Software Requirements Specification

Food Manufacturing - Quality Attributes Analysis Tool

1. Functional Requirements

Data Input:
- The software must load an Excel file (Food_Manufacturing_Quality_Attributes.xlsx) containing numerical data for quality attributes and a target column named Defects(count).

Visualization:
- Generate the following visualizations:
  - Correlation Heatmap of all numeric attributes.
  - Defect Counts Distribution Plot.
  - Boxplots of each quality attribute vs. Defects(count).
  - Feature Importance Bar Chart for defect prediction.

Machine Learning Model:
- Perform binary classification to predict defective products based on quality attributes.
- Display a Confusion Matrix and Classification Report.

User Output:
- All generated plots must be displayed sequentially.
- Print model evaluation metrics (confusion matrix, classification report) in the console.

2. Non-Functional Requirements

Platform:
- OS: Windows, macOS, or Linux
- Environment: Python 3.7+

Performance:
- Should efficiently handle small to medium-sized datasets (up to ~100,000 rows).

Usability:
- Run via a Python script in Jupyter Notebook or IDE (VS Code, PyCharm).
- Outputs displayed as static images and console text.

3. Software Requirements

Programming Language:
- Python 3.7 or higher

Libraries/Dependencies (installable via pip):
pip install pandas matplotlib seaborn scikit-learn

Required Python Libraries:
- pandas: Data loading and manipulation
- matplotlib: Plotting graphs and charts
- seaborn: Enhanced statistical visualizations
- scikit-learn: Machine learning model and evaluation metrics

4. Input File Format

- Excel file named Food_Manufacturing_Quality_Attributes.xlsx
- Must contain numeric columns:
  - MoistureContent(%)
  - Weight(g)
  - Color(L*)
  - Texture(N)
  - Sweetness(Brix)
- One target column named Defects(count)