**NYC TAXI TRIPS**

This dataset was chosen because of the well-defined data dictionary and the richness of the data across multiple years. The total size of the dataset was too huge to be analyzed by Excel. Therefore, the first step was to create a SQL database, create a table and insert the data. This provided a perfect opportunity to showcase SQL and PowerBI skills. It’s going to be challenging but fun.

Steps.

1. Study the data dictionary to understand the data types.

* The data types were identified based on the data dictionary definitions and observations from the data tables to place best match.
* Assumption: The description from the data dictionary table was correctly defined.

1. Create SQL Server database
2. Create a table NYC\_Trips and Taxi\_Zones
3. Import the data
4. Explore the data in SQL Server database
5. Aggregate the data before importing it into PowerBI
6. Crete data models

Challenges & Solutions

|  |  |  |
| --- | --- | --- |
| Serial No | Challenge | Solution |
| 1. | On Insert into Taxi\_Zones table, fields ‘Borough’, Zone\_Name’ & ‘Service\_Zone’ had limited characters as it was created as a VARCHAR only | Created an ALTER Query to increase the size of characters to 500 |
| 2. | Altering multiple columns gave an error | Created separate ALTER query for each field. |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |