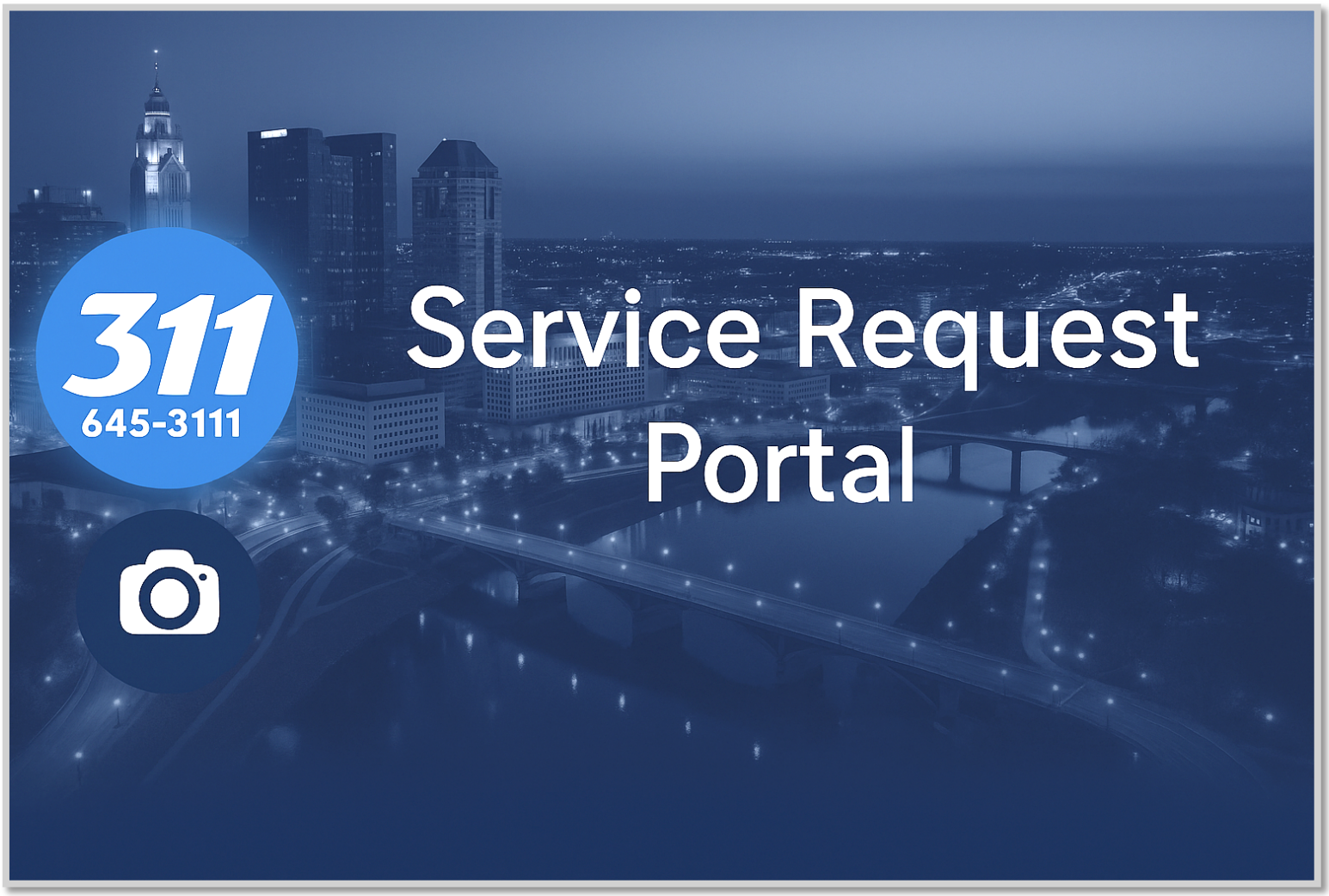

City 311 Service Requests Dashboard

DataWalkers | Dabot.ai – Data Analyst Challenge

Presented by: Neharika Dara

Tools Used: Python (Pandas): Data Cleaning & Pre-Processing

Power BI: Data Modeling, Visualization & Automation Alerts

Executive Summary

This project analyzes **150 K 311 Service Requests** recorded between **December 2017 and March 2018**.
The goal is to understand when and where citizens most frequently submit requests, how quickly departments resolve them, and where automation or AI-based alerts could improve efficiency.

Using **Python** for cleaning and **Power BI** for visualization, the project converts raw civic data into a live dashboard that highlights patterns, bottlenecks, and opportunities for predictive action.

Objectives

- Identify trends by **time of day**, **weekday**, and **month**.
- Evaluate **department efficiency** using average resolution time.
- Compare **open vs. closed cases** to gauge backlog.
- Build a foundation for **automation & AI-driven alerts** within Power BI.

Dataset Overview

Attribute	Description
Source	311 City Service Requests – Datawalker.ai
Records Processed	**150 244 unique cases** after cleaning
Time Period Covered	December 2017 – March 2018
Main Columns (24 total)	Case ID, Opened, Closed, Updated, Status, Responsible Agency, Category, Request Type, Neighborhood, Latitude, Longitude, Resolution Time (Days), Case Status, etc.
Geographic Coverage	Entire city area with ≈ 99 % mapped records
Open vs Closed Cases	25 % open 75 % closed
Average Resolution Time	≈ 2 days (across closed cases)
Longest Delays	≈ 2 400 cases > 24 days to close

Data Preparation & Quality Observations

Overview

Before creating the Power BI dashboard, the raw 311 dataset was thoroughly cleaned and validated using **Python (Pandas)**.
This ensured that every timestamp, category, and geographic coordinate was accurate and consistent for analysis. Proper cleaning was essential to generate the 150K+ record dataset used in your final dashboard KPIs.

Data quality is the foundation for reliable insights.
Therefore, each field was systematically reviewed to identify missing values, invalid entries, duplicates, and inconsistent timestamps.

Data Quality Summary

After cleaning, the dataset contained **150,244 valid records** with **no duplicate Case IDs** and **no invalid or future dates**. Nearly **99% of all service requests were geocoded**, ensuring strong spatial accuracy for mapping. Over **37,000 missing resolution fields** were logically replaced with `-1` to represent open cases, allowing these records to remain part of the analysis. In total, **24 key columns were validated**, resulting in a clean, reliable dataset that was fully ready for Power BI visualization and performance analytics.

Step	Action Taken	Purpose
1. **Missing Value Handling**	Replaced nulls in *Resolution_Time* and related columns with -1 to represent open/unresolved cases.	Keeps open requests in the dataset without distorting averages.
2. **Date Conversion**	Converted Opened, Closed, and Updated columns to datetime format.	Enables accurate calculation of resolution durations and trends by time.
3. **Resolution Metrics**	Created Resolution_Time_days = (Closed - Opened).	Forms the base for KPI measures like "Avg Resolution Days."
4. **Flag Creation**	Added Is_Open (1=Open, 0=Closed) and Geo_Status (Mapped/Not Mapped).	Simplifies filtering in Power BI for case status and mapping visuals.
5. **Latitude & Longitude**	Missing coordinates were labeled as "Not Mapped" instead of removing records.	Maintains 99% geographic coverage for mapping visuals.
6. **Duplicate Check**	Verified using df.duplicated() — none found.	Ensures dataset uniqueness and accuracy.

Key Observation

By assigning logical placeholders like `-1` instead of removing incomplete data, every case remains part of the analysis. This allows the dashboard KPIs to accurately distinguish between **open**, **closed**, and **delayed** cases — producing a more realistic operational view for decision-makers.

Feature Engineering & Data Modeling

Overview

After cleaning the dataset in Python, the next step was to enhance it with analytical features and model it effectively in **Power BI**.
Feature engineering added calculated fields that power the KPIs, filters, and trend visuals, while the data model ensures accurate relationships between dates, requests, and agencies.

New Column	Logic / Formula	Purpose
Year	Year = YEAR([Opened])	Enables yearly filtering and comparisons.
Month / Month-Year	FORMAT([Opened], "MMM YYYY")	Allows trend visuals by month and clear timeline breakdown.
Hour	HOUR([Opened])	Supports the "Requests by Hour" visual.
Weekday	FORMAT([Opened], "dddd")	Enables weekday pattern analysis.
Resolution_Time_days	[Closed] – [Opened] (calculated in Python)	Forms the base for Average Resolution metrics.
Is_Open	IF(LOWER([Status]) = "closed", 0, 1)	Binary indicator for open or closed cases.

Data Model Structure

- **Fact Table:** 311_Cases_Final_Cleaned
  → Contains all service requests and derived metrics.
- **Date Table (DAX Generated):** Date_Table = CALENDAR(MIN('311_Cases_Final_Cleaned'[Opened]), MAX('311_Cases_Final_Cleaned'[Opened]))
- Marked as the official **Date Table** for all time intelligence functions.
- **Relationships:**
- Date_Table[Date] → 311_Cases_Final_Cleaned[Opened]
- Ensures consistent filtering for all visuals using date hierarchy (Year, Month, Day, Hour).

Model Validation

- Verified relationship integrity (no "many-to-many" joins).
- Confirmed DAX measures compute correctly across filters.
- Checked total row counts between Power BI and Python output (150 244).
- Ensured no circular dependencies.

Outcome

The dataset was transformed into a **clean, dynamic data model** ready for Power BI visualizations.
This model supports all KPI cards, date-based filters, and interactive slicers (Responsible Agency, Year, Month) — ensuring seamless analysis and automation setup.

KPI Design & Summary Metrics

Overview

To summarize performance and provide instant insights, seven **Key Performance Indicators (KPIs)** were created in Power BI using **DAX formulas**. These metrics automatically update when filters such as *Agency*, *Year*, or *Month* are applied, ensuring dynamic and accurate performance tracking.

KPI Name	DAX Formula (Simplified)	Purpose / Interpretation
KPI Name	**DAX Formula (Simplified)**	**Purpose / Insight**
Total Requests	COUNTROWS('311_Cases_Final_Cleaned')	Displays the total volume of service requests (150K).
Open Requests	CALCULATE(COUNTROWS('311_Cases_Final_Cleaned'), '311_Cases_Final_Cleaned'[Is_Open] = 1)	Shows all active, unresolved requests (37K).
Closed Requests	CALCULATE(COUNTROWS('311_Cases_Final_Cleaned'), '311_Cases_Final_Cleaned'[Is_Open] = 0)	Represents successfully resolved requests (113K).
Average Resolution Days	AVERAGE('311_Cases_Final_Cleaned'[Resolution_Time_days])	Measures mean turnaround time for all closed cases (3 days).
Delayed Cases (>24 Days)	CALCULATE(COUNTROWS('311_Cases_Final_Cleaned'), '311_Cases_Final_Cleaned'[Resolution_Time_days] > 24)	Identifies long-duration or bottleneck cases (2,507).
Open >7 Days	CALCULATE(COUNTROWS('311_Cases_Final_Cleaned'), DATEDIFF('311_Cases_Final_Cleaned'[Opened], TODAY(), DAY) > 7)	Highlights aging open requests (37K).

Insights from KPIs

- The system processed **150K total requests**, closing **75%** of them.
- Only **3 average days** are needed to close a typical request — strong performance.
- However, **25% remain open**, including **2.5K delayed** and **37K aged (7+ days)**, showing improvement opportunities in case follow-up and automation.

Outcome

- The KPI layer serves as the analytical backbone of the dashboard.
  It provides real-time operational intelligence — helping city supervisors monitor request loads, detect aging cases, and ensure service teams maintain efficiency.

Time Series & Trend Analysis Visuals

Overview

Time-based analysis is essential to understand when citizens are most active in submitting service requests.
By examining request patterns by **month**, **weekday**, and **hour of day**, city teams can plan staff schedules and allocate resources proactively. The visuals in this section provide a complete time-series breakdown to highlight demand fluctuations throughout the dataset period.

Requests by Month (Pie Chart)

- **Visual Type:** Donut Chart
- **Field:** `Month_Year`
- **Values:** Count of `CaseID`
- **Colors:**
  - Dec → Light Blue (#A8DADC)
  - Jan → Teal (#008080)
  - Feb → Sky Blue (#00B3B3)
  - Mar → Navy Blue (#003366)

Insight

Request volume peaks in **January 2018 (≈ 32%)**, followed by **February (27%)** and **December (26%)**.
This pattern suggests seasonal demand — likely linked to post-holiday maintenance, cleanup, and weather-related reports.
Requests start declining by **March**, indicating reduced issue reporting toward spring.

Requests by Weekday (Bar Chart)

- **X-Axis:** `Weekday` (Mon–Sun)
- **Y-Axis:** Count of `CaseID`
- **Bar Color:** Teal Blue (#008080)
- **Labels:** White font inside bars
- **Sort Order:** Monday → Sunday

Insight

Most requests are submitted on **Tuesdays, Wednesdays, and Thursdays** (~25K each), while **weekends** show a 30% drop.
This reveals that weekday field staff capacity directly affects response rates, as citizens tend to log service issues during workdays.

Requests by Hour of Day (Line Chart)

- **X**-Axis: **Hour (0–23)**
- Y-Axis: **Count of CaseID**
- Line Style: **Teal line with circular markers (6 px)** Background: **Transparent, light gridlines**

Insight

Reporting activity starts increasing after **6 AM**, peaking between **8 AM and 11 AM**, where over **13.9K requests** were logged around **10 AM**.
After noon, the volume gradually declines, reaching its lowest at midnight.
This shows that citizens primarily engage during **morning hours**, aligning with field team schedules and daily commute times.

Overall Time-Based Insights

- **Peak Activity:** Early mornings and mid-week days.
- **Low Activity:** Weekends and late nights.
- **Operational Insight:** Staff should be allocated heavily in morning shifts (8–12 PM) and mid-week to handle the majority of cases efficiently.

Outcome

These visuals provide time-intelligence for planning, workload prediction, and future forecasting.
By integrating this analysis with Dabot.ai automation, the city can preemptively adjust resources during known high-demand periods — improving response efficiency and citizen satisfaction.

Agency & Category Performance Analysis

Overview

This page analyzes how city departments (agencies) and service categories perform in handling citizen 311 requests. By comparing **total request volume** and **average resolution time**, it highlights which departments are efficient, which face delays, and where operational improvements or automation could be introduced.

Top Responsible Agencies (Horizontal Bar Chart)

- **X-Axis:** Count of `CaseID`
- **Y-Axis:** `Responsible Agency` (sorted descending)
- **Bar Color:** Navy Blue (#003366)
- **Data Labels:** White text inside bars
- **Tooltip:** Agency Name, Total Requests, Avg Resolution (Days)

Insight

The **DPW Ops Queue** manages the **highest workload (~63 K cases)**, focusing on street cleaning and illegal dumping. Other key agencies include **Recology Abandoned (25 K)**, **DPT Abandoned Vehicles (8 K)**, and **Muni Work Queue (4 K)**. This clearly establishes **DPW Ops** as the city's operational backbone and signals a potential need for greater resource allocation or automation support to maintain performance under heavy demand.

Average Resolution Days by Agency (Column Chart)

- **X-Axis:** `Responsible Agency`
- **Y-Axis:** Average of `Resolution_Time_days`
- **Column Color:** Teal Blue (#008080)
- **Data Labels:** On (gray font)
- **Title:** "Average Resolution Time by Agency"

Insight

While most agencies resolve cases within **2–3 days**, the **DPH Homeless Outreach Team** records the **longest average resolution (~60 days)**, followed by the **DPT Paint Shop (~53 days)** and **DPT Meters (~45 days)**. These extended durations suggest complex case types—often requiring coordination across multiple departments or external partners.

Top Service Categories (Bar Chart)

- **X-Axis:** Count of `CaseID`
- **Y-Axis:** `Category` (sorted descending)
- **Bar Color:** Teal Blue (#008080)
- **Tooltip:** Category, Count, Avg Resolution (Days)

Insight

The most reported categories are **Street and Sidewalk Cleaning (~50 K)**, **Graffiti (~22 K)**, **Encampments (~16 K)**, and **Homeless Concerns (~12 K)**.
Together, these make up nearly **60 %** of total 311 requests—highlighting citizens' top priorities around **cleanliness, sanitation, and public safety**.

Average Resolution by Category (Column Chart)

- **X-Axis:** `Category`
- **Y-Axis:** Average of `Resolution_Time_days`
- **Color Gradient:** Teal → Orange (for longer delays)
- **Title:** "Average Resolution by Service Category"

Insight

Administrative categories such as **General Request – Economics and Workforce Development (~40 days)**, **Rent Board (~35 days)**, and **Temporary Sign Request (~34 days)** take the longest to close.
Field-response categories like **Street and Sidewalk Cleaning** and **Graffiti** resolve quickly, usually within **1–3 days**, demonstrating strong on-ground efficiency.

Summary Observation

- **DPW Ops Queue** dominates case volume yet maintains good efficiency.
- **Homeless Outreach and permit-related departments** require process optimization to reduce backlogs.
- **Administrative categories** show the longest resolution times due to approval complexities.
- The city can enhance performance through automation of routine tasks and better resource distribution toward high-impact areas.

Dashboard Summary & Key Insights

Overview

The **311 Services Analysis & Automation Dashboard** provides a complete view of how citizen service requests are received, processed, and resolved across city departments.
Using data from **December 2017 to March 2018**, the dashboard combines real-time metrics, time-based patterns, and performance comparisons to evaluate city responsiveness.
Each visual automatically updates with filters for **Agency**, **Year**, and **Month**, enabling quick analysis and decision-making.

KPI Highlights

These KPIs create an at-a-glance summary of operational performance.
The **low average resolution time** shows efficiency, while the **25% open ratio** highlights areas for ongoing improvement.

KPI	Value	Meaning
Total Requests	150 K	Overall volume of service requests submitted.
Open Requests	37 K	Cases that remain unresolved or in progress.
Closed Requests	113 K	Successfully completed requests.
Average Resolution Time	3 days	Average number of days to resolve a request.
Delayed Cases (>24 days)	2,507	Long-duration cases indicating bottlenecks.
Open >7 Days	37 K	Ongoing issues taking over a week to resolve.
% Open Requests	24.83 %	Share of currently open cases in total workload.

Trend Analysis Insights

Requests by Hour

- Activity rises sharply after **7 AM**, peaking at **10 AM (≈ 13.9 K requests)**.
- Late-night reporting is minimal (<2%), showing citizen engagement mostly during daytime.
- Insight → Morning hours should have maximum support staff to address new cases faster.

Requests by Weekday

- **Tuesday to Thursday** is busiest (≈ 25–26 K each).
- Weekends drop significantly (Saturday/Sunday ≈ 16–17 K).
- Insight → Scheduling more field teams' mid-week would optimize service responsiveness.

Requests by Month

- **January 2018** saw the highest activity (≈ 32% of total cases), followed by **February (27%)**.
- **December** shows early ramp-up post-holiday, while **March** begins declining.
- Insight → Seasonal spikes in early-year months may relate to post-holiday maintenance needs

Performance Insights by Agency & Category

- **DPW Ops Queue** is the city's busiest and most efficient agency, resolving thousands of cases rapidly.
- **Homeless Outreach** and **Permitting departments** have longer resolution cycles (30–60 days), signaling a need for process streamlining.
- **Street Cleaning** and **Graffiti Removal** dominate volume (50 K+ requests), emphasizing cleanliness as the city's top citizen concern.

Automation & Alert Integration

The dashboard includes alert mechanisms that can be expanded for automation:

- ⚠️ Highlight delayed cases (>24 days) for real-time escalation.
- 📩 Trigger notifications to supervisors when open cases exceed weekly thresholds.
- 🔄 Enable AI-based forecasting for expected request volume by hour/day.

This automation framework aligns with **Dabot.ai's mission** turning static reports into dynamic, intelligent systems that detect issues before they escalate.

Conclusion

The 311 Services Dashboard successfully transforms raw data into actionable insights for public service improvement.
It enables the city to **monitor workload, track efficiency, predict trends, and proactively resolve issues**—all in one interactive platform.
This project demonstrates how **Python data cleaning + Power BI automation** can work together to build scalable, AI-ready analytics solutions.

Recommendations & Future Enhancements

Overview

The 311 Services Dashboard reveals strong opportunities to enhance **efficiency, automation, and predictive analytics** within city operations.
These recommendations aim to help departments act faster, reduce backlogs, and make data-driven decisions through Power BI and AI integration.

Operational Improvements

- **Optimize Scheduling:** Deploy more field staff during peak days (Tuesday–Thursday) and morning hours (8–11 AM) when request volume is highest.
- **Faster Coordination:** Automate inter-department workflows—especially between **DPH Homeless Outreach** and **DPW Ops**—to shorten long-resolution cases.
- **Digital Tracking:** Use automated status updates for permitting and administrative requests to minimize manual follow-ups.
- **Efficient Routing:** Assign incoming service requests automatically to specialized DPW and Recology teams to prevent backlogs.

Automation & Alert Mechanisms

- **Delay Alerts:** Power BI can trigger notifications when cases remain open beyond 7 or 24 days.
- **Geo Alerts:** Highlight neighborhoods with high unresolved request clusters.
- **Demand Spike Detection:** Send alerts when daily requests exceed normal volume thresholds.
- **Chatbot Integration:** Through Dabot.ai, enable citizens to track or update their requests using an AI assistant.

AI & Predictive Opportunities

- **Forecasting:** Use historical data to predict daily/hourly request volumes for proactive scheduling.
- **Anomaly Detection:** Identify sudden increases in specific categories like illegal dumping or graffiti.
- **Geo Clustering:** Map problem hotspots to improve area-specific planning.
- **Sentiment Insights:** Analyze citizen comments in “Status Notes” for satisfaction or recurring issues.

Future Vision

Integrating Power BI automation with **Dabot.ai’s AI capabilities** will transform the 311 system from reactive reporting to proactive management.
Real-time dashboards, predictive alerts, and automated workflows will help the city resolve issues faster, improve transparency, and strengthen citizen trust.