**TTC Delay Analysing Project for 2022 & 2023.**

Below are the fields available in the datasets:

A screenshot of a computer

Description automatically generated

Criteria 1: Number of delay incidents in each route (column: Route, Row: count).

**Data preparation:**

Steps followed to prepare the data fields.

1. Right click on field ‘Route’.
2. Select ‘Convert to Dimension’.

A screenshot of a computer

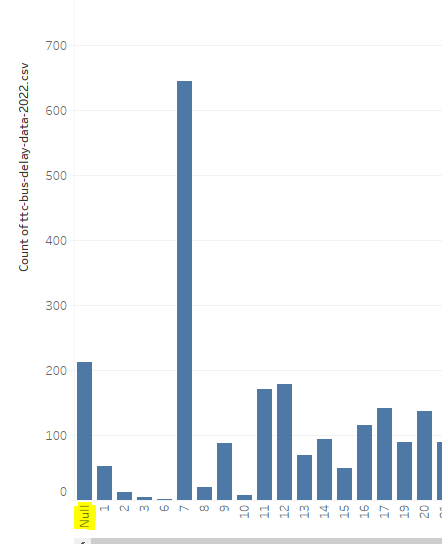
Description automatically generated

1. Drag the ‘Route’ field into Columns and Count variables(ttc-bus-delay-data-2022.csv) into Rows.

A screenshot of a computer

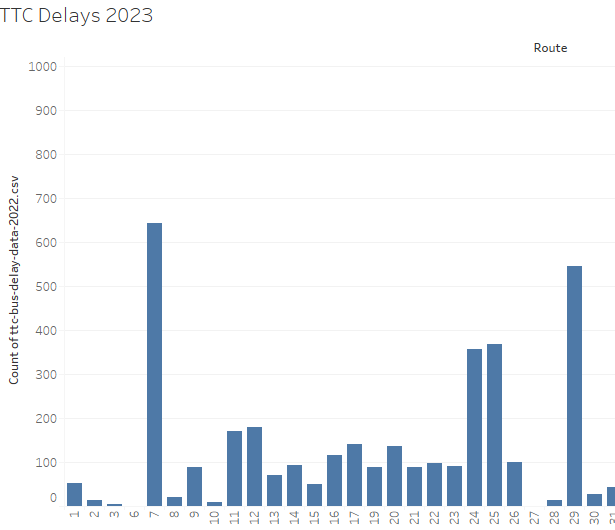
Description automatically generated

1. When you arrange the data, you can see that there is a ‘Null’ field at the beginning of the list of columns. This occurs because the ‘Route’ was not entered for some records when recording the data. These records are as bad data. Because there is no way that we can identify what are the data that should be entered to the missing ‘Route’ field. Therefore, these records must be removed from the data set, or it would not be useful in the process.



1. We can remove the null values, by adding ‘Route’ into Filters box. Then uncheck the ‘Null’. Then the ‘Null’ values will be removed from the data representation. A screenshot of a computer

   Description automatically generated
2. Then we will get a fine dataset without NULL values for the Route criteria.



1. Now we need to perform the same steps to the dataset of the year 2023.
2. In addition to that put the ‘Month’ field into ‘Filers box’ and then select ‘Month’. Then select only January to June from the checkbox. Because we have data from Jan to June for 2022.