A REPORT OF ONE MONTH TRAINING

at

NETMAX TECHNOLOGIES

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD

OF THE DEGREE OF

BACHELOR OF TECHNOLOGY

(COMPUTER SCIENCE & ENGINEERING)

JUNE-JULY ,2025

SUBMITTED BY:

Kirat

URN:2302587

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

GURU NANAK DEV ENGINEERING COLLEGE LUDHIANA

(An Autonomous College Under UGC ACT 1956)

CERTIFICATE BY COMPANY

GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA

CANDIDATE'S DECLARATION

I Kirat hereby declare that I have undertaken one month training **NETMAX TECHNOLOGIES** during a period from June 2025 to July 2025 in partial fulfilment of requirements for the award of degree of B.Tech (Computer Science and Engineering) at GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA. The work which is being presented in the training report submitted to Department of Computer Science and Engineering at GURU NANAK DEV ENGINEERING COLLEGE, LUDHIANA is an authentic record of training work.

Signature of the Student

The one month industrial training Viva–Voce Examination of_____ has been held on _____ and accepted.

Signature of Internal Examiner Signature of External Examiner

ABSTRACT

This project was completed as part of a 4 week industrial training program in Data Analytics at Netmax Technologies, Chandigarh. The primary objective of the training was to gain practical exposure to data extraction, analysis, and visualization using Power BI and SQL.

Two major projects were undertaken during the training — Adventure Works Sales Dashboard using Power BI and Restaurant Sales Analysis using SQL. The Adventure Works Dashboard involved designing an interactive Power BI report to analyze global sales performance, profit distribution, and product category trends. It incorporated key performance indicators (KPIs) such as total sales, cost of goods sold (COGS), and profit margin, supported by visuals like bar charts, maps, and pie charts for intuitive analysis.

The Restaurant Sales Analysis project focused on SQL-based data processing and reporting. Various queries were executed for data cleaning, filtering, and aggregation to identify top-selling products, customer purchase patterns, and revenue distribution across time periods.

This training enhanced my understanding of data analytics concepts, ETL (Extract, Transform, Load) processes, and business intelligence (BI) tools. It strengthened my technical skills in Power BI, SQL Server, while improving my ability to interpret and visualize business data. Overall, this project demonstrates the practical application of analytics tools in solving real-world business problems and emphasizes the value of data-driven decision-making in modern enterprises.

ACKNOWLEDGE

I am pleased to express my sincere gratitude to all who supported the successful completion of my project as part of my 4 week industrial training in Data Analytics using Power BI and SQL at Netmax Technologies, Chandigarh.

I would like to extend my heartfelt thanks to Netmax Technologies Pvt Ltd**.** for providing the platform, resources, and guidance to work on practical projects such as the Adventure Works Sales Dashboard in Power BI and Restaurant Sales Analysis using SQL. The training, combining theoretical knowledge with hands-on sessions, greatly enhanced my understanding of data analytics, data visualization, and business intelligence concepts.

I am especially grateful to my trainers and mentors Shubham Sir at Netmax Technologies for their constant support, valuable insights, and constructive feedback, which helped me develop interactive dashboards, perform data cleaning, and extract meaningful insights from raw datasets.

I also thank my academic institution and faculty members for their encouragement, my peers for their collaboration and discussions, and my family and friends for their continuous support throughout this training.

Finally, I acknowledge the contributions of open-source communities and documentation platforms, whose resources made this learning experience and project work possible.

About The Company

Netmax Technologies Pvt. Ltd., located in Chandigarh, India, is a well-established industrial training institute and IT company offering a wide range of technical programs for engineering and computer science students. Netmax provides specialized courses in Data Analytics, Power BI, SQL, Machine Learning, Python Programming, Artificial Intelligence, Data Science, Web Development, Networking, and Embedded Systems.

The 30-day Data Analytics training program at Netmax is designed to bridge the gap between theoretical knowledge and real-world applications. During the training, students gain practical experience in data extraction, cleaning, visualization, and reporting.

A key strength of Netmax is its project-based approach, allowing students to work on real-time applications such as interactive dashboards, sales analyses, and data-driven reporting. The institute also emphasizes critical thinking, problem-solving, and self-learning to prepare students for future roles in the tech industry.

Additionally, Netmax supports students through career counselling, resume building, interview preparation, and placement assistance. Its training certifications are widely recognized and have helped many students secure roles in startups and established IT companies.

Overall, my experience with Netmax Technologies was highly enriching and industry-relevant. The training enhanced my skills in Power BI, SQL, and data analytics, and enabled me to successfully implement practical projects like the Adventure Works Sales Dashboard and Restaurant Sales Analysis

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEFINATIONS

Data Analytics : The process of systematically analyze datasets to draw conclusions, identify trends, and support decision-making through statistical and computational techniques.

Data Science: A multidisciplinary field that uses scientific methods, algorithms, and systems to extract knowledge and insights from structured and unstructured data.

Business Intelligence (BI): A set of technologies and processes used to collect, store, and analyze data to support strategic business decisions.

Power BI: A Microsoft-developed business analytics tool that allows users to visualize data, share insights, and create interactive dashboards and reports.

SQL: Structured Query Language – a standard language used to communicate with and manage data stored in relational databases.

Database: An organized collection of structured data that can be easily accessed, managed, and updated.

Dataset: A collection of data points organized in a tabular or structured format, used for analysis and visualization.

ETL: Extract, Transform, Load – the process of extracting raw data from various sources, transforming it into a clean, usable format, and loading it into a database or visualization tool.

Data Cleaning: The process of identifying and correcting inaccuracies, missing values, and inconsistencies in data to improve quality and reliability.

Data Transformation: The process of converting raw data into a structured format suitable for analysis, often involving aggregation, normalization, and filtering.

Data Modeling: The process of creating relationships between data tables and defining logical structures for efficient analysis in tools like Power BI.

Data Visualization: The graphical representation of data through charts, graphs, and dashboards to simplify complex information and highlight insights.

Dashboard: An interactive visual display that presents key performance indicators (KPIs), metrics, and data summaries for quick analysis.

KPI: Key Performance Indicator – a quantifiable measure that indicates how effectively a company is achieving key business objectives.

DAX: Data Analysis Expressions – a formula and expression language used in Power BI for custom calculations and data modelling.

Power Query Editor: A built-in tool in Power BI used for importing, cleaning, transforming, and preparing data before visualization.

SQL Query: A command written in SQL syntax used to retrieve, insert, update, or delete data in a database.

Database Management System (DBMS): Software used to create, manage, and manipulate databases efficiently. Examples include MySQL, SQL Server, Oracle, and PostgreSQL.

Relational Database Management System (RDBMS): A type of DBMS that organizes data into tables connected by relationships.

Primary Key: A unique identifier for each record in a database table, ensuring data integrity.

Foreign Key: A field in one table that links to the primary key of another table, establishing a relationship between the two.

CSV: Comma-Separated Values – a plain text format for storing tabular data, often used as an input source for analytics tools.

Power BI Service: The cloud-based platform for publishing, sharing, and collaborating on Power BI reports and dashboards.

Power BI Desktop: The local application for creating, designing, and editing Power BI reports before publishing them online.

Data Pipeline: A series of data processing steps that move data from raw sources to analytics tools for reporting.

Filter: A feature used to display specific subsets of data in a dashboard

Slicer: An interactive filtering control in Power BI that allows users to segment data dynamically

Measure: A calculated field in Power BI used to aggregate data values dynamically.

Fact Table: A table that stores quantitative data for analysis, often linked to multiple dimension tables.

Dimension: A descriptive attribute or category in data used to analyze a fact table

ACRONYMS AND ABBREVIATIONS

AI: Artificial Intelligence

API: Application programme Interface

BI: Business Intelligence

CSV: Comma Separated Values

DAX: Data Analysis Expression

KPI: Key Performance Index

SQL: Structured Query Languages

UI: User Interface

PK: Primary Key

FK: Foreign Key

DBMS: Database Management System

RDBMS: Relational Database Management System

OLAP: Online Analytical Processing

CHAPTER 1

INTRODUCTION

1.1 Introduction to Data Analytics Training

The rapid advancement of technology and the exponential growth of data in every sector have given rise to a new discipline known as Data Analytics. It represents the convergence of mathematics, statistics, programming, and business intelligence — all directed toward one fundamental goal: extracting meaningful insights from raw data to support decision-making. Organizations today generate massive amounts of information through digital transactions, customer interactions, and online activities. Data analytics enables them to utilize this vast data effectively to improve performance, optimize operations, and gain competitive advantages.

The Data Analytics training undertaken during the industrial program at *Netmax Technologies, Chandigarh* aimed to equip trainees with both the conceptual foundation and practical experience required to work with real-world datasets. The training emphasized understanding the complete data analytics lifecycle — from data collection to visualization — while developing problem- solving, analytical, and technical skills essential for today's data-driven industries.

During the training, students were introduced to the importance of data as the new form of organizational capital. Companies across domains such as finance, healthcare, e-commerce, manufacturing, and education rely heavily on data analytics to understand customer behavior, detect trends, forecast outcomes, and make evidence-based decisions. Through a well-structured curriculum, the training helped participants comprehend how analytical thinking transforms raw, unorganized information into actionable insights.

The course began with an introduction to the fundamentals of data analytics, explaining various types of data — numerical, categorical, and textual — and the importance of data preprocessing.

Students learned about the challenges associated with raw data, such as missing values, duplicates, and inconsistencies, and how to clean and prepare data for analysis. Emphasis was placed on the fact that nearly 70–80% of the analytical process involves data cleaning and preparation, making it one of the most critical stages in any analytics project.

Once the basics were established, participants were trained to explore datasets using exploratory data analysis (EDA) techniques. EDA helps in understanding the structure, distribution, and relationships between different variables. By using visualization tools, learners could identify hidden trends, outliers, and correlations within datasets. The training encouraged an inquisitive approach — asking the right questions about data before attempting to build models or draw conclusions. This stage built the analytical mindset necessary for interpreting real-world data problems effectively.

A key focus of the training was on the use of modern analytical tools and software that are standard in the industry. The participants worked extensively with Python and its libraries such as Pandas, NumPy, Matplotlib, and Seaborn, which form the backbone of most data analytics workflows. Python's versatility allowed trainees to perform data manipulation, statistical analysis, and visualization seamlessly within a single environment.

Alongside Python, SQL/MySQL was used to handle structured databases, query large tables efficiently, and integrate multiple data sources. Students also learned to perform preliminary analysis in Microsoft Excel, utilizing pivot tables, charts, and formulas for summary statistics. To develop visualization and reporting skills, tools such as Power BI and Tableau were introduced, enabling participants to create professional dashboards and communicate analytical results effectively to non-technical audiences.

Another important aspect of the training was feature engineering and transformation, where raw data is converted into a more informative format suitable for analysis or modeling. Trainees

practiced deriving new features, normalizing data, and encoding categorical variables. Through hands-on sessions, they understood how these preprocessing steps directly influence the quality of insights obtained from the data.

Collaboration was an integral part of the 4-week Data Analytics training at Netmax Technologies Pvt. Ltd., Chandigarh. Trainees worked in teams to discuss analytical challenges, share ideas, and validate results obtained from SQL queries and Power BI dashboards. This collaborative environment closely reflected real-world industry practices, where teamwork and cross-functional coordination are essential for solving data-driven business problems. Regular mentor guidance and feedback sessions provided technical clarity and direction, helping participants refine their analytical approach and improve the accuracy of their insights.

Throughout the training period, the program followed a structured, progressive learning approach. Each week focused on a specific stage of the data analytics lifecycle — starting with

data collection and understanding, followed by data cleaning and transformation using SQL, then visualization and dashboard creation in Power BI. By the end of the training, participants were able to perform complete end-to-end data analysis independently.

The final phase of the training emphasized applying all acquired knowledge to practical business projects. Students developed and presented dashboards such as the Adventure Works Analysis Dashboard, Sales Analysis Dashboard, and Restaurant Analysis using SQL, integrating multiple datasets and highlighting key business performance metrics. These projects served as capstone exercises, combining theoretical understanding, data processing, and visualization into a single cohesive analysis.

Beyond technical proficiency, the training also focused on interpretation and communication of analytical results. Trainees learned that analytics is not just about numbers, but about presenting a meaningful story through data. Skills such as report writing, data storytelling, and creating impactful

visualizations using Power BI were emphasized. Participants were encouraged to document findings clearly and make logical, data-driven recommendations based on evidence and trends observed in their analysis.

Overall, the Data Analytics using Power BI and SQL training was a comprehensive and skill-enhancing experience. It strengthened both technical and analytical thinking capabilities while developing confidence to handle real-world data challenges. By the end of the program, students were proficient in extracting data using SQL, building professional dashboards in Power BI, and deriving actionable insights to support informed business decisions.**[1]**

1.2 Background of Data Analytics

In today's digital era, data has emerged as the most valuable resource. Organizations generate massive volumes of data daily through business transactions, customer interactions, supply chain operations, healthcare systems, and numerous other sources. However, raw data in itself does not hold value unless it is processed, analyzed, and converted into meaningful insights. This is precisely where Data Analytics plays a significant role.

Data Analytics refers to the systematic computational analysis of data to identify meaningful patterns, trends, and correlations. It involves applying statistical techniques, algorithms, and modern computational tools to extract actionable information from raw data.

Figure 1.1 Flow of Data Analytics

Historically, analytics began with manual calculations and simple tabulations. Over time, with

the rise of computers and business intelligence tools, analytics became more structured and sophisticated. Today, the availability of large-scale data, coupled with advancements in machine learning and artificial intelligence, has transformed data analytics into a cornerstone of digital innovation.

The importance of data analytics can be seen across multiple domains:

- **Healthcare**: Predicting disease outbreaks, recommending treatments.
- **Finance**: Fraud detection, risk assessment, portfolio optimization.
- **Retail and E-commerce**: Customer segmentation, personalized recommendations, inventory optimization.
- **Transportation**: Route optimization, predictive maintenance, demand forecasting.
- **Social Media and Marketing**: Sentiment analysis, campaign effectiveness measurement.

Thus, the scope of data analytics is immense, making it a highly relevant and in-demand skill in the present and future job markets**.[2]**

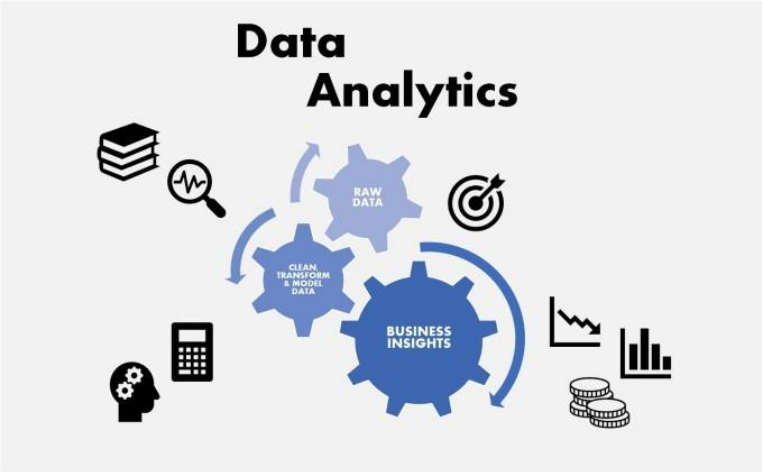

Figure 1.2: Transformation of Data into Insights through Analytics

1.3 Importance of Data Analytics in Industry

In modern enterprises, data is often called the “new oil” because of its transformative potential. With proper analysis, organizations can turn raw data into a strategic asset that drives innovation and competitiveness. Some of the key roles of data analytics in industries include:

1. **Data-driven Decision Making**: Instead of relying solely on intuition, managers and policymakers now use analytics-driven dashboards and reports to take evidence-based decisions.
2. **Trend Forecasting**: Predictive analytics enables businesses to forecast sales, customer demand, and market fluctuations.
3. **Operational Efficiency**: By identifying inefficiencies, companies can reduce costs, improve productivity, and optimize resource allocation.
4. **Customer Insights**: Analytics reveals customer preferences, buying behaviors, and satisfaction levels, which helps in designing better products and services.
5. **Risk Management**: In sectors like banking and insurance, analytics helps in fraud detection, risk assessment, and compliance monitoring.

The above applications highlight that analytics is not limited to IT companies but is equally important in every industry where data is generated. Therefore, undergoing industrial training in this field provides students with a competitive advantage in their career.

1.4 Objectives of the Training

The main purpose of this 30-day training program was to gain theoretical knowledge and practical exposure in the field of data analytics. The objectives of the training were:

- To understand the concepts and processes of data analytics.
- To learn the different types of analytics: Descriptive, Diagnostic, Predictive, and Prescriptive.
- To acquire hands-on skills in widely used analytics tools such as Python, SQL, and Excel.
- To practice data preprocessing, cleaning, and visualization techniques on real datasets.

1.5 Theoretical Explanation of Data Analytics Concepts

Data Analytics is not a single-step activity; rather, it is a structured process that involves multiple stages. The major components include:

1. **Data Collection** – Gathering raw data from databases, sensors, online platforms, or business transactions.
2. **Data Cleaning and Preprocessing** – Handling missing values, removing duplicates, normalizing formats, and ensuring quality.).
3. **Data Analysis** – Applying statistical models, algorithms, and programming techniques to identify insights.
4. **Data Visualization and Reporting** – Representing results in the form of dashboards, graphs, and summaries for decision-making.
5. **Data Transformation** – Converting data into meaningful structures (tables, charts, graphs) for analysis.

Figure 1.3 Components of Data Analytics

1.51 Types of Data Analytics:

Prescriptive Analytics: Suggests "what should be done."

Descriptive Analytics: Explains "what happened" in the past.

Diagnostic Analytics: Explains "why it happened."

Predictive Analytics: Predicts "what is likely to happen."

This theoretical foundation formed the basis for the practical training modules undertaken at Netmax Technology.

Table 1.1: Types of Data Analytics

Type of Analytics	Purpose	Example
Descriptive	Explains what happened	Monthly sales reports
Diagnostic	Explains why it happened	Drop in sales due to high prices
Predictive	Predicts what will happen	Forecasting next quarter's revenue
Prescriptive	Suggests what should be done	Recommending promotional strategies

1.6 Software Tools Learned

During the industrial training at Excellence Technology, Mohali, several essential tools and software platforms were learned and utilized. Each played a vital role in understanding, analyzing, and visualizing data effectively. The combination of these technologies provided a strong foundation in data analytics, enabling efficient data handling, visualization, and interpretation.

The following tools were integral to the training process:

1.61 Python Programming language

Python is one of the most widely used programming languages in data analytics and machine learning. Its simplicity, versatility, and vast library support make it an ideal choice for performing end-to-end data analysis tasks.

During the training, Python was used extensively for data cleaning, transformation, visualization, and basic statistical analysis. Several popular libraries were explored:

- **NumPy (Numerical Python):**

  NumPy provides efficient array and matrix operations. It supports mathematical logical operations on large datasets and is the foundation for most data manipulation tasks in Python. It was used for operations such as handling missing data, performing aggregations, and conducting mathematical computations.

- **Pandas (Python Data Analysis Library):**

  Pandas was the core library used for importing, cleaning, and transforming datasets. It provides high-level data structures like DataFrames, which simplify data manipulation tasks such as filtering, grouping, merging, and reshaping. The

Titanic dataset project heavily relied on Pandas for exploratory data analysis and preprocessing.

- **Matplotlib:**

  Matplotlib is a powerful data visualization library used to create a variety of static, animated, and interactive plots. It was employed to draw bar charts, histograms, and line graphs, which helped to visualize relationships and trends in the dataset clearly.

- **Seaborn:**

  Seaborn is built on top of Matplotlib and provides a more attractive, high-level interface for drawing informative and aesthetic statistical graphics. It was primarily used to create correlation heatmaps, violin plots, and boxplots for the Titanic dataset.

1.62 Power Bi / Tableau

To enhance the visualization and presentation of analytical results, Power BI and Tableau were introduced during the training as professional business intelligence tools. These tools are used for dashboard creation, data visualization, and interactive reporting.

Key learnings included:

- Connecting datasets from Excel, CSV, or SQL databases.
- Cleaning and transforming data using built-in tools.
- Designing interactive dashboards with filters, slicers, and dynamic visuals.
- Using charts, maps, and KPI indicators to summarize data effectively.

These tools provided hands-on experience in presenting analytical results in a visually engaging and professional manner, which is crucial for decision-making in real-world business scenarios. Power BI and Tableau bridged the gap between technical analysis and business reporting, demonstrating how data analytics supports actionable insights.

1.63 SQL (Structured Query Language)

During the training, SQL (Structured Query Language) was introduced as the fundamental tool for data extraction, manipulation, and analysis. It served as the backbone for preparing datasets used in Power BI dashboards and analytical reports. SQL enabled participants to interact directly with relational databases and derive meaningful insights through queries.

Key learnings included:

• Understanding database concepts such as tables, relationships, primary and foreign keys.
• Using SELECT, WHERE, and ORDER BY statements to retrieve and filter data efficiently.
• Applying JOIN operations to combine data from multiple tables for comprehensive analysis.
• Using GROUP BY and aggregate functions (SUM, AVG, COUNT, MAX, MIN) for
• Implementing subqueries, views, and conditions to simplify complex analytical tasks.
• Performing data cleaning and transformation using SQL commands to prepare data for visualization.

SQL played a crucial role in building the foundation of data analytics by allowing participants to handle real-world datasets such as those used in the Sales Analysis and Restaurant Analysis projects. Through hands-on practice, trainees learned how to extract key performance indicators (KPIs), and understand how structured data supports data-driven decision-making.

This component of the training strengthened the logical and analytical thinking required for querying, managing, and preparing data — a skill essential for any data analyst working in professional environments.

CHAPTER 2

TRAINING WORK UNDERTAKEN

2.1 Overview of Industrial Training

As a part of the academic requirement and professional development initiative, I undertook a 30-day industrial training program at Netmax Technologies Pvt. Ltd., a renowned IT training and development company based in Chandigarh, India. This training aimed to provide a structured and immersive experience into the field of Data Analytics using Power BI and SQL one of the most in-demand domains in the current technological landscape.

The program was designed to serve as a bridge between academic learning and real-world applications. It provided exposure to industry expectations, domain-specific knowledge, and conceptual clarity in key areas related to data and machine intelligence. The experience not only deepened my understanding of machine learning fundamentals but also introduced me to the broader ecosystem of artificial intelligence and data science.

Figure 2.1 Netmax Technologies

2.2 About the Training Provider – Netmax Technologies

Netmax Technologies Pvt. Ltd. is one of the leading technology training institutes in North India, with a strong presence in the Chandigarh region. Established with the goal of equipping students and professionals with industry-relevant skills, Netmax has earned a reputation for delivering high-quality, practical, and industry-aligned training programs across various IT domains.

The institute is recognized for its project-based learning approach, highly qualified trainers, and an updated curriculum that reflects modern industry requirements. Over the years, Netmax has trained thousands of students, engineers, and working professionals, enabling them to become proficient in cutting-edge technologies. Training programs are offered in multiple formats, including short-term, long-term, online, and classroom-based sessions.

2.21 Training Duration:

- Total Duration: 30 Days
- Mode: Offline / Online (as applicable)
- Daily Sessions: 2–3 hours per day (Theory + Practical)

2.22 Objectives of the Training

- Build a strong foundation in data analytics concepts and business intelligence tools.
- Teach practical implementation using SQL and Power BI.
- Develop skills to analyze real-world datasets and generate insights.
- Prepare participants for end-to-end data analytics projects, from data extraction to dashboard reporting.

2.23 Skills and Tools Acquired During Training

Table 2.1: Skills and Tools Used

TOOLS AND TECHNOLOGIES	PURPOSE
SQL	Querying and managing relational databases; extracting, filtering, and aggregating data.
Power BI	Creating interactive dashboards, visualizations, and business intelligence reports.
Excel	Data entry, preliminary analysis, and working with CSV datasets.
Power Query Editor	Cleaning, transforming, and preparing data for analysis in Power BI.
DAX (Data Analysis Expressions)	Writing formulas and calculations for measures, KPIs, and aggregations in Power BI.
Tables, Relationships, Keys	Structuring data models, defining primary and foreign keys, creating fact and dimension tables.
Charts & Visuals (Bar, Pie, Line, Map, KPI Cards)	Representing data graphically to highlight trends, patterns, and insights.
Slicers & Filters	Enabling interactive dashboards with dynamic user control
Data Cleaning & Transformation Techniques	Handling missing values, duplicates, formatting errors, and creating derived fields.
Version Control / GitHub	Managing project files, collaborative work,

	and version tracking (if applicable).
Data Analysis Skills	Understanding trends, correlations,

2.3 Key Learning Outcomes

2.31 Technical Skills: □

- Developed expertise in Power BI for creating interactive dashboards, reports, and visualizations.
- Learned data cleaning and transformation techniques to prepare accurate and reliable datasets.

2.32 **Analytical & Problem-Solving Skills:**

- Developed the ability to analyze large datasets, identify trends, patterns, and anomalies.
- Learned to derive actionable insights that support decision-making in business scenarios.
- Practiced performing end-to-end analytics workflows, from data extraction to visualization and reporting.

2.33 **Data Visualization & Reporting Skills:**

- Learned to design professional dashboards with charts, maps, KPIs, slicers, and interactive visuals.
- Gained the ability to present data stories effectively, using visuals to convey key business insights.
- Developed skills to document and report findings, providing logical narratives for

stakeholders.

2.34 Business Intelligence & Decision-Making:

- Understood the role of data analytics in business intelligence and decision-making processes.
- Learned to track key performance indicators (KPIs) and evaluate business performance

2.4 Power BI: Tools, Functions, and Applications:

Power BI, developed by Microsoft, is a powerful business intelligence and data visualization tool used for analyzing data and presenting insights through interactive dashboards and reports. During the 4-week training, Power BI was extensively used to transform raw datasets into meaningful visual stories, combining SQL-based data extraction with advanced visualization techniques**.**

Figure 2.2 Power BI Introduction

2.41 Key Features and Functions of Power BI:

Power BI offers several features that make it a professional tool for data analytics:

Data Connectivity

- Ability to connect to multiple data sources: SQL Server, Excel, CSV, Web APIs, and online

datasets.

- Supports both **direct query** and **import mode** for data handling

Data Cleaning and Transformation (Power Query)

- **Remove duplicates and errors** from datasets.
- **Split, merge, and pivot** columns for structured analysis.
- Change **data types** and format numeric/date fields.
- Create **calculated columns** for additional metrics.

Data Modeling

- Establish **relationships between tables** (one-to-one, one-to-many).
- Use **fact tables and dimension tables** for analytical modeling.
- Define **hierarchies** for drill-down visualizations.

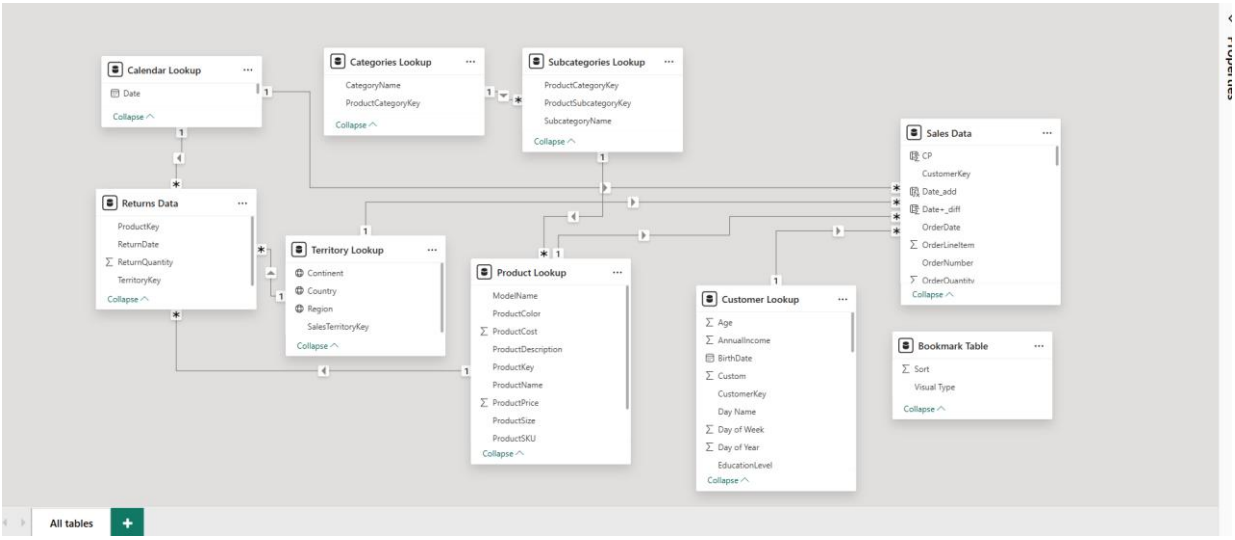

Figure 2.3 Data Modelling

DAX (Data Analysis Expressions)

- A formula language used to create **measures, calculated columns, and KPIs**.
- Common DAX functions used in training:

- SUM (), AVERAGE (), COUNT **()** – Aggregate numeric data.
- CALCULATE **()** – Compute conditional aggregations.
- IF (), SWITCH () – Conditional logic in measures.
- DATEADD (), YEAR (), MONTH () – Time intelligence for trend analysis.
- RELATED () – Retrieve values from related tables.

Visualizations and Graphs

- **Bar and Column Charts:** Compare categories over time or regions.
- **Line Charts:** Trend analysis over periods (e.g., sales, revenue).
- **Pie Charts / Donut Charts:** Show proportions or percentage contributions.
- **Area Charts:** Visualize cumulative trends or totals.
- **Tree Maps:** Represent hierarchical data, such as product categories.
- **KPI Cards:** Display key metrics like total sales, revenue, or profit margins.
- **Maps:** Geographic visualization of regional data.
- **Slicers:** Interactive filters for selecting data subsets dynamically.
- **Drill-through & Drill-down:** Navigate from summarized metrics to detailed records.

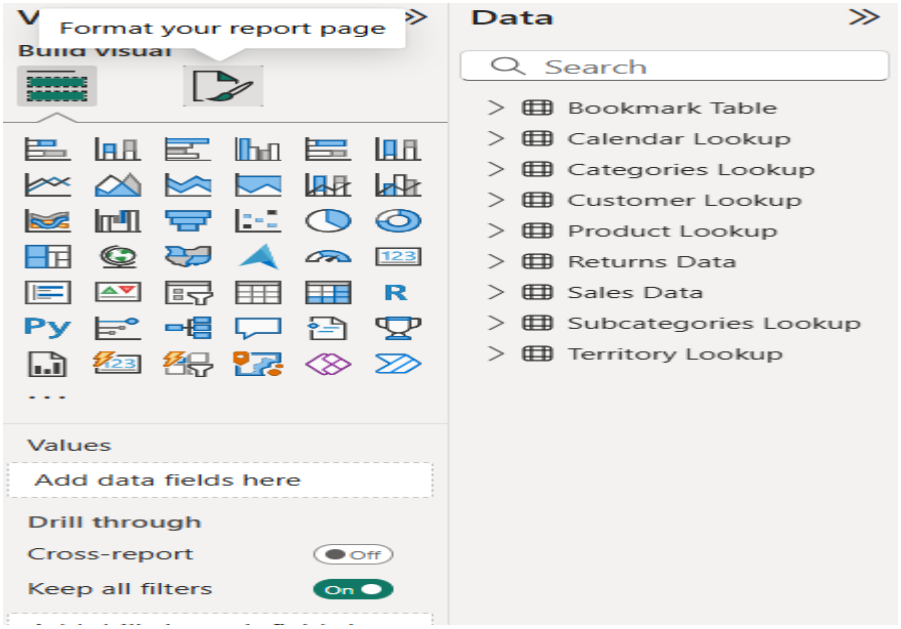

Figure 2.4 Power Bi Visualization

Reports and Dashboards

- Combine multiple visuals into **interactive dashboards** for business insights.
- Create **storytelling reports** that highlight patterns, outliers, and KPIs.
- Use **bookmarks and navigation buttons** for guided report exploration.

Sharing and Collaboration

- Publish dashboards to **Power BI Service (cloud platform)** for team collaboration.
- Set up **scheduled refresh** for live datasets.
- Export visuals and reports in **PDF, PPT, or Excel** for stakeholders.

2.5 Real-World Applications of Power BI:

Table 2.2 Applications of power BI

Power BI Feature	Description / Function	Real Application

Data Connectivity	Connect to multiple sources: SQL Server, Excel, CSV, APIs	Retail companies integrating sales, inventory, and customer data for holistic reporting
Data Modelling & Relationships	Define relationships between tables, create hierarchies	E-commerce platforms linking customer, order, and product tables for detailed analytics
DAX Functions	Aggregation (SUM, AVG), Conditional (IF, SWITCH), Time Intelligence (DATEADD, TOTALYTD)	Calculating KPIs, yearly revenue growth, and product performance metrics in sales analytics
Real-Time Analytics	Streaming data visuals, refreshable dashboards	IoT devices monitoring real-time metrics (e.g., sensor data, stock levels)
Reports & Dashboards	Multi-page interactive dashboards combining visuals	Executives monitoring business health, revenue, and KPIs in real-time

2.6 Structured Query Language (SQL)

SQL (Structured Query Language) is a standard programming language used for managing and manipulating relational databases. During the 4-week training, SQL was a core focus, allowing participants to interact with datasets, perform analysis, and extract actionable insights.

2.6.1 Types of SQL Commands:

SQL commands are broadly categorized into Data Definition, Data Manipulation, Data Querying, Data Control, and Transaction Control.

1) **Data Definition Language (DDL):**

- Purpose: Define and manage database structures.
- Common Commands:
    - CREATE DATABASE – Create a new database.
    - CREATE TABLE – Define a new table with columns and data types.
    - ALTER TABLE – Modify table structure
    - DROP TABLE – Delete an existing table.
    - Example:

```sql
CREATE TABLE Orders (
    OrderID INT PRIMARY KEY,
    CustomerID INT,
    OrderDate DATE,
    TotalAmount DECIMAL(10,2)
);
```

Figure 2.5 DDL SQL Query

2) **Data Manipulation Language (DML):**

Purpose: Insert, update, or delete data in tables.

- **Common Commands:**
    - INSERT INTO – Add new records.

- UPDATE – Modify existing records.
- DELETE – Remove records

```sql
INSERT INTO Orders (OrderID, CustomerID, OrderDate, TotalAmount)
VALUES (1, 101, '2025-10-01', 500.00);
```

Figure 2.6 DML SQL Query

. **3)Data Querying / Retrieval (DQL):**

- Purpose: Extract and analyze data.
- Common Commands:
  - SELECT – Retrieve data from one or more tables.
  - WHERE – Filter rows based on conditions.
  - ORDER BY – Sort data.
  - DISTINCT – Remove duplicate records.
- Example:

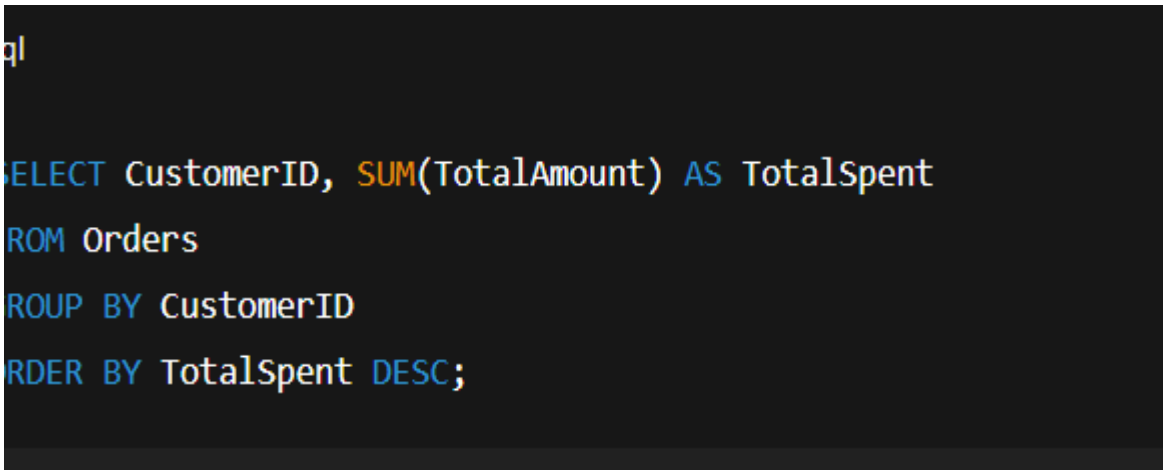

Figure2.7 DQL SQL Query

4). **Aggregate Functions and Grouping:**

- Purpose: Summarize data for analysis.
- Functions:
    - SUM () , AVG (), COUNT (), MIN (), MAX ()
- Grouping: GROUP BY to aggregate based on a column.
- Example:

```sql
SELECT MenuItem, COUNT(*) AS OrderCount
FROM Orders
GROUP BY MenuItem
HAVING COUNT(*) > 50;
```

Figure 2.8 Aggregate Functions

5). **Joins and Relationships**

- Purpose: Combine data from multiple tables for detailed insights.
- Types of Joins:
    - INNER JOIN – Only matching records from both tables.
    - LEFT JOIN – All records from left table + matching records from right table.
    - RIGHT JOIN – All records from right table + matching from left table.
    - FULL OUTER JOIN – All records from both tables.
- Example:

```sql
SELECT o.OrderID, c.CustomerName, o.TotalAmount
FROM Orders o
INNER JOIN Customers c ON o.CustomerID = c.CustomerID;
```

figure 2.9 Joins

6. Subqueries / Nested Queries

- **Purpose:** Perform complex queries inside other queries.
- **Example:** Find customers who spent more than average:

```sql
SELECT CustomerID, TotalAmount
FROM Orders
WHERE TotalAmount > (SELECT AVG(TotalAmount) FROM Orders);
```

Fiureg 2.10 Subqueries

7. Constraints

- Purpose: Ensure data integrity and prevent errors.
- Common Constraints:
    - PRIMARY KEY, FOREIGN KEY, UNIQUE, NOT NULL, CHECK, DEFAULT

8. Views

- Purpose: Create reusable query results as virtual tables.

2.7 Applications of SQL in Training Projects:

1. **Restaurant Database Analysis**

- Extracted **most popular dishes**, **total revenue by category**, and **peak hours**.
- Used **aggregations, joins, and subqueries** for detailed reporting.

2. **Sales Analysis**
   - Calculated **monthly and yearly sales trends**, revenue, and order counts.
   - Joined multiple tables to analyze customer behavior and product performance.
3. **End-to-End Workflow**
   - Designed database schema (tables, keys, relationships).
   - Inserted and cleaned raw data.
   - Queried, aggregated, and prepared data for Power BI dashboards.

CHAPTER 3

RESULT AND DISCUSSION

3.1 Overview of Project Outcomes

The main objective of the Adventure Works Analysis project was to develop a comprehensive sales and business intelligence dashboard by integrating multiple tables from the Adventure Works database. The project aimed to extract meaningful insights regarding sales performance, product categories, regional trends.

After the implementation and testing phases, the project successfully achieved its goals. Key outcomes of the project include:

- Integration of Multiple Tables: Data from various tables such as Calendar Table, Category Table, Product Lookup Table, Returns Data , Sales Data , subcategory Lookup, Territory Lookup were joined and transformed using SQL queries. This enabled a unified dataset suitable for analytics and visualization.
- Data Cleaning and Preparation: Missing values, duplicates, and inconsistencies in order and product data were identified and corrected. Derived metrics such as total sales, profit, and quantity sold were calculated using SQL and DAX.
- Interactive Dashboards in Power BI: The processed data was used to create interactive dashboards, which included:
    - Regional Sales Analysis: Comparing revenue across different geographical regions.
    - Product Performance: Identifying top-performing products and product categories..

- Trend Analysis: Monthly and yearly sales trends using line charts and KPI cards.

- Analytical Insights: The dashboards provided actionable insights for business decision-making, such as identifying high-revenue products, seasonal trends, and areas requiring attention for sales improvement.

- Dynamic and Interactive Features: Users could filter dashboards by region, product category, or time period, enabling a flexible and user-driven exploration of sales data.

Overall, the project demonstrated the practical application of SQL for data extraction and cleaning, and Power BI for visualization and business intelligence, allowing participants to gain hands-on experience in end-to-end data analytics workflows**.[4]**

3.2 Understanding the Dataset:

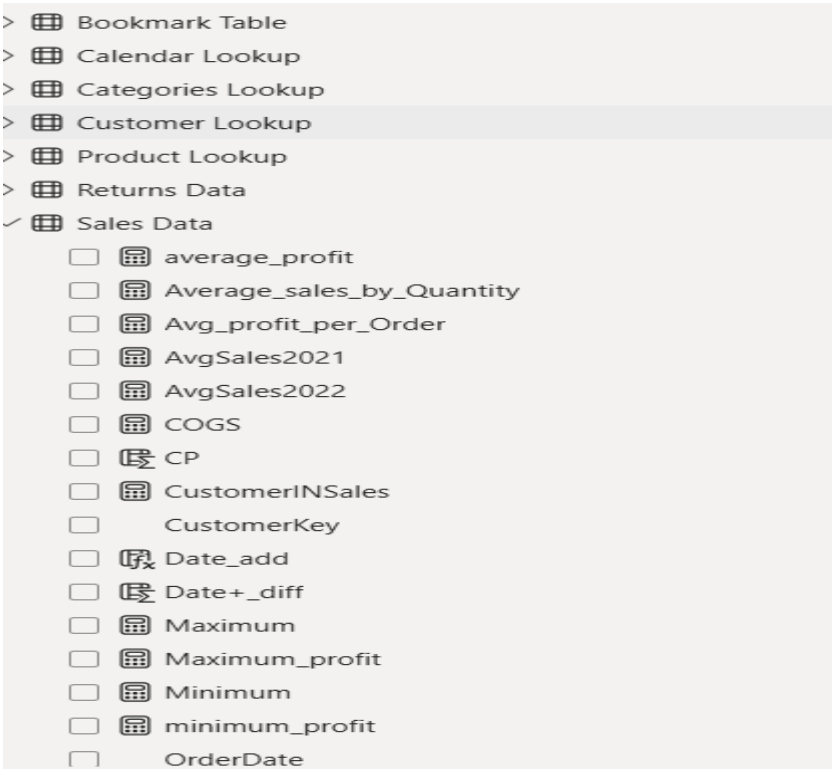

Figure 3.1 Dataset

This Dataset is being provided by the Institute itself as it has the several tables in it. With which we are used to do the cleaning, modelling, transformation and the data visualization.

Key tables used in the project:

a. **Calendar Table:** It Give us the data and the information about the Year, Month, Day, Day name
b. **Categories Lookup**: It tells about the types of Categories, and type of Product Categories in the table.
c. **Customer Lookup**: it is the table that contain the all the information about the customer, Name (Full Name, Last Name), Salary, Sales, Occupation Martial Status, Gender, Home Owner, Total Children.
d. **Product Lookup**: It tells about the product features, style, Colour, Key, Description, Subcategory etc.
e. **Return Data**: It tells about the Return Data of the Product.
f. **Sales Data:** it is one of the important Table among all of the table as this is the backbone for the Power Bi Visualization contains the all of the measures that are responsible for the Transformation and the Data modelling of the Data.
g. **Subcategories Lookup**: Defines the Behaviour of the Subcategories Product.
h. **Territory Lookup:** it Define the Geographical Location, and helpful in the Data Analysis.

3.3 Data Extraction Using SQL:

he Adventure Works database is a sample business database provided by Microsoft that simulates the operations of a manufacturing company named *Adventure Works Cycles*, which produces and sells bicycles, components, and accessories.

Before writing SQL queries, it was essential to understand:

- The schema structure (Sales, Production, Person, Purchasing, etc.)
- Table relationships (Primary and Foreign Keys)
- The type of data each table holds (textual, numeric, date/time)

- Business entities like Customer, Product, Sales Order, and Territory

: Combining Product and Category Information

```sql
SELECT
    p.ProductID,
    p.Name AS ProductName,
    pc.Name AS Category,
    p.StandardCost,
    p.ListPrice
FROM Production.Product AS p
INNER JOIN Production.ProductCategory AS pc
    ON p.ProductCategoryID = pc.ProductCategoryID;
```

Figure 3.2 Combining Information

3.4 Data Cleaning and Transformation:

Transform Data refers to the process of cleaning, shaping, and preparing data before using it in your visualizations. This is done in the Power Query Editor, which opens when you click on Home > Transform Data in Power BI Desktop. Think of it like preparing ingredients before cooking — if your data is clean and organized, your reports will be faster, more accurate, and easier to build**.**

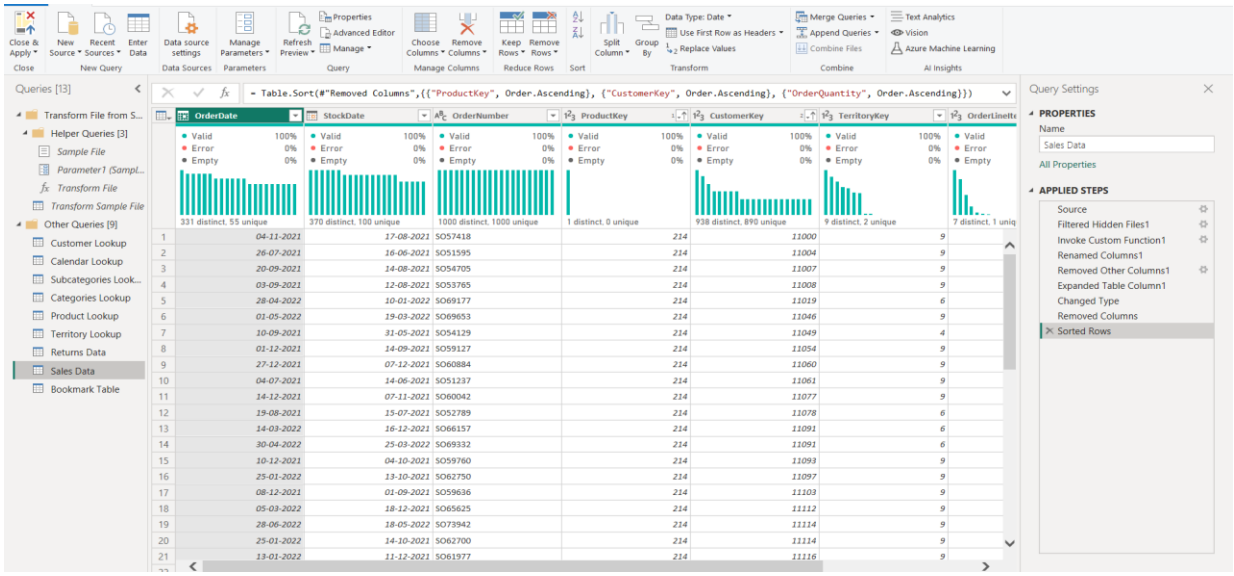

Figure 3.3 Transform Data

What Can You Do with Transform Data (Power Query)?

Here are some common data transformation tasks:

1. Remove Unnecessary Columns/Rows • Delete columns you don't need to make the model lighter and easier to use.

2. Filter Rows • Keep only the rows you want — for example, filter out blank values or specific categories.

3. Change Data Types • Set correct data types (text, number, date, etc.) so Power BI treats your data correctly.

4. Rename Columns • Give columns meaningful names to make your model more readable.

5. Split Columns • Break a single column into multiple ones (e.g., split "Full Name" into "First Name" and "Last Name").

6. Merge Columns • Combine multiple columns into one (e.g., combine "City" and "State").

7. Group By • Aggregate data by grouping — for example, group sales by region and calculate the total.

8. Replace Values • Replace wrong or missing values (e.g., change "N/A" to "0").

9. Pivot and Unpivot • Change the layout of data: o Pivot: Turn rows into columns. o Unpivot: Turn columns into rows (often used to normalize data).

10. Remove Duplicates • Keep only unique values from a column or table.

11. Add Custom Columns • Use formulas to create new columns (Power Query uses M language for this, different from DAX).

12. Append Queries • Combine two or more tables with the same columns (e.g., monthly sales files).

13. Merge Queries • Join two tables (like a VLOOKUP) to bring in columns from another table.

3.5 Data Modelling in Power BI:

What is Data Modelling in Power BI? Data modelling is the process of:

• Importing data from various sources (Excel, SQL Server, SharePoint, etc.)

• Transforming and cleaning that data (using Power Query)

• Creating relationships between different tables

• Defining measures, calculated columns, and hierarchies using DAX (Data Analysis Expressions)

• Designing a structure that makes reporting easy, accurate, and efficient**[5]**

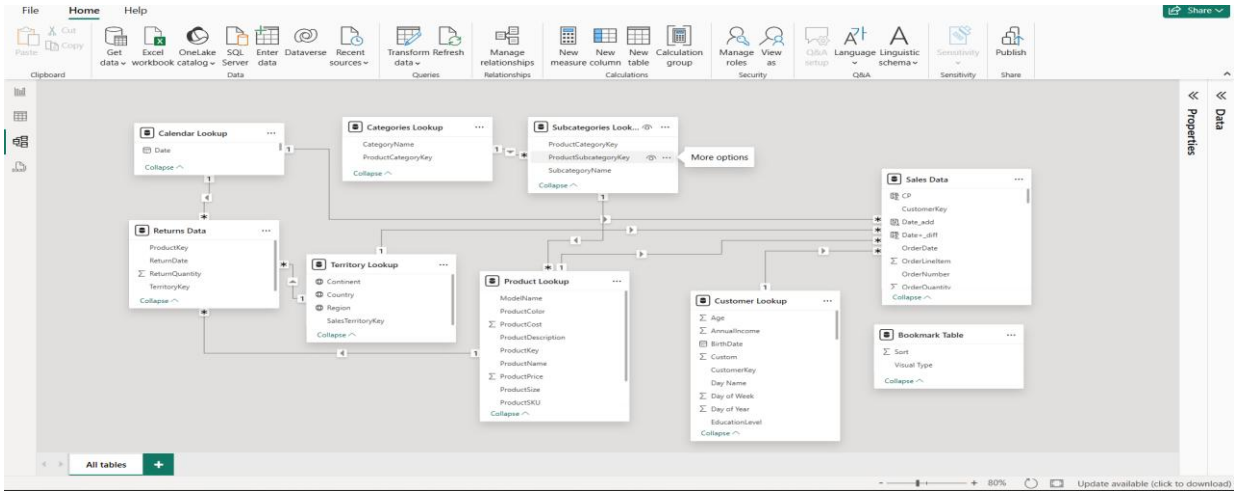

Figure 3.4 Data Modelling

Components of Data Modelling in Power BI :

1. Tables

• Raw data from sources

• Can be fact tables (e.g., Sales) or dimension tables (e.g., Date, Product)

2. Relationships

• Links between tables

• Usually defined using primary key and foreign key

• Can be:

- One-to-many (most common)
- Many-to-many
- One-to-one

• Helps Power BI understand how data connects (e.g., which customer bought which product)

3.Star Schema vs Snowflake Schema

• Star schema is preferred in Power BI (Fact table at center, surrounded by dimension tables)

• It improves performance and usability

4.Measures

- Calculations that are computed on the fly (dynamic)
- Very efficient and powerful • Example: Total Sales = SUM(Sales[Amount])

5.Hierarchies

• Logical grouping, like Year > Quarter > Month > Day

• Used for easy navigation and drill-down in visuals

3.6 VISUALIZATION IN POWER BI

Visualizations (or visuals) are graphical representations of your data, used to uncover trends, patterns, and insights.

Table 3.1 Visualization in Power BI

Visual Type	Use Cases
Bar/Column Chart	Compare values across categories
Line Chart	Show trends over time
Pie Chart/Donut Chart	Show proportions of a whole
Tree Map	Show hierarchical data as nested rectangles
Map	Geographical data visualization
Scatter Plot	Show progress toward a goal

Table/Matrix	Show performance against a target
KPI Indicator	Show relationship between two numerical values
Gauge	Show detailed data in rows and columns

3.61 Purpose of Visualizations

1. Simplify Complex Data Convert large tables into easily understood charts

2. Highlight Key Metrics Track performance (e.g., KPIs, targets, goals)

3. Detect Trends and Patterns Understand changes over time or across categories

4. Enable Interaction Let users drill down, filter, or slice data dynamically

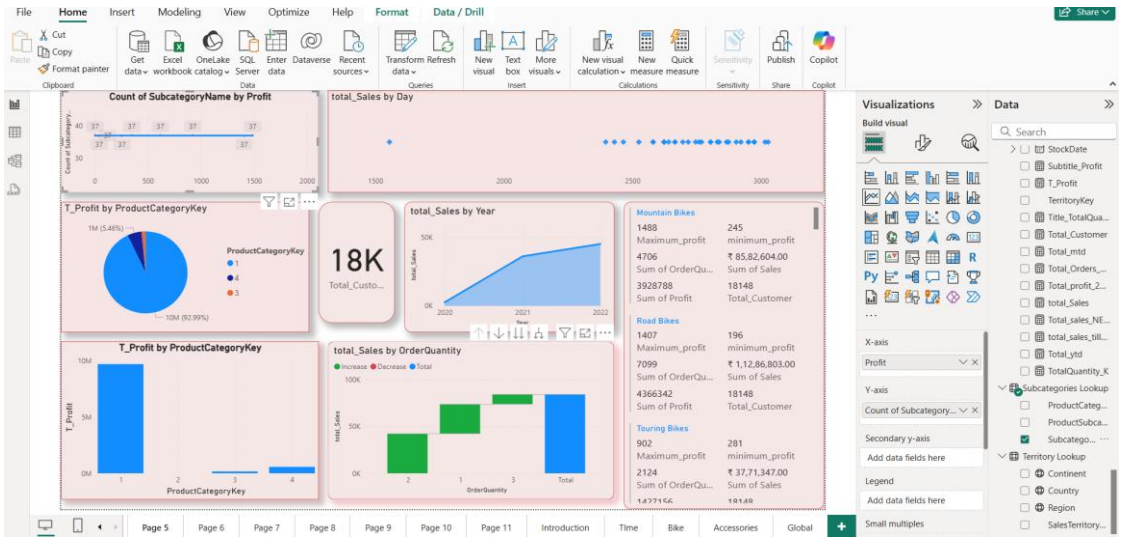

Figure 3.5 Visualization in Power BI

3.7 Analysis and Insights:

- Identified top-selling products and categories.
- Analyzed regional sales trends to highlight strong and weak markets.

- Detected seasonal patterns in sales.
- Derived customer insights, such as frequent buyers and high-value customers.
- Created a summary report to communicate actionable insights for decision-making.

figure 3.6 Introduction to project

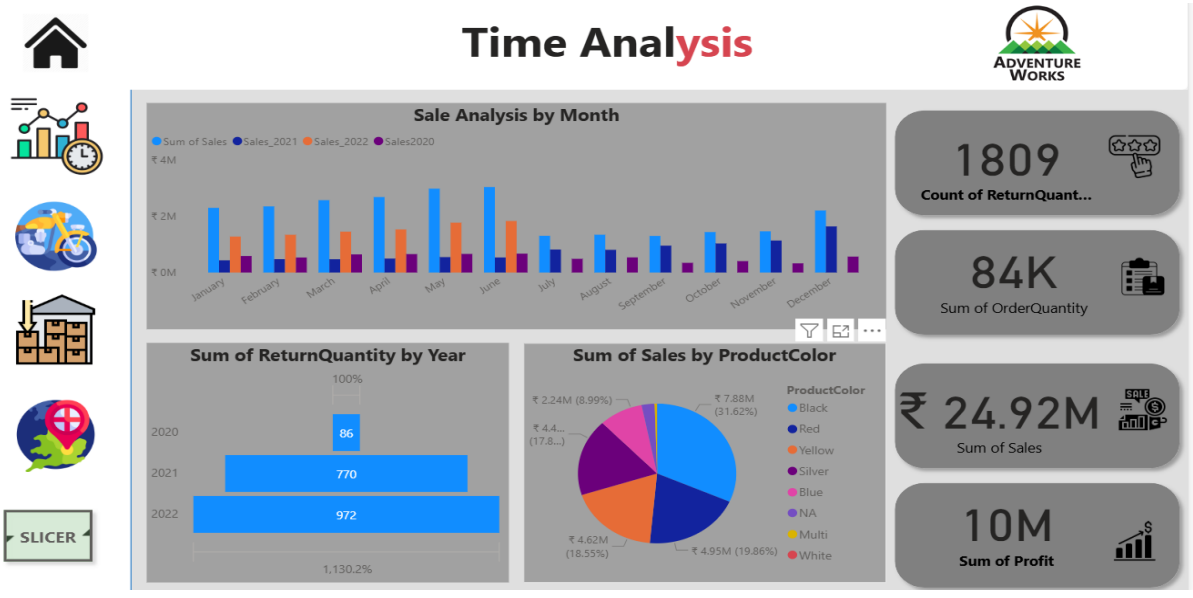

figure 3.7 Time Analysis of Project

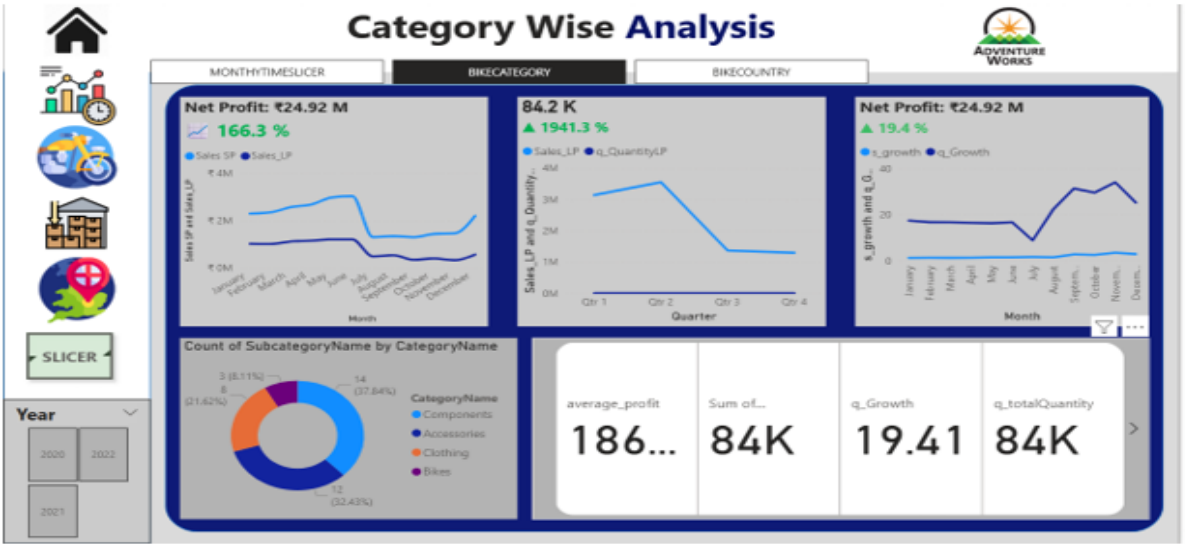

Figure 3.8 Category Wise Analysis

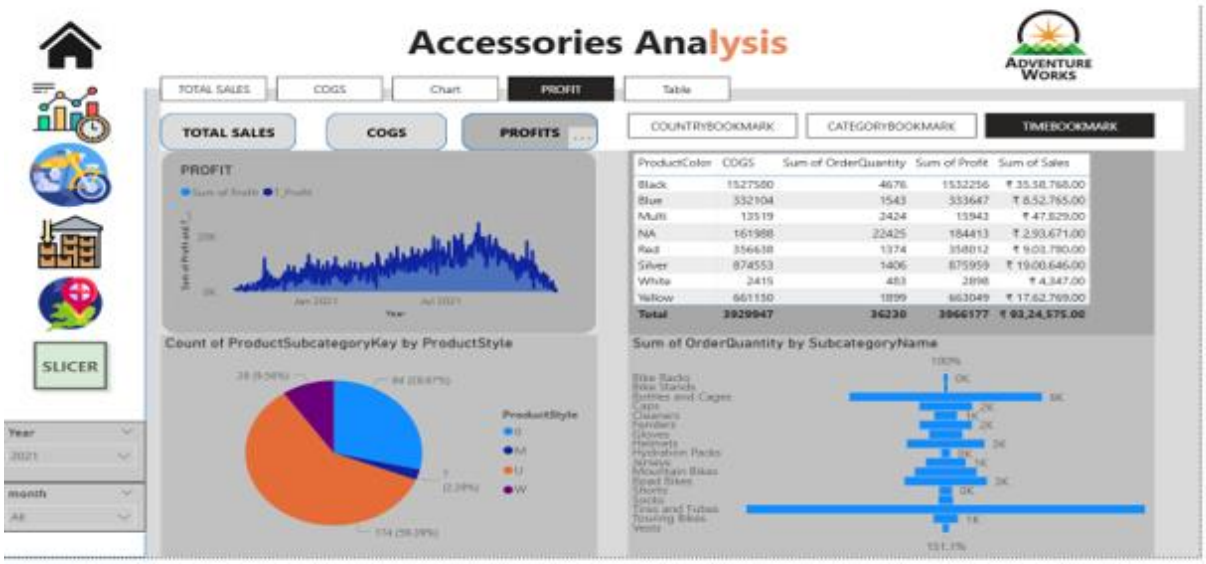

Figure 3.9 Accessories Analysis

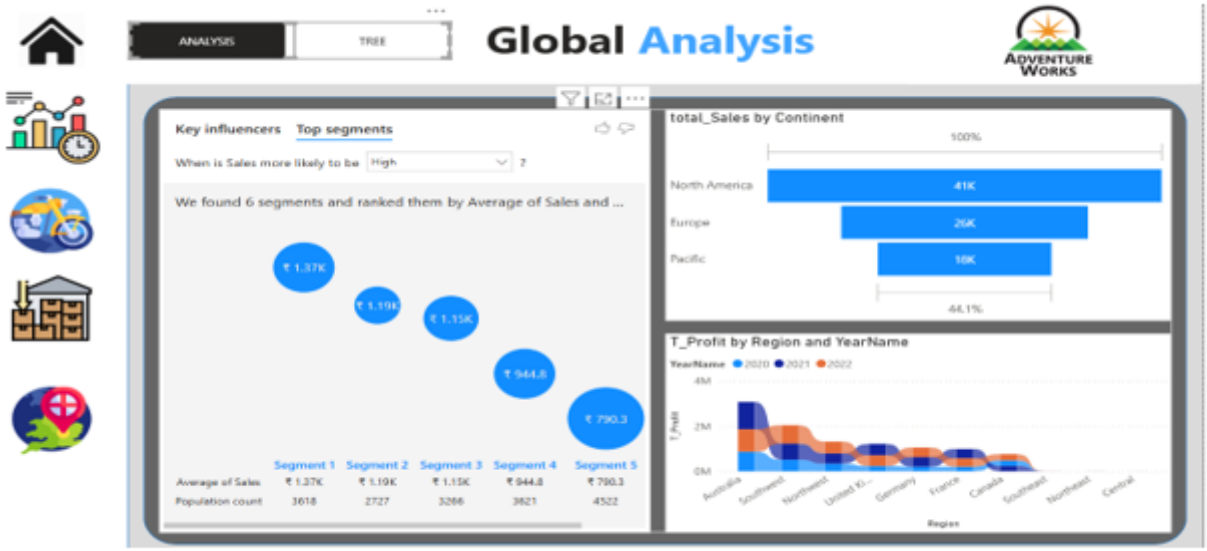

Figure 3.10 Global Analysis

CHAPTER 4

CONCLUSION AND FUTURE SCOPE

4.1 Conclusion

The Adventure Works Analysis Project provided an in-depth understanding of how data analytics can be applied to derive meaningful business insights from raw data. Throughout the project, SQL and Power BI were effectively used to extract, transform, analyze, and visualize data from the Adventure Works database — a simulated enterprise dataset representing a global manufacturing company.

The project began with data extraction and cleaning using SQL, where complex queries were written to join multiple relational tables and create a consolidated dataset. This step helped in understanding data modeling, relationships, and normalization in relational databases.

Through this project, strong technical and analytical skills were developed:

- Proficiency in writing optimized SQL queries for data extraction.
- Understanding data relationships and schema design.
- Hands-on experience in Power BI dashboard creation using DAX functions, slicers, and KPIs.

Overall, the project successfully demonstrated the end-to-end workflow of data analytics — from data sourcing and processing to visualization and insight generation. It bridged the gap between theoretical learning and practical application, aligning with industry-oriented problem-solving.

4.2 FUTURE SCOPE

While the current analysis focused primarily on sales performance and customer insights, there are several potential directions in which the project can be expanded:

Table 4.1 Future Scope

Area	Future Enhancement / Scope
Advanced Analytics	Integrate predictive modelling and forecasting using tools like Python or Power BI AI visuals to predict future sales trends.
Data Integration	Connect real-time or external data sources (e.g., Excel, APIs, or CRM tools) for more dynamic dashboards and live analytics
Automation	Automate data refresh and reporting processes using Power BI Service and SQL Scheduled Jobs.
Enhanced Visualization	Add advanced visuals such as decomposition trees, heatmaps, or drill-through reports for deeper insights.
Performance Optimization	Apply SQL query optimization techniques to handle larger datasets efficiently.
Machine Learning Integration	Extend the project to include customer segmentation, sales prediction, or recommendation systems using ML algorithms.
Deployment and sharing	Publish dashboards on the Power BI Cloud Service and create role-based access controls for business teams.

REFRENCES

Books

[1] W. McKinney, *Python for Data Analysis*, 2nd ed. Sebastopol, USA: O'Reilly Media, 2017.

[2] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd eda, Power BI. Sebastopol, USA: O'Reilly Media, 2019.

Conference Paper

[3] H. Chen, S. C. Laroiya, and M. Adithan, "Precision Machining of Advanced Ceramics," *International Conference on Advanced Manufacturing Technology (ICMAT-94)*, Johor Bahru, Malaysia, 1994, pp. 203–210.

Online Source

[4] https://www.microsoft.com/en-us/power-platform/products/power-bi.

[5] https://www.godaddy.com/en-in/help/what-is-microsoft-power-bi-42272

Periodic/Journals

[6]https://datascience.codata.org/

[7] R. E. Kalman, "New results in linear filtering and prediction theory," *Journal of SQL.*

[8] https://www.inderscience.com/

APPENDIX

Appendix A: EDA (Exploratory Data Analysis Code).

```
top_customers = df.groupby('CustomerKey')['OrderQuantity'].sum().nlargest(10)
plt.figure(figsize=(10,4))
top_customers.plot(kind='bar', color='orange')
plt.title("👤 Top 10 Customers by Order Quantity")
plt.xlabel("CustomerKey")
plt.ylabel("Total Quantity")
plt.show()
```

```
plt.figure(figsize=(8,5))
sns.scatterplot(x='ProductCost', y='ProductPrice', data=pdt, color='teal')
plt.title("Product Cost vs Product Price")
plt.xlabel("Product Cost")
plt.ylabel("Product Price")
plt.show()
```

```
plt.figure(figsize=(10,5))
sns.barplot(x='ProductName', y='Profit', data=top_profit, palette='viridis')
plt.title("Top 10 Profitable Products")
plt.xticks(rotation=45, ha='right')
plt.show()
```

Figure 4.1 Eda Code For Ananlysis

APPENDIX B: DAX FUNCTIONS

```
q_GrowthFormated = IF([q_Growth] >= 0,
    FORMAT([q_Growth], "▲ #0.0 %"), FORMAT([q_Growth], "▼ #0.0 %") )
```

```
CALCULATE(
    [q_totalQuantity],
    FILTER(
        'Sales Data',
        'Sales Data'[OrderDate] IN
            SELECTCOLUMNS(
                ADDCOLUMNS(
                    VALUES('Calendar Lookup'[Date]),
                    "PrevYearDate", DATEADD('Calendar Lookup'[Date], -1, YEAR)
                ),
                "Date", [PrevYearDate]
            )
    ))
```

Figure 4.2 Dax Functions