

Customer Analytics using Business Intelligence

Detailed Project Report

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

CONTENT

Detailed Project Report..... 1

1. Introduction..... 4

2. Problem Statement..... 4

3. Objectives..... 5

4. Dataset Description..... 5

5. Methodology..... 6

5.1 Data Preprocessing in Python..... 6

5.2 Dashboarding in Power BI..... 6

5.3 Insight Generation..... 7

6. Architecture..... 7

7. Dashboard Summary..... 7

KPIs..... 7

Branch Performance..... 8

Complaints vs Commendations per Issue Category..... 8

Top Issues..... 8

Trends Over Time..... 8

Complaints vs Commendations (Pie Chart)..... 8

8. Key Insights..... 8

9. Recommendations..... 9

10. Tools & Technologies..... 9

11. Project Outcomes..... 10

12. Conclusion..... 10

1. Introduction

Customer satisfaction is the backbone of public transportation systems. Agencies such as NYC Buses, Subways, Long Island Rail Road (LIRR), and Metro-North Railroad (MNR) collectively serve millions of commuters daily. With this vast customer base, feedback in the form of complaints and commendations plays a vital role in improving operational efficiency and enhancing passenger experiences.

This project leverages Business Intelligence (BI) and Data Analytics to analyze large-scale customer feedback data collected by the MTA (Metropolitan Transportation Authority). Using Python for data preprocessing and Power BI for visualization, the project delivers interactive dashboards that highlight customer pain points, employee performance, and service trends.

In addition, the project lays the foundation for future predictive modeling to automate access control for employees, thereby reducing manual intervention and operational costs.

2. Problem Statement

Despite collecting customer feedback for years, most organizations lack effective systems to analyze this data systematically. Feedback often gets lost in spreadsheets or siloed reports, resulting in missed opportunities for service improvement.

At the same time, companies face challenges in granting/revoking employee access to IT resources. Manual approval workflows create inefficiencies and delays, reducing employee productivity.

Thus, the problem has two aspects:

1. **Customer Feedback Gap** → Understanding customer concerns, categorizing issues, and identifying improvement areas.
2. **Employee Access Gap** → Automating role-based access management to eliminate repetitive manual work.

3. Objectives

The objectives of this project are twofold:

1. **Customer Analytics (Primary Focus):**

- Clean and analyze MTA customer feedback data.
- Identify trends, patterns, and recurring issues.
- Build interactive Power BI dashboards for decision-makers.

2. **Employee Access Modeling (Future Scope):**

- Use historical employee access data to train machine learning models.
- Predict whether access should be automatically granted/revoked when employees change roles.

4. Dataset Description

- Source: MTA Customer Feedback Data (2014 onwards).
- Volume: ~183,887 records.
- Features:
 - Agency → NYC Buses, Subways, LIRR, MNR
 - Commendation or Complaint → Type of feedback
 - Subject Matter → Broad area (Employees, Trains, Safety, etc.)
 - Subject Detail → Specific element (e.g., CSR - Ambassador, Staff Behavior)

- Issue Detail → Customer's reported problem/praise (e.g., "Very Helpful", "Delay")
- Year, Quarter → Time of feedback
- Branch/Line/Route → Affected branch/line
- **Distribution:**
 - Complaints: ~96%
 - Commendations: ~4%

This imbalance highlights that customers are more likely to report negative experiences than positive ones.

5. Methodology

5.1 Data Preprocessing in Python

- **Step 1:** Loaded raw dataset in Jupyter Notebook.
- **Step 2:** Replaced "No Value" with missing values (NaN).
- **Step 3:** Removed duplicates.
- **Step 4:** Standardized text (e.g., "late/delay" → "Late/Delay").
- **Step 5:** Exported a cleaned CSV for Power BI dashboarding.

5.2 Dashboarding in Power BI

- Created KPI cards for high-level metrics.
- Designed bar charts, line charts, and pie charts.

- Added interactive filters for Quarter and Agency.
- Aligned dashboard with corporate-style wireframe.

5.3 Insight Generation

- Compared complaints vs commendations.
- Identified top recurring issues.
- Monitored trends over years and across branches.

6. Architecture

The project follows a **modular architecture**:

1. **Data Source:** Customer feedback CSV.
2. **Data Processing Layer:** Python (pandas, numpy).
3. **Visualization Layer:** Power BI dashboards.
4. **Storage:** GitHub repo for collaboration & documentation.
5. **Future AI Module:** Predictive access model using historical employee data.

7. Dashboard Summary

KPIs

- **Total Feedbacks:** 188
- **Complaints:** 16.54K

- **Commendations:** 814

Branch Performance

- Some branches like New Haven and Port Jefferson received more complaints.

Complaints vs Commendations per Issue Category

- **Top Complaint Drivers:** Staff behavior, delays, maintenance, and safety.
- **Commendations:** Mostly around “very helpful/friendly” employees.

Top Issues

- Negative: rude behavior, reckless driving, delays.
- Positive: very helpful/friendly.

Trends Over Time

- Complaints remained consistently high across years.
- Commendations showed no significant growth.

Complaints vs Commendations (Pie Chart)

- ~95% complaints, ~5% commendations → extremely skewed sentiment.

8. Key Insights

1. **Service Quality Gap:** Most complaints are linked to delays, scheduling, and maintenance failures.

2. **Employee Behavior Issues:** Rude/inappropriate staff interactions significantly impact customer satisfaction.
3. **Positive Highlights:** Commendations, though rare, point to **individual employees exceeding expectations.**
4. **Systemic Challenges:** Trends show persistent complaint levels, indicating deeper systemic issues.

9. Recommendations

- **Operational Improvements:** Prioritize fixing frequent service breakdowns and delays.
- **Employee Training:** Launch soft-skills training programs to improve customer interaction.
- **Recognition Programs:** Reward commendable employees to encourage positive behavior.
- **Digital Transformation:** Enable real-time feedback collection through mobile apps.
- **Predictive Analytics (Future Scope):** Use AI/ML to proactively identify high-risk service areas and automate employee access management.

10. Tools & Technologies

- **Python** → Data cleaning, preprocessing
- **Pandas, Numpy** → Data wrangling
- **Matplotlib, Seaborn, WordCloud** → EDA visualizations

- **Power BI** → Dashboard creation
- **GitHub** → Version control, documentation, project portfolio

11. Project Outcomes

- Built interactive BI dashboards with real insights.
- Identified pain points in customer service across agencies.
- Developed a blueprint for predictive access automation.
- Created end-to-end project documentation for academic/professional submission.

12. Conclusion

The project demonstrates how Business Intelligence can transform raw data into actionable insights. By analyzing customer feedback, agencies gain a clearer understanding of service gaps and areas needing improvement.

Although complaints dominate the dataset, commendations reveal valuable opportunities to promote a culture of recognition. With the future integration of machine learning, the system can also reduce manual workloads by automating employee access rights.

This comprehensive approach not only improves customer satisfaction but also enhances operational efficiency within the organization.