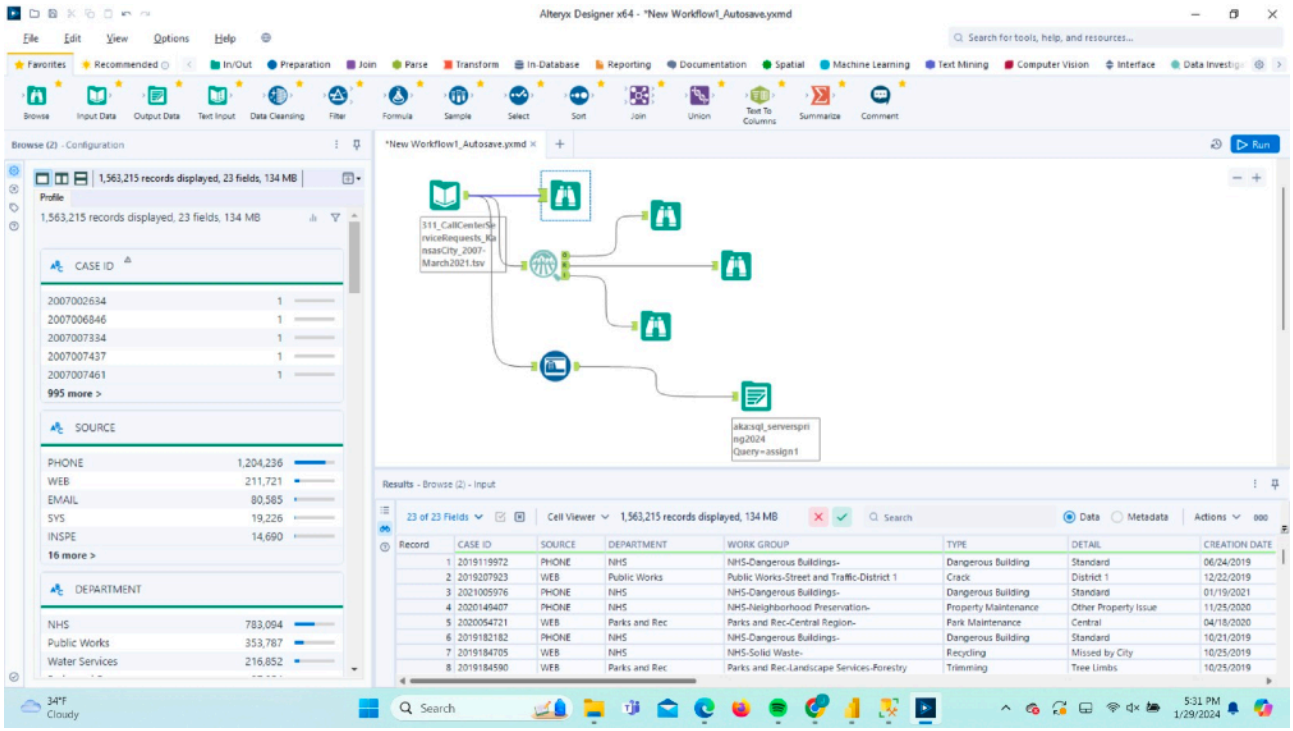
VAISHNAVI PATEL 002242426

DAMG 7370 Designing Advanced Data Architectures for Business Intelligence

ALTERYX Data Profiling -

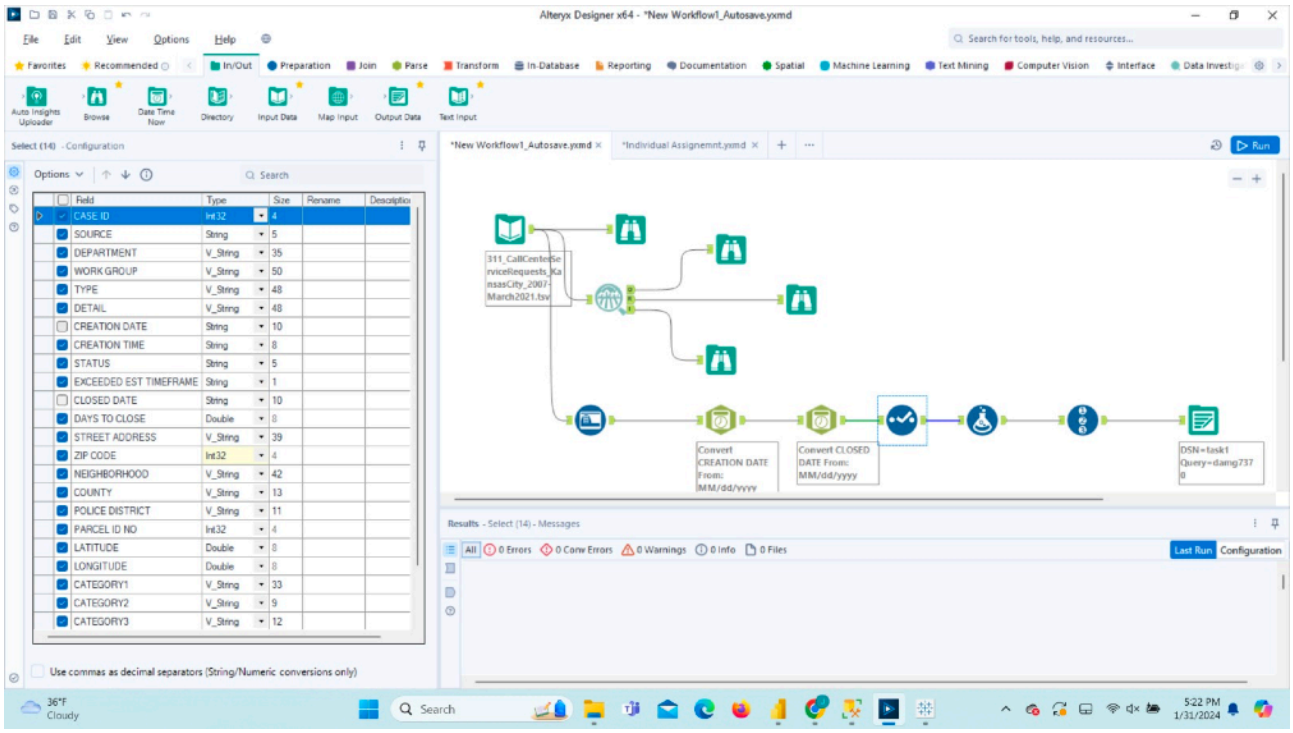


Observations Made - Columns Containing Null Values

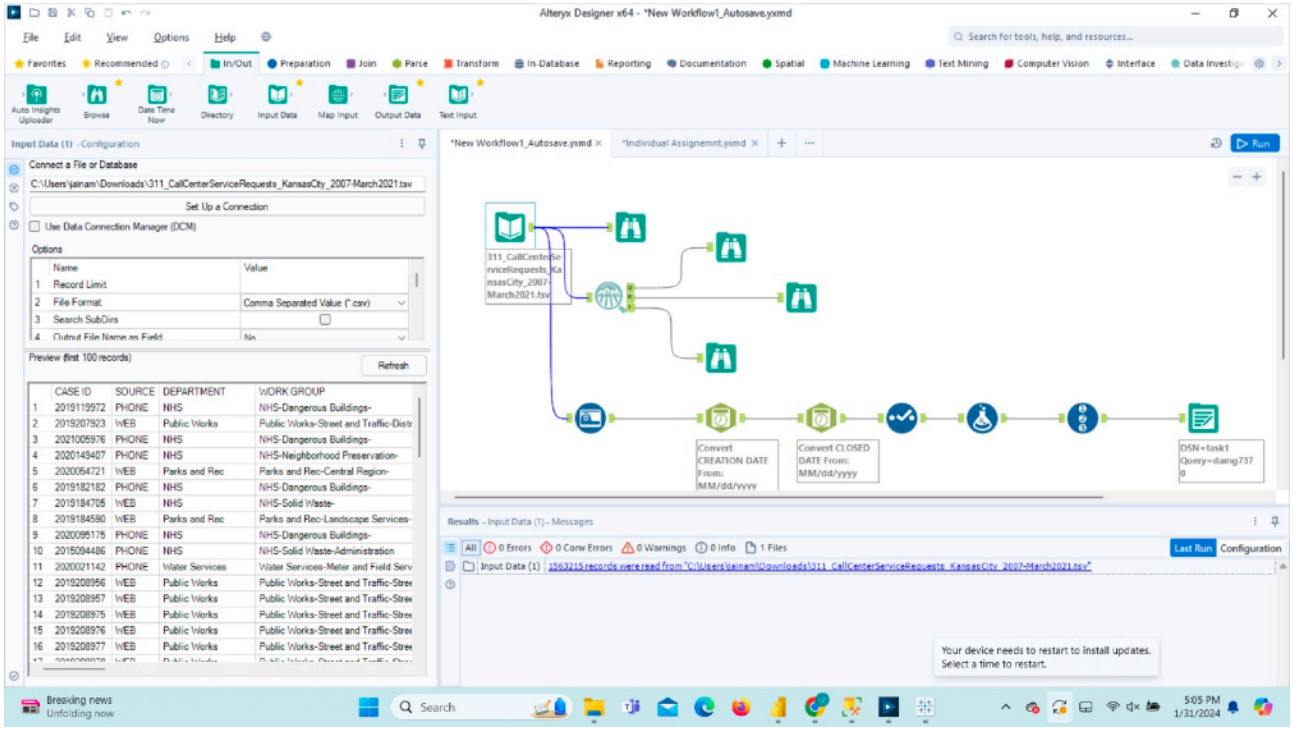
Column Name	Null Values
Closed Date	12,702
Street Address	24
Zip Code	826
Neighbourhood	46,106
County	66,956
Police District	32,265
Category 2	10,01,657
Category 3	14,09,943

VAISHNAVI PATEL 002242426

Field with wrongly assigned values -

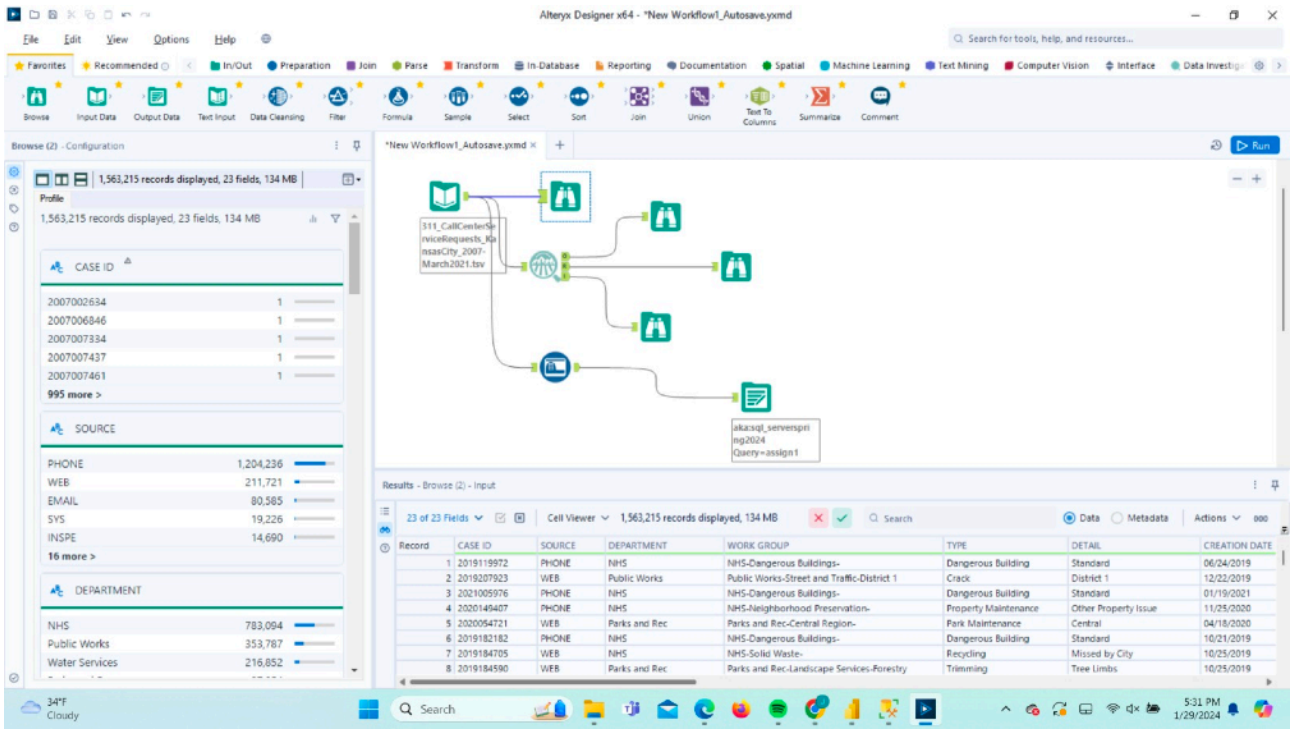


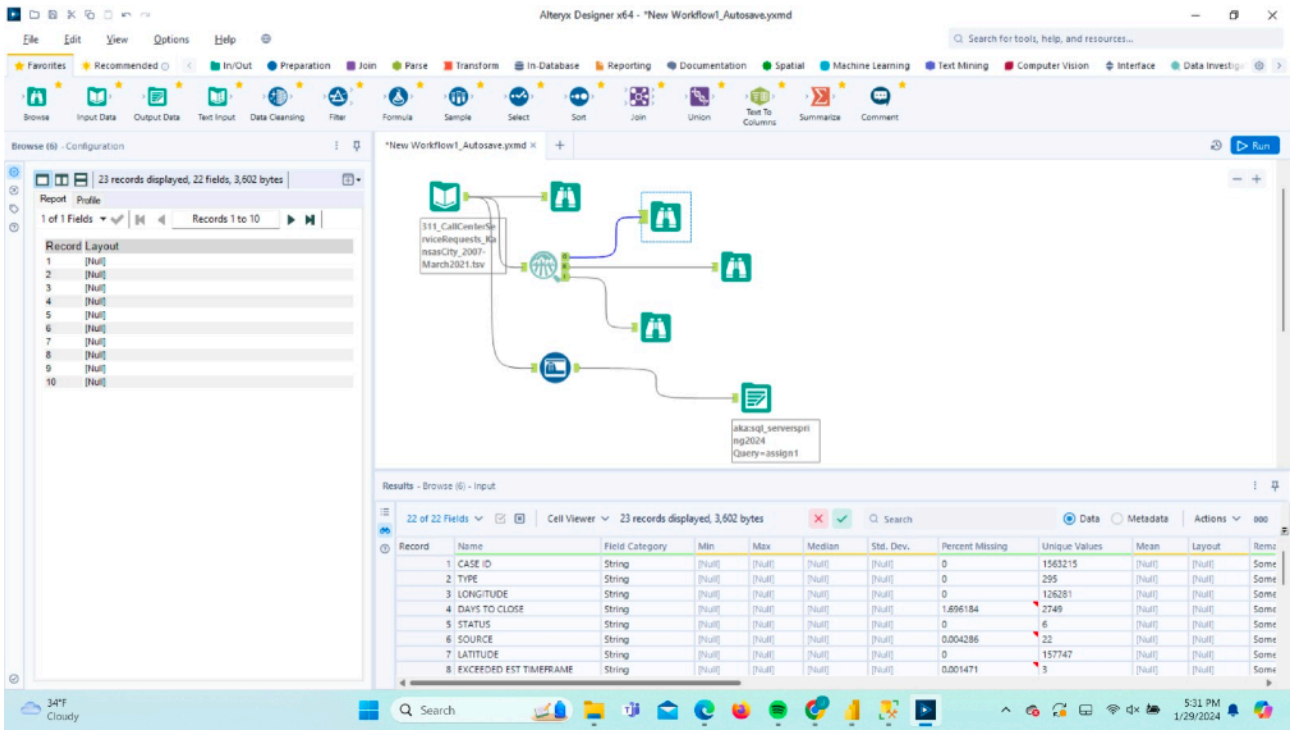
Creation Date - String data type to date type
Creation Time- String to Time data type
Closed date- String to Date data type
ZIP code-Double to Integer type
Days to Close-to be changed to Int data type



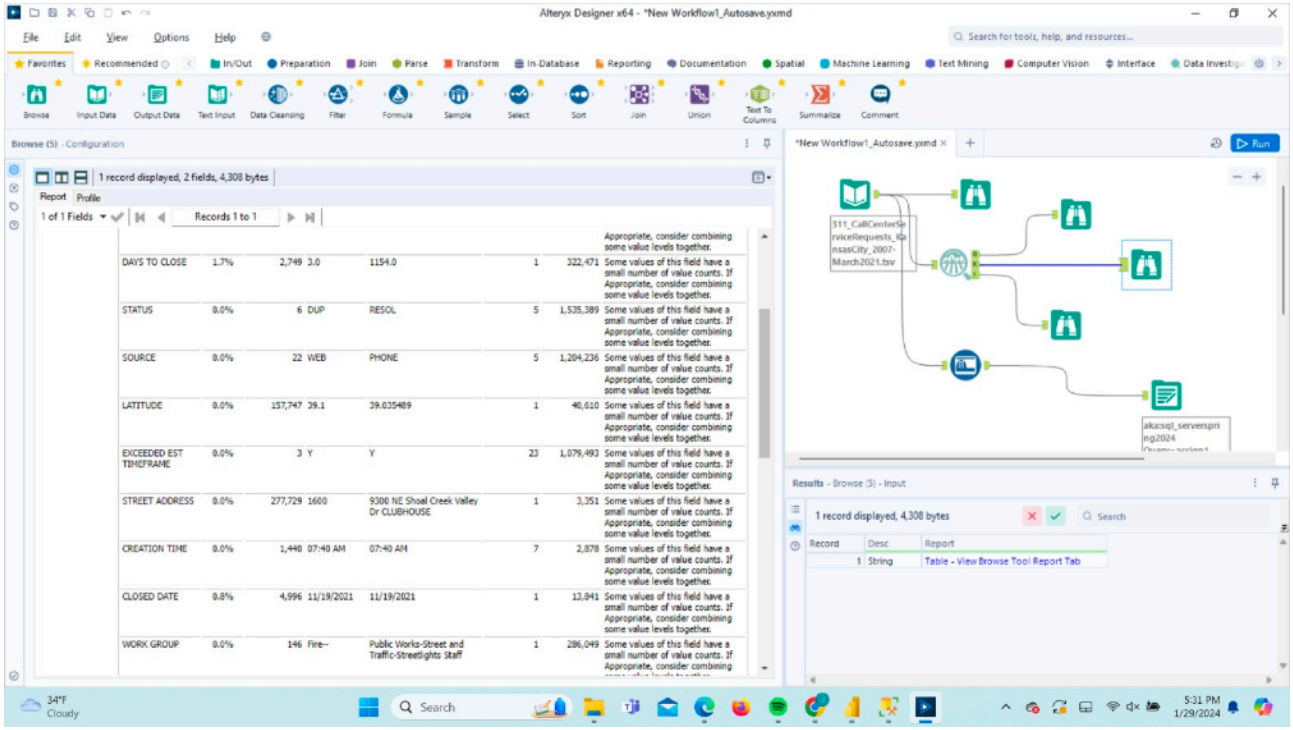
VAISHNAVI PATEL 002242426

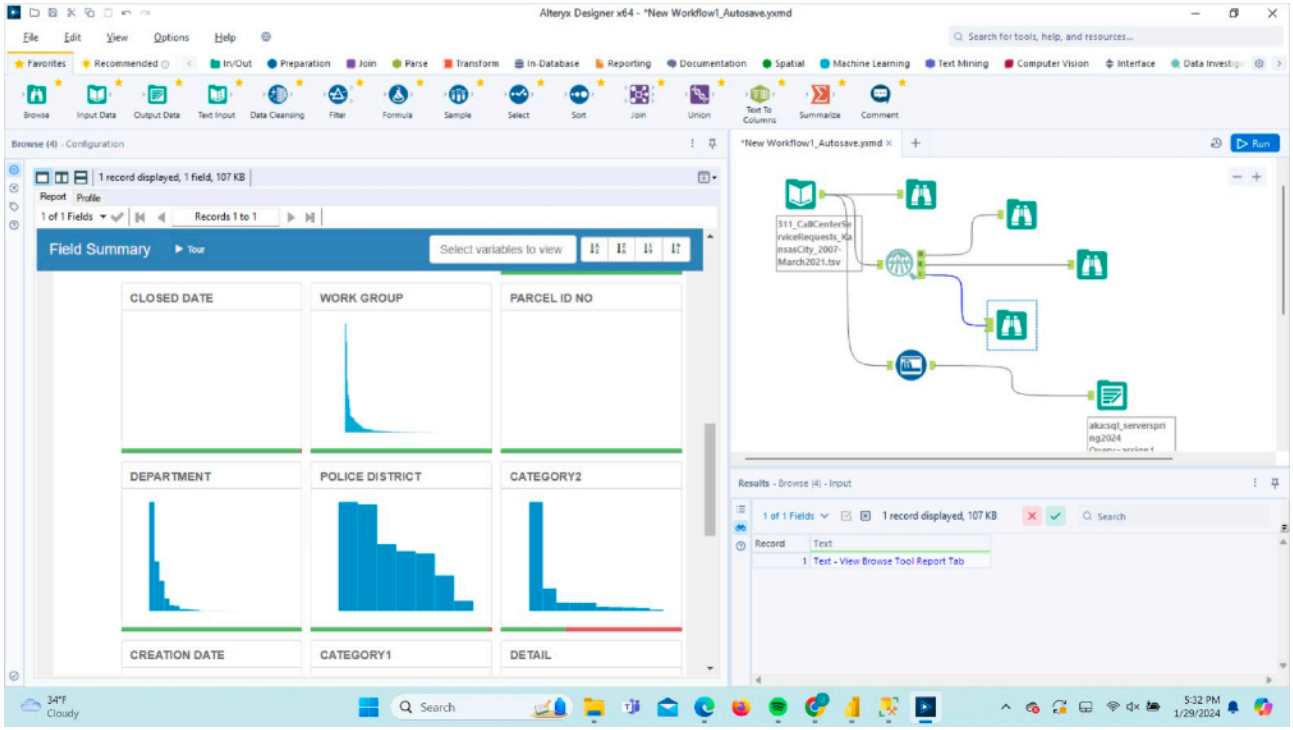
Reading and Profiling data -





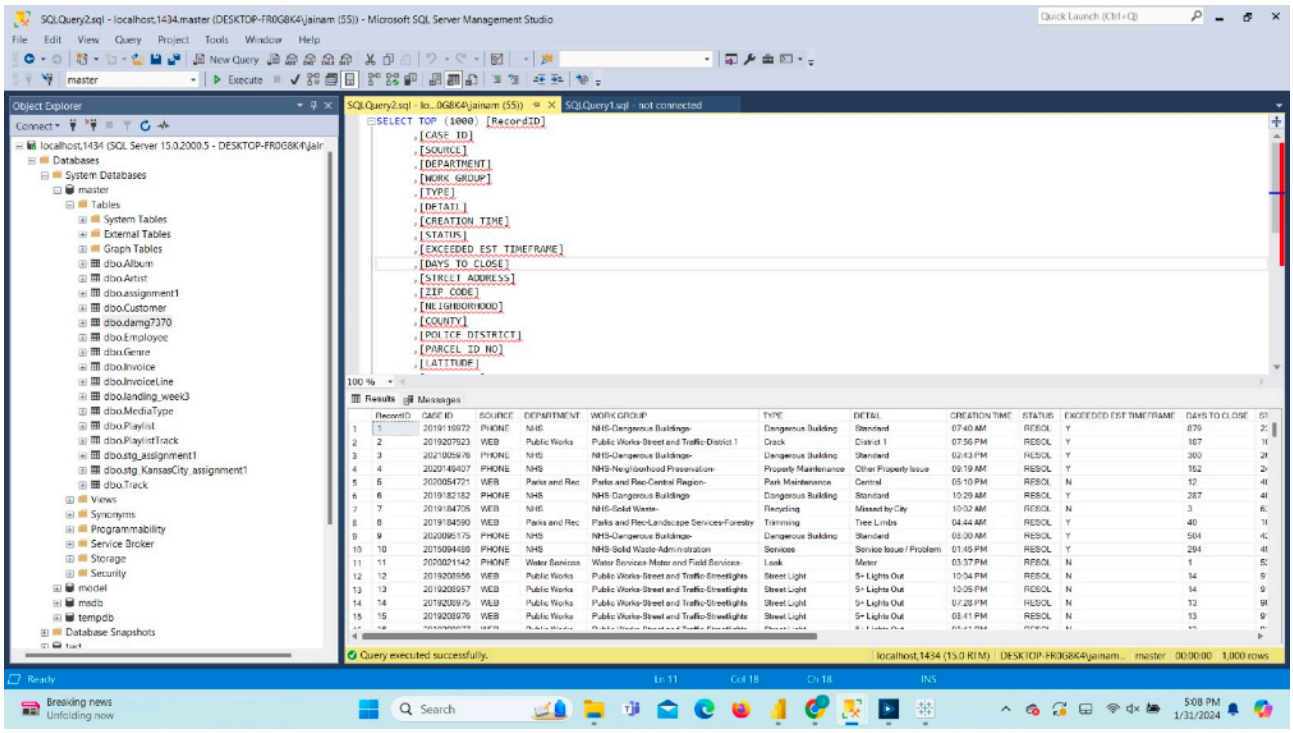
VAISHNAVI PATEL 002242426





VAISHNAVI PATEL 002242426

Populating Data on SQL Server



SQL Queries-

(1) Service Requests Over Time:

• What is the overall trend in Service Requests over the years 2018-2021?
• How have Service Requests changed on a monthly basis?

```
SELECT YEAR([CREATION TIME]) AS Year, COUNT(*) AS TotalRequests
FROM dbo.damg7370
WHERE YEAR([CREATION TIME]) BETWEEN 2018 AND 2021
GROUP BY YEAR([CREATION TIME])
ORDER BY Year;
```

```
SELECT YEAR([CREATION TIME]) AS Year, MONTH([CREATION TIME]) AS Month, COUNT(*) AS MonthlyRequests
FROM dbo.damg7370
WHERE YEAR([CREATION TIME]) BETWEEN 2018 AND 2021
GROUP BY YEAR([CREATION TIME]), MONTH([CREATION TIME])
ORDER BY Year, Month;
```

(2) Volume of service requests received from different sources:

• What is the overall trend in Service Requests over Sources?

```
SELECT [SOURCE], COUNT(*) AS TotalRequests

FROM dbo.damg7370

GROUP BY [SOURCE]

ORDER BY TotalRequests DESC;
```

(3) Volume of service requests received by Department:

• What is the overall trend in Service Requests received by Departments?

```
SELECT [DEPARTMENT], COUNT(*) AS TotalRequests

FROM dbo.damg7370

GROUP BY [DEPARTMENT]

ORDER BY TotalRequests DESC;
```

(4) Top 10 Performance Metrics (Response Time) per CATEGORY and Type of Request:

• What are the top 10 cases whose response time was fastest? Categorize it with Category1 and Type of Request.

```
SELECT TOP 10 [CATEGORY1], [TYPE], DATEDIFF(day, [CREATION TIME], [closed_date]) AS ResponseTime

FROM dbo.damg7370

ORDER BY ResponseTime;
```

(5) Geographical Visualization:

• What are the Top 10 areas where most number of request were raised?

```
SELECT TOP 10 [NEIGHBORHOOD], COUNT(*) AS TotalRequests
```

```
FROM dbo.damg7370

GROUP BY [NEIGHBORHOOD]

ORDER BY TotalRequests DESC;
```

(6) Departmental Workload Comparison:

• How does the workload vary among different departments and work groups? Create a visual representation to highlight the distribution.

```
SELECT [DEPARTMENT], [WORK GROUP], COUNT(*) AS TotalRequests

FROM dbo.damg7370

GROUP BY [DEPARTMENT], [WORK GROUP];
```

(7) Response Time Analysis:

• Visualize the distribution of response times for each department. Are there any outliers or patterns in response times?

```
SELECT [DEPARTMENT], DATEDIFF(day, [CREATION TIME], [closed_date]) AS ResponseTime

FROM dbo.damg7370;
```

(8) Service Request Status Composition:

• Create a visualization to show the composition of service request statuses (open, closed, in progress). How has this composition changed over the years 2018-2021?

```
SELECT [STATUS], YEAR([CREATION TIME]) AS Year, COUNT(*) AS TotalRequests

FROM dbo.damg7370

WHERE YEAR([CREATION TIME]) BETWEEN 2018 AND 2021

GROUP BY [STATUS], YEAR([CREATION TIME])

ORDER BY Year, [STATUS];
```

(9) Time to Closure Analysis:

• Visualize the average days to close service requests for each category1. Are there categories with consistently longer closure times?

SELECT [CATEGORY1], AVG(DATEDIFF(day, [CREATION TIME], [closed_date])) AS AvgDaysToClose

FROM dbo.damg7370

GROUP BY [CATEGORY1];

• Show top 10 (If you need help on how to restrict top 10 contact us and we can guide / help you)

SELECT TOP 10 [CATEGORY1], AVG(DATEDIFF(day, [CREATION TIME], [closed_date])) AS AvgDaysToClose

FROM dbo.damg7370

GROUP BY [CATEGORY1]

ORDER BY AvgDaysToClose DESC;

(10) Workload Efficiency:

• Create a visualization to show the relationship between workload (number of service requests) and efficiency (days to close) for each department?

SELECT [DEPARTMENT], COUNT(*) AS TotalRequests, AVG(DATEDIFF(day, [CREATION TIME], [closed_date])) AS AvgDaysToClose

FROM dbo.damg7370

GROUP BY [DEPARTMENT];