

311 Service Requests Analysis – Kansas City, MO (2007–2021)

An end-to-end BI implementation: Data Quality Assessment, ETL, and Insights

Tools: Alteryx, SQL (MySQL), Tableau, Power BI

TABLE OF CONTENTS

1. Executive Summary.....	3
2. Data Quality Assessment.....	3
General Data Cleaning Strategies:.....	4
3. ETL & Data Preparation.....	6
4. Key Insights & Strategic Recommendations.....	7
5. Appendix.....	8
5.1 Tableau Dashboards:.....	9
5.2 PowerBI Dashboards:.....	12
5.3 SQL Query:.....	16

1. Executive Summary

This report analyzes over 1.5 million records from Kansas City's 311 Call Center to uncover patterns in civic service operations. The project includes data profiling, ETL design, and dashboard development to answer key business questions around workload, efficiency, and geographic distribution of service requests.

2. Data Quality Assessment

Dataset Overview:

This dataset comprises 1.5M+ service requests submitted to Kansas City's 311 Call Center. Each entry includes fields such as case ID, creation/closure dates, request type, department, status, and geospatial coordinates. A detailed profiling of these fields revealed several data quality issues requiring attention prior to analysis.

Initial profiling of the dataset revealed several issues including missing values, inconsistent formats, and outliers. Below is a summary of the most critical data quality concerns and proposed resolutions:

DAYS TO CLOSE	1.7% missing values, outliers (e.g., 1154)	Impute with median or flag/remove outliers
CLOSED DATE	0.8% missing, inconsistent formats	Standardize format, impute if necessary
CATEGORY2	64.1% missing	Assess value, drop or impute
CATEGORY3	89.9% missing	Exclude from analysis due to high sparsity
ZIP CODE	0.1% missing, inconsistent types	Normalize format, impute
POLICE DISTRICT	2.3% missing, inconsistent entries	Standardize or impute
NEIGHBORHOOD	2.9% missing	Impute with most frequent value or predictive modeling

COUNTY	4.3% missing	Exclude or use secondary lookup
--------	--------------	---------------------------------

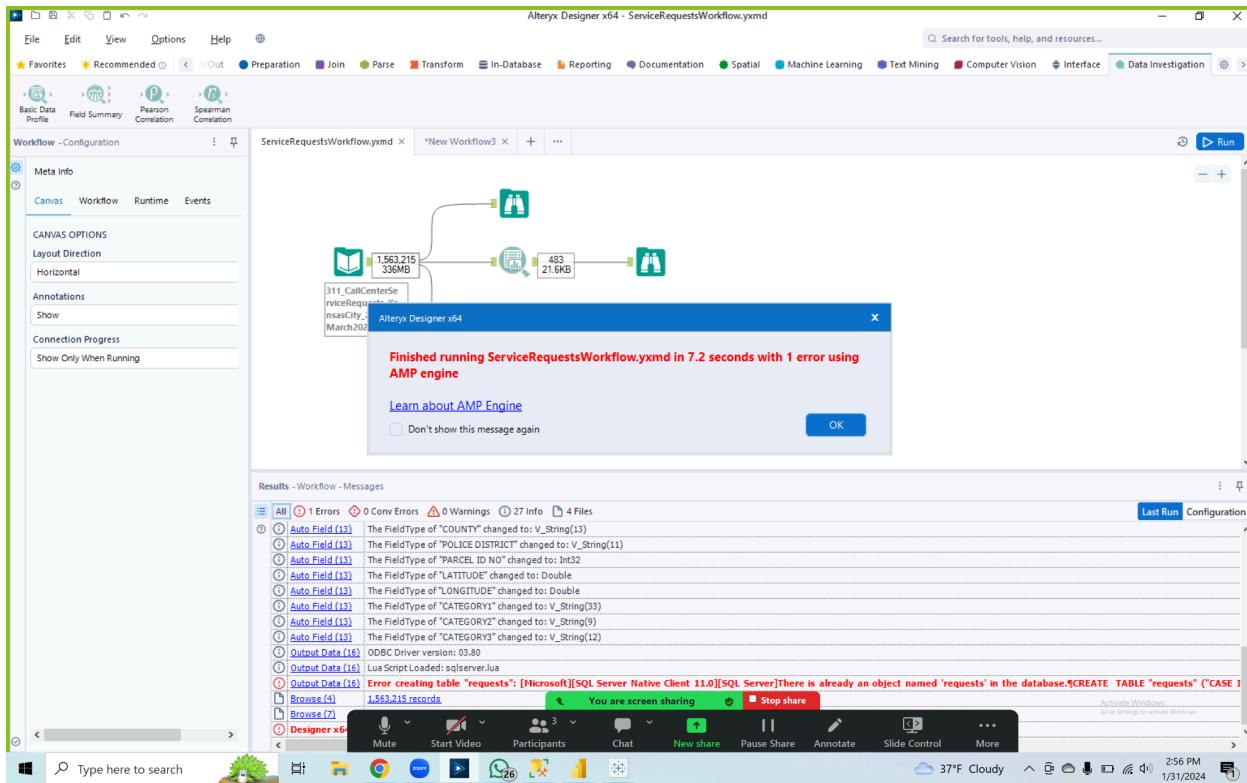
General Data Cleaning Strategies:

- Normalize string fields and eliminate leading/trailing whitespace
- Convert all date/time fields to ISO-8601 standard format
- Remove or flag duplicate case IDs
- Investigate low-frequency categorical values and consolidate if appropriate

In conclusion, the dataset presented multiple data integrity challenges including missing data, inconsistent formats, and outliers. A structured data cleaning plan, particularly focused on temporal and categorical attributes, was critical to ensuring downstream reporting and analytics reliability.

The screenshot shows the Alteryx Designer interface with the following components:

- Toolbar:** File, Edit, View, Options, Help, etc.
- Top Bar:** Search for tools, help, and resources...; Data Investigation tab.
- Left Panel:**
 - Favorites, Recommended, Out, Preparation, Join, Parse, Transform, In-Database, Reporting, Documentation, Spatial, Machine Learning, Text Mining, Computer Vision, Interface.
 - Browse (4) - Configuration: Shows 1,563,215 records displayed, 23 fields, 133 MB. Includes a table for CASE ID and a summary for SOURCE (PHONE: 1,204,236, WEB: 211,721, EMAIL: 80,585, SYS: 19,226, INSPE: 14,690).
 - DEPARTMENT: NHS (783,094), Public Works (353,787), Water Services (216,852), Parks and Rec (87,954), Health (39,543).
 - WORK GROUP: 22 more >.
- Middle Panel:** ServiceRequestsWorkflow.ymd workflow diagram. It starts with a "311_CallCenterServiceRequests_KansasCity_2007-March2021.tsv" input file, which is connected to a "Parse" tool. The output of the Parse tool is then connected to a "Join" tool, which is followed by a "Transform" tool. Finally, the output of the Transform tool is connected to a "dcm:ServiceRequests Table=requests" destination.
- Bottom Panel:** Results - Browse (4) - Input. A table showing 23 of 23 Fields with 1,563,215 records displayed, 133 MB. The columns include Record, CASE ID, SOURCE, DEPARTMENT, WORK GROUP, TYPE, DETAIL, CREATION DATE, and CREATION TIME. The data includes various service requests like Dangerous Buildings, Parks Maintenance, Recycling, and Street Light issues.



3. ETL & Data Preparation

Using Alteryx, the raw TSV data was profiled, validated, and loaded into a MySQL relational database. The ETL process included staging tables, field transformations, and format standardization to support ad-hoc reporting and BI visualization.

The screenshot shows the Microsoft SQL Server Management Studio interface. On the left, the Object Explorer displays the database structure, including the ServiceRequests database and its tables. In the center, a query window titled 'SQLQuery2.sql - not connected' contains a T-SQL script. The script starts with a comment block, followed by a select statement to find the top 10 cases for 'Animal' type bites. It then alters the 'requests' table to change the 'CREATION_DATE' column to a DATETIME type. Subsequent queries show the creation date distribution, a grouped count by longitude and latitude, and a total row count. At the bottom, a results grid titled 'Results of Messages' shows the execution of the query, displaying 21 rows of data from the 'requests' table. The columns include CASE_ID, SOURCE, DEPARTMENT, WORK_GROUP, TYPE, DETAIL, CREATION_DATE, CREATION_TIME, STATUS, EXCEEDED EST TIMEFRAME, CLOSED_DATE, and DAYS TO CLOSE.

CASE_ID	SOURCE	DEPARTMENT	WORK_GROUP	TYPE	DETAIL	CREATION_DATE	CREATION_TIME	STATUS	EXCEEDED EST TIMEFRAME	CLOSED_DATE	DAYS_TO_CLOSE
1	2019119972	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2019-06-24 00:00:00.000	07:40 AM	RESOL Y	11/19/2021	879
2	2019207923	WEB	Public Works	Public Works-Street and Traffic-District 1	Crack	District 1	2019-12-22 00:00:00.000	07:56 PM	RESOL Y	06/26/2020	187
3	2021005978	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2021-01-19 00:00:00.000	02:43 AM	RESOL Y	11/19/2021	300
4	2020054713	PHONE	NHS	NHS-Environmental Protection-	Property Maintenance	Standard	2020-04-18 00:00:00.000	07:40 AM	RESOL Y	04/20/2021	152
5	2020054721	WEB	Parks and Rec	Parks and Rec-Central Region	Park Maintenance	Central	2020-04-18 00:00:00.000	05:10 AM	RESOL N	04/30/2020	12
6	2019182182	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2019-10-21 00:00:00.000	10:29 AM	RESOL Y	08/03/2020	287
7	2019184705	WEB	NHS	NHS-Solid Waste-	Recycling	Missed by City	2019-10-25 00:00:00.000	10:02 AM	RESOL N	10/28/2019	3
8	2019184590	WEB	Parks and Rec	Parks and Rec-Landscape Services-Forestry	Trimming	Tree Limbs	2019-10-25 00:00:00.000	04:44 AM	RESOL Y	12/04/2019	40
9	2020054493	PHONE	NHS	NHS-Environmental Protection-	Dangerous Building	Standard	2020-04-07 00:00:00.000	06:45 AM	RESOL Y	11/29/2021	504
10	2019094468	PHONE	NHS	NHS-Solid Waste-Administration	Services	Service issue / Problem	2019-09-07 00:00:00.000	07:45 PM	RESOL Y	08/27/2018	294
11	2020021142	PHONE	Water Services	Water Services-Meer and Field Services-	Leak	Meer	2020-02-11 00:00:00.000	03:37 PM	RESOL N	02/12/2020	1
12	2019208958	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-24 00:00:00.000	10:04 PM	RESOL N	01/07/2020	14
13	2019208957	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-24 00:00:00.000	10:05 PM	RESOL N	01/07/2020	14
14	2019208956	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-25 00:00:00.000	07:28 PM	RESOL N	01/07/2020	13
15	2019208959	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-25 00:00:00.000	07:30 PM	RESOL N	01/07/2020	13
16	2019208977	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-25 00:00:00.000	08:41 PM	RESOL N	01/07/2020	13
17	2019208978	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-25 00:00:00.000	08:42 PM	RESOL N	01/07/2020	13
18	2019208979	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-12-25 00:00:00.000	08:42 PM	RESOL N	01/07/2020	13
19	2019055892	PHONE	NHS	NHS-National Park Preservation-	Property Maintenance	Other Property Issue	2019-05-15 00:00:00.000	07:19 PM	RESOL Y	09/17/2020	548
20	2020054492	PHONE	NHS	NHS-Neighborhood Preservation-Open Entry	Dangerous Building	To Entry	2020-05-15 00:00:00.000	04:05 PM	RESOL Y	08/27/2018	11
21	2019048659	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2019-04-11 00:00:00.000	07:33 AM	RESOL Y	08/03/2021	876

The screenshot shows the Microsoft SQL Server Management Studio (SSMS) interface. In the Object Explorer, a database named 'ServiceRequests' is selected. In the center pane, a query window titled 'SQLQuery2.sql - DESKTOP-DCV4P31.ServiceRequests (DESKTOP-DCV4P31\yuchi (63)) - Microsoft SQL Server Management Studio' contains the following T-SQL code:

```

use ServiceRequests;

-- DESKTOP-DCV4P31
select top 10 [CASE ID], ([DAYS TO CLOSE]) from requests where CATEGORY1 = 'Animal' AND TYPE = 'Bite' order by [DAYS TO CLOSE] ASC
exec sp_rename 'requests:[CREATION_DATE]', '[CREATION_DATE]', 'COLUMN';

ALTER TABLE requests
ALTER COLUMN CREATION_DATE DATETIME;

select COUNT(*) from requests;

select * from requests order by [DAYS TO CLOSE];

select top 10 * from requests group by LATITUDE,LONGITUDE order by(count([CASE ID])) desc;

SELECT [CASE ID], COUNT(*) FROM REQUESTS GROUP BY LONGITUDE, LATITUDE;

select count([case id]) AS 'TOTAL ROWS' from requests;

```

The results pane shows a single row of data with the column 'TOTAL ROWS' containing the value '1563215'. At the bottom of the screen, a status bar indicates 'Query executed successfully.' and other session details.

4. Key Insights & Strategic Recommendations

This section answers the following key business questions:

1. Service Request Trends Over Time

- **Request volumes peaked in 2019**, followed by a decline likely due to COVID-19's impact in 2020–2021.
- Month-wise analysis revealed **seasonal spikes** during summer months (May–August).

2. Request Source Breakdown

- **Phone and Web were the dominant channels**, with web usage increasing year-over-year—indicating a shift toward digital service engagement.

3. Departmental & Workgroup Load

- The **Public Works** and **Neighborhood Services** departments consistently received the highest number of requests.
- Disparity in workload across workgroups highlights potential resource reallocation needs.

4. Response Time & Efficiency

- **Average days to close** varied significantly by category and department.
- Categories such as "Street Maintenance" and "Illegal Dumping" had higher closure times, suggesting complex follow-up processes.
- **Top-performing departments** closed cases in under 2 days, whereas outliers exceeded 1000 days—highlighting potential backlogs or data entry issues.

5. Geographic & Neighborhood Analysis

- Certain ZIP codes (e.g., 64130, 64131) and council districts reported consistently high service request volumes.
- Tableau heatmaps revealed **concentrated clusters of activity** that can inform future infrastructure or staffing needs.

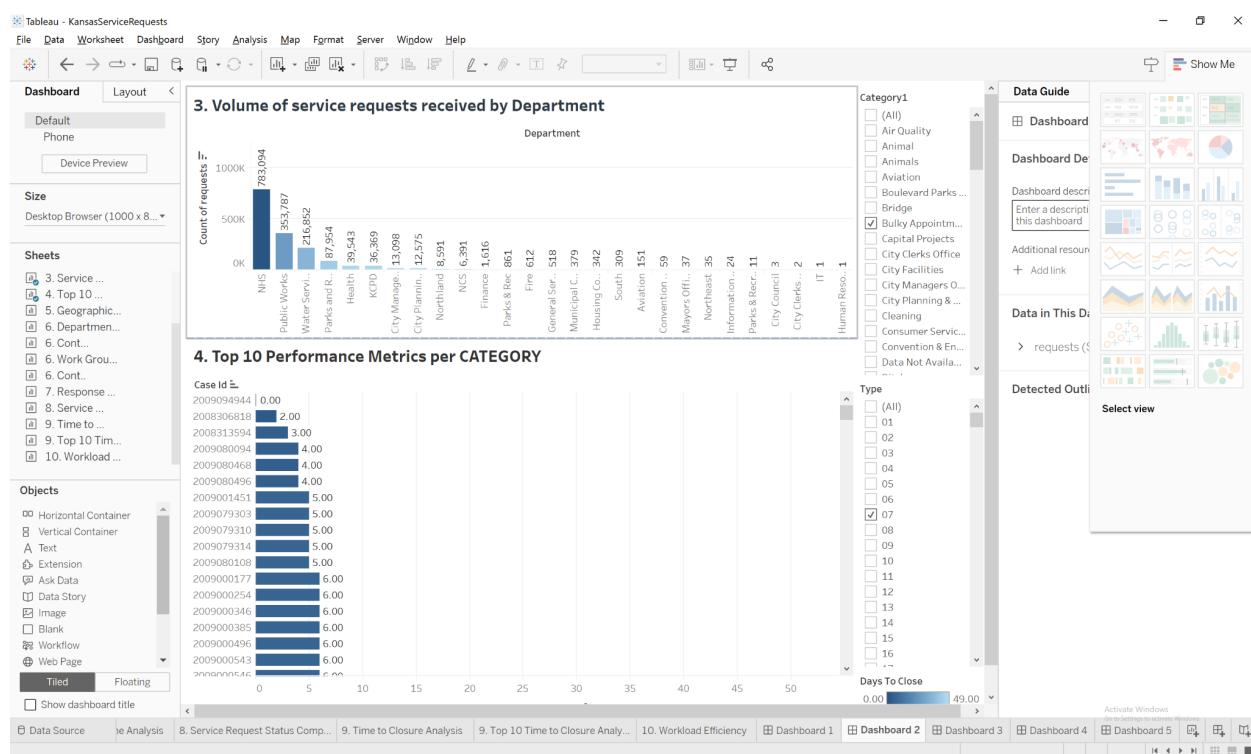
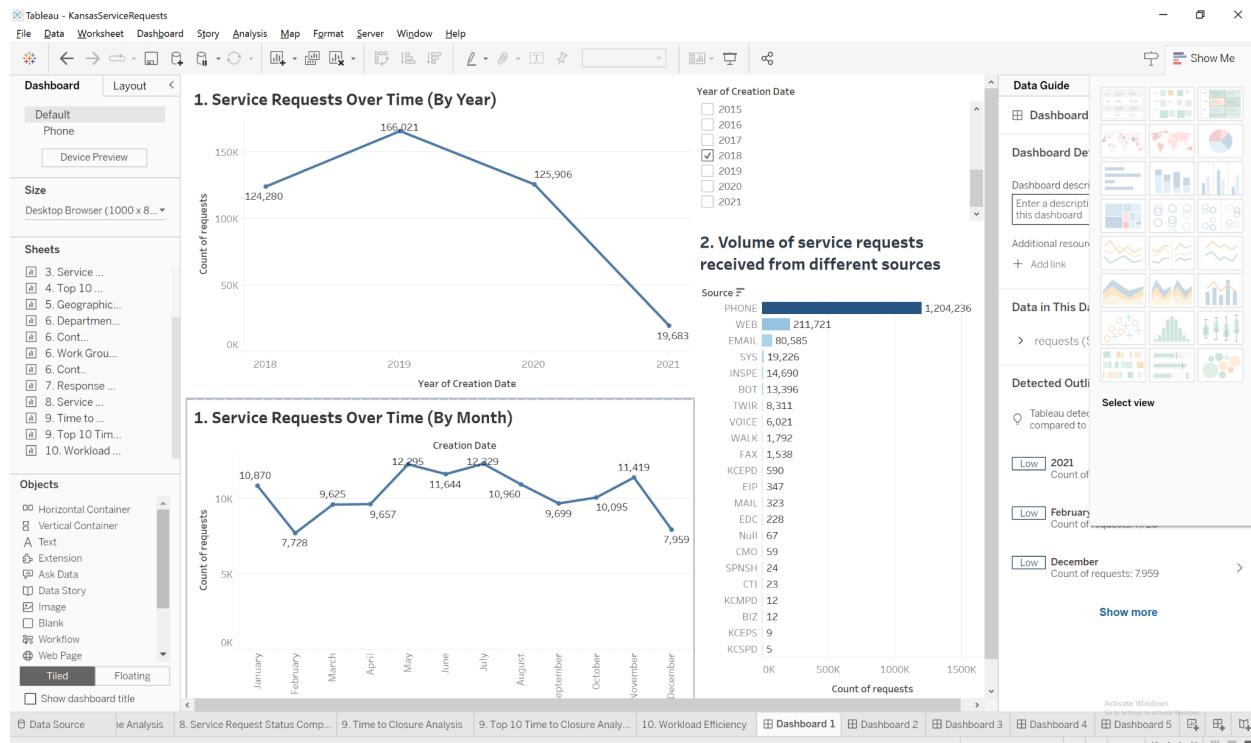
6. Status Composition & Resolution Rates

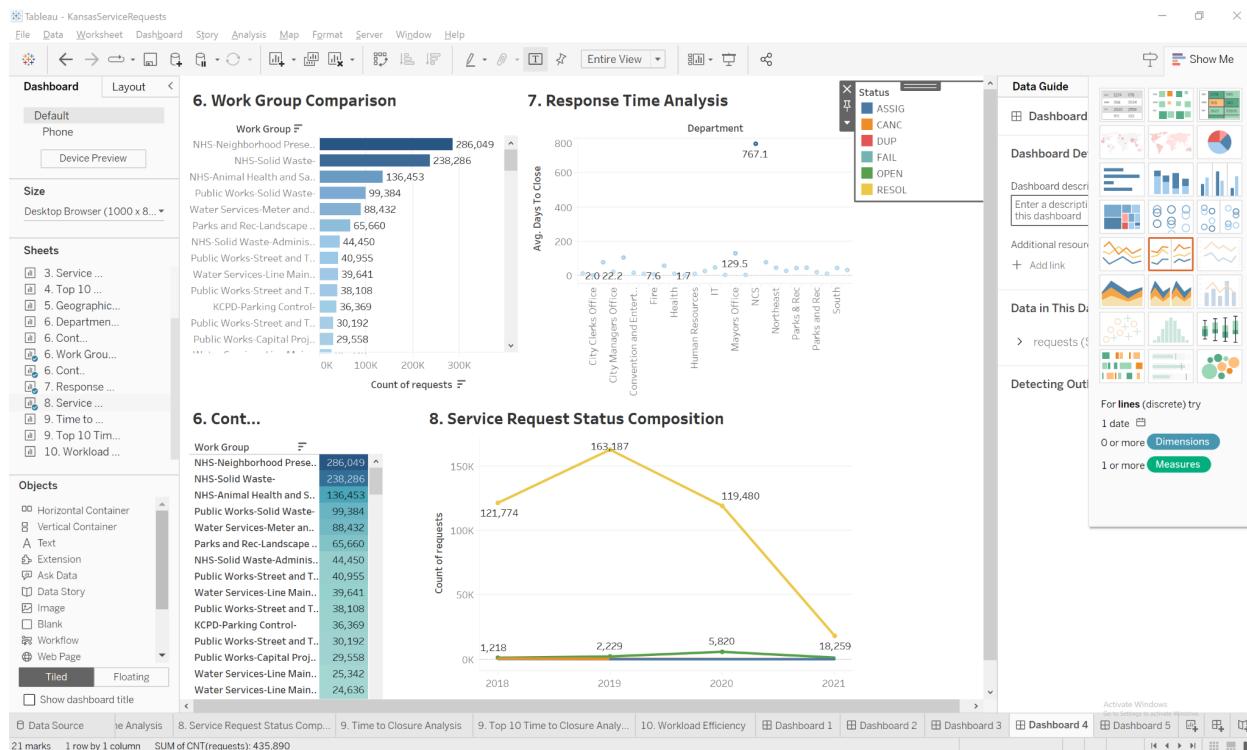
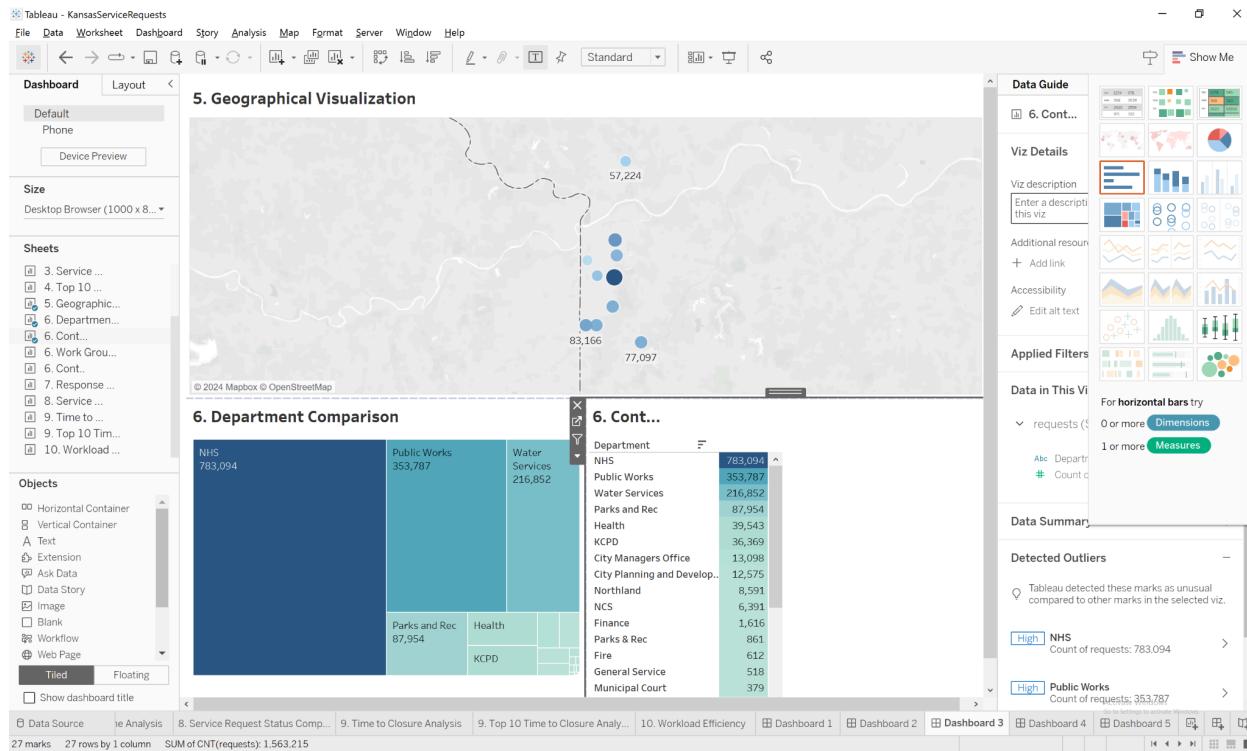
- Over the years, the proportion of **closed vs. open requests improved**, but the volume of long-pending open requests in certain neighborhoods remains an area for concern.

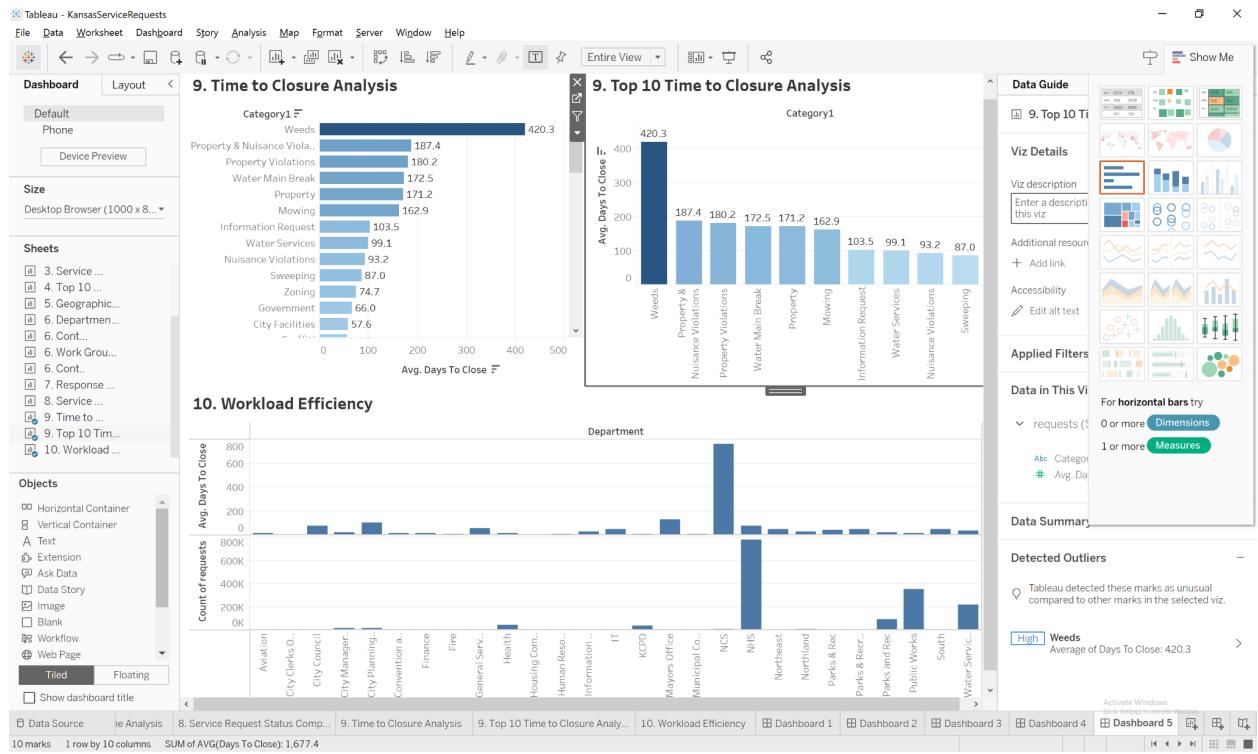
5. Appendix

The appendix includes SQL queries and screenshots of dashboards published on Tableau Public and Power BI.

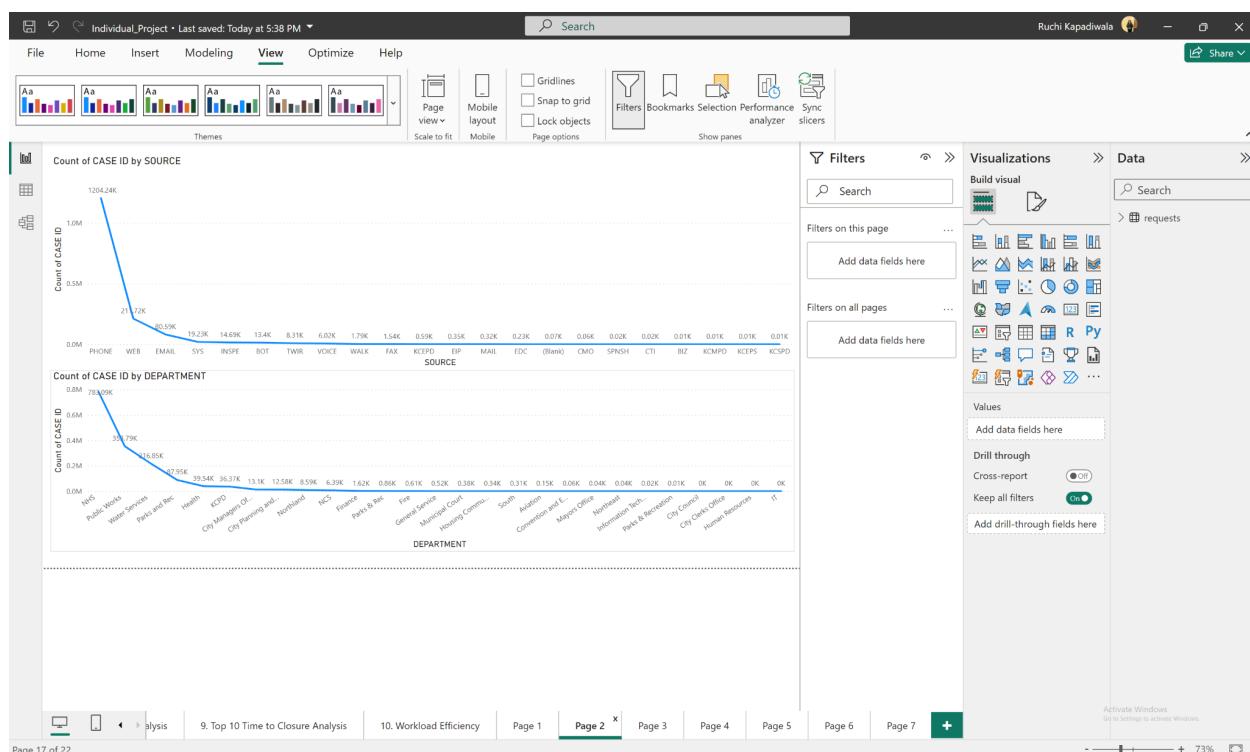
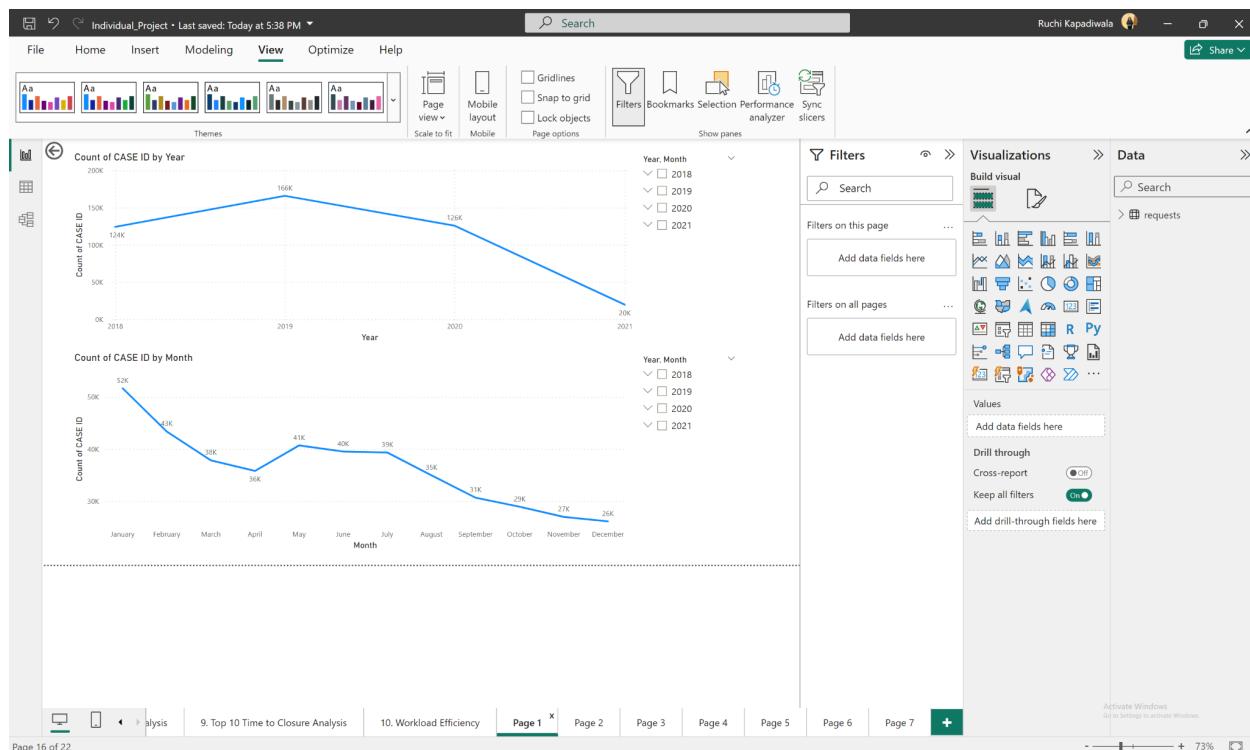
5.1 Tableau Dashboards:

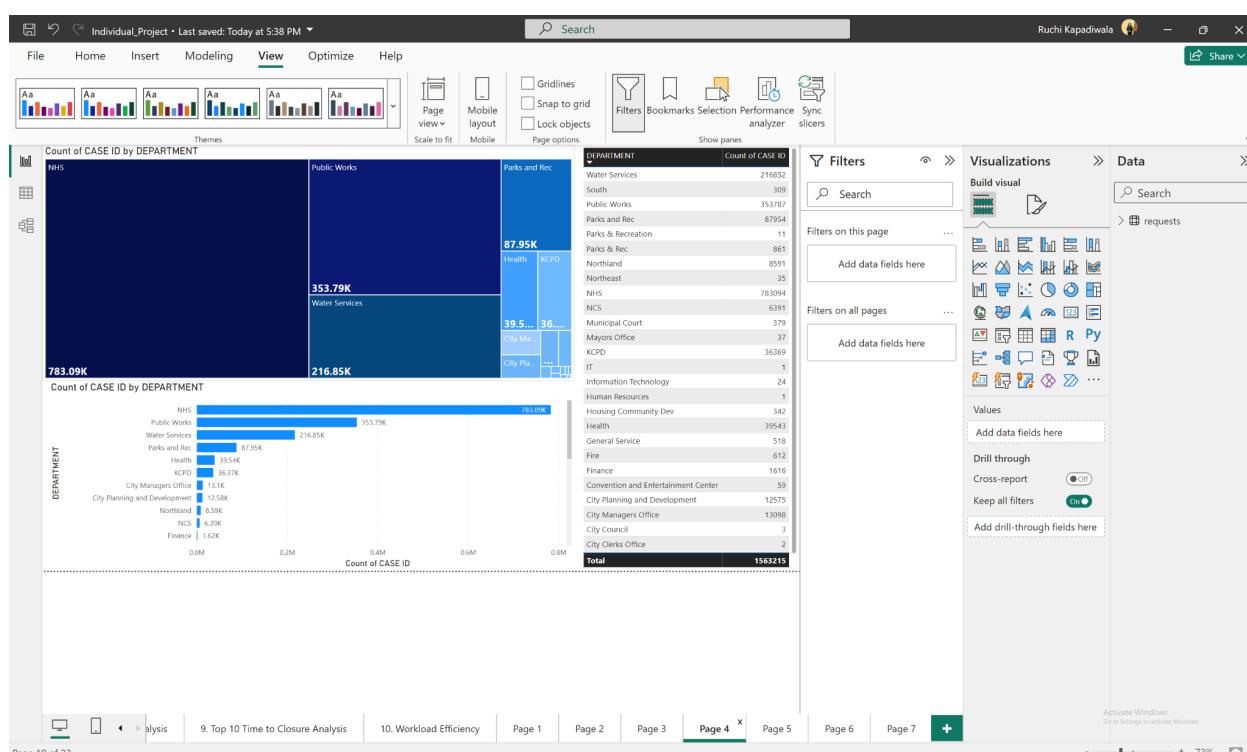
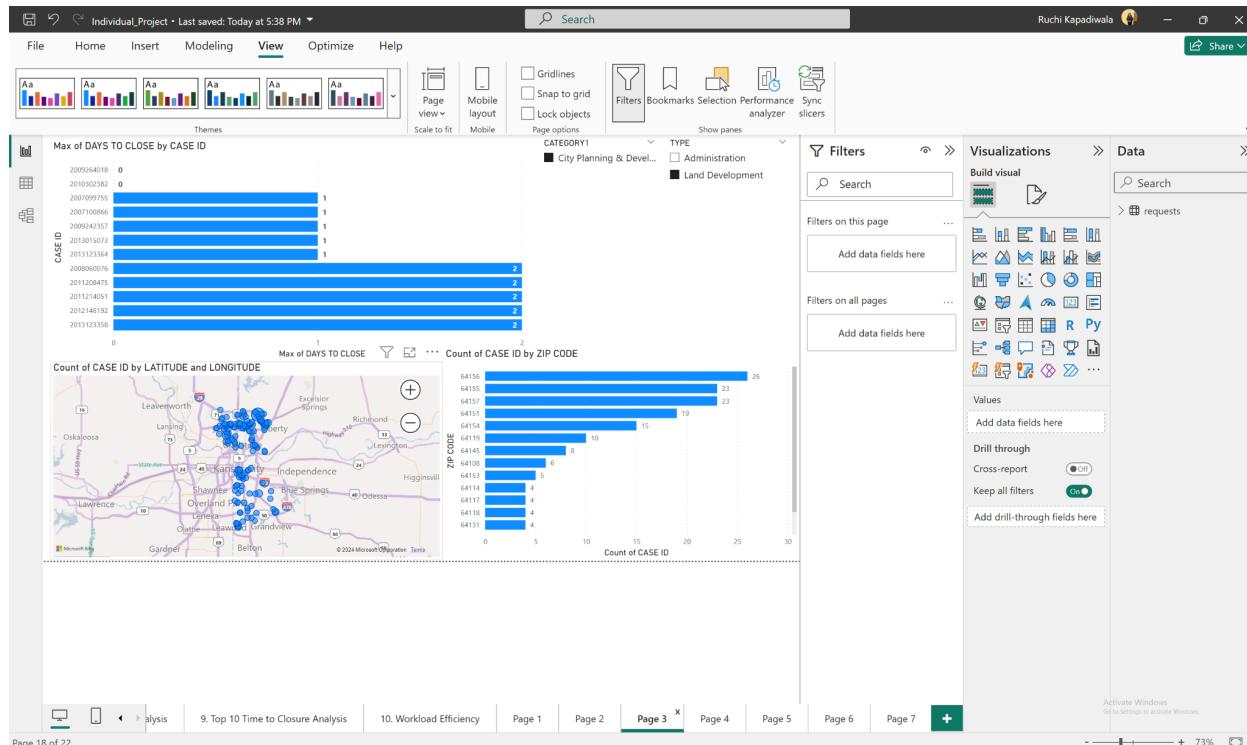


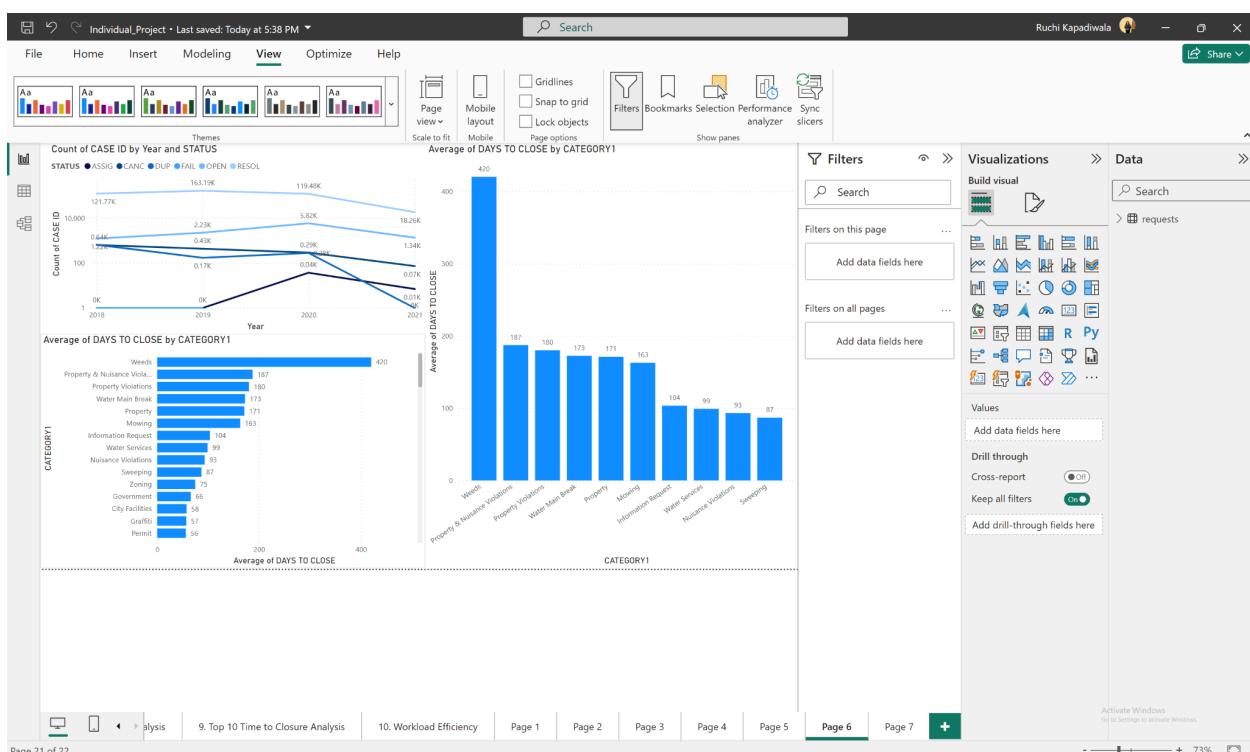
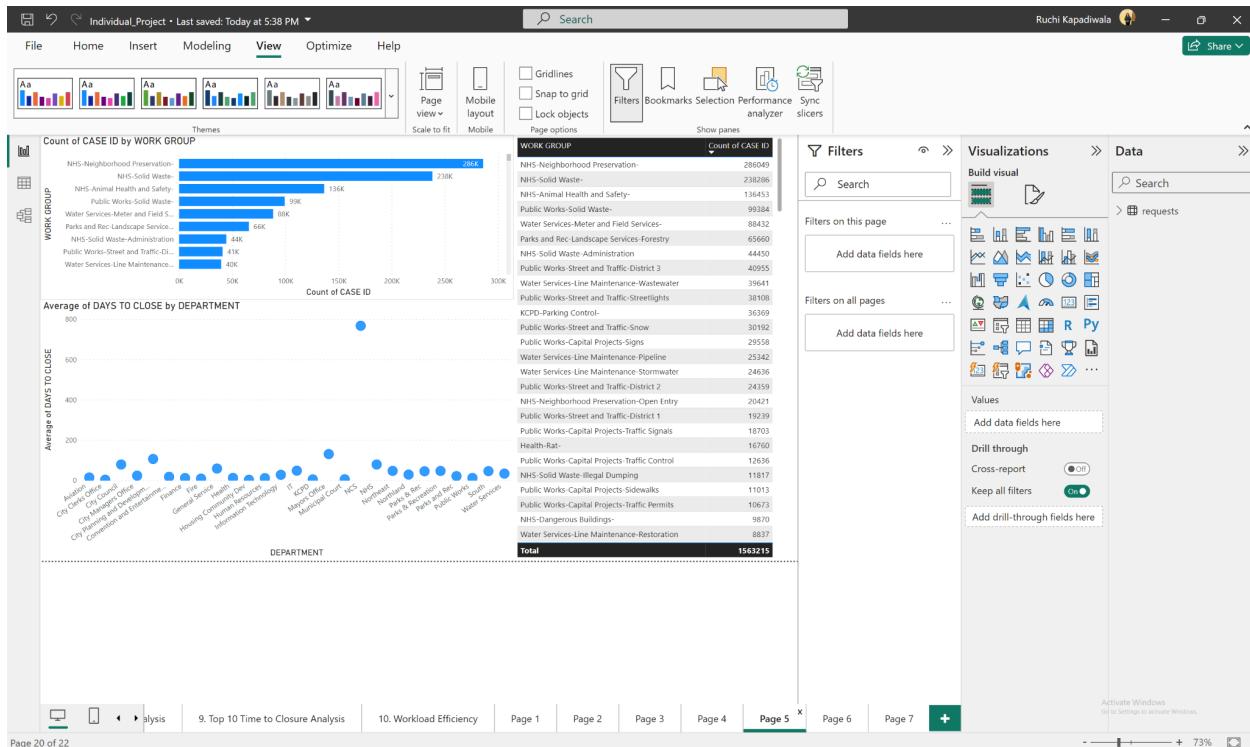


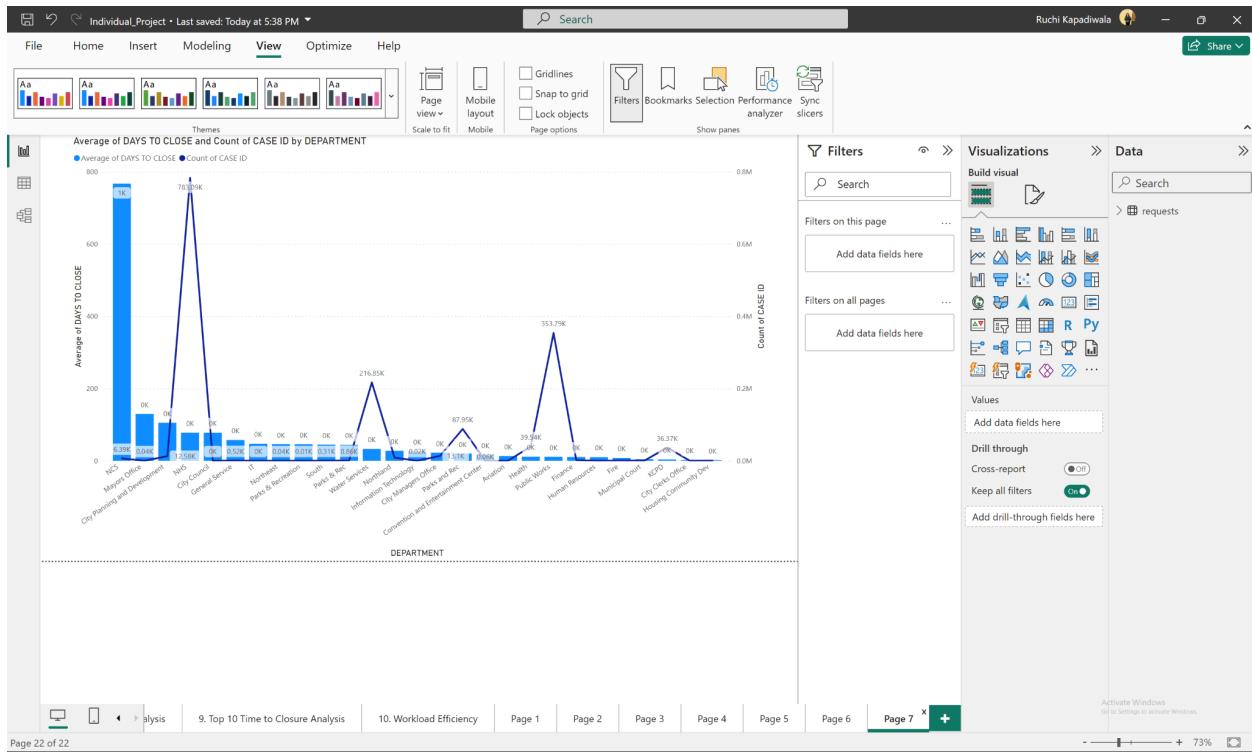


5.2 PowerBI Dashboards:









5.3 SQL Query:

1. SELECT YEAR(CREATION_DATE) AS 'YEAR', COUNT([CASE ID]) AS 'SERVICE REQUESTS' FROM requests GROUP BY YEAR(CREATION_DATE) HAVING YEAR(CREATION_DATE) IN (2018, 2019, 2020, 2021);

```
SELECT MONTH(CREATION_DATE) AS 'MONTH', YEAR(CREATION_DATE) AS 'YEAR ', COUNT([CASE ID]) AS 'SERVICE REQUESTS' FROM requests GROUP BY YEAR(CREATION_DATE), MONTH(CREATION_DATE) HAVING YEAR(CREATION_DATE ) IN (2018, 2019, 2020, 2021) ORDER BY YEAR(CREATION_DATE), MONTH(CREATION_DATE);
```

2. SELECT SOURCE, COUNT([CASE ID]) AS RequestCount FROM requests GROUP BY Source ORDER BY RequestCount DESC;

3. SELECT DEPARTMENT, COUNT([CASE ID]) AS RequestCount FROM requests GROUP BY DEPARTMENT ORDER BY RequestCount DESC;

4. SELECT DEPARTMENT, COUNT([CASE ID]) AS Workload FROM requests GROUP BY DEPARTMENT ORDER BY Workload DESC;

```
SELECT [WORK GROUP], COUNT([CASE ID]) AS Workload FROM requests GROUP BY [WORK GROUP] ORDER BY Workload DESC;
```

5. SELECT TOP 10 [ZIP CODE], COUNT([CASE ID]) AS RequestCount FROM Requests GROUP BY [ZIP CODE] ORDER BY RequestCount DESC;

6. SELECT DEPARTMENT, COUNT([CASE ID]) AS Workload FROM requests GROUP BY DEPARTMENT ORDER BY Workload DESC;

SELECT [WORK GROUP], COUNT([CASE ID]) AS Workload FROM requests GROUP BY [WORK GROUP] ORDER BY Workload DESC;
7. SELECT DEPARTMENT, AVG([DAYS TO CLOSE]) FROM requests GROUP BY DEPARTMENT ORDER BY DEPARTMENT ASC;
8. SELECT YEAR([CREATION_DATE]) AS Year, STATUS, COUNT([CASE ID]) AS StatusCount FROM requests WHERE YEAR(CREATION_DATE) BETWEEN 2018 AND 2021 GROUP BY Year(CREATION_DATE), STATUS ORDER BY YEAR ASC;
9. SELECT CATEGORY1, AVG([DAYS TO CLOSE]) AS AvgDaysToClose FROM requests WHERE CATEGORY1 != 'Data Not Available' GROUP BY CATEGORY1 ORDER BY AvgDaysToClose DESC;

SELECT TOP 10 CATEGORY1, AVG([DAYS TO CLOSE]) AS AvgDaysToClose FROM requests WHERE CATEGORY1 != 'Data Not Available' GROUP BY CATEGORY1 ORDER BY AvgDaysToClose DESC;
10. SELECT DEPARTMENT, COUNT([CASE ID]) AS Workload, AVG([DAYS TO CLOSE]) AS AvgEfficiency FROM requests GROUP BY DEPARTMENT ORDER BY DEPARTMENT;