

Data Quality Report

Dataset Overview:

The dataset contains service request data from the 311 call center in Kansas City, MO. It consists of various fields related to Service Requests such as case ID, type, location coordinates, closure details, and categorical attributes like department, neighborhood, and category. Each field has been assessed for missing values, uniqueness, and range of values.

Issues Identified:

Missing Values:

- Fields with Missing Values: The fields "CATEGORY2" and "CATEGORY3" exhibit a significant proportion of missing values, with 64.1% and 89.9% missing respectively. This suggests a potential data integrity issue that needs to be addressed.

Inconsistent Data Types:

- Some fields like "CASE ID" and "ZIP CODE" are marked as strings but contain numeric values. This inconsistency may lead to issues during analysis or processing.

Outliers:

- In "DAYS TO CLOSE", there are outliers with extremely high values (e.g., 1154.0) that may skew analysis results.

Inaccurate Values:

- There are discrepancies in values within fields like "POLICE DISTRICT" where values such as "East" and "Shoal Creek" seem inconsistent or incomplete.

Data Cleaning Plan:

Handling Missing Values:

- For fields with a high proportion of missing values (e.g., "CATEGORY2" and "CATEGORY3"), consider imputation techniques like mean, median, or mode depending on data distribution.

Standardizing Data Types:

- Convert fields with inconsistent data types (e.g., "CASE ID" and "ZIP CODE") to the appropriate data type (numeric or string) to ensure uniformity and prevent potential errors.

Addressing Outliers:

- Evaluate outliers in "DAYS TO CLOSE" to determine if they are valid data points. If not, consider removing them or applying a transformation to mitigate their impact on analysis.

Cleaning Inaccurate Values:

- Review and correct inconsistencies in fields like "POLICE DISTRICT" by referring to a reliable source for accurate data.

Conclusion:

The dataset exhibits several data quality issues, including missing values, inconsistent data types, outliers, and inaccuracies. Addressing these issues through data cleaning techniques will improve the reliability and usability of the dataset for analysis and decision-making purposes.

Alteryx Designer x64 - ServiceRequestsWorkflow.ymd

File Edit View Options Help

Favorites Recommended >Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Search for tools, help, and resources...

Browse (4) - Configuration

Profile
1,563,215 records displayed, 23 fields, 133 MB

CASE ID

- 2,007,002,634 1
- 2,007,006,846 1
- 2,007,007,334 1
- 2,007,007,437 1
- 2,007,007,461 1
- 995 more >

SOURCE

PHONE	1,204,236
WEB	211,721
EMAIL	60,585
SYS	19,226
INSP	14,690
16 more >	

DEPARTMENT

NHS	783,094
Public Works	353,787
Water Services	216,852
Parks and Rec	87,954
Health	39,543
22 more >	

Results - Browse (4) - Input

Record	CASE ID	SOURCE	DEPARTMENT	WORK GROUP	TYPE	DETAIL	CREATION DATE	CREATION TIME
1	2019119972	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	06/24/2019	0740 AM
2	2019207923	WEB	Public Works	Public Works-Street and Traffic-District 1	Crack	District 1	12/22/2019	0756 PM
3	2021005976	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	01/19/2021	0243 PM
4	2020149407	PHONE	NHS	NHS-Neighborhood Preservation-	Property Maintenance	Other Property Issue	11/25/2020	0919 AM
5	2020054721	WEB	Parks and Rec	Parks and Rec-Central Region-	Park Maintenance	Central	04/18/2020	0510 PM
6	2019182182	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	10/21/2019	1029 AM
7	2019184705	WEB	NHS	NHS-Solid Waste-	Recycling	Missed by City	10/25/2019	1002 AM
8	2019184590	WEB	Parks and Rec	Parks and Rec-Landscape Services-Forestry	Trimming	Tree Limbs	10/25/2019	0944 AM
9	2020095175	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	07/13/2020	0800 AM
10	2015094486	PHONE	NHS	NHS-Solid Waste-Administration	Services	Service Issue / Problem	08/07/2015	0145 PM
11	2020001142	PHONE	Water Services	Water Services-Meter and Field Services-	Leak	Meter	02/11/2020	0337 PM
12	2019208956	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	St+ Lights Out	12/24/2019	1004 PM
13	2014208947	WFR	Public Works	Public Works-Clean and Traffic-Cleanable	Clean Light	Clean Light Out	12/24/2019	1007 PM

```

graph LR
    Source[311 CallCenterServiceRequests_KansasCity_2007-March2021.tsv] --> Filter1{ }
    Filter1 --> Sort1[ ]
    Sort1 --> Filter2{ }
    Filter2 --> Sort2[ ]
    Sort2 --> Filter3{ }
    Filter3 --> Sort3[ ]
    Sort3 --> Output[dcm:ServiceRequests Table=requests]
  
```

Alteryx Designer x64 - ServiceRequestsWorkflow.ymd

File Edit View Options Help

Favorites Recommended >Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Search for tools, help, and resources...

Workflow - Configuration

Meta Info

Canvas Workflow Runtime Events

CANVAS OPTIONS

Layout Direction
Horizontal

Annotations
Show

Connection Progress
Show Only When Running

Workflow - Configuration

ServiceRequestsWorkflow.ymd

1,563,215 336MB

483 21.6KB

Finished running ServiceRequestsWorkflow.ymd in 7.2 seconds with 1 error using AMP engine

Learn about AMP Engine

Don't show this message again

OK

Results - Workflow - Messages

All	1 Errors	0 Conv Errors	0 Warnings	27 Info	4 Files
Auto Field (13)	The FieldType of "CATEGORY1" changed to: V_String(13)				
Auto Field (13)	The FieldType of "POLICE DISTRICT" changed to: V_String(11)				
Auto Field (13)	The FieldType of "PARCEL_ID_N" changed to: Int32				
Auto Field (13)	The FieldType of "LATITUDE" changed to: Double				
Auto Field (13)	The FieldType of "LONGITUDE" changed to: Double				
Auto Field (13)	The FieldType of "CATEGORY1" changed to: V_String(33)				
Auto Field (13)	The FieldType of "CATEGORY2" changed to: V_String(9)				
Auto Field (13)	The FieldType of "CATEGORY3" changed to: V_String(12)				
Output Data (16)	ODBC Driver version: 03.80				
Output Data (16)	Lua Script Loaded: sqldriver.lua				
Error creating table "requests": [Microsoft][SQL Server Native Client 11.0][SQL Server]There is already an object named 'requests' in the database.	CREATE TABLE "requests" ("CASE_ID" int, "SOURCE" varchar(50), "DEPARTMENT" varchar(50), "WORK_GROUP" varchar(50), "TYPE" varchar(50), "DETAIL" varchar(50), "CREATION_DATE" date, "CREATION_TIME" time)				

Last Run Configuration

Browse (4) 1,563,215 records

You are screen sharing Stop share

Designer x64

Mute Start Video Participants Chat New share Pause Share Annotate Slide Control More

Type here to search

37°F Cloudy 2:56 PM 1/31/2024

SQLQuery1.sql - DESKTOP-DCV4P31.ServiceRequests (DESKTOP-DCV4P31\uchi (63)) Executing... - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Quick Launch (Ctrl+Q) X

ServiceRequests | Execute |

Object Explorer

Connect ▾ ↻

Databases

- System Databases
- Database Snapshots
- ServiceRequests
- Table
- System Tables
- FileTables
- External Tables
- Graph Tables
- dbo.requests
- Columns
- Keys
- Constraints
- Triggers
- Indexes
- Statistics
- Views

SQLQuery2.sql - not connected SQLQuery1.sql - DES... (63) Executing... X

```

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 * 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;

```

Results Activate Window

CASE ID	SOURCE	DEPARTMENT	WORK GROUP	TYPE	DETAIL	CREATION DATE	CREATION TIME	STATUS	EXCEEDED EST TIMEFRAME	CLOSED DATE	DAYS TO CLOSE
1	2019119932	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2019-04-24 00:00:00.000	07:40 AM	RESOL Y	11/19/2021	87
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 PM	RESOL Y	11/15/2021	300
4	2020149407	PHONE	NHS	NHS-Neighborhood Preservation-	Property Maintenance	Other Property Issue	2020-11-25 00:00:00.000	09:19 AM	RESOL Y	04/26/2021	152
5	2020054721	WEB	Parks and Rec.	Parks and Rec-Central Region	Park Maintenance	Central	2020-11-21 00:00:00.000	08:45 PM	RESOL N	04/30/2020	12
6	2019184705	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2019-11-21 00:00:00.000	09:28 AM	RESOL Y	08/03/2020	297
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	202009175	PHONE	NHS	NHS-Dangerous Buildings-	Dangerous Building	Standard	2020-07-13 00:00:00.000	08:00 AM	RESOL Y	11/29/2021	504
10	2019184590	PHONE	NHS	NHS-Neighborhood Preservation-	Property Maintenance	Other Property Issue / Problem	2019-10-24 00:00:00.000	01:45 PM	RESOL Y	06/27/2019	294
11	2020021142	PHONE	Water Services	Water Services-Main and Field Services	Leak	Service Issue / Problem	2020-02-11 00:00:00.000	01:45 PM	RESOL N	03/12/2020	1
12	2019208956	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	PHONE	NHS	NHS-Dangerous Buildings-	Street Light	5+ Lights Out	2019-12-24 00:00:00.000	10:05 PM	RESOL N	01/07/2020	14
14	2019208975	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-10-25 00:00:00.000	07:28 PM	RESOL N	01/07/2020	13
15	2019208976	WEB	Public Works	Public Works-Street and Traffic-Streetlights	Street Light	5+ Lights Out	2019-10-25 00:00:00.000	08:41 PM	RESOL N	01/07/2020	13
16	2019208977	PHONE	NHS	NHS-Dangerous Buildings-	Street Light	5+ Lights Out	2019-10-25 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	2019-10-25 00:00:00.000	08:42 PM	RESOL N	01/07/2020	13
18	2019208979	PHONE	NHS	NHS-Neighborhood Preservation-	Street Light	5+ Lights Out	2019-10-25 00:00:00.000	08:42 PM	RESOL N	01/07/2020	13
19	2019655892	PHONE	NHS	NHS-Neighborhood Preservation-	Property Maintenance	Other Property Issue	2019-09-19 00:00:00.000	01:20 PM	RESOL Y	09/17/2020	548
20	2020082884	PHONE	NHS	NHS-Neighborhood Preservation-Open Entry	Open To Entry	Open To Entry	2020-06-15 00:00:00.000	04:03 PM	RESOL Y	06/26/2020	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	875

LN 12 Col 1 Ch 1 INS

DESKTOP-DCV4P31 (15.0 RTM) DESKTOP-DCV4P31\uchi ... ServiceRequests 00:00:00 0 rows

Ready

SQLQuery1.sql - DESKTOP-DCV4P31.ServiceRequests (DESKTOP-DCV4P31\uchi (63)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Quick Launch (Ctrl+Q) X

ServiceRequests | Execute |

Object Explorer

Connect ▾ ↻

Databases

- System Databases
- Database Snapshots
- ServiceRequests
- Table
- System Tables
- FileTables
- External Tables
- Graph Tables
- dbo.requests
- Columns
- Keys
- Constraints
- Triggers
- Indexes
- Statistics
- Views

SQLQuery1.sql - not connected SQLQuery1.sql - DES... (63) * Executing... X

```

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;

```

Results Activate Window

TOTAL ROWS
1 1563215

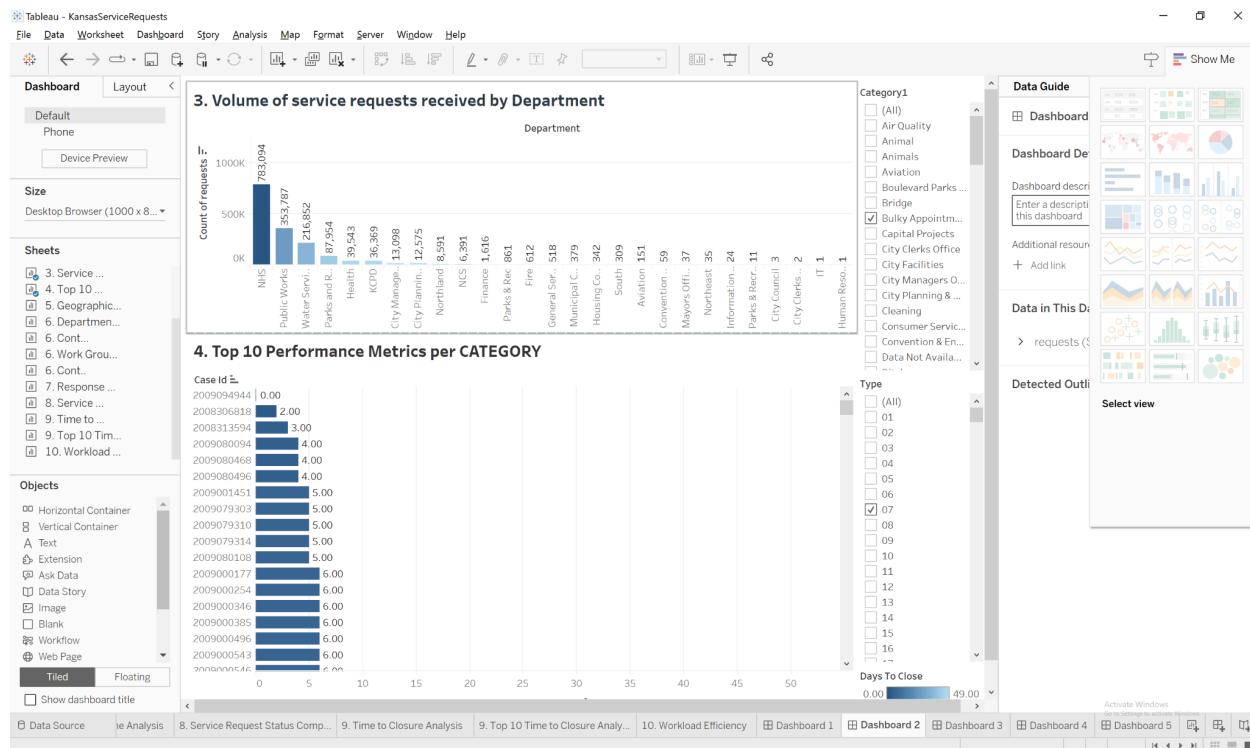
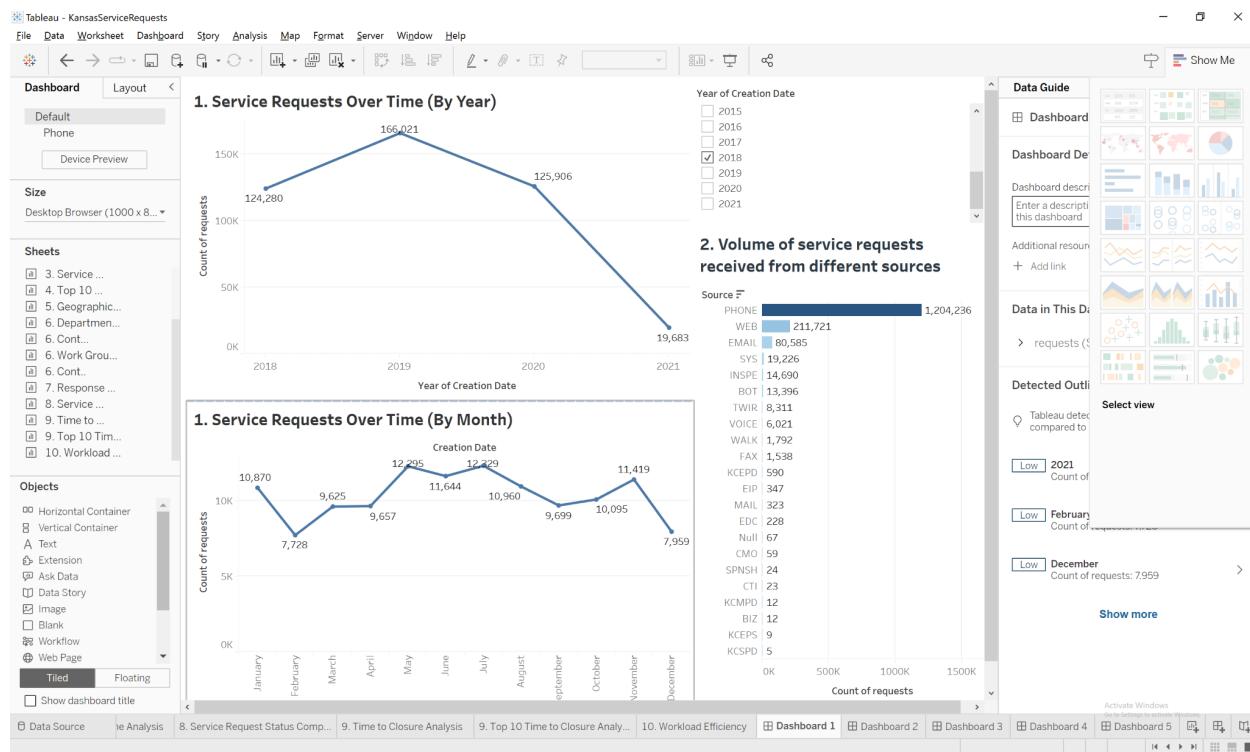
Query executed successfully.

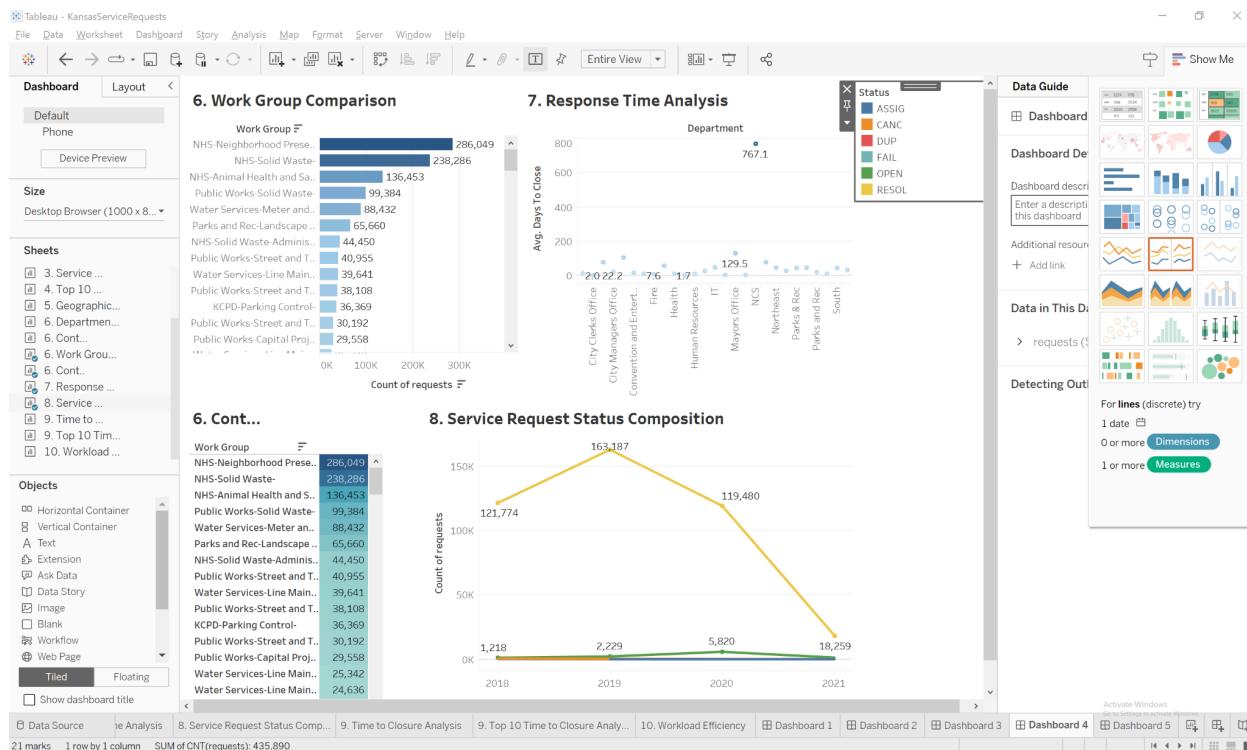
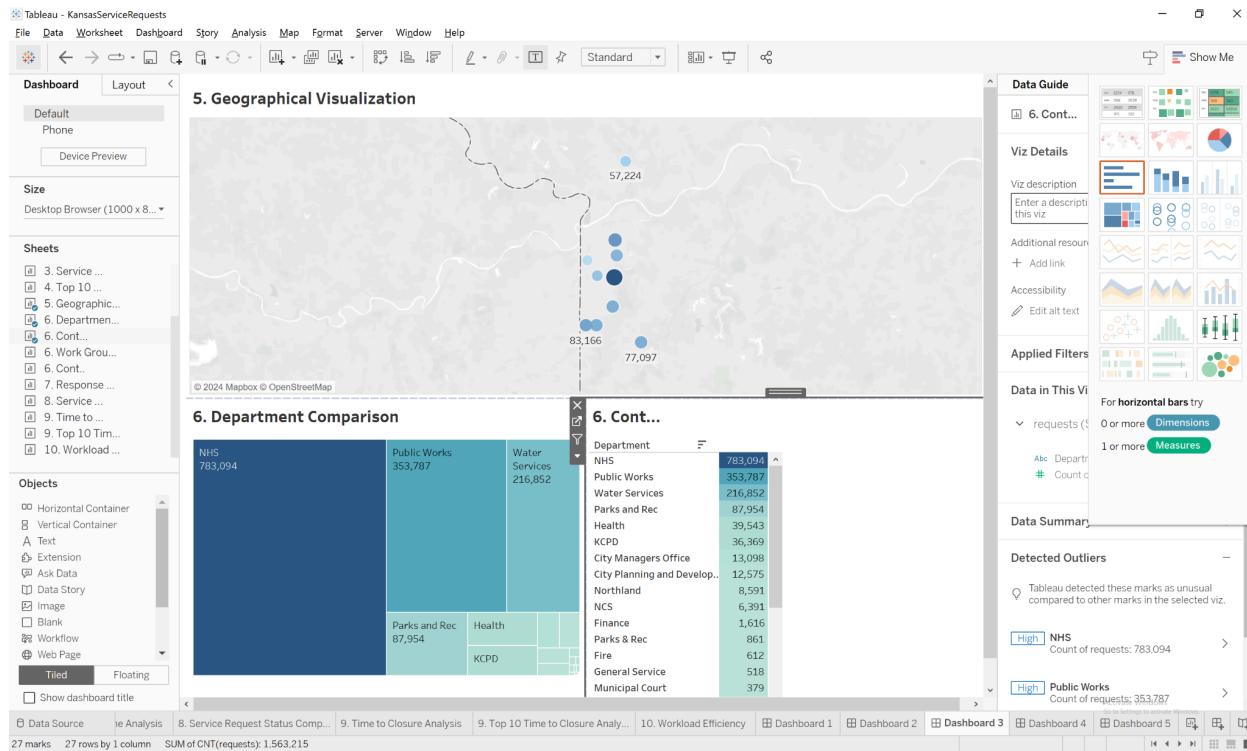
LN 21 Col 1 Ch 1 INS

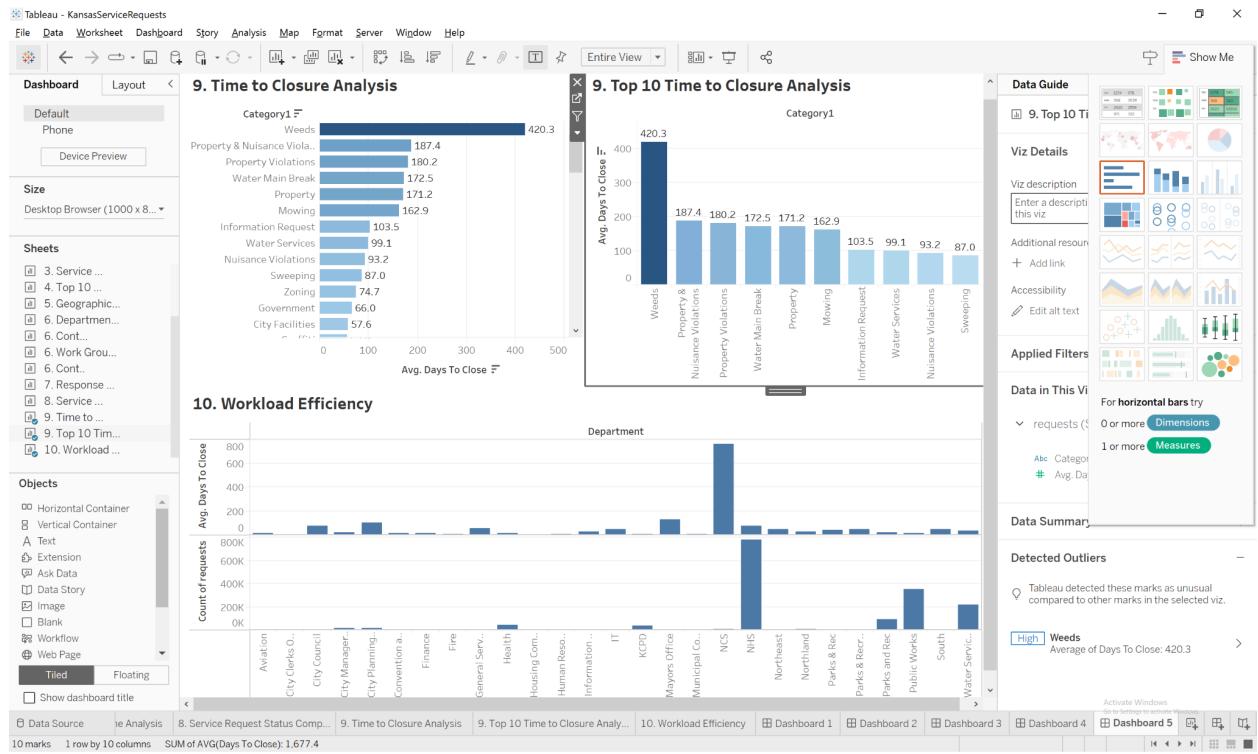
DESKTOP-DCV4P31 (15.0 RTM) DESKTOP-DCV4P31\uchi ... ServiceRequests 00:00:00 1 rows

Ready

TABLEAU DASHBOARDS:







POWER BI DASHBOARDS:

Individual Project • Last saved: Today at 5:38 PM

File Home Insert Modeling View Optimize Help

Themes Page view Scale to fit Mobile layout Page options Gridlines Snap to grid Lock objects Show panes

Filters Bookmarks Selection Performance analyzer Sync slicers

Count of CASE ID by Year

Year	Count of CASE ID
2018	124K
2019	166K
2020	126K
2021	20K

Count of CASE ID by Month

Month	Count of CASE ID
January	52K
February	36K
March	38K
April	36K
May	41K
June	40K
July	39K
August	35K
September	31K
October	29K
November	27K
December	26K

Filters Visualizations Data

Build visual Search

Filters on this page Add data fields here

Filters on all pages Add data fields here

Values Add data fields here

Drill through Cross-report Off

Keep all filters On

Add drill-through fields here

Analysis 9. Top 10 Time to Closure Analysis 10. Workload Efficiency Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7 +

Activate Windows Go to settings to activate Windows

Page 16 of 22

Individual Project • Last saved: Today at 5:38 PM

File Home Insert Modeling View Optimize Help

Themes Page view Scale to fit Mobile layout Page options Gridlines Snap to grid Lock objects Show panes

Filters Bookmarks Selection Performance analyzer Sync slicers

Count of CASE ID by SOURCE

SOURCE	Count of CASE ID
PHONE	1,204,24K
WEB	21.72K
EMAIL	0.05K
SYS	19.23K
INSPE	14.69K
BOT	13.4K
FWR	8.31K
VOICE	6.02K
WALK	1.79K
FAX	1.54K
KCPD	0.55K
EIR	0.35K
MAIL	0.32K
EOC	0.23K
E8m	0.07K
CMO	0.06K
SPNSH	0.02K
CTI	0.01K
BIZ	0.01K
KCPD	0.01K
KCEPS	0.01K
KCPD	0.01K

Count of CASE ID by DEPARTMENT

DEPARTMENT	Count of CASE ID
NYC	79,400K
Public Works	35,79K
Water Services	11,85K
Parks and Rec.	12,95K
Health	12,54K
City Manager's Of.	13,3K
City Planning and...	17,50K
PCS	8,59K
Finance	6,29K
Public Ref.	1,62K
General Services	0,61K
Municipal Court	0,52K
Housing Commis.	0,38K
South	0,34K
Aviation	0,31K
Mayor's Office	0,15K
Northeast	0,06K
Information Tech.	0,04K
Parks & Recreation	0,04K
City Clerks Office	0K
Human Resources	0K

Filters Visualizations Data

Build visual Search

Filters on this page Add data fields here

Filters on all pages Add data fields here

Values Add data fields here

Drill through Cross-report Off

Keep all filters On

Add drill-through fields here

Analysis 9. Top 10 Time to Closure Analysis 10. Workload Efficiency Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7 +

Activate Windows Go to settings to activate Windows

Page 17 of 22

Individual_Project - Last saved: Today at 5:38 PM

File Home Insert Modeling View Optimize Help

Themes Scale to fit Mobile layout Mobile

Gridlines Snap to grid Lock objects Page view Show panes

Filters Bookmarks Selection Performance analyzer Sync slicers

CATEGORY1
■ City Planning & Devel... □ Administration ■ Land Development

Max of DAYS TO CLOSE by CASE ID

CASE ID	Max of DAYS TO CLOSE
2009294018	0
201930382	0
2007099755	1
2007100664	1
2009240573	1
2013015073	1
2013123364	1
2008060076	2
2011200475	2
201214051	2
2012144192	2
2013123358	2

Count of CASE ID by LATITUDE and LONGITUDE

ZIP CODE Count of CASE ID

ZIP CODE	Count of CASE ID
64115	26
64117	23
64113	19
64114	15
64119	10
64105	8
64112	6
64116	5
64111	4
64110	4
64118	4
64111	4

Analysis 9. Top 10 Time to Closure Analysis 10. Workload Efficiency Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7 +

Activate Windows Go to settings to activate Windows

Page 18 of 22

Individual_Project - Last saved: Today at 5:38 PM

File Home Insert Modeling View Optimize Help

Themes Scale to fit Mobile layout Mobile

Gridlines Snap to grid Lock objects Page view Show panes

Filters Bookmarks Selection Performance analyzer Sync slicers

DEPARTMENT Count of CASE ID

NHS
Public Works
Water Services
Health
KCPD
City Man
City Pla
783.09K
353.79K
216.85K
87.95K
39.5...
36...
783.09K

Count of CASE ID by DEPARTMENT

DEPARTMENT	Count of CASE ID
NHS	783.09K
Public Works	353.79K
Water Services	216.85K
Health	87.95K
KCPD	39.5...
City Man	36...
City Pla	783.09K

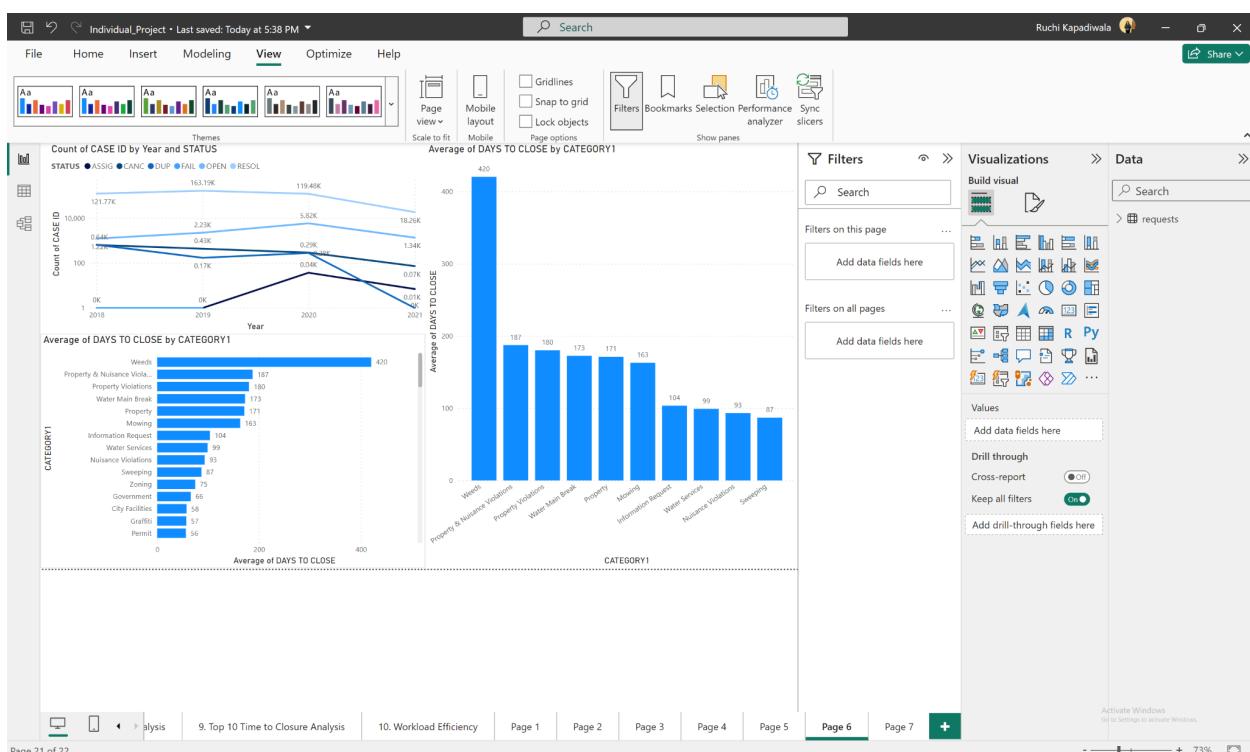
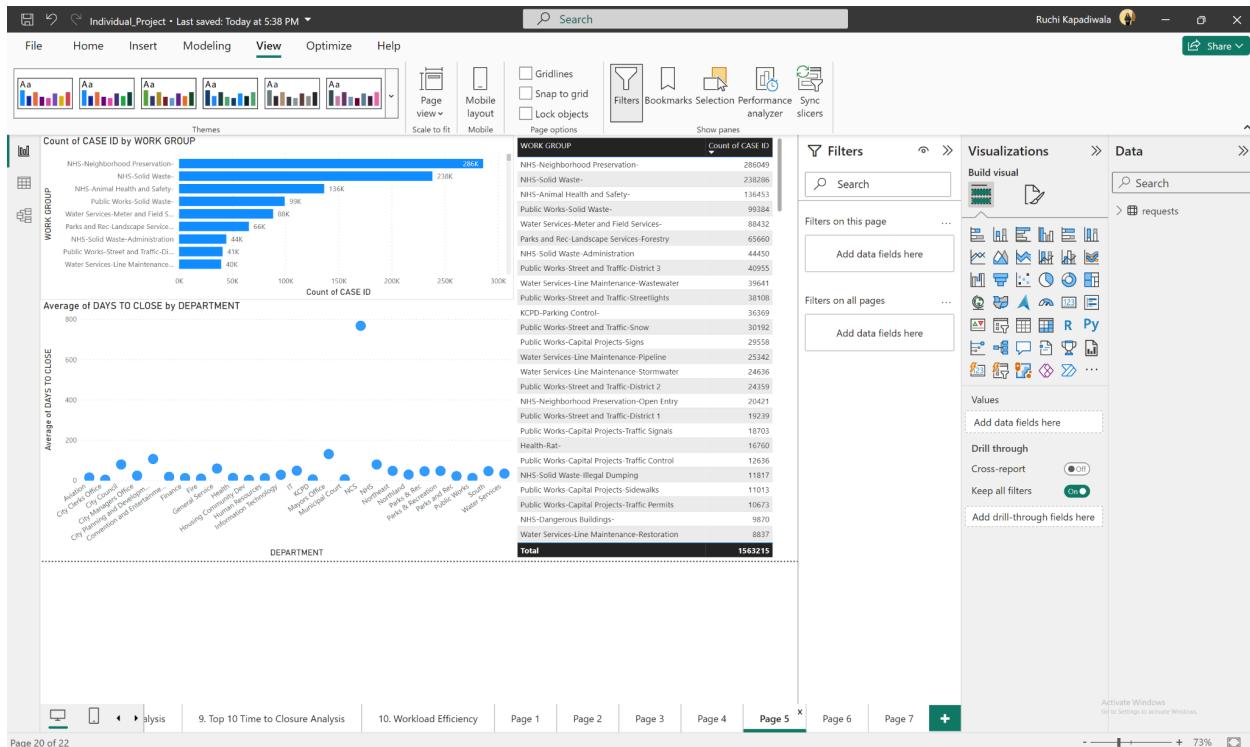
DEPARTMENT Count of CASE ID

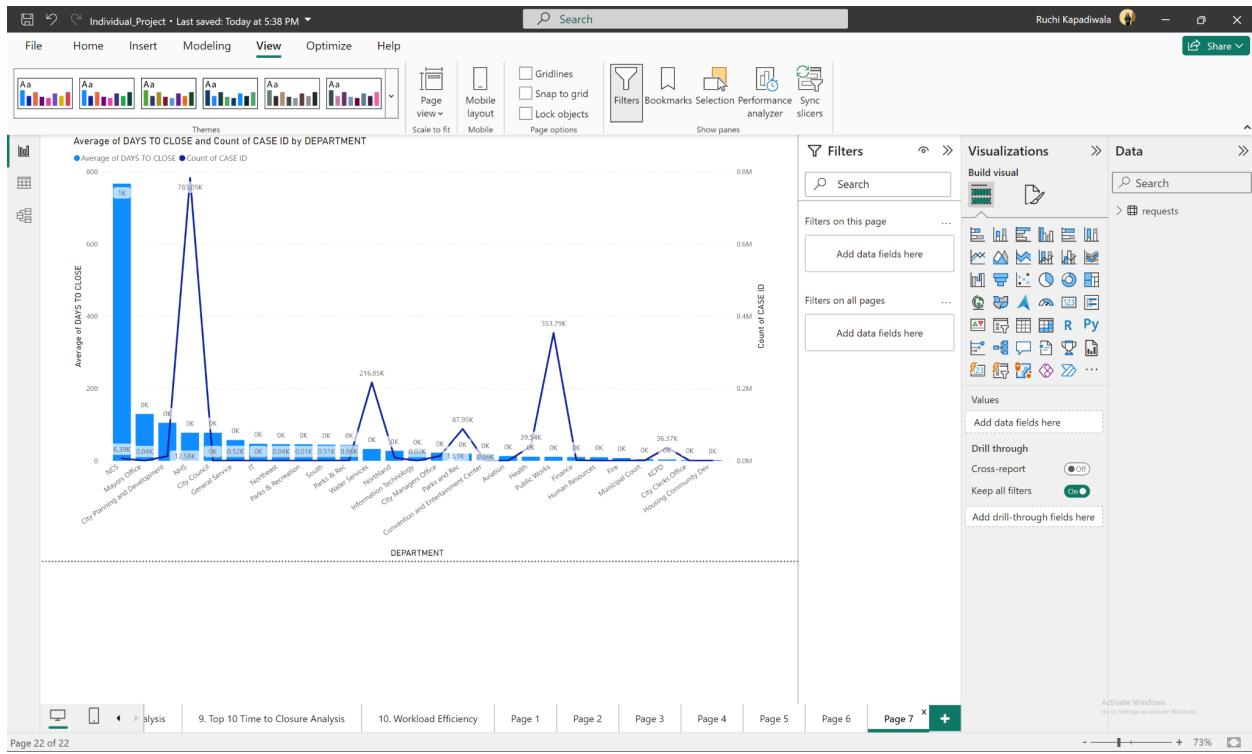
DEPARTMENT	Count of CASE ID
NHS	783.09K
Public Works	353.79K
Water Services	216.85K
Parks and Rec	87.95K
Health	39.5...
KCPD	36...
City Managers Office	13.1K
City Planning and Development	12.58K
Northland	12.5K
NCS	11.9K
Finance	11.6K
Total	1563215

Analysis 9. Top 10 Time to Closure Analysis 10. Workload Efficiency Page 1 Page 2 Page 3 Page 4 Page 5 Page 6 Page 7 +

Activate Windows Go to settings to activate Windows

Page 19 of 22





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;