NYC Parking Project

NYC Open Data

- Open Parking and Camera Violations
- Parking Violations Issued



BUSINESS INTELLIGENCE GUIDEBOOK

From Data Integration to Analytics



RICK SHERMAN

FOREWORD BY **CLAUDIA IMHOFF**PRESIDENT OF INTELLIGENT SOLUTIONS. INC.

Deliverables

Perform Data Profiling

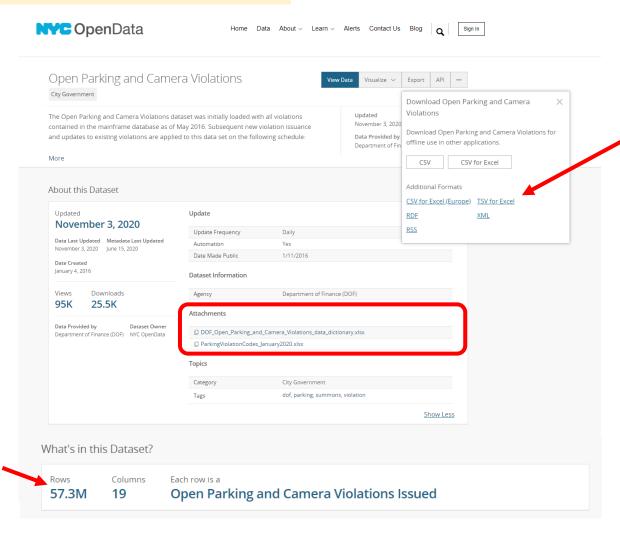
- Load data into stage & reject tables:
 - NYC Open Parking and Camera Violations data
 - NYC Parking Violations data for 2014 to 2021

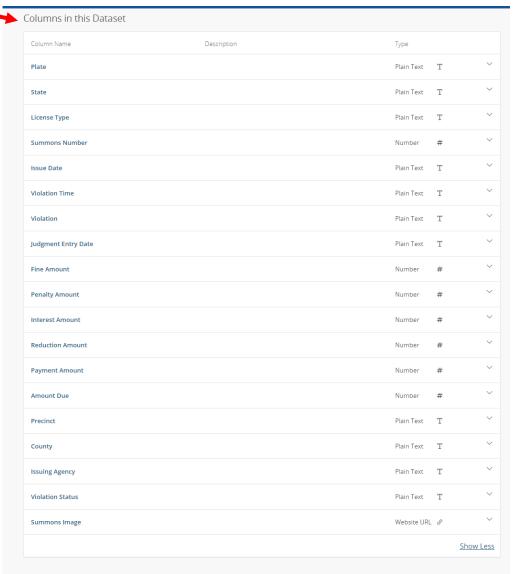
- Create dashboards
 - NYC Open Parking and Camera Violations
 - NYC Parking Violations



Open Parking and Camera Violations

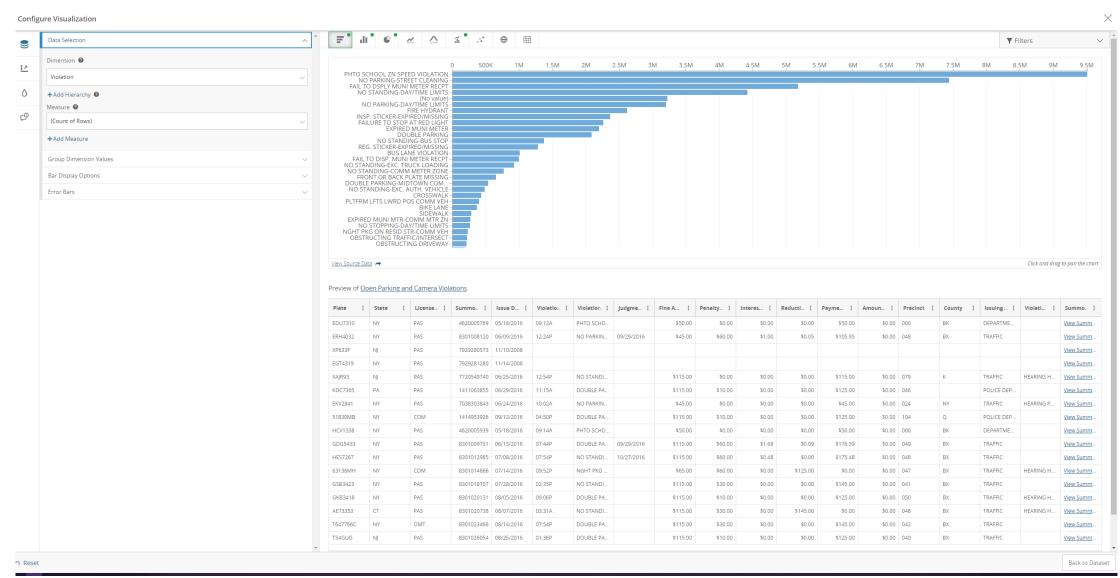
Note: image linked to NYC Open Data



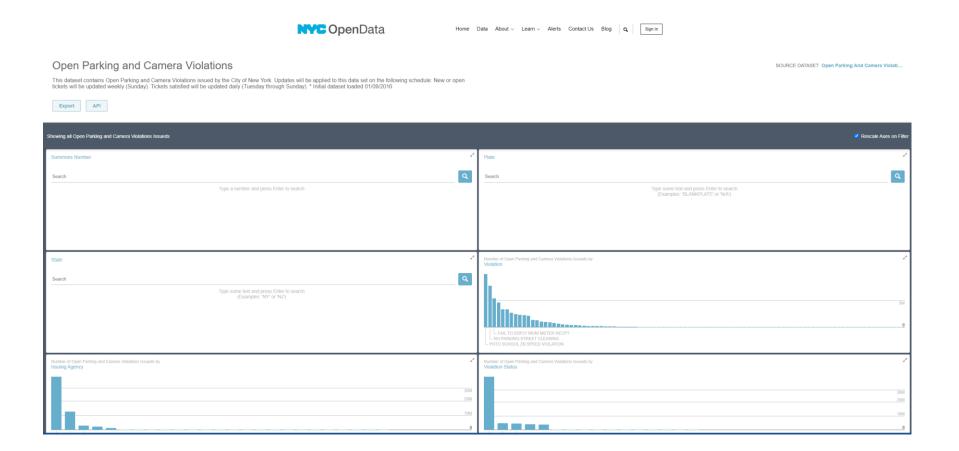


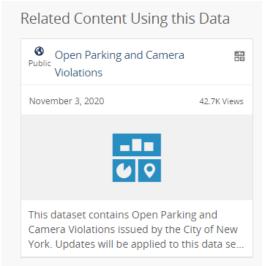


Open Parking and Camera Violations



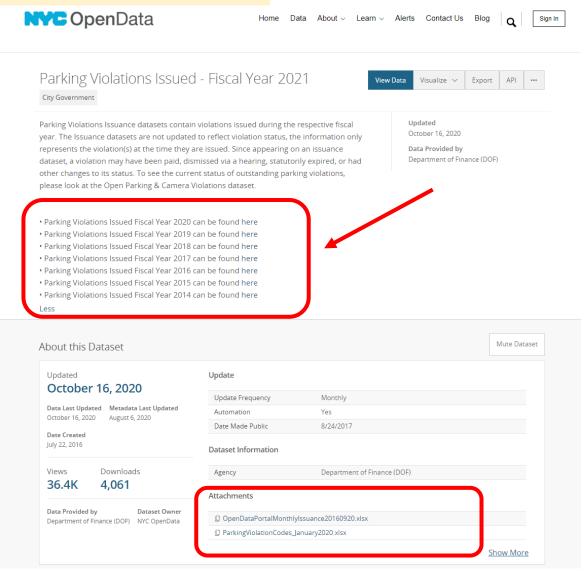
Open Parking and Camera Violations

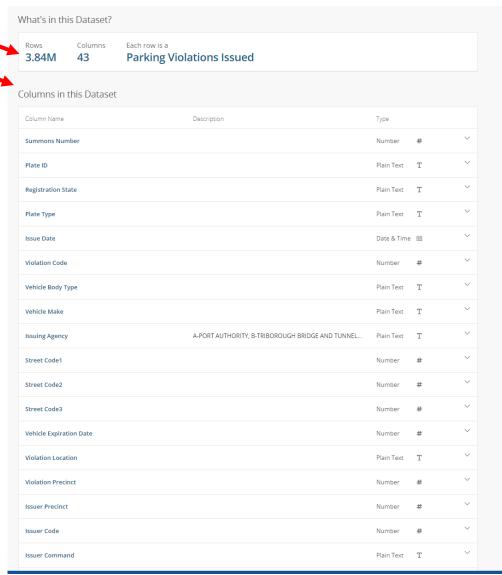




Parking Violations Issued - Fiscal Year 2021

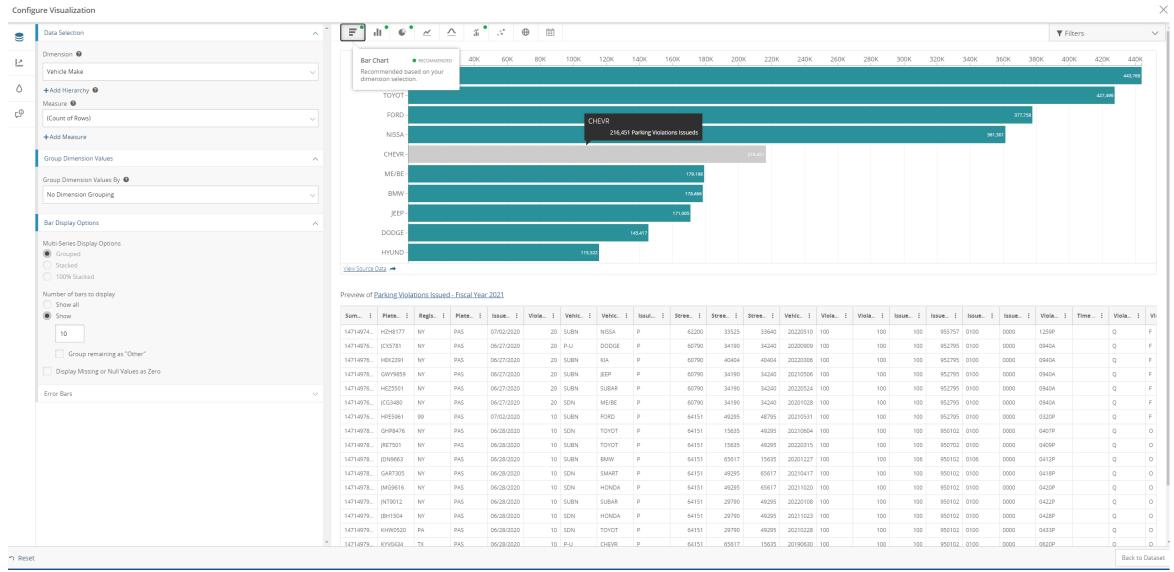
Note: image linked to NYC Open Data







Parking Violations Issued



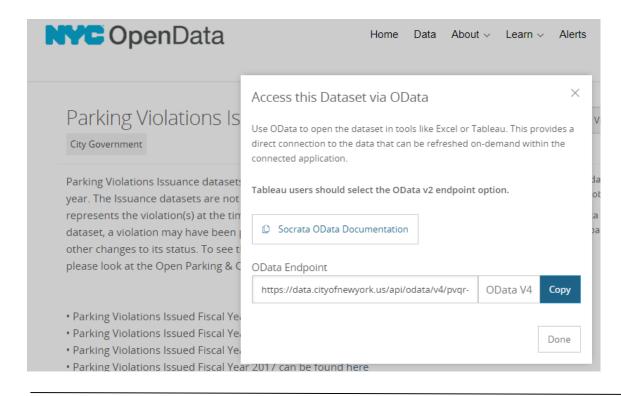
OData Connections

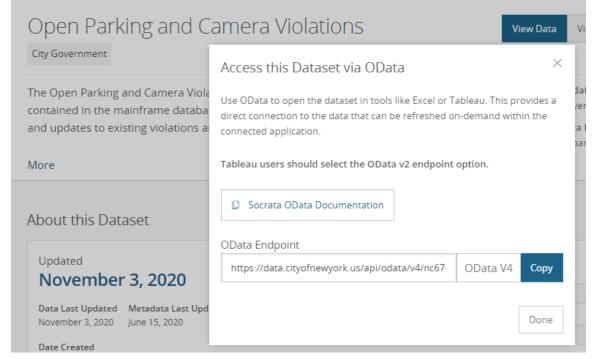
OData

https://data.cityofnewyork.us/api/odata/v4/pvqr -7yc4

OData

https://data.cityofnewyork.us/api/odata/v4/nc67 -uf89







Deliverables: NYC Open Parking and Camera Violations

- Load NYC Open Parking and Camera Violations data 57.3M rows
 - OneDrive \Data NYC Open Parking and Camera
 Violations\Open_Parking_and_Camera_Violations.tsv
 - Sql script: create_STG_open_parking_and_camera_violations .sql
 - STG_open_parking_and_camera_violations
 - STG_open_parking_and_camera_violations_rejects
 - Use Talend
- Upload
 - table counts total & by issue_date year
 - 2. Exported Talend jobs
 - 3. screenshots of completed Talend job
- Totals Rows: 57.3M



Deliverables: NYC Open Parking and Camera Violations

- Date data type: Issue_Date & Judgement_Entry_Date
- Time data type: Violation_Time
- Data quality checks
 - Issue_Date:
 - Invalid dates in source, i.e. 1/36/2014
 - Invalid data entry for dates, i.e. date greater than today's date (use for this example)
 - Judgement_Entry_Date:
 - o Invalid dates in source, i.e. 1/36/2014
 - Invalid data entry for dates, i.e. date greater than today's date (use for this example)
- Data cleansing
 - Zero length strings (or blanks) should be NULLs



Deliverables: NYC Parking Violations data for 2014 to 2020

- Load NYC Parking Violations data for 2014 to 2021
 - OneDrive \Data NYC Parking Violations
 - Parking_Violations_Issued_-_Fiscal_Year_2021.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2020.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2019.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2018.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2017.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2016.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2015.tsv
 - Parking_Violations_Issued_-_Fiscal_Year_2014.tsv
 - Sql script: create stg_parking_violations and rejects.sql
 - Use Talend

- Upload
 - table counts total & by issue_date year
 - 2. Exported Talend jobs
 - screenshots of completed talend jobs



Deliverables: NYC Parking Violations data for 2014 to 2020

- Load NYC Parking Violations
 - Issue_Date is date data type
 - Vehicle_Expiration_Date & Date_First_Observed are an integer in YYYYMMDD format except for 2015 (and you need to convert to integer)
 - Several columns are an integer data type and have 0 values that should become NULL
- Data quality checks
 - Issue_Date:
 - o Invalid dates in source, i.e. 1/36/2014
 - Invalid data entry for dates, i.e. date greater than today's date (use for this example)
 - Vehicle_Expiration_Date & Date_First_Observed as an integer data type with YYYYMMDD values
 need to have valid YYYYMMDD values
 - There may be a column with an integer data type that has a string as a value in the source
 - There may be a couple of columns that are strings that have lengths that do not fit into target table and appear to be invalid data

