

Implementation

Implementation

-The implementation of the AdventureWorks Business Intelligence Project followed a structured, phase-based approach to ensure accuracy, consistency, and high-quality analytical outputs. The process included data preparation, modeling, exploration, visualization, and dashboard development. The steps below describe the full technical workflow used during implementation.

1. Data Acquisition & Importing

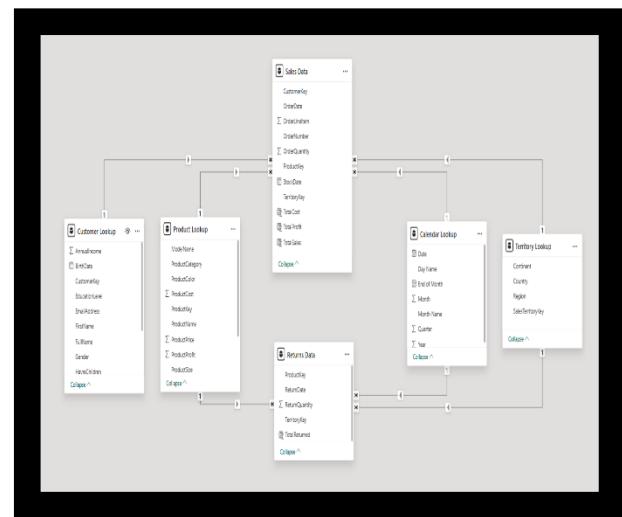
The first step was to collect and import all relevant tables from the **AdventureWorksDW** database.

- **Fact Tables:**

- FactInternetSales
- FactResellerSales

- **Dimension Tables:**

- DimProduct
- DimProductSubcategory
- DimProductCategory
- DimCustomer
- DimGeography
- DimDate



Schema which Identifies the
Tables relationships

Relationships of tables are already more explained within the Schema

2. Data Cleaning & Transformation

1. Structural Optimization

- Column Removal: Eliminated irrelevant/unused columns (Pritik, email, source columns) to reduce dataset size and complexity

- Column Management: Performed data type conversions and registered new columns while managing old ones

- Column Creation: Added new calculated columns for business metrics and analysis

2. Data Quality Enhancement

- Error Handling: Addressed missing values, duplicates, and common data errors

- Value Standardization:
 - Standardized gender field values

 - Replaced placeholder/underscored values

 - Handled unknown values in ProductSize & ProductStyle fields

- Format Consistency: Standardized date formats and numeric fields throughout the dataset

3. Text Data Processing

- Field Reorganization: Split, merged, and trimmed text fields as needed
- Name Field Standardization: Formatted and combined name columns for consistency
- Text Cleaning: Applied trimming and formatting to ensure text data uniformity

4. Business Logic & Calculations

- Metric Creation: Added calculated columns for product profit and other business metrics
- Data Validation: Verified numeric columns used in calculations (Sales Amount, Cost, Profit fields)
- Conditional Transformations: Applied business rules for handling specific data scenarios

5. Data Integration

- Calendar Table Preparation: Processed and standardized date dimension tables
- Automated Merging: Implemented folder merging automation for efficient data consolidation
- This ensured high data quality and prepared the dataset for modeling.

-Some changes which illustrated found [Here](#)

3. Merging Product Category & Subcategory Tables

To simplify reporting and improve analysis readiness:

- **DimProductCategory** and **DimProductSubcategory** were merged using ProductCategoryKey.
- The resulting table contained:
 - Product Category
 - Product Subcategory
 - Product Details

Screenshots provided clearly show:

- Category table before merging
- Subcategory table before merging
- Final product table with merged fields

This step streamlined product-related KPIs and ensured consistent category hierarchy.

4. Data Modeling

A relational star-schema model was designed to link fact and dimension tables. Using Power BI's *Model View*, relationships were created based on primary and foreign keys.

Key relationships included:

- **FactInternetSales → DimProduct** via ProductKey
- **DimProduct → DimProductSubcategory → DimProductCategory**
- **FactInternetSales → DimDate** via OrderDateKey
- **DimCustomer → DimGeography**

The final ER-style model ensured:

- One-to-many cardinalities in all directions
- Correct filtering context across dashboards
- Optimized performance for DAX calculations

5. Creating Calculated Columns & Measures (DAX)

To support KPIs and dashboards, several DAX expressions were created:

Examples:

- **Total Sales** = SUM(FactInternetSales[SalesAmount])
- **Total Cost** = SUM(FactInternetSales[TotalProductCost])
- **Total Profit** = [Total Sales] – [Total Cost]
- **Profit Margin %** = DIVIDE([Total Profit], [Total Sales])

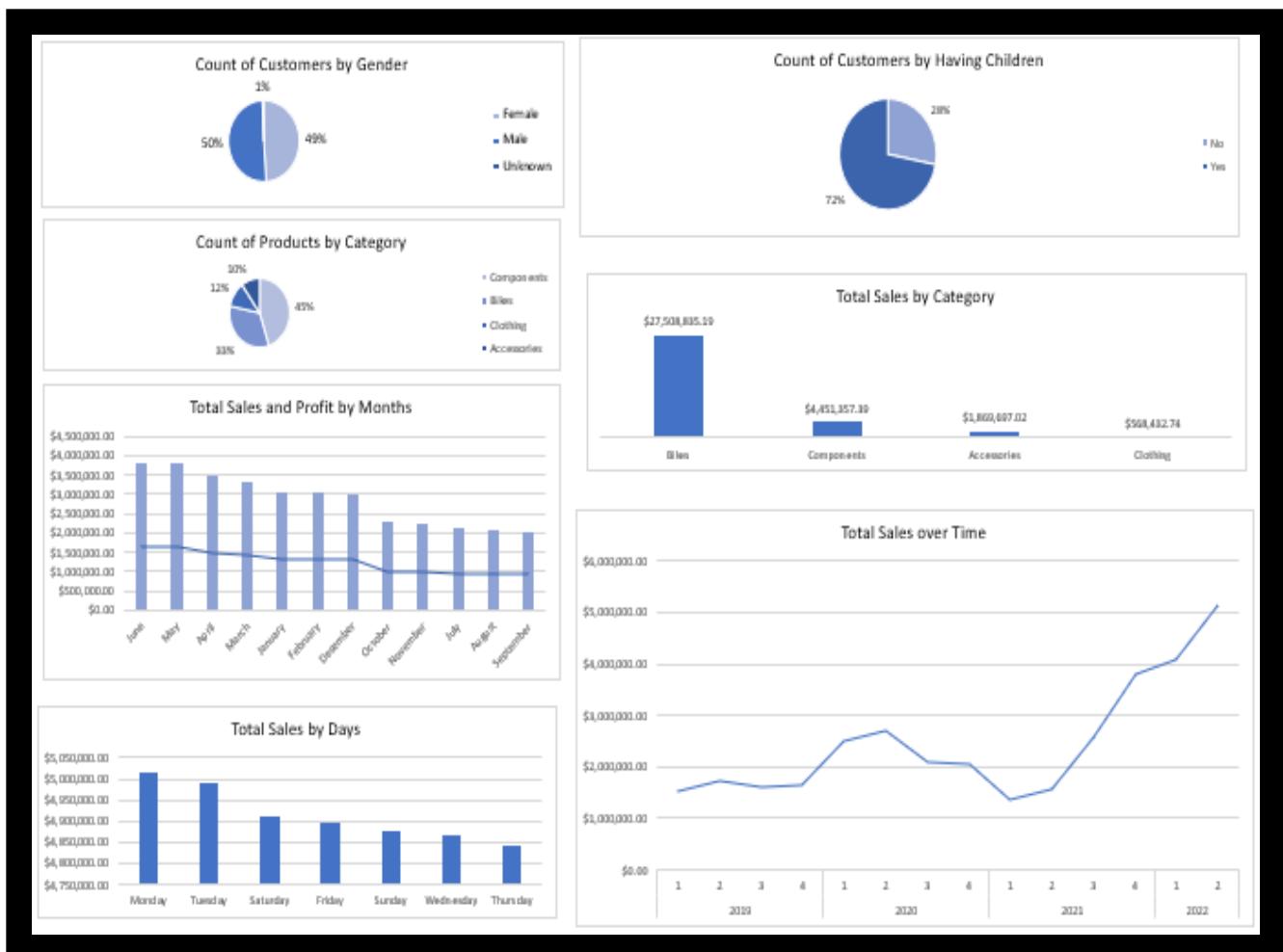
6-Exploratory Data Analysis (EDA)

Before building dashboards, EDA was conducted to understand key patterns:

- Sales distribution by country
- Top-selling customers
- Category-level profitability
- Quarterly sales fluctuations
- Return rates by product

This step helped guide KPI selection and dashboard layout decisions.

As Shown in Following Graphs , most insightful KPIs



7- Building the Interactive Dashboard

Using Power BI Desktop, the visual dashboard was designed based on UI/UX best practices.

The dashboard includes:

Top Metrics Cards

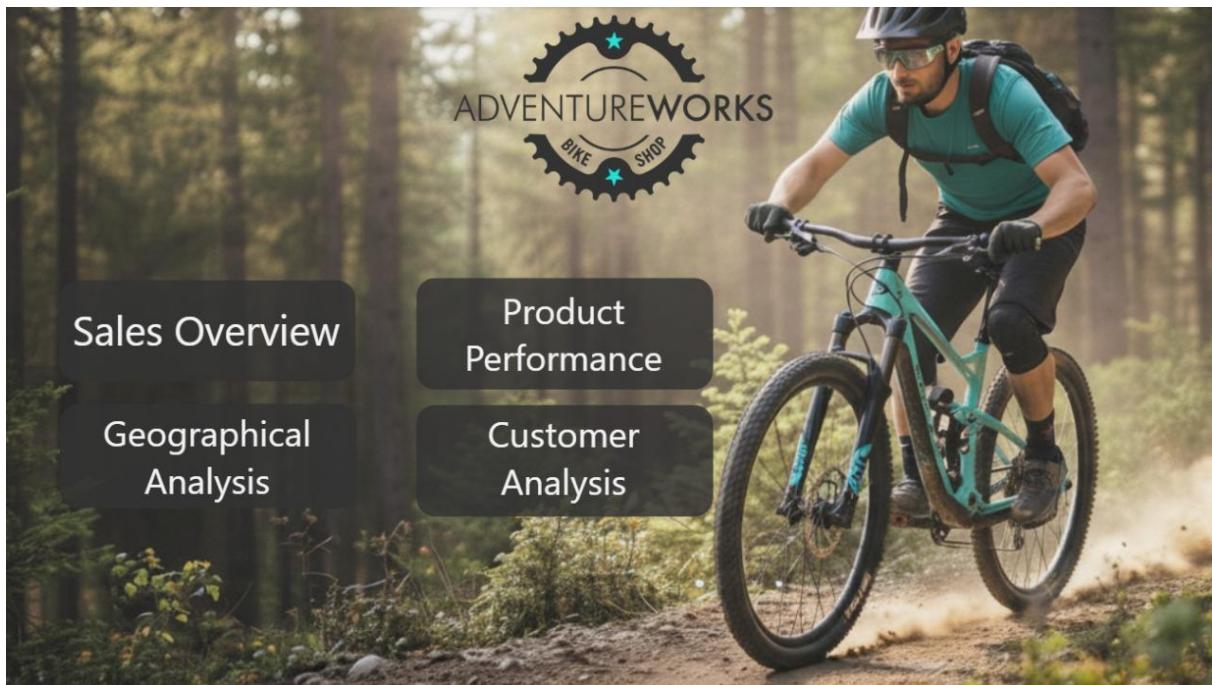
- Total Sales
- Total Cost
- Total Profit
- Profit Margin



Visual Components

- **Line Chart:** Sales Trend Over Time
- **Map Visualization:** Sales by Country
- **Donut/Pie Chart:** Profit per Category
- **Bar Chart:** Top 10 Customers

A page of the Dashboard



Main board of the Dashboard

8. Validation & Performance Testing

The dashboard was thoroughly tested to ensure accuracy and reliability:

- DAX measures validated against sample calculations
- Relationship filters tested for correctness
- Visual interactions checked (cross-filtering, drill-down)
- Performance optimized using:
 - Column reduction
 - Efficient star schema
 - Disabling unnecessary auto-date hierarchies

Data Cleaning

This screenshot shows the Microsoft Power BI Data Editor interface. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. A sidebar on the left lists several cleaning steps that have been applied to the data:

- OrderNumber: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- ProductKey: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- CustomerID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TerritoryID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- OrderQuantity: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TotalSales: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- Telephone: Clean & Filter, Remove Duplicates, Add Column, Add Comment.

This screenshot shows the Microsoft Power BI Data Editor interface with a focus on data distribution analysis. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. To the right, there are four histograms showing the distribution of data for each column. The histograms are color-coded by value count (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10+). The histograms are labeled as follows:

- OrderNumber: 100% Valid, 0% Null, 0% Error, 0% Empty, 0% Bad.
- ProductKey: 100% Valid, 0% Null, 0% Error, 0% Empty, 0% Bad.
- CustomerID: 100% Valid, 0% Null, 0% Error, 0% Empty, 0% Bad.
- TerritoryID: 100% Valid, 0% Null, 0% Error, 0% Empty, 0% Bad.

This screenshot shows the Microsoft Power BI Data Editor interface. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. A sidebar on the left lists one cleaning step applied to the data:

- OrderNumber: Clean & Filter, Remove Duplicates, Add Column, Add Comment.

This screenshot shows the Microsoft Power BI Data Editor interface. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. A sidebar on the left lists numerous cleaning steps applied to the data:

- OrderNumber: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- ProductKey: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- CustomerID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TerritoryID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- OrderQuantity: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TotalSales: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- Telephone: Clean & Filter, Remove Duplicates, Add Column, Add Comment.

This screenshot shows the Microsoft Power BI Data Editor interface. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. A sidebar on the left lists many cleaning steps applied to the data:

- OrderNumber: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- ProductKey: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- CustomerID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TerritoryID: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- OrderQuantity: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- TotalSales: Clean & Filter, Remove Duplicates, Add Column, Add Comment.
- Telephone: Clean & Filter, Remove Duplicates, Add Column, Add Comment.

This screenshot shows the Microsoft Power BI Data Editor interface. The main area displays a table of sales data with columns including OrderNumber, ProductKey, CustomerID, TerritoryID, OrderQuantity, TotalSales, and Telephone. A sidebar on the left lists one cleaning step applied to the data:

- OrderNumber: Clean & Filter, Remove Duplicates, Add Column, Add Comment.