Sanjivani Rural Education Society's
Sanjivani College of Engineering, Kopargaon-423603
(An Autonomous Institute Affiliated to Savitribai Phule Pune University, Pune)
NAAC 'A' Grade Accredited, ISO 9001:2015 Certified
Department of Information Technology
NBA Accredited-UG Prpgram

Class: S.Y. B. Tech Semester: IV
Subject: Data Visualization and Story Telling (SEIT261)

Practical No.: 02

Title:
Implementing data preparation and data modeling in power BI.

Software Requirements:

- Power BI Desktop

Theory:

<u>Data Preparation in Power BI</u>

Objective:

To understand how to import, clean, and transform data using Power BI's Power Query Editor.

Steps:

- **Load Data into Power BI:**
  - Open Power BI Desktop.
  - Click on Home > Get Data.
  - Choose Excel, CSV, SQL Server, or another data source.
  - Click Load to import the dataset.
- **Use Power Query for Data Cleaning:**
  - Click on Transform Data to open Power Query Editor.
  - Perform data cleaning tasks:
  - Remove duplicates
  - Handle missing values (replace nulls, remove rows, etc.)
  - Change data types (text, number, date, etc.)
  - Split and merge columns
  - Add calculated columns
  - Click Close & Apply to load cleaned data.
- **Create Relationships Between Tables:**

- Go to the Model View.
- Drag and drop to connect related tables using primary and foreign keys.
- Define relationships (one-to-many, many-to-one).

- **Data Visualization**

  Use Power BI visuals to present insights effectively:

  - **Bar, Line, and Pie Charts** for trends and comparisons.
  - **Tables & Matrices** for detailed views.
  - **Maps** for geographic analysis.
  - **KPI Cards & Gauges** for key performance indicators.

- **Interactive Dashboards & Reports**
  - Create **interactive dashboards** using slicers, filters, and drill-through options.
  - Implement **bookmarks and tooltips** for enhanced user experience.
  - Publish reports to **Power BI Service** for sharing and collaboration

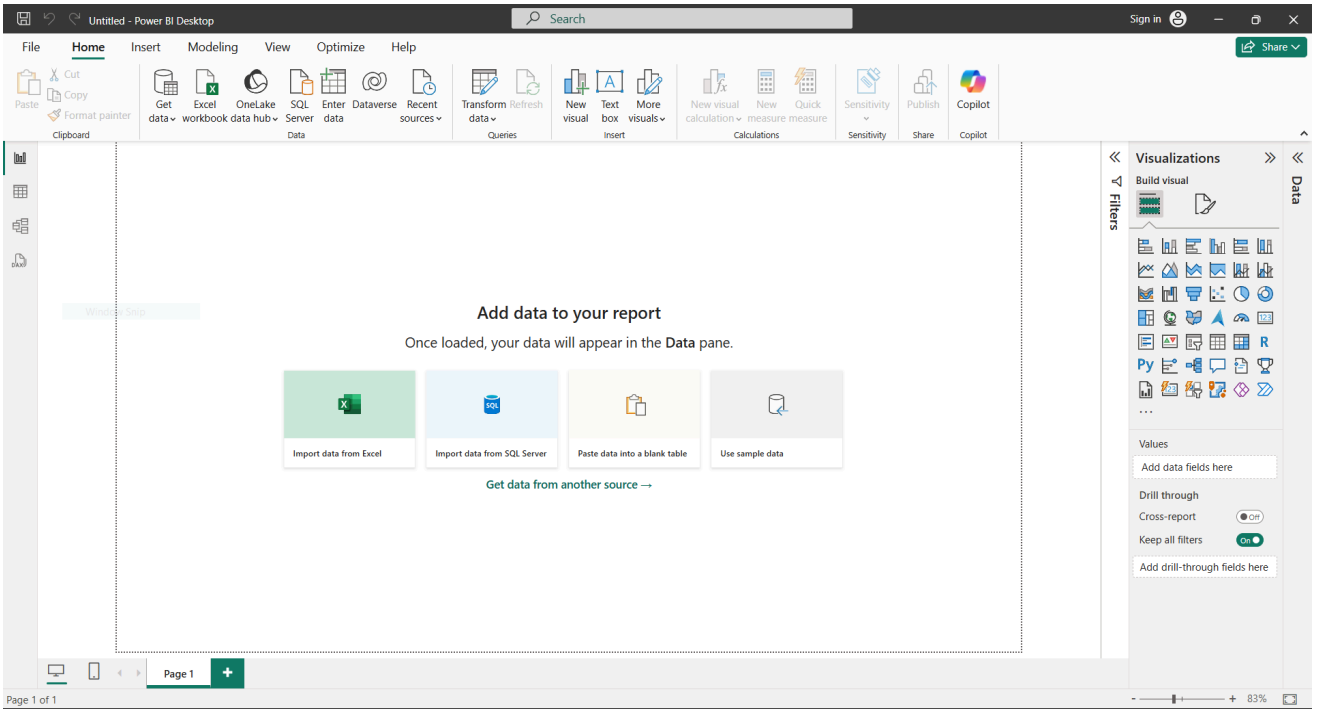


Fig1. Power BI Workspace

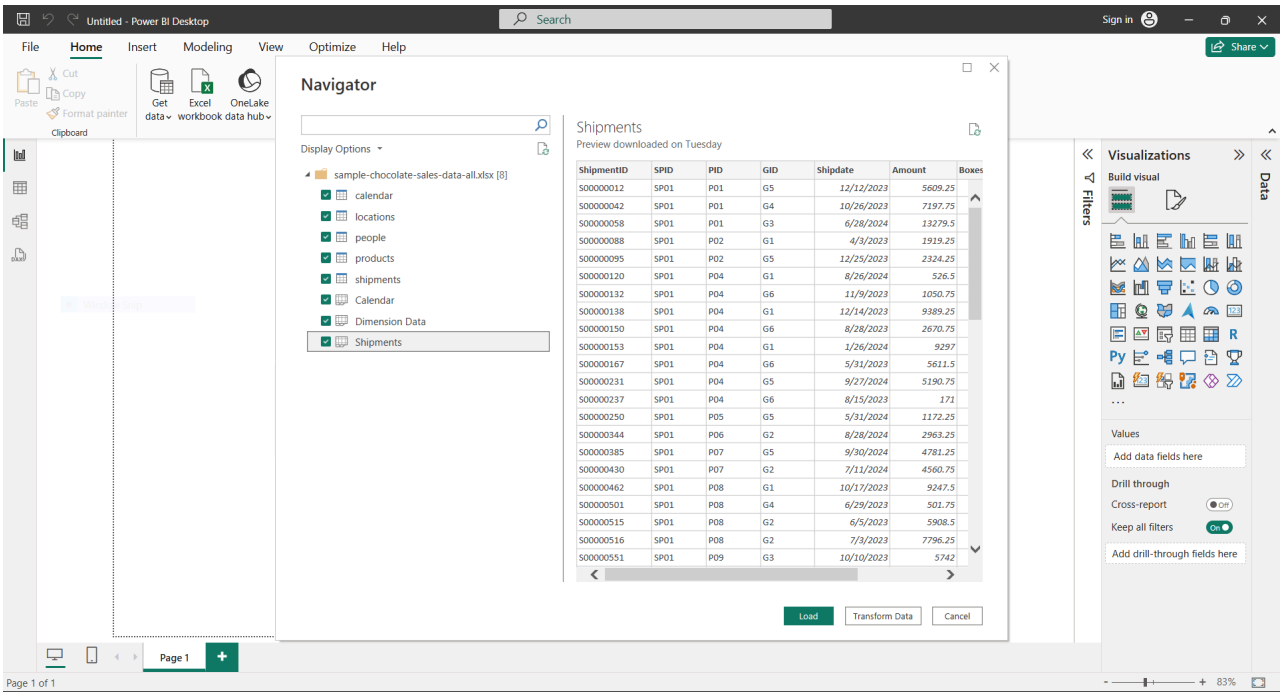


Fig 2. Import Data

What is Data Modeling?

Data modeling is the process of creating visual representations of the connections between data structures, with information about the individual attributes contained within those data structures.

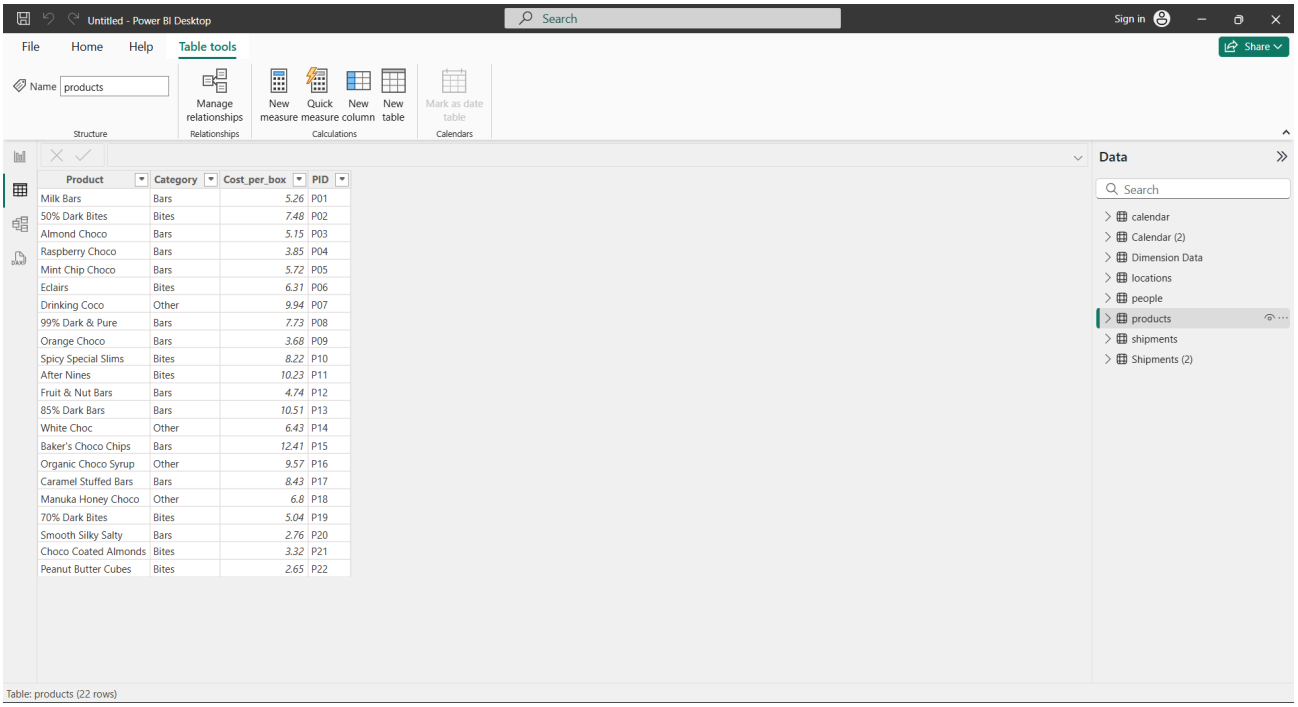
When talking about data modeling in general, the term you will hear most often is the star schema. This is a widely adopted approach to designing data warehouses and relational databases and is the recommended approach to take in **Power BI** as well.

There are two main benefits of using a star schema:

- **Usability:** A star schema makes your data model cleaner and more organized, and your report will be easier to use.
- **Performance:** Star schemas allow you to easily scale your report to very large volumes of data. DAX measures also calculate faster and the Power BI report refreshes faster overall.

In Power BI, **Table View**, **Model View**, and **Report View** serve different purposes for data analysis and visualization. **Table View** allows users to see the raw data in a structured format, similar to an Excel spreadsheet, where they can inspect, filter, and manipulate data within tables. **Model View** provides a graphical representation of the relationships between different tables, enabling users to define relationships, create hierarchies, and manage data modeling aspects like primary and foreign keys. **Report View** is the main workspace where users create and design interactive reports using visuals such as charts, graphs, and

slicers to present insights effectively. These three views collectively help users in data preparation, modeling, and visualization within Power BI.



Product	Category	Cost_per_box	PID
Milk Bars	Bars	5.26	P01
50% Dark Bites	Bites	7.48	P02
Almond Choco	Bars	5.15	P03
Raspberry Choco	Bars	3.85	P04
Mint Chip Choco	Bars	5.72	P05
Eclairs	Bites	6.31	P06
Drinking Coco	Other	9.94	P07
99% Dark & Pure	Bars	7.73	P08
Orange Choco	Bars	3.68	P09
Spicy Special Slims	Bites	8.22	P10
After Nines	Bites	10.23	P11
Fruit & Nut Bars	Bars	4.74	P12
85% Dark Bars	Bars	10.51	P13
White Choc	Other	6.43	P14
Baker's Choco Chips	Bars	12.41	P15
Organic Choco Syrup	Other	9.57	P16
Caramel Stuffed Bars	Bars	8.43	P17
Manuka Honey Choco	Other	6.8	P18
70% Dark Bites	Bites	5.04	P19
Smooth Silky Salty	Bars	2.76	P20
Choco Coated Almonds	Bites	3.32	P21
Peanut Butter Cubes	Bites	2.65	P22

Fig 3. Table View

How to Create Relationships in Power BI

There are two ways you can create a relationship in Power BI:

1. Select a field from one table and drag it onto the field in the second table with which you want the relationship to form.
2. Select Manage Relationships from the ribbon and select “New” to add a relationship using the same window that we will be discussing next (except that it will start as blank).

By default, Power BI will try to infer a relationship between tables; it doesn't always get this right, so you may wish to turn this feature off in the settings or delete any relationships that are created automatically. To edit the relationship, right-click the connecting line between them and select “Properties”.

Cardinality in Power BI

Cardinality in Power BI defines the **relationship type** between tables based on how many records in one table match records in another table. It helps optimize data models and ensures accurate calculations.

Types of Cardinality:

1. **One-to-Many (1:*)** – The most common relationship type. One record in the primary table corresponds to multiple records in the related table.
    - Example: A **Customers** table (one customer) links to a **Sales** table (many purchases).
2. **Many-to-One (*:1)** – Equivalent to **One-to-Many**, just from the perspective of the related table.
3. **Many-to-Many (*:*)** – When multiple records in both tables are related to each other. Power BI uses an intermediary table or **bidirectional filtering** to handle this.
    - Example: **Students** and **Courses** (one student enrolls in many courses, and one course has many students).
4. **One-to-One (1:1)** – Rarely used, where each record in Table A has exactly one corresponding record in Table B.

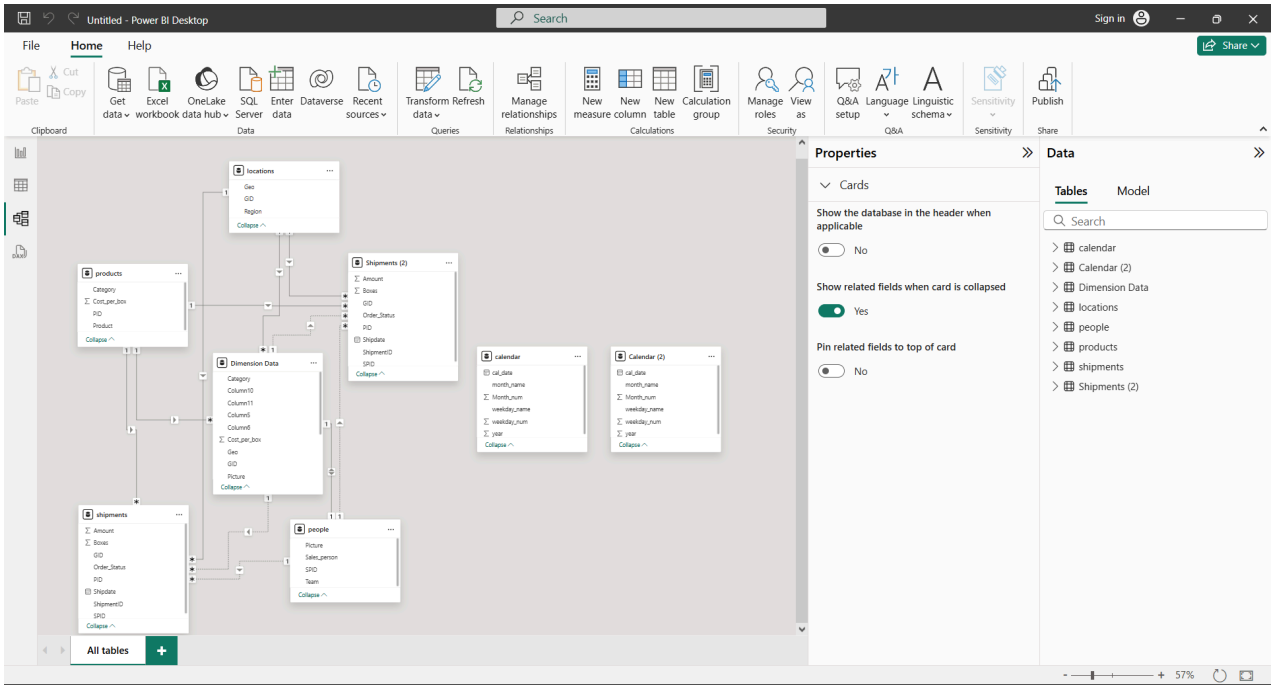


Fig 4. Model View

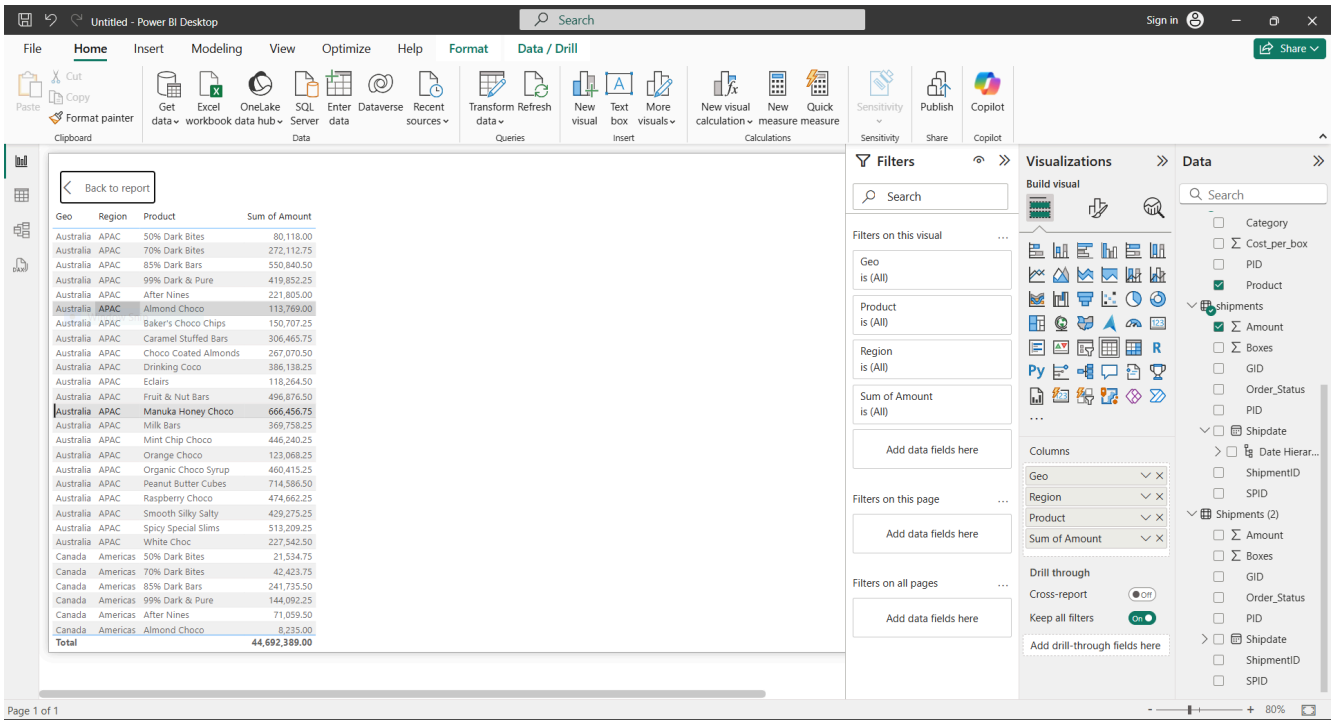


Fig 5. Report View

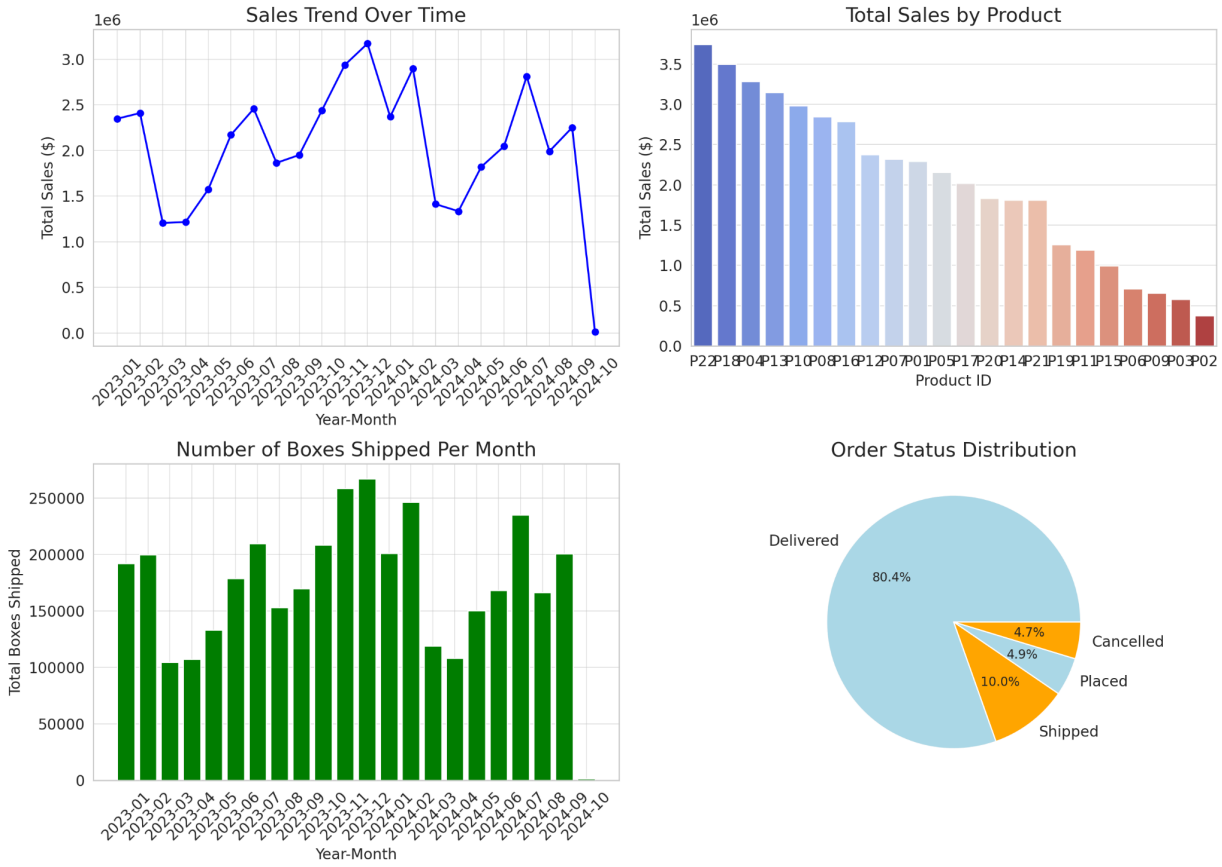


Fig 6. Dashboard

Best Practices:

- **Use One-to-Many whenever possible** for efficient relationships.
- **Avoid Many-to-Many relationships** unless necessary, as they can impact performance.
- **Optimize Cross-filter direction** (Single or Both) based on analytical needs.
- **Check relationship properties** in **Model View** to ensure correct cardinality.

Frequently Asked Questions:

- What are the key steps involved in data preparation in Power BI, and why is it important?
- Explain the difference between star schema and snowflake schema in Power BI data modeling. When would you use each?
- How does Power Query help in transforming and cleaning data before loading it into Power BI?
- What are the different types of relationships in Power BI data modeling, and how do they affect report performance?
- What are calculated columns and measures in Power BI? How do they differ in terms of storage and performance impact?

Conclusion:

Thus, we have successfully explored the concepts of Data Preparation and Data Modeling in Power BI. We began by understanding the importance of data preparation in ensuring clean and structured data for analysis. We then explored the fundamentals of data modeling, including relationships, schemas, and the use of DAX for calculations. Additionally, we discussed how Power Query facilitates data transformation before loading it into Power BI. By mastering these concepts, we can efficiently build optimized and insightful Power BI reports for data-driven decision-making.

References:

- Practical Tableau: 100 Tips, Tutorials, and Strategies from a Tableau Zen Master, Ryan Sleeper, Oreilly Publications, 2018
- Data Visualization with R: 111 Examples by Thomas Rahlf, Springer, 2020.
- Learning Microsoft Power BI - Jeremey Arnold, 2022
- Learn Power BI: A Comprehensive, Step-by-step Guide for Beginners to Learn Real-worldBusiness Intelligence - Greg Deckler, 2022.
- Tableau 10 Complete Reference: Joshua N Milligan.

Prepared By,
Dr. R. D. Chintamani
Mr. U. B. Sangule
Ms. P. G. Thakre
Subject In-charge

Approved By,
Dr. M. A. Jawale
HOD-IT