

WEEK 1 REPORT

Topic: Power BI – Data Import & Data Cleaning (Power Query)

Intern Name: Harika Vavilapalli

Internship Domain: Data Analytics / Data Science

Tool Used: Microsoft Power BI

Focus: Data Loading, Data Cleaning, Data Transformation using Power Query

1. Introduction

In the first week of my internship, I focused on learning the fundamentals of **Microsoft Power BI**, mainly concentrating on **data importing and data preparation**. This week helped me understand how raw datasets are brought into Power BI and how they can be cleaned and transformed before analysis.

Since real-world datasets often contain missing values, unwanted columns, incorrect formats, or duplicate entries, data cleaning becomes a crucial step. I explored Power BI's built-in data transformation tool called **Power Query**, which is used to prepare data effectively for further analysis.

This week was mainly about building a strong base in Power BI so that I can apply it in later stages such as exploratory data analysis, reporting, and working with real-time/live data.

2. Objectives of Week 1

The main goals for Week 1 were:

- To understand the basic purpose and features of **Power BI**
- To explore the Power BI interface and workflow
- To learn how to **import datasets** into Power BI Desktop
- To perform **data cleaning and transformation** using **Power Query**
- To understand the importance of clean and structured data for analysis

3. Tasks Completed

3.1 Understanding Power BI Basics

During the initial sessions, I learned what Power BI is used for and how it supports data-driven decision making. I explored the environment and understood the main components such as:

- **Power BI Desktop workspace**
- Report view, Data view, and Model view (basic understanding)
- The overall process flow:
Load Data → Transform Data → Prepare Dataset

This gave me clarity on how Power BI is structured and how users work with data inside it.

3.2 Importing Data into Power BI

Next, I practiced importing datasets into Power BI. I learned how to:

- Load data from sources like:
 - Excel files
 - CSV files
 - Text files (basic understanding)
- Preview data before loading it
- Check column names and data structure during import
- Use the “Transform Data” option to open Power Query for preprocessing

This step helped me understand that importing is not just loading the file, but also verifying whether the data is usable and consistent.

3.3 Data Cleaning Using Power Query

A major part of Week 1 was working with **Power Query Editor**, where I learned different ways to clean and transform data. The following tasks were practiced:

a) Removing Unnecessary Data

- Removed unwanted columns that were not required
- Removed extra rows such as blank rows or irrelevant entries
- Filtered data when needed to focus on useful records

b) Handling Missing Values

- Identified missing or null values in columns
- Understood how missing values can affect analysis later
- Learned different ways to handle them:
 - Replacing null values
 - Removing rows with missing data (when necessary)

c) Correcting Data Types

- Checked column data types (Text, Number, Date, etc.)
- Converted incorrect data types into the correct format
- Understood why correct data types are important for future calculations

d) Removing Duplicates

- Detected duplicate values in the dataset
- Removed duplicates to ensure accuracy and consistency

e) Renaming and Organizing Columns

- Renamed column headers to make them clearer and easier to understand
- Organized data to improve readability and structure

f) Basic Transformations

- Sorting data based on values
- Splitting columns (example: splitting a combined field into separate fields)
- Replacing values for consistency (example: correcting spelling variations)

Power Query helped me understand that data preparation is a step-by-step process, and every transformation is recorded so it can be tracked and updated when data changes.

4. Key Learnings

By the end of Week 1, I gained strong understanding of the data preparation process. My key learnings include:

- Power BI is not only for visualization, but also for **data preparation**
- **Power Query** is a powerful tool for cleaning and transforming datasets
- Clean data is necessary for accurate results in future analysis
- Real-world datasets often require preprocessing before analysis
- Data types, missing values, and duplicates must be handled carefully

5. Challenges Faced

Some challenges I faced during this week were:

- Understanding which cleaning method is best for different data issues
- Managing missing values without losing important information
- Identifying the correct data type for some columns
- Getting comfortable with Power Query steps and transformations

However, with practice and guidance, I was able to understand the process better and improve my confidence.

6. Outcome of Week 1

At the end of Week 1, I was able to:

- ✓ Understand Power BI workflow and environment
- ✓ Import datasets successfully into Power BI Desktop
- ✓ Use Power Query for cleaning and transformations
- ✓ Prepare structured and cleaned datasets for further analysis

This week helped me build a strong foundation for the next steps of the internship.

7. Plan for Week 2

In Week 2, I will focus on learning and working with **VS Code**, including:

- Setting up coding environment
- Understanding file structure and scripts
- Loading datasets using code
- Preparing for Exploratory Data Analysis (EDA)

Conclusion

Week 1 was focused on learning Power BI fundamentals, especially data importing and cleaning using Power Query. This week helped me understand the importance of preparing clean and structured data before moving into deeper analysis. The skills learned in Week 1 will support upcoming internship tasks such as EDA and working with live data.