



FINAL PROJECT REPORT

Amazon sales data analytics

SUBMITTED BY
KRISHNANUNNI BR

INTRODUCTION

In today's highly competitive business landscape, effective sales management has become crucial for organizations aiming to optimize distribution methods, minimize costs, and enhance profitability. With the exponential growth of e-commerce platforms like Amazon, analyzing sales data has emerged as a pivotal strategy for understanding consumer behavior, identifying market trends, and making informed decisions.

This project aims to perform an ETL (Extract-Transform-Load) analysis on an Amazon sales dataset to uncover valuable insights related to sales trends. The focus will be on analyzing sales data on a month-wise and year-wise basis, as well as a detailed yearly_monthly analysis. By transforming raw sales data into a structured format, we will extract meaningful metrics and relationships among various attributes that influence sales performance.

By understanding sales trends and identifying key factors that drive performance, businesses can refine their sales strategies, enhance customer targeting, and optimize inventory management. The insights derived from this analysis will empower stakeholders to make data-driven decisions, ultimately contributing to improved sales outcomes and sustained competitive advantage in the marketplace. This project will leverage advanced analytical techniques and tools to provide a comprehensive overview of Amazon's sales dynamics, offering actionable insights for sales management and strategic planning.

Extract-Transform-Load

- Data profiling to find errors, empty records etc
- Removing empty records
- Replacing errors- e.g., null records with mean or median depending on categorical or numerical column
- Removing duplicates
- Formatting - data types- e.g., to date tyme, currency, whole number, decimals, integer, text etc. and helps while performing Dax operation for numerical columns

Extract-Transform-Load on amazon sales data

Extract: Sources of data (e.g., 100 Sales Records, Products, customers, region, division, Amazon Food Category etc.).

Transform: Data cleaning and shaping in Power Query.

- promoted Headers and changed data type of columns
- performed complex steps to modify order date column into required date column format
- renamed country name from "Cote d'Ivoire " to "Cote d voire ", to match currency table with newly added exchange rate column inorder to create relationships in data model
- relaced null values with 0 in discount amount column in amazon food category table
- created new column named 'Discount amount modified' and performed queries to removed negative values from the discount amount column

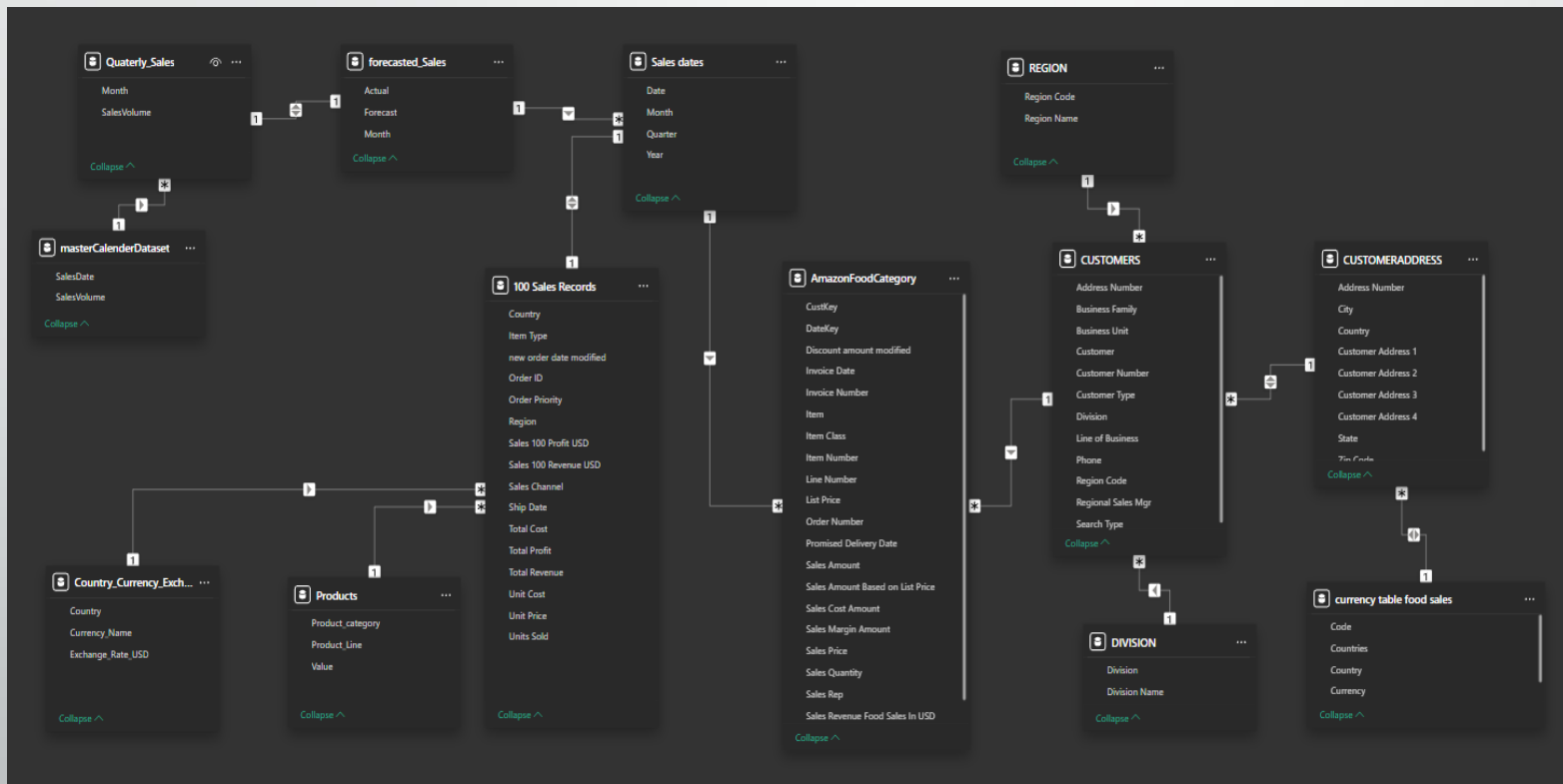
Load: Loading transformed data into Power BI for analysis

DATA MODELLING

Data Modeling (Schema Setup and Dax):

- Creating tables (e.g., date tables, currency tables etc., other dimensional tables),
 - Define measures and calculated columns using DAX (Data Analysis Expressions) for custom calculations. (E.g., Revenue Amazon food is a measure created)
 - Setting data types and define primary/foreign keys(unique identifiers) to establish relationships btw table using cardinality (one-to-one, one-to-many, or many-to-many) and cross filter function (single directional or bi-directional for data flow from one table to other).
 - Configuring table and column properties under right and top section in model view for tables. e.g., summarization and categorization under column tools
 - Build hierarchies (e.g., date column- year, month, day etc.), calculated columns, and measures
 - Ensure the model reflects the correct structure for the analysis, aligning dimension tables with fact tables using flat or star or snowflake schemas relevant for the table
- Snowflake schema has been used in this amazon sales data model due to its complexity.

Amazon Sales-Snowflake Schema- Data Model



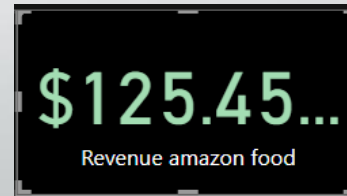
DATA VISUALIZATION

Report Building (Visualization):

- The purpose of data visualization is to provide a clear and concise overview of Amazon sales data enabling users to quickly identify trends, patterns & insights.
- Use the data model to create visuals, such as charts, tables, and graphs.
- Applying filters, slicers, and conditional formatting to enhance insights.

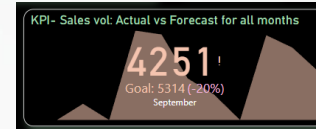
Key Visualizations

- Cards- for showing measure values of sales quantity and sales revenue

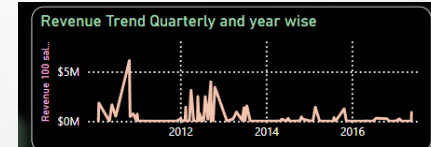


DATA VISUALIZATION

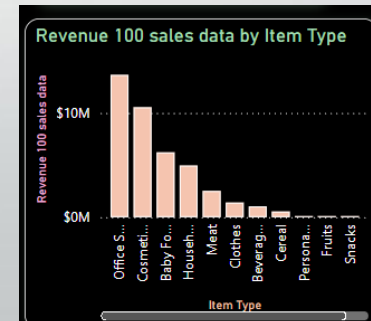
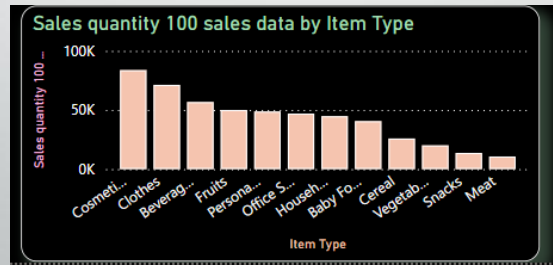
- KPIs – Sales vs forecast for all months



- Line chart- for showing revenue trend quarterly and yearwise

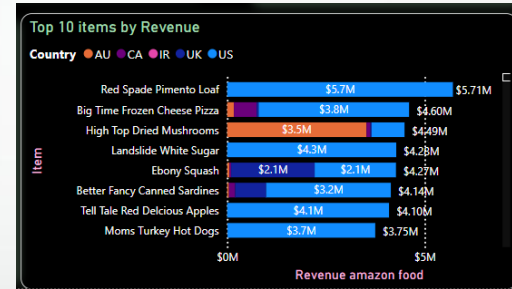
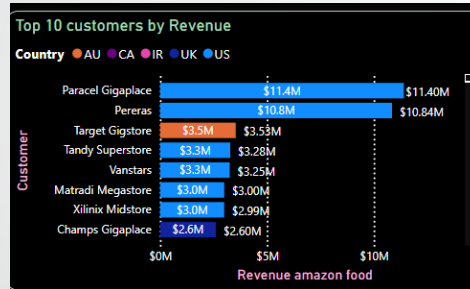


- Clustered column chart- for showing revenue sales data by item type and sales quantity by item type

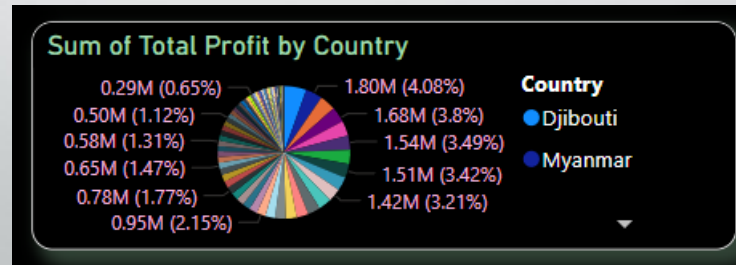


DATA VISUALIZATION

- Stacked bar chart- For displaying top 10 customers amazon food by revenue with legend as country and also top 10 items amazon food by revenue with legend as country

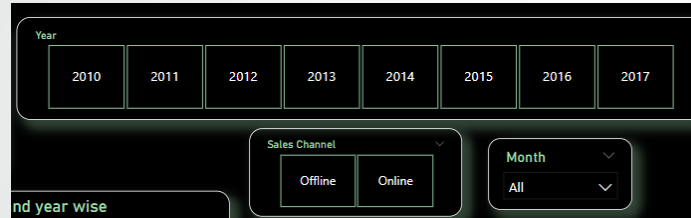


- Pie chart- for showing sales profit country wise

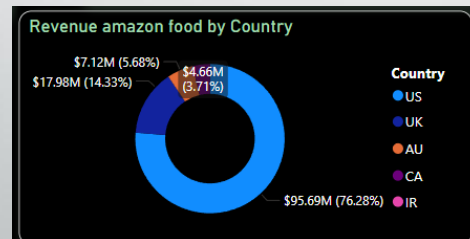


DATA VISUALIZATION

- Slicers – for filtering years and months, sales channel etc.



- Donut chart- for showing Revenue of amazon food by country and sales quantity by region name



DATA VISUALIZATION

- Map visual – for spotting states with sales quantity for amazon food and also sum of sales quantity by region name



POWER BI REPORT GENERATION

KEY TRENDS AND INSIGHTS

- Across all 7 Years, 100Sales Amount ranged from 1.32 M to 18.45M.
- 2012 accounted for 44.7% of Sales Revenue. 2012 had the highest Sales Revenue at 18.45 M and the most sold product by item type was office supplies which was around 5.3M.
- Europe accounted for highest sales by region with 23.2M followed by Australia and Oceania at 8.3M from 2010- 2017. The country with maximum sales was Djibouti which accounted for 4.08% of overall revenue.
- Online sales accounted for 45.19% of total revenue and offline sales accounted for 54.8% of sales revenue throughout the period of 8 years.

POWER BI REPORT GENERATION

Amazon Food Sales analysis:

- US accounted for 76.28% of total sales revenue of all countries and came top with 95.69M and CA came last with 4.66M which was 3.71% of total sales from 2017-2019
- Q1 acquired the highest number of total sales quarter wise across 3 years with 39.9M peak revenue and Q2 came last with 26.2M.
- Red spade Pimento loaf was the top selling food with 5.3M sales revenue followed by frozen cheese pizza with 4.6M and Paracel Giga place in US was the top performing customer with 11.4M sales revenue.

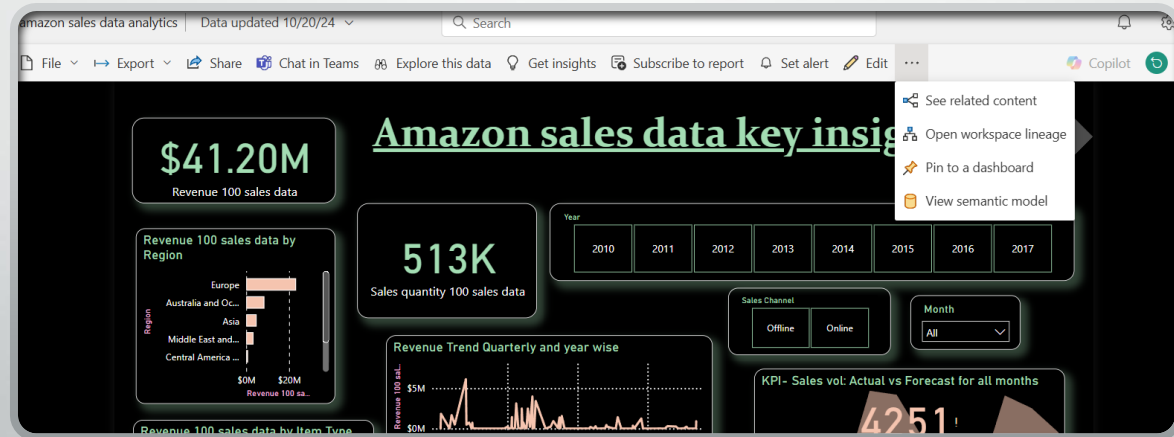
POWER BI SERVICE EMBEDDED REPORT GENERATION



POWER BI SERVICE

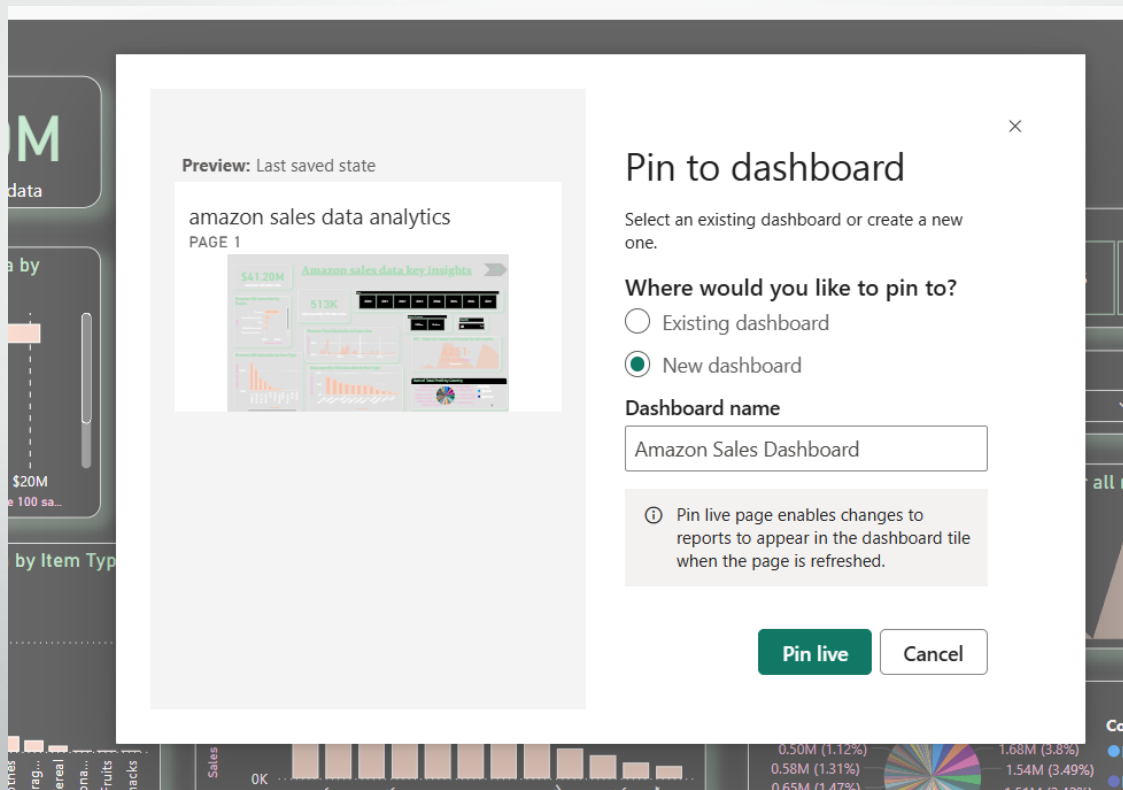
Publishing and Sharing:

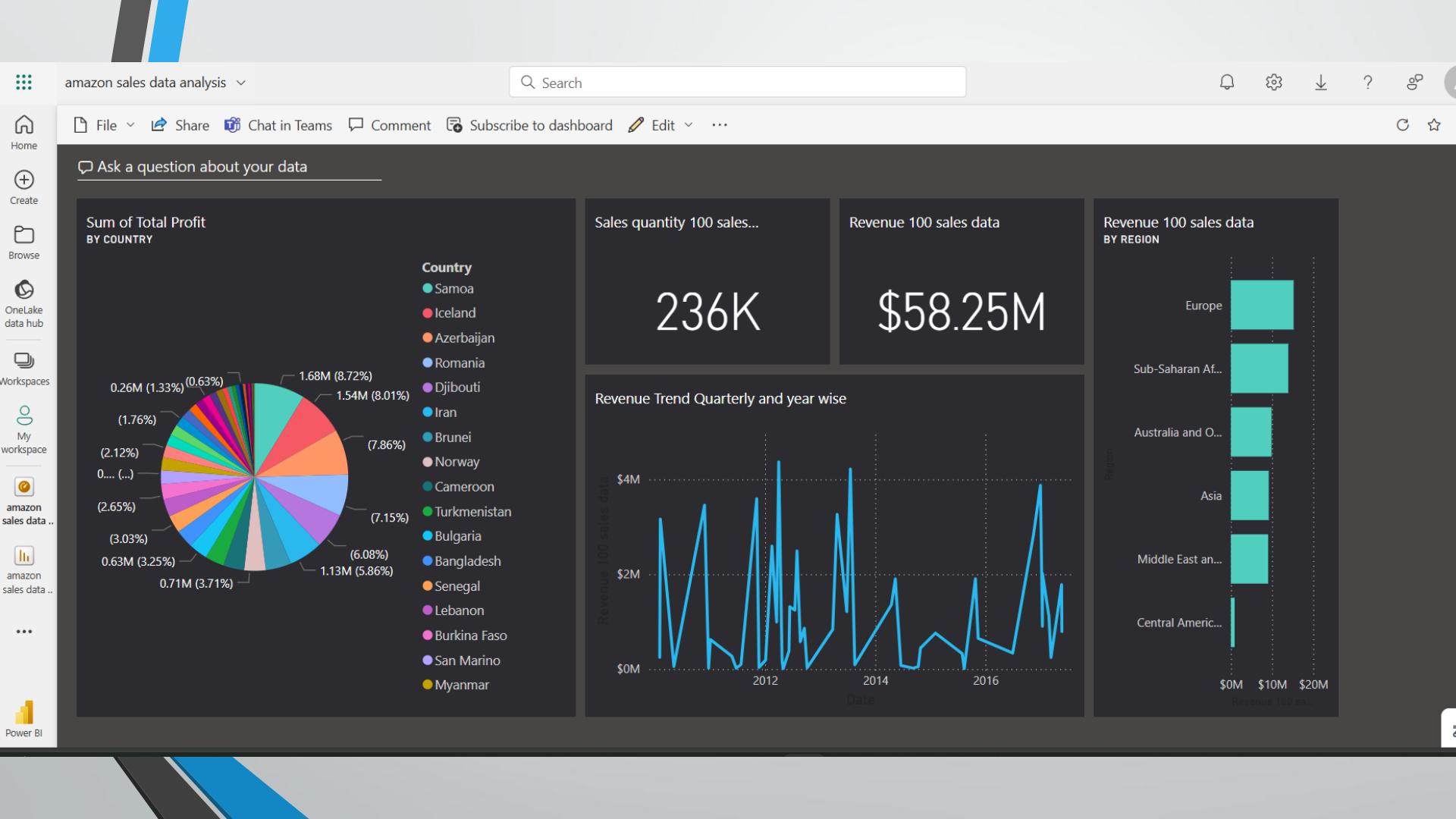
- Publish the report to Power BI Service for sharing and collaboration



POWER BI DASHBOARD CREATION

- pinning and creating dashboards in powerbi service







CONCLUSION

In conclusion, the data analysis of amazon sales was done on 3 different levels –sales key insights, food sales insights and food sales market. Extract-Transform-Load has been done and several key metrics, factors ,attributes and relationships have been generated using data modelling and data visualization in power bi desktop. The developed report was published to power bi service and corresponding dashboard was created



THANK YOU