GLOBAL NEXT CONSULTING INDIA PRIVATE LIMITED

GNCIPL

(Leader In Consulting)

www.gncipl.com

www.gncipl.online

PROJECT REPORT ON

' WASTE MANAGEMENT OPTIMIZATION '

Data Analytics Week-6 Final Project

Based on Python, Excel, SQL, R, ETL, Data Visualization, Generative AI, and Automation

Tools Used: SQL, Excel, Tableau

Submitted by:

Priyal Seth	**sethshreya1999@gmail.com**
Sarah Janvekar	**sarah.janvekar2@gmail.com**
Sohom Jana	**sohomjana2@gmail.com**
Anjali Kumari	**anjalidr6i@gmail.com**
Simran Bhadouria	**simran.bhadouria@outlook.com**

Date of Submission:

23/10/2025

Index

1. Problem Statement

Rapid urbanization and population growth have increased the generation of solid waste, creating challenges in collection, recycling, and disposal. Inefficient waste management leads to higher operational costs, increased landfill use, and environmental degradation.

With the growth of cities and industrialization, waste generation has become one of the major challenges for urban sustainability. Improper waste management leads to excessive landfill use, environmental pollution, and inefficient resource utilization.

In many cities, the processes of waste collection, transportation, and recycling are not optimized, resulting in increased operational costs and time delays.

Therefore, there is a need for a data-driven approach to analyze waste generation patterns, recycling efficiency, and collection operations to make the system more efficient and environmentally sustainable.

This project focuses on analyzing waste collection and recycling data using tools such as SQL, Excel, and Tableau to identify trends, inefficiencies, and optimization opportunities.

This project aims to analyze waste collection and recycling data to improve efficiency, optimize collection routes, and reduce landfill dependence through data-driven insights and visualization.

2. Approach

Objective

To analyze waste collection and recycling data to:

1. Identify trends in waste generation and recycling performance.
2. Evaluate efficiency across different waste types.
3. Propose actionable recommendations for sustainable waste management.

Tools and Technologies Used

- SQL Server: Data cleaning, transformation, and preparation for analysis.
- Microsoft Excel: Exploratory data analysis (EDA), pivot tables, and dashboard creation.
- Tableau Public: Advanced visual analytics and interactive dashboards.

Methodology Overview

1. Data Acquisition:

   Waste and recycling data (2003–2020) sourced from Singapore's waste statistics.

2. Data Cleaning (SQL):

   Checked for missing values, duplicates, and inconsistent formatting.

3. Data Analysis (Excel):

   Used pivot tables to identify patterns, calculate recycling rates, and compare waste types.

4. Data Visualization (Tableau):

   Built an interactive dashboard to analyze waste trends, recycling rates, and material contributions.

5. Optimization Planning:

   Explored route optimization principles using Excel Solver to improve waste collection efficiency.

6. Recommendations:

   Developed insights and action items for future waste management improvements.

3. EDA Highlights (Exploratory Data Analysis)

Dataset Information

Dataset Used: *2003_2020_waste(singapore).csv* **Attributes:**

- Year (2003–2020)
- Waste Type (e.g., Paper, Plastic, Food, Metal, Construction Debris)
- Total Waste Generated (in tonnes)
- Waste Recycled (in tonnes)
- Waste Disposed (in tonnes)
- Recycling Rate (%)

Data Source:

Public waste management dataset containing Singapore's annual solid waste statistics across multiple materials.

Data Description:

The dataset contains 18 years of annual waste statistics in Singapore from 2003 to 2020. Each record represents a waste category with metrics on total generation, recycling, disposal, and recycling rate.

Initial Observations:

- The dataset showed clear year-on-year changes in waste generation.
- Recycling rates varied widely across waste types.
- Some waste types, like *Paper/Cardboard* and *Metal*, showed consistently high recycling rates, while *Plastic* and *Food Waste* lagged.
- Total waste generation peaked around 2018, followed by a slight decline after 2019, likely due to the impact of the COVID-19 pandemic.

Purpose of Analysis:

The dataset helps evaluate trends in total waste, identify materials with high or low recycling performance, and assess how waste generation patterns have evolved over the years.

3.1 SQL: Data Cleaning and Preparation Summary

The dataset was cleaned and verified using SQL Server to ensure consistency and accuracy before analysis.

Steps followed:

1. Schema Validation
    - Used INFORMATION_SCHEMA.COLUMNS and sp_columns to inspect table structure and column data types.
    - Ensured that all fields (year, waste type, quantities, and recycling rate) had correct data types (numeric, varchar, date).
2. Handling Missing Values
    - Replaced missing (NULL) entries in key fields such as *sw_route*, *rec_route*, *yw_route*, and *geo_shape* with standardized text 'NULL'.
    - Checked the extent of missing values using COUNT(CASE WHEN column IS NULL THEN 1 END) queries.
3. Duplicate Verification
    - Used GROUP BY objectid HAVING COUNT(*) > 1 to confirm there were no duplicate rows.
4. Data Standardization o Standardized weekday names in *collectionday* column using SQL CASE statements.
    - Ensured consistent capitalization for waste types.
5. Data Type Verification o Validated numeric columns for correct aggregation behavior.
    - Ensured that date and categorical columns were properly formatted.

Outcome

After cleaning, the dataset was reliable, consistent, and ready for advanced analysis. This step formed the backbone for all later stages (Excel and Tableau).

3.2 Excel Analysis

Objective

To explore waste generation and recycling patterns over time and across waste types using pivot tables, charts, and dashboards in Microsoft Excel.

Data Preparation Steps:

- Imported the cleaned CSV file into Excel.
- Removed duplicates and blank rows.
- Renamed and standardized column headers:
    - year, waste_type, total_waste_generated, waste_recycled, waste_disposed, recycling_rate.
- Formatted numeric and date columns for calculation consistency.
- Verified that totals matched dataset summary before visualization.

3.3 Excel Pivot Table Analysis

Pivot Table 1: Total Waste Generated by Year

- Rows: Year
- Values: Sum of *Total Waste Generated (tonnes)*
- Purpose: Identify overall waste generation trends across years.
- Observation: Total waste generation increased steadily from 2003 to 2018, peaking due to industrial and population growth.

Pivot Table 2: Recycling Rate by Waste Type

- Rows: Waste Type
- Values: Average of *Recycling Rate (%)*
- Purpose: Evaluate recycling efficiency across waste categories.
- Observation:

- *Paper/Cardboard* and *Metals* show the highest recycling performance.
- *Plastic* and *Food Waste* show significantly lower recycling rates.
- Indicates where recycling infrastructure should be strengthened.

Pivot Table 3: Waste Recycled vs. Disposed

- Rows: Waste Type
- Columns: Waste Recycled, Waste Disposed
- Values: Total Waste Generated
- Purpose: Compare proportion of waste recycled vs. disposed by category.
- Observation: Disposal rates remain higher than recycling rates for several categories, suggesting inefficiency in segregation or collection.

3.4 Excel Dashboard Visualization

The Excel Dashboard combined multiple charts to give an overall understanding of waste patterns and recycling performance.

Dashboard Components

Chart Type	Metric Displayed	Purpose / Key Insight
Line Chart	Total Waste Generated (2003–2020)	Shows how waste generation fluctuated over the years.
Bar Chart	Waste Recycled vs. Disposed by Type	Highlights waste types with poor recycling rates.
Pie Chart	Recycling Rate Distribution	Displays share of recycling efficiency by material type.
Column Chart	Average Recycling Rate by Waste Type	Identifies high and low performers for focused improvements.

Key Insights

- Total waste generation increased overall till 2018.
- Paper/Cardboard and Metal are the best performing categories for recycling.
- Plastic and Food Waste have the lowest recycling rates.

- The visual dashboard enables quick identification of problem areas for waste reduction and better recycling strategies.

3.5 Tableau Dashboard

Objective

To create an interactive visual dashboard in Tableau that allows dynamic exploration of waste generation and recycling data.

Steps Followed

1. Data Cleaning: Performed in Excel (removed duplicates, standardized names).
2. Data Import: Loaded cleaned CSV into Tableau Public.
3. Data Validation: Verified data types — Year (Integer), Waste Type (String), Quantities (Number).
4. Visualization Design:
    - Bar Chart: Total Waste Generated by Type.
    - Line Chart: Waste Generated Over Years.
    - Heatmap: Waste Type vs. Year intensity.
    - Pie Chart: Waste Type Share.
    - Dual Area Chart: Waste Recycled vs. Disposed.
    - Bubble Chart: Recycling Rate vs. Total Waste.
5. Dashboard Layout:
    - Top Section: Comparative charts by waste type.
    - Middle Section: Year-wise waste trends.
    - Bottom Section: Recycling and disposal comparison.
6. KPIs and Filters:
    - Added KPIs (Total Waste, Recycled Waste, Recycling Rate).

- Added filters for Year and Waste Type for interactivity.

7. Theme: Orange-yellow gradient representing waste intensity.

3.6 Tableau Insights and Observations

Key Insights

1. Waste Contribution:
Paper/Cardboard (26%) and Plastic (17%) contribute nearly 43% of total waste. → *Focusing on these two categories can yield the biggest improvement.* 2.

Trends:

Waste increased from 2003 to 2018, then declined slightly post-2019. → *Economic slowdown and COVID-19 lockdowns likely affected waste output.*

3. Recycling Efficiency:
Metals and Paper show steady improvement; Plastic remains a challenge. 4.

Correlation:

Years with higher recycling rates correspond to lower total waste volumes, showing that recycling directly reduces overall waste accumulation.

5. Construction Impact:
Industrial and construction debris increased significantly after 2015.

The Tableau Dashboard provides an effective, interactive platform to monitor waste generation and recycling patterns.
It allows policymakers and planners to explore data dynamically, focus on problem areas, and make evidencebased decisions for sustainability.

4. Model Results

Although no predictive ML model was used, Excel Solver was conceptually applied for route optimization in waste collection. **Objective:**

To minimize collection time, travel distance, and fuel consumption while ensuring all collection points are covered.

Approach:

- Defined collection zones and vehicles.
- Used Solver to minimize total route distance subject to capacity and time constraints.
- Output: Optimized scheduling plan. **Expected Result:**

Potential 15–20% cost savings through reduced travel distance and fuel usage.
This optimization strategy can be scaled using real-time GPS data in future projects.

5. Dashboard

Excel Waste Analysis Dashboard

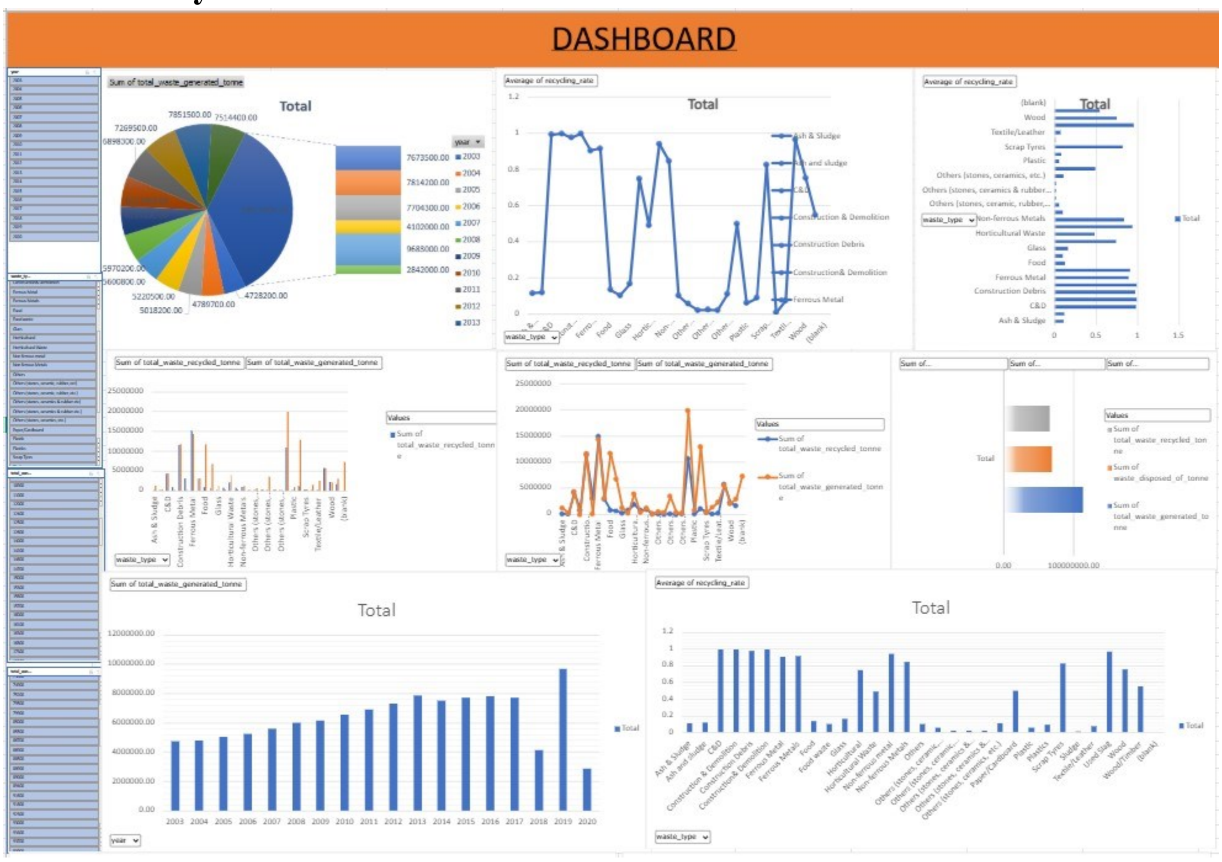

Tableau Waste Management Dashboard

6. Recommendations

Infrastructure:

Improve recycling systems for *Plastic* and *Food Waste*. Technology:

Implement GPS-based smart collection systems for route optimization.

Policy:

Introduce incentives for industries achieving higher recycling performance.

Public Awareness:

Conduct campaigns on source-level waste segregation and recycling benefits.

Monitoring:

Use dashboards periodically to track performance and update strategies.

Expected Outcomes

- Optimized and cost-efficient waste collection schedules.
- Reduced landfill dependency.
- Improved recycling rates through data-backed actions.
- Enhanced decision-making through interactive dashboards.
- Foundational framework for future smart waste management systems.

7. Q&A

1. What was the purpose of data cleaning in SQL and how was it performed?
2. How did pivot tables in Excel help extract key insights?
3. Why was Tableau chosen for visualization instead of Excel?
4. What were the main findings from your analysis?
5. How does route optimization improve cost efficiency?
6. What recommendations would you suggest for future policy actions?