

---

# Data Analytics Mini Project Guidelines

## Project Objective

The mini project aims to assess your ability to perform real-world data analysis using **Excel** for data cleaning and transformation, and **Power BI** for data visualization and report building.

This carries a total of 25 marks: 15 marks for the mini project and 10 marks for the intermediate assessment (MCQs).

## Project Scope

Your project will involve the following key components:

### 1. Data Pre-processing using Excel

- Perform cleaning and transformation tasks on raw data such as:
  - Removing duplicates
  - Handling missing values
  - Standardizing formats
  - Creating calculated fields (if required)
  - Filtering and sorting
  - Creating Pivot Tables (if required)

### 2. Data Visualization using Power BI

- Build a report/dashboard with:
  - Multiple visualizations that communicate insights clearly
  - Appropriate use of charts (bar, line, pie, slicers, cards, etc.)
  - Clear and interactive layout (using filters, slicers, drill-downs, etc.)
  - Proper titles, labels, and legends

# Project Domains

Here are some sample domain ideas to help you get started.  
However, you are free to choose any domain or topic of your interest.

## 1. Sales & E-commerce

- Sales performance dashboard
- Customer segmentation based on region or product category
- Trend and seasonality analysis of sales data

## 2. Finance

- Expense tracking dashboard
- Budget vs. Actual analysis
- Profitability by product or region

## 3. Healthcare

- Hospital/clinic patient records analysis
- Disease count trends across months/years
- Treatment cost analysis

## 4. Retail

- Inventory and stock analysis
- Customer purchasing patterns
- Store-wise performance comparison

## 5. Education

- Student performance dashboard
- Course enrollment and completion analysis
- Feedback and satisfaction trends

## 6. Human Resources (HR)

- Employee attrition or retention reports
- Department-wise performance or salary comparison
- Leave and attendance trends

## 7. Environmental

- Traffic pattern analysis across regions or time periods
- Air quality monitoring and pollutant trends
- Water consumption or waste generation tracking

## 8. Social Media & Public Sentiment

- Sentiment analysis of social media posts or movie reviews (e.g., Twitter, IMDB)
  - Trending hashtags or topics analysis
  - Public opinion analysis on a campaign or event
- 

## Dataset Requirements

The dataset should have:

- **Minimum of 500 records**
  - **At least 10 features/columns** (preferably a mix of numerical and categorical data)
- 

## Data Sources

Here are some reliable sources where you can find datasets for your project:

1. **Kaggle:** [Kaggle Datasets](#)
  2. **UCI Machine Learning Repository:** [UCI Repository](#)
  3. **Google Dataset Search:** [Google Dataset Search](#)
  4. **Data.gov (U.S.):** [Data.gov](#)
  5. **Data.gov,in (India):** [Data.gov.in](#)
  6. **World Health Organization (WHO):** [WHO Data](#)
  7. **India Data Portal:** [India Data Portal](#)
  8. **Awesome Public Datasets:** [Awesome Public Datasets](#)
  9. **OpenML:** [OpenML](#)
  10. **FiveThirtyEight:** [FiveThirtyEight Datasets](#)
-

## Submission Requirements

You must submit the following files:

1. **Original Dataset Source**

- Provide the dataset link from where the raw data was sourced.

2. **Cleaned Dataset**

- Submit the **.xlsx** Excel file after performing data cleaning and transformation.

3. **Power BI Report (PDF Export)**

- Export and submit your Power BI report/dashboard as a **PDF file (.pdf)**.
- 

## Mini-Project Rubrics [Total: 15 marks]

**TASK 1:** Dataset Selection (3 marks)

- Relevance, size, and source of the dataset.

**TASK 2:** Data Pre-processing using Excel (5 marks)

- Cleaning, transformation, and Excel features used.

**TASK 3:** Data Visualization using Power BI (7 marks)

- Visualizations, insights, and design of the report/dashboard.
- 

## Intermediate Assessment [10 marks]

- Platform: **HackerEarth**
  - Eligibility: No pending assignments and completed module-end assignment
  - You will receive an email with the test link.
  - Format:
    - **5 MCQs from Excel**
    - **5 MCQs from Power BI**
    - Total: **10 questions**
-