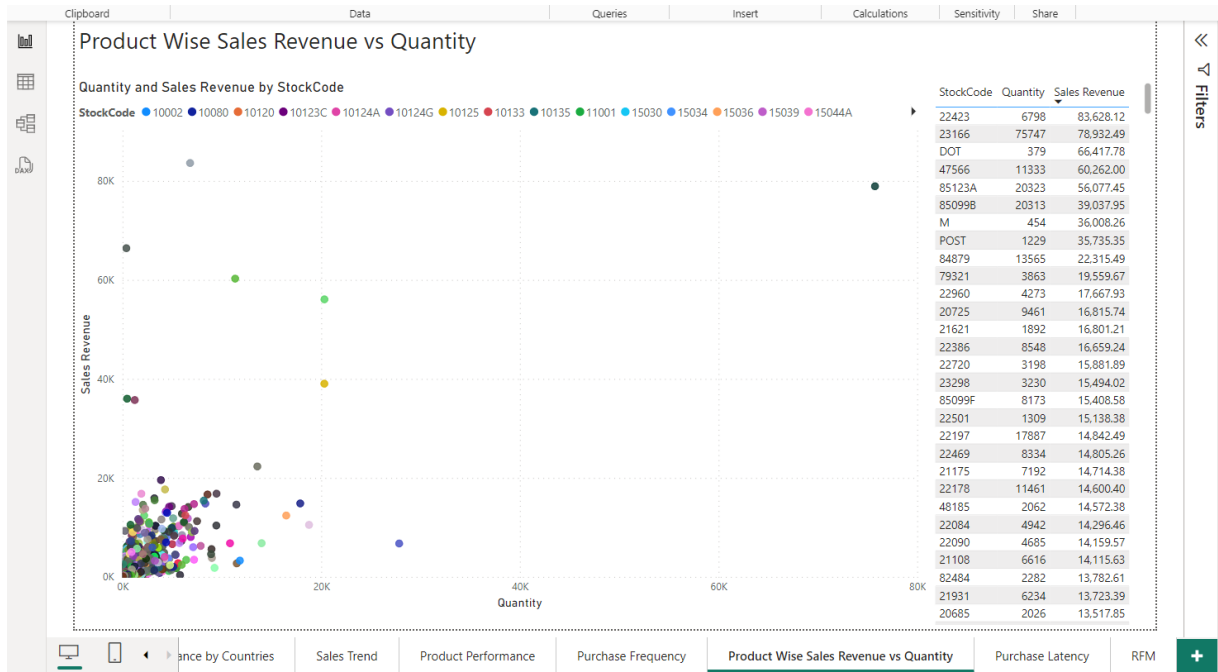
→Product Performance



→Purchase Frequency



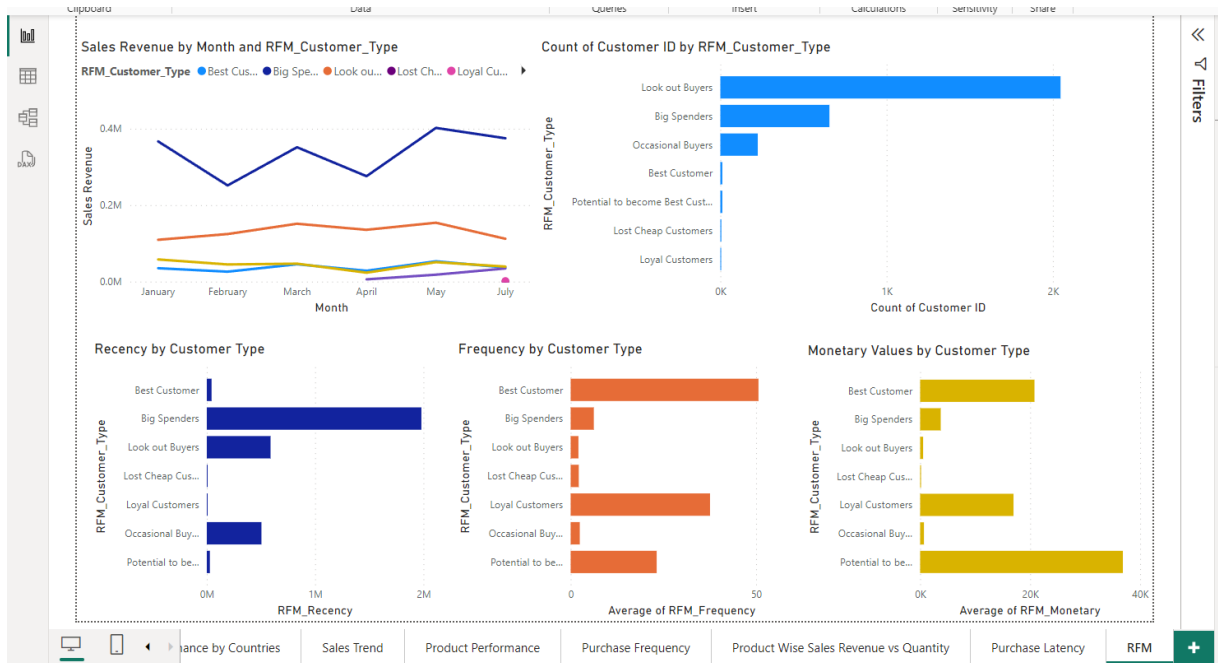
→Product wise sales revenue vs quantity



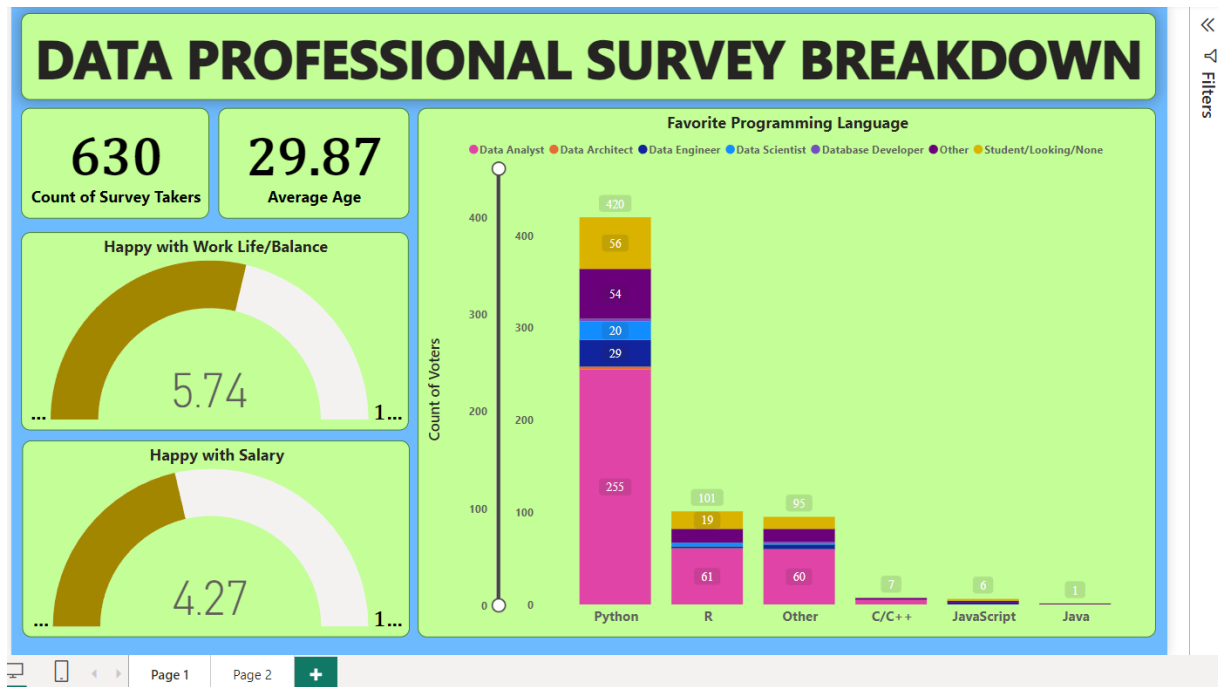
→ Purchase Latency

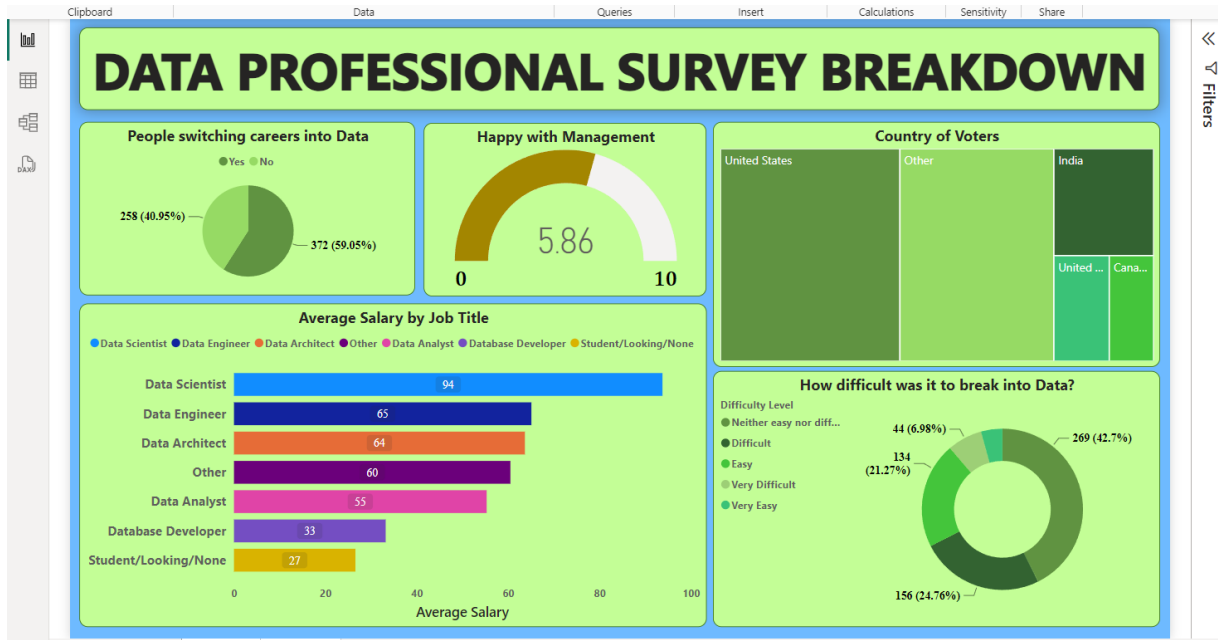
Month	0	1	2	3	4	5	6
January	741	152					
February	496	150	262				
March	539	201	290	145			
April	361	170	304	120	71		
May	319	280	368	221	94	54	
July	260	225		430	119	80	36
Total	2716	687	648	713	274	134	36

→RFM

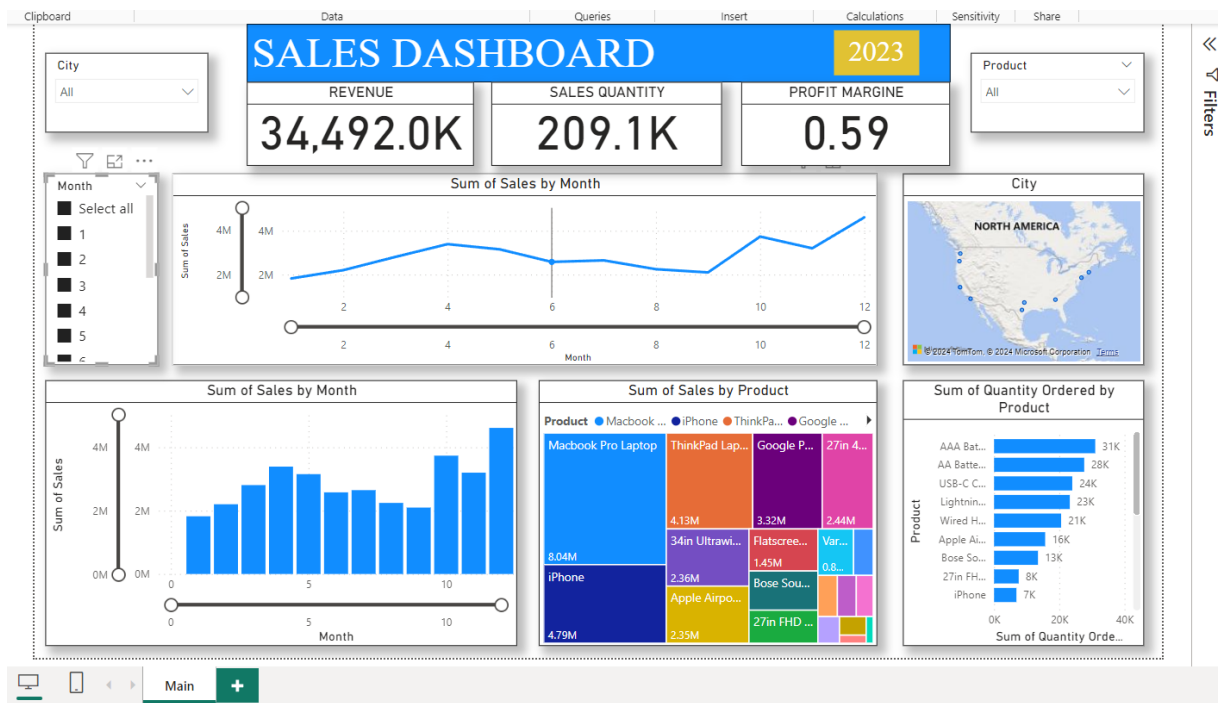


4.) Data professional survey break down:



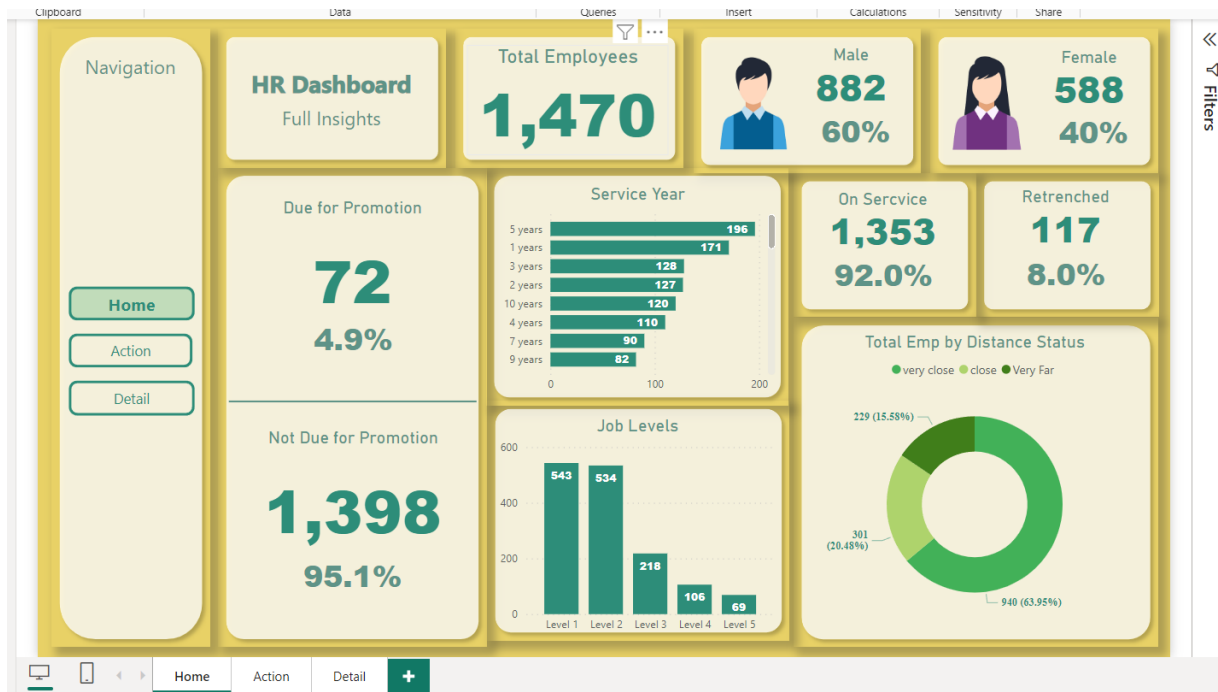


5.) Sales Dashboard

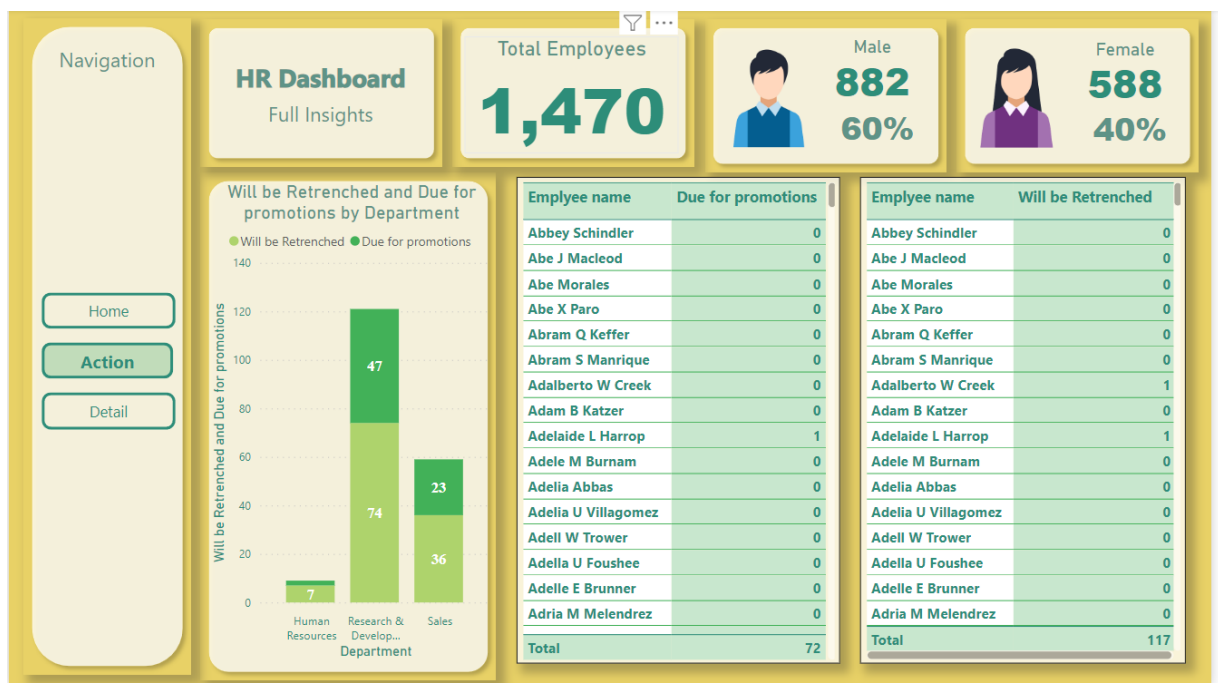


6.) Human Resource Dashboard

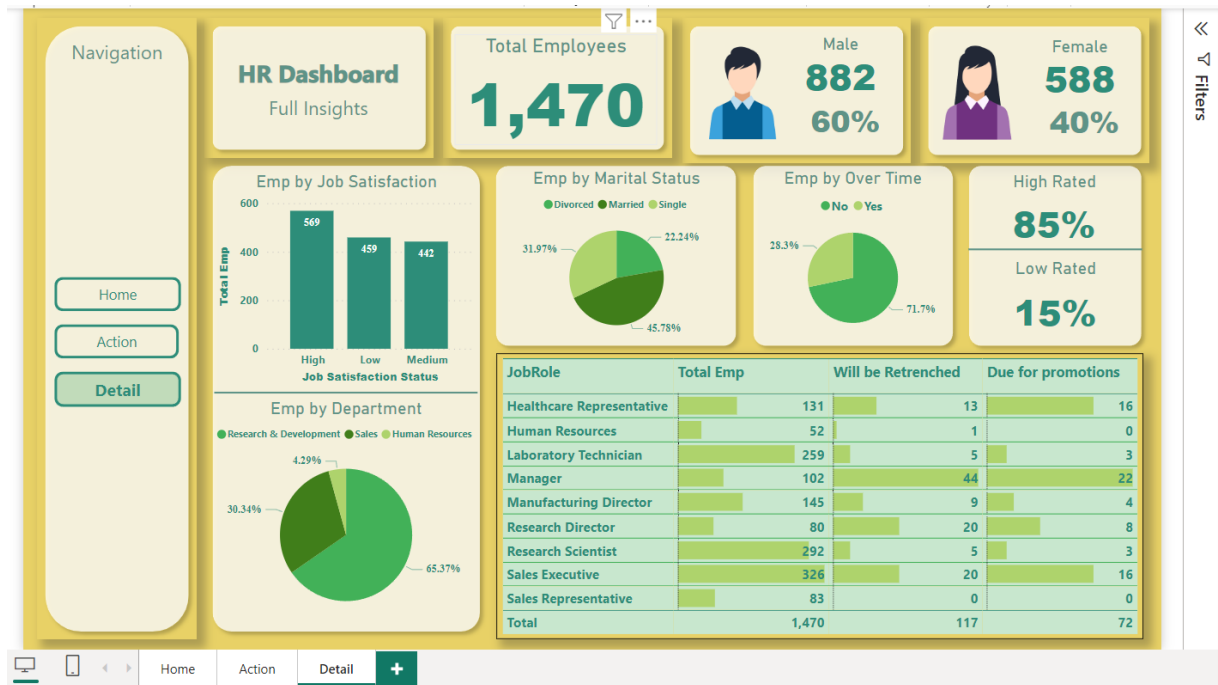
→ Home



→ Action



→ Details



Conclusion:

Building a Strong Foundation for Data Analysis with Power BI Data Modelling

Effective data modelling in Power BI is the cornerstone of successful data analysis. It lays the groundwork for clear, insightful reports and facilitates efficient exploration of your business data. By understanding your data sources, meticulously transforming and structuring your data, and establishing clear relationships between tables, you can create a robust data model.

This well-defined model offers numerous advantages:

Simplified data exploration: Users can easily navigate and analyze data through intuitive visualizations and filters.

Improved performance: A well-structured model reduces query complexity and ensures faster data retrieval for interactive analysis.

Enhanced data quality: By normalizing dimension tables and minimizing redundancy, you promote data integrity and consistency.

Scalability and maintainability: A well-designed model can accommodate future data growth and simplifies modifications as your business needs evolve.

Investing time and effort in crafting a well-structured data model in Power BI is an investment in the future of your data analysis. It empowers users to glean valuable insights from your data, ultimately driving better decision-making across your organization. Remember, data modelling is an iterative process. As you gain experience and your data needs evolve, you can refine your model to ensure it continues to meet your analytical requirements.

Certificate:

