

Customer Analytics using Business Intelligence

Date Issued	Project title	Author
09-09-2025	Customer Analytics	Daizy Gupta

CONTENT

Dataset Description.....	4
Data Cleaning Plan (Python).....	4
Exploratory Data Analysis (EDA).....	5
Python (Jupyter Notebook: analysis.ipynb).....	5
Power BI Dashboard Design.....	5
Pages / Visuals:.....	6
Filters (Slicers):.....	6
Data Flow.....	6
Deliverables.....	7

Dataset Description

Source: Customer feedback dataset (CSV).

<https://github.com/DAIZYGUPTA/customer-analytics-bi-project/blob/main/data/mta-customer-feedback-data-beginning-2014-1.csv>

Column Name	Description
Agency	Name of the transit agency (e.g., Long Island Rail Road).
Commendation or Complaint	Type of feedback: Commendation (positive) or Complaint (negative).
Subject Matter	General category (e.g., Employees, Service, etc.).
Subject Detail	Specific subcategory (e.g., CSR - Ambassador, Information Booth).
Issue Detail	Text describing the feedback (e.g., "Very Helpful / Friendly").
Year	Year when the feedback was received.
Quarter	Quarter of the year (1–4).
Branch/Line/Route	Geographic or service route related to feedback.

Data Cleaning Plan (Python)

1. Missing Values:

- Replace "No Value" with NaN.
- Decide whether to drop or impute missing rows depending on frequency.

2. Duplicates:

- Remove duplicate rows.

3. Text Normalization:

- Standardize text in *Issue Detail* (e.g., unify "Very Helpful / Friendly" and "Very Helpful/Friendly").
- Convert all text to proper case for consistency.

4. Export Clean Data:

- Save as `cleaned_customer_feedback.csv` for Power BI.

Exploratory Data Analysis (EDA)

Python (Jupyter Notebook: `analysis.ipynb`)

Planned analysis:

1. Commendations vs Complaints

- Count by year & quarter.
- Line chart to show trends.

2. Subject Matter Distribution

- Bar chart: Which categories (Employees, Service, etc.) receive the most feedback.

3. Subject Detail & Issue Detail

- Top 10 most frequent issues.
- Word cloud visualization of *Issue Detail*.

4. Route/Branch Analysis

- Heatmap of complaints/commendations by route.
- KPI: Which route gets the most complaints vs commendations.

Power BI Dashboard Design

Pages / Visuals:

1. Dashboard Home (KPI Summary)

- Total Feedbacks
- % Complaints vs Commendations
- Most Common Issue

2. Trend Analysis

- Line chart: Complaints vs Commendations by Year/Quarter.
- Stacked bar: Feedback type per year.

3. Subject Matter Analysis

- Tree Map or Bar Chart: Subject Matter distribution.
- Table: Subject Detail vs Issue Count.

4. Route/Branch Performance

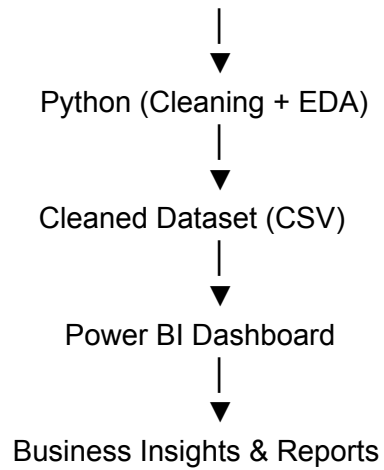
- Map/Heatmap: Feedback by Route.
- KPI: Best & Worst performing routes.

Filters (Slicers):

- Year
- Quarter
- Subject Matter
- Route

Data Flow

Raw Dataset (CSV)



Deliverables

- **analysis.ipynb** → Python code for data cleaning & EDA.
- **cleaned_customer_feedback.csv** → Exported dataset for Power BI.
- **CustomerAnalytics.pbix** → Power BI dashboard file.
- **Visualizations & Reports:** Included in docs /.