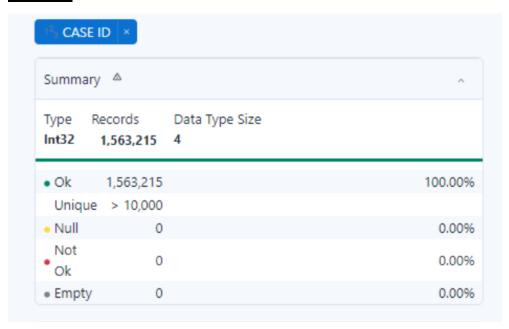
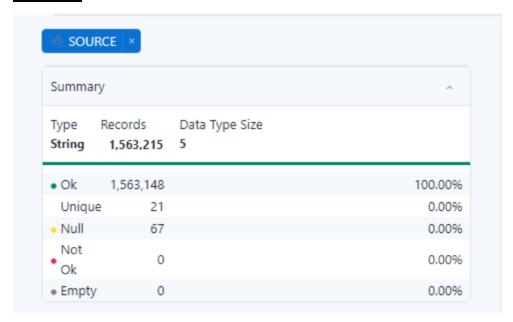
Observations made on the dataset.

CASE ID



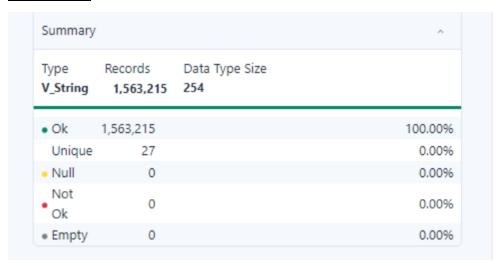
There are no null values in the CASE ID field and over $10,\!000$ unique values. This column does not include any null values. The data type was originally V_String, which was changed to int 32 after considering the maximum value of $2,\!021,\!033,\!686$. The size of int 32 is between - 2147483648 and +2147483647, which is well inside the range.

Source ID



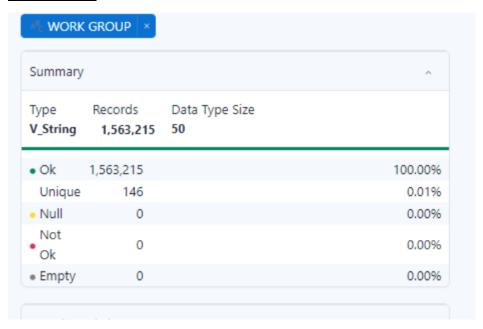
As we can see, there are 67 null values and 21 unique values in this particular case. This was initially of the datatype V_String, but it was subsequently changed to String as it requires a less size 8 than V_String, which requires 254.

Department



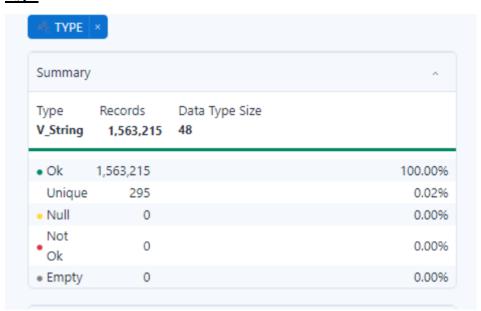
The department categories make up the 27 distinct values in this column, none of which are null. The National Health Service (NHS) comes first, then public works, etc.

Work Group



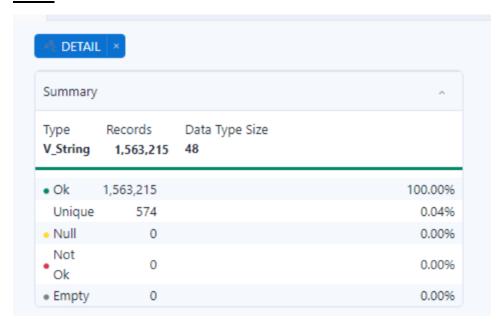
With a datatype size of 50, this column has 146 unique values and no null entries.

Type



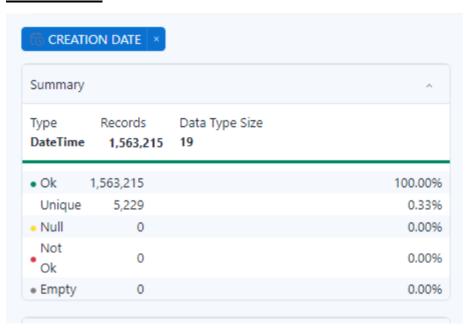
There are 295 unique values in this column of data type V_String , and there are no null values.

Detail



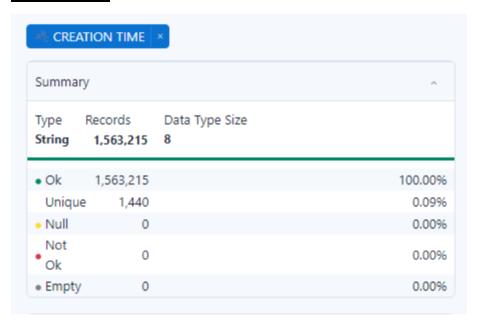
The previous example has no null values, a data type of V_String, and 574 unique values.

Creation Date



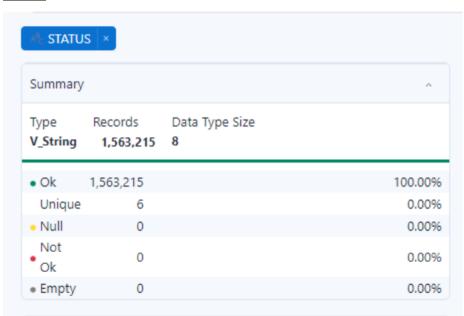
This column has a data type of DateTime and no null values. It has around 5229 distinct values.

Creation Time



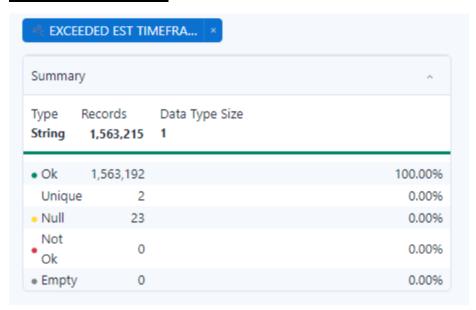
This column is of data type string and contains no null values. The unique values for this are 1440. This may be changed from its text value to a Date Time datatype.

Status



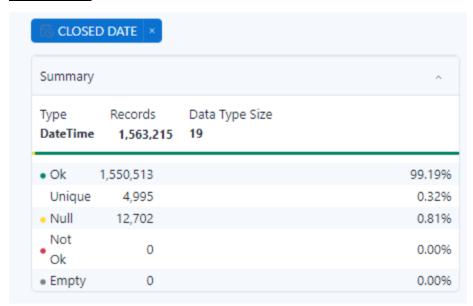
There are just six distinct values of the data type V String in this instance; null values are absent.

Exceed Est TimeFrame



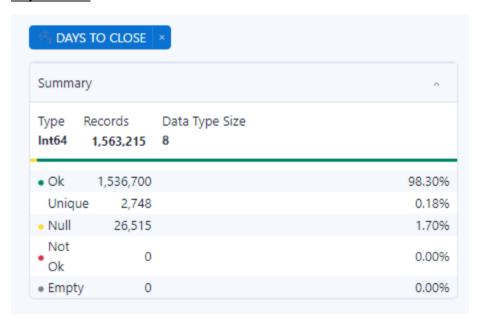
There are 23 null values and 2 unique values in the column that follows. Here, the data type is a string.

Closed Date



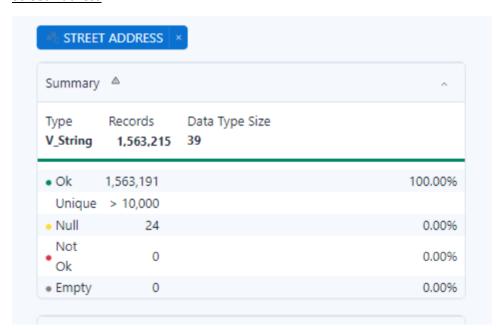
There are 12702 null values and 4995 unique values in the column above; parsing mistake is most likely the reason of these values.

Days to close



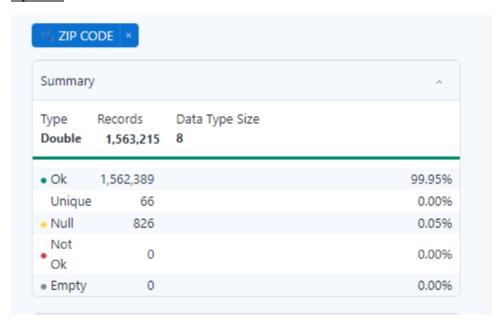
There are 2748 values and a large number of null values in this column. This can be the result of the majority of cases not yet being closed.

Street Address



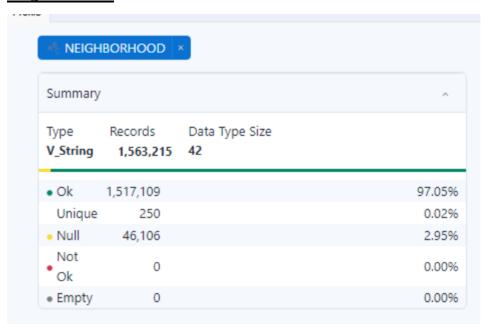
There are 24 null values despite the large number of unique values in this. The data type for this is V_S tring.

Zip Code



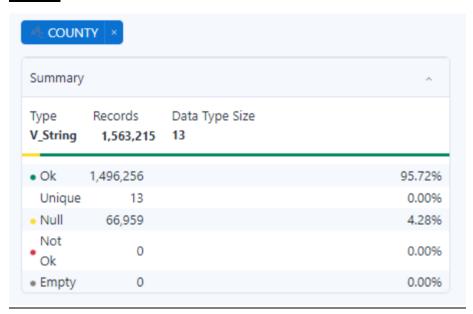
The data type of this column is "Double," and its size is 8. Its unique values are 66. The count of null values is 826.

Neighbourhood



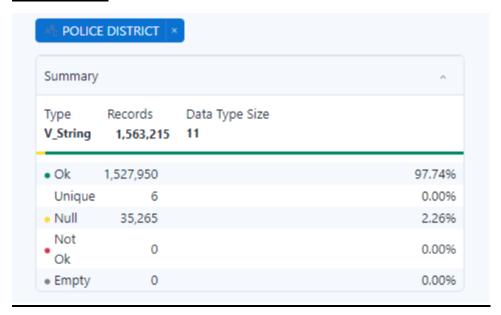
In addition to several null values, this field has 250 unique values.

County



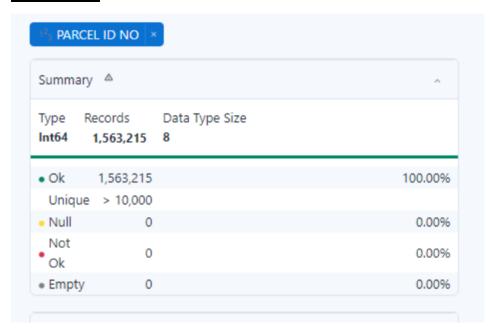
There are several null values and 13 distinct values in this column nation.

Police District



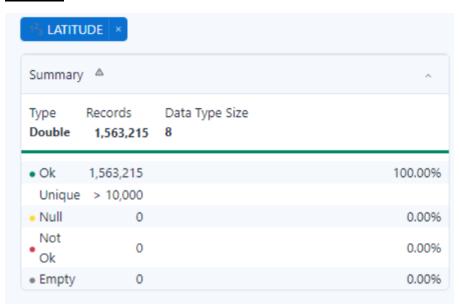
There are several null values in this column along with six unique values. The data type of this column is V String.

Parcel ID No



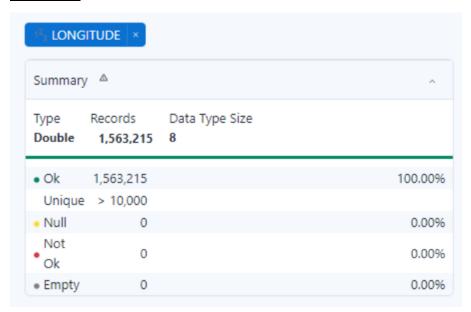
There are several distinct values in this column—none of them are null. It uses int64 as its data type.

Latitude



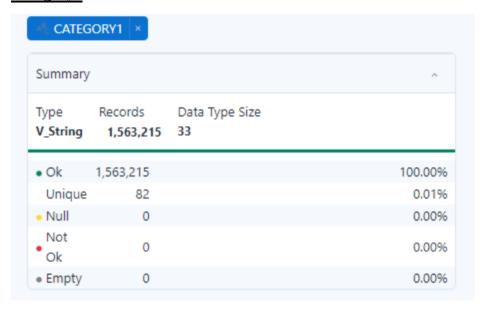
This column has a lot of unique values. There are 0 null values and the data type of this is Double.

Longitude



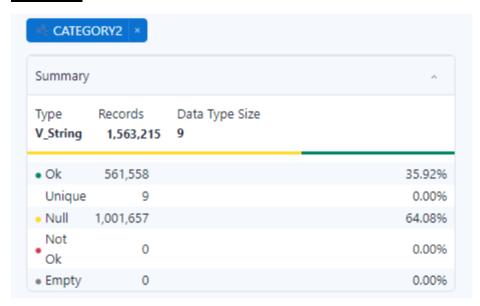
There are many unique values in this column and no null values. Double is the data type for this.

Category1



There are 82 distinct values in this column and no null values. The data type for this is V_String.

Category2



This has 9 unique values and a lot of null values.

Category3

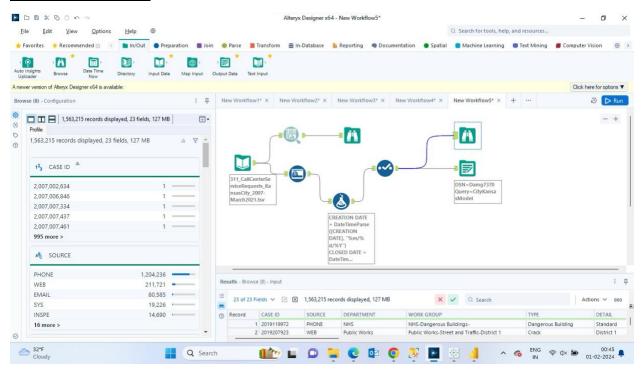


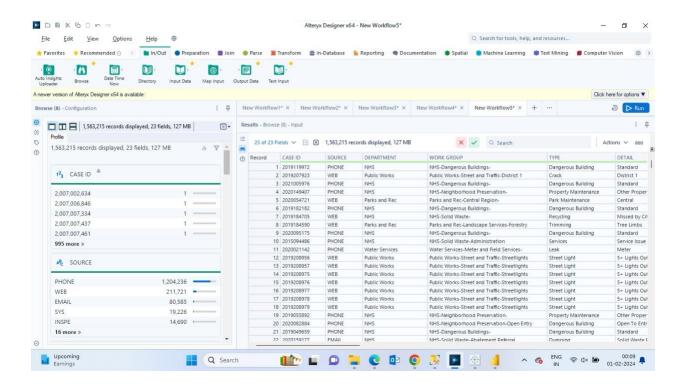
This column has 3 unique values and a lot of null values. This is of the data type V_String. approaches to data set cleaning.

- 1. Rows with null values can be eliminated: We can easily eliminate rows that are not necessary for analysis.
- 2. Replace null values with placeholders: If we wish to keep the rows that have null values but still show that there is no data, we may substitute placeholder values like "Unknown" for the null values.

3. Input Missing numbers: We can approximate missing numbers by using additional information. For instance, if you know the address or zip code, we may be able to determine the neighborhood.

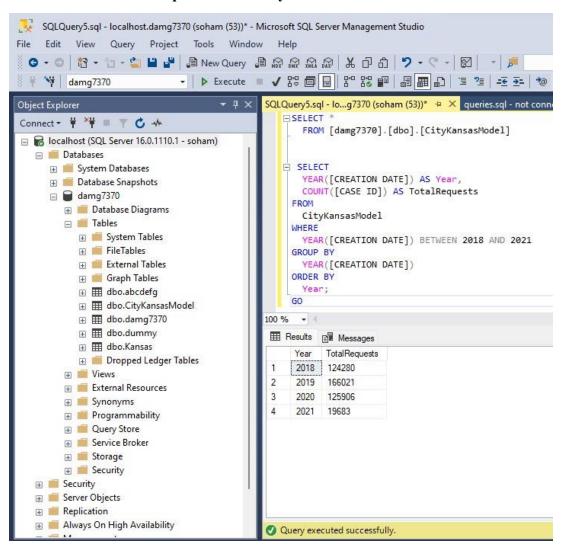
Workflow of Alteryx



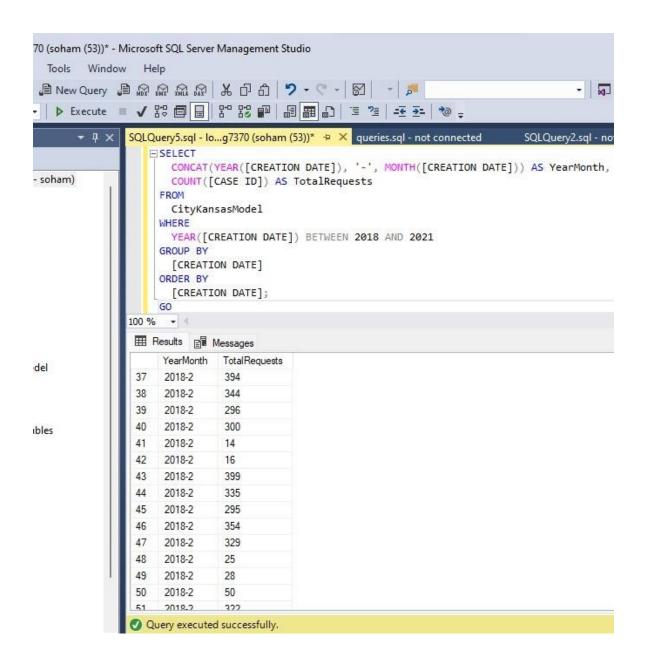


SQL Queries for verification.

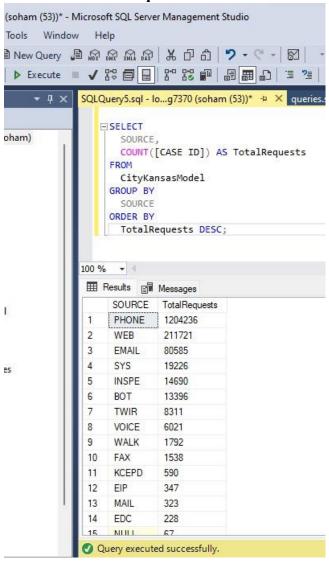
1. Total Service Requests Over the year 2018-2021



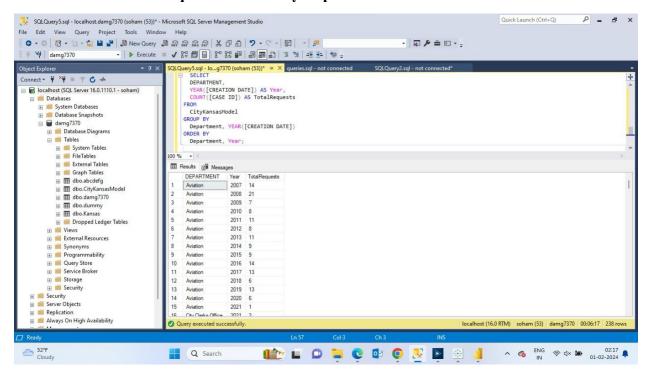
Service Requests monthly



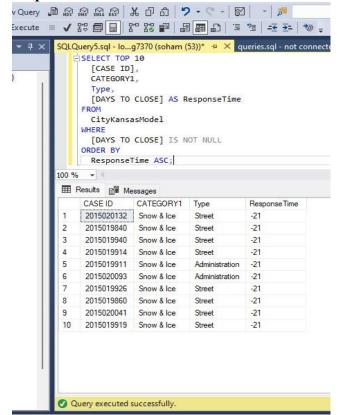
2. Volume of service requests received from different sources.



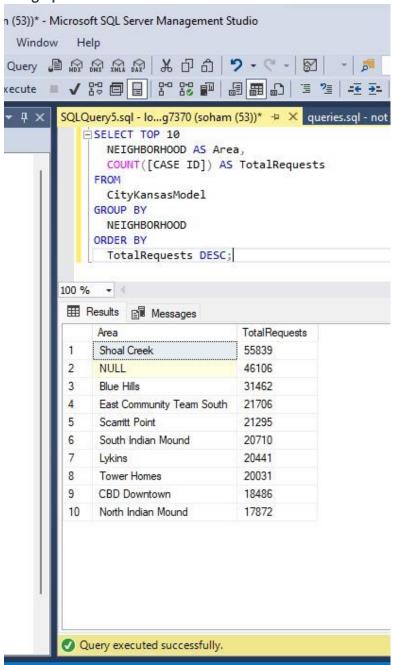
3. Volume of service requests received by Department.



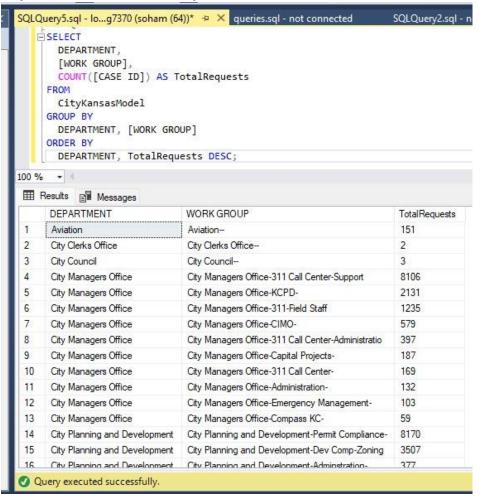
4. Top 10 Performance Metrics (Response Time) per CATEGORY and Type of Request:



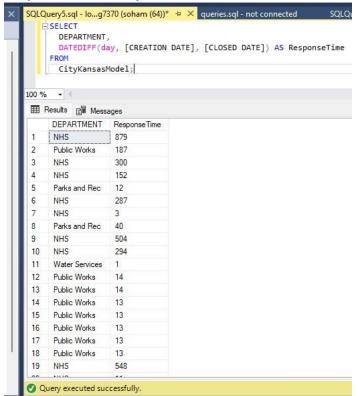
5. Geographical Visualization:



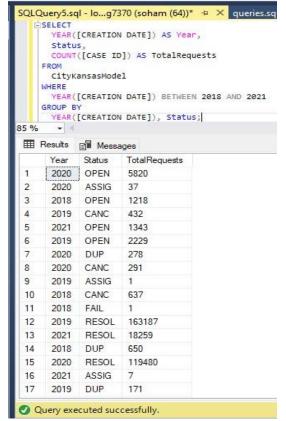
6. Departmental Workload Comparison:



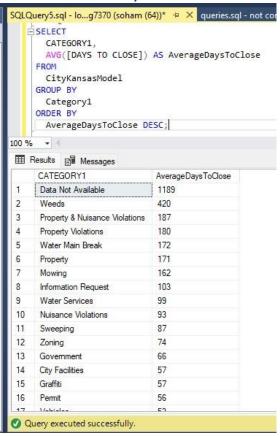
7. Response Time Analysis:



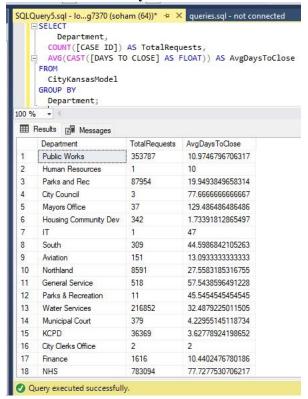
8. Service Request Status Composition:



9. Time to Closure Analysis:



10. Workload Efficiency



Row counts in sql.

