

Food Inspections Assignment

A Data Analyst Journey

- Dealing with Food Establishment Inspections in three cities
 - Chicago
 - Dallas
 - New York
- All three cities are recording data on food inspections
 - Note that data differs in **content** & **schema**



Food Establishment Inspections

- Deliverables
 - Part 1: Week 1
 - Get data
 - Load data into Stage table(s) (stg_)
 - Perform data profiling
 - Part 2: Week 1
 - Design Dimensional Model
 - Create DDL SQL script
 - Part 3: Week 2
 - Load dimensional model into Integration tables (int_) from Stage table(s)
 - Part 4: Week 3
 - Create BI dashboards

Food Establishment Inspections

- Deliverables: Part 1
 - Get data from downloaded files from Open Data
 - Load data into Stage table(s) (stg_)
 - Select one for target database: MySQL, SQL Server
 - Follow Staging Guidelines
 - Perform data profiling
- Stage Table
 - Add **DI_CreateDate** – date & time row loaded
 - Add **DI_WorkflowFileName** – the file name of your Alteryx workflow
 - If any dates are stored as text, add a column where that text is converted to date
- Submit
 - Screenshot of Alteryx workflow completed
 - List time your job took to complete
 - DDL script for Stage table(s)

Food Establishment Inspections

- Deliverables: Part 2
 - Identify Dimensions & Facts
 - Create a Dimensional Data Model (ER/Studio or Navicat)
 - Create DDL for any database being used in this class
 - Create schema in chosen database
- Select one: MySQL, SQL Server
- Submit
 - Screenshots of each of the above
 - ER/Studio dm1 file, DDL scrips





Food Establishment Inspections


- Deliverables: Part 3
 - Create data preparation workflow(s) to load data into Integration Schema
 - i.e., dimensional model using Alteryx
 - Load data
- Stage Table
 - Add **DI_CreateDate** – date & time row loaded
 - Add **DI_WorkflowFileName** – the file name of your Alteryx workflow
 - All tables must have a surrogate key (SK)
 - Dates or Datetimes must be in date or datetime data type(s) not text
 - you can also have a date or datetime column in text if that is how came from source and there are errors in dates in source
 - Dates also must have a column representing a date SK i.e. YYYYMMDD
- Submit
 - Screenshots of each of the above
 - Alteryx file(s)
 - Table row counts

Food Establishment Inspections


- Deliverables: Part 4
 - Create BI dashboards in both
 - Power BI Desktop published in PBI Service
 - Tableau Desktop published in Tableau Online

NYC Food Inspections

 <https://opendata.cityofnewyork.us>   

 OpenData

311 | [Search all NYC.gov websites](#)

 OpenData

[Home](#) [Data](#) [About](#) [Learn](#) [Contact Us](#) | [Sign In](#)

Open Data for All New Yorkers

Open Data is free public data published by New York City agencies and other partners. **Learn from past Open Data Week events**, or **sign up for the NYC Open Data mailing list** to find training opportunities and upcoming events.

Search Open Data for things like 311, Buildings, Trees

NYC Open Data

2022 Progress Report

[Acknowledgements](#) [About](#) [Contact](#)

Home

Introduction

Strategic Plan Update


Updates to Data Quality and Documentation Standards

Dataset Highlights

Open Data by the Numbers

Compliance Plan

NYC Open Data: 2022 Progress Report



Learn about the next decade of NYC Open Data, and read our [2022 Report](#)

NYC Food Inspection DATA

- Data URL
 - [DOHMH New York City Restaurant Inspection Results | NYC Open Data \(cityofnewyork.us\)](https://cityofnewyork.us)



DOHMH New York City Restaurant Inspection Results

Health

The dataset contains every sustained or not yet adjudicated violation citation from every full or special program inspection conducted up to three years prior to the most recent inspection for restaurants and college cafeterias in an active status on the RECORD DATE (date of the data pull). When an inspection results in more than one violation, values for associated fields are repeated for each additional violation record. Establishments are uniquely identified by their CAMIS (record ID) number. Keep in mind that thousands of restaurants start business and go out of business every year; only restaurants in an active status are included in the dataset. Records are also included for each restaurant that has applied for a permit but has not yet been inspected and for inspections resulting in no violations. Establishments with inspection date of 1/1/1900 are new establishments that have not yet received an inspection. Restaurants that received no violations are represented by a single row and coded as having no violations using the ACTION field.

Because this dataset is compiled from several large administrative data systems, it contains some illogical values that could be a result of data entry or transfer errors. Data may also be missing.

This dataset and the information on the Health Department's Restaurant Grading website come from the same data source. The Health Department's Restaurant Grading website is here: <http://www1.nyc.gov/site/doh/services/restaurant-grades.page>

See the data dictionary file in the Attachments section of the OpenData website for a summary of data fields and allowable values.

Updated
June 11, 2023

Data Provided by
Department of Health and Mental Hygiene (DOHMH)

About this Dataset

Mute Dataset

Updated
June 11, 2023

Data Last Updated
June 11, 2023

Metadata Last Updated
June 11, 2023

Date Created
August 1, 2014

Views
93.4K

Downloads
47.7K

Data Provided by
Department of Health and Mental Hygiene (DOHMH)

Dataset
Owner
NYC
OpenData

Dataset Information

Agency	Department of Health and Mental Hygiene (DOHMH)
Update	
Update Frequency	Daily
Automation	Yes
Date Made Public	7/10/2015

Attachments

- [RestaurantInspectionDataDictionary_09242018.xlsx](#)
- [About_NYC_Restaurant_Inspection_Data_on_NYC_OpenData_050222.docx](#)

Topics

Category	Health
Tags	restaurant, inspection, violation, grade, adjudication, fines, 2018od4a-report, 2018od4a-video, food safety

What's in this Dataset?

Rows
198K

Columns
32

Each row is a
Restaurant Citations

NYC Food Inspection DATA

DOHMH New York City Restaurant Inspection Results

Health

The dataset contains every sustained or not yet adjudicated violation citation from every full or special program inspection conducted up to three years prior to the most recent inspection for restaurants and college cafeterias in an active status on the RECORD DATE (date of the data pull). When an inspection results in more than one violation, values for associated fields are

[More](#)

Updated
June 11, 2023

Data Provided by
Department of Health
(DOHMH)

View Data

Visualize

Export

API

...

Download DOHMH New York City Restaurant Inspection Results

Download DOHMH New York City Restaurant Inspection Results for offline use in other applications.

CSV

KML

Shapefile

Additional Formats

[CSV for Excel](#)

[KMZ](#)

[TSV for Excel](#)

[CSV for Excel \(Europe\)](#)

[RDF](#)

[XML](#)

[GEOJSON](#)

[RSS](#)

About this Dataset

Updated

June 11, 2023

Dataset Information

NYC Food Inspection DATA – Columns

Columns in this Dataset

Column Name	Description	Type		
CAMIS	This is an unique identifier for the entity (restaurant); 10-digit i...	Plain Text	T	▼
DBA	This field represents the name (doing business as) of the entit...	Plain Text	T	▼
BORO	Borough in which the entity (restaurant) is located.;• 1 = MAN...	Plain Text	T	▼
BUILDING	Building number for establishment (restaurant) location	Plain Text	T	▼
STREET	Street name for establishment (restaurant) location	Plain Text	T	▼
ZIPCODE	Zip code of establishment (restaurant) location	Plain Text	T	▼
PHONE	Phone Number; Phone number provided by restaurant owner/...	Plain Text	T	▼
CUISINE DESCRIPTION	This field describes the entity (restaurant) cuisine. ; Optional fi...	Plain Text	T	▼
INSPECTION DATE	This field represents the date of inspection; NOTE: Inspection ...	Date & Time	📅	▼
ACTION	This field represents the actions that is associated with each r...	Plain Text	T	▼
VIOLATION CODE	Violation code associated with an establishment (restaurant) i...	Plain Text	T	▼
VIOLATION DESCRIPTION	Violation description associated with an establishment (resta...	Plain Text	T	▼
CRITICAL FLAG	Indicator of critical violation; "• Critical • Not Critical • Not Appli...	Plain Text	T	▼

SCORE	Total score for a particular inspection; Scores are updated bas...	Number	#	▼
GRADE	Grade associated with the inspection; • N = Not Yet Graded• A ...	Plain Text	T	▼
GRADE DATE	The date when the current grade was issued to the entity (rest...	Date & Time	📅	▼
RECORD DATE	The date when the extract was run to produce this data set	Date & Time	📅	▼
INSPECTION TYPE	A combination of the inspection program and the type of inspe...	Plain Text	T	▼
Latitude		Number	#	▼
Longitude		Number	#	▼
Community Board		Plain Text	T	▼
Council District		Plain Text	T	▼
Census Tract		Plain Text	T	▼
BIN		Plain Text	T	▼
BBL		Plain Text	T	▼
NTA		Plain Text	T	▼
Location Point1		Point	📍	▼
Zip Codes	This column was automatically created in order to record in w...	Number	#	▼
Community Districts	This column was automatically created in order to record in w...	Number	#	▼
Borough Boundaries	This column was automatically created in order to record in w...	Number	#	▼
City Council Districts	This column was automatically created in order to record in w...	Number	#	▼
Police Precincts	This column was automatically created in order to record in w...	Number	#	▼

NYC Food Inspection DATA – Columns

Related Content Using this Data

 DOHMH New York City
Restaurant Inspection Results

June 11, 2023 24.7K Views



This dataset provides restaurant inspections, violations, grades and adjudication information

NYC OpenData

[Home](#) [Data](#) [About](#) [Learn](#) [Alerts](#) [Contact Us](#) [Blog](#) [Q](#)

[Sign In](#)

DOHMH New York City Restaurant Inspection Results

This dataset provides restaurant inspections, violations, grades and adjudication information

[Export](#)

[API](#)

SOURCE DATASET [DOHMH New York City Restaurant I...](#)

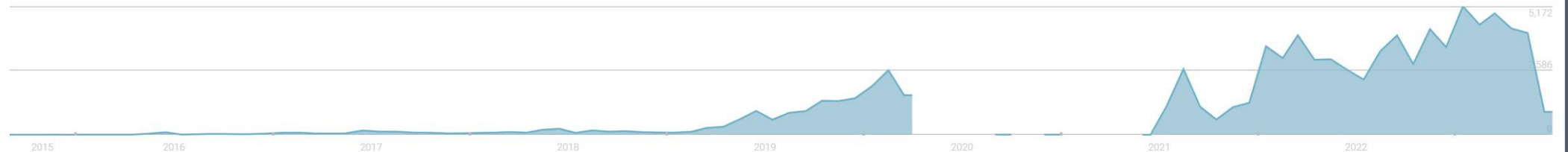
Showing all Restaurant Citations

☒ Rescale Axes on Filter

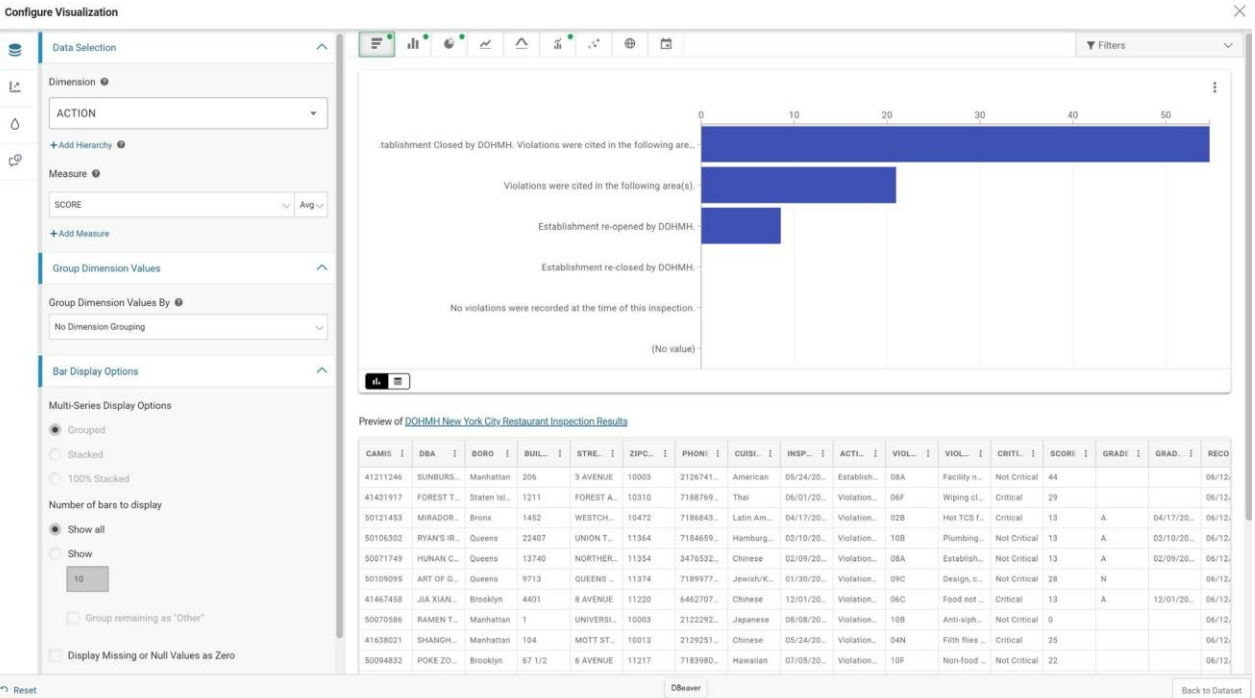
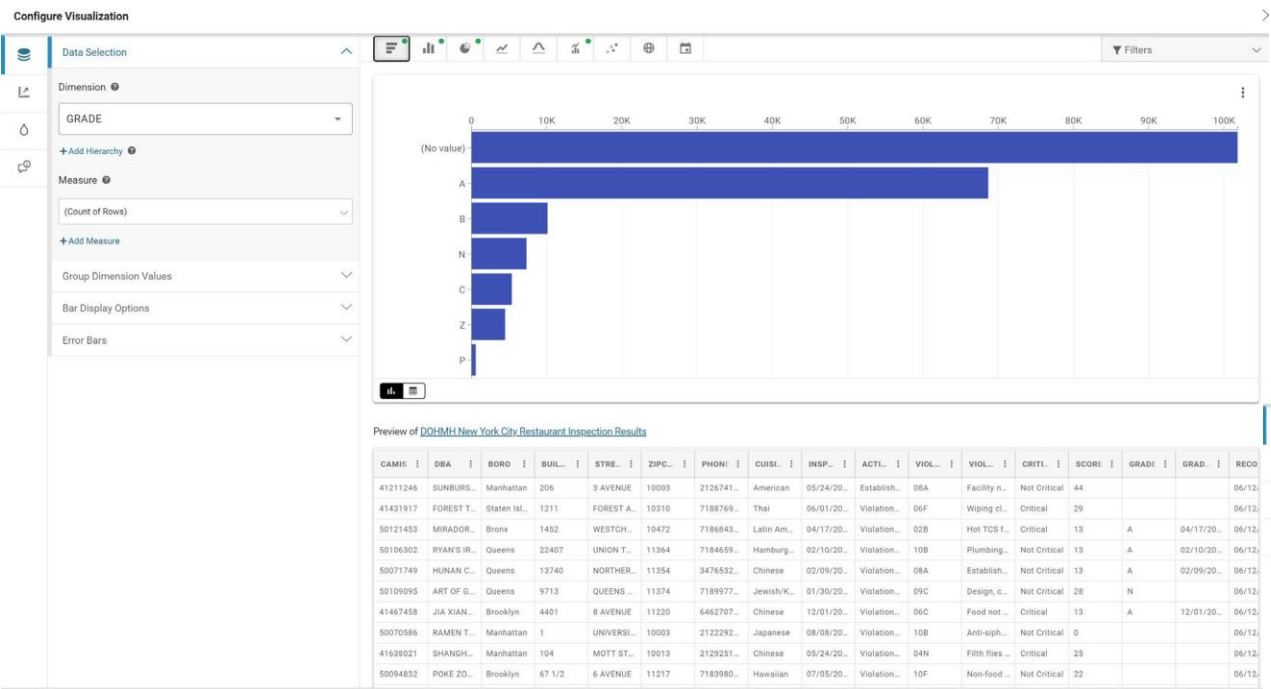
Number of Restaurant Citations by
GRADE DATE

The date when the current grade was issued to the entity
(restaurant)

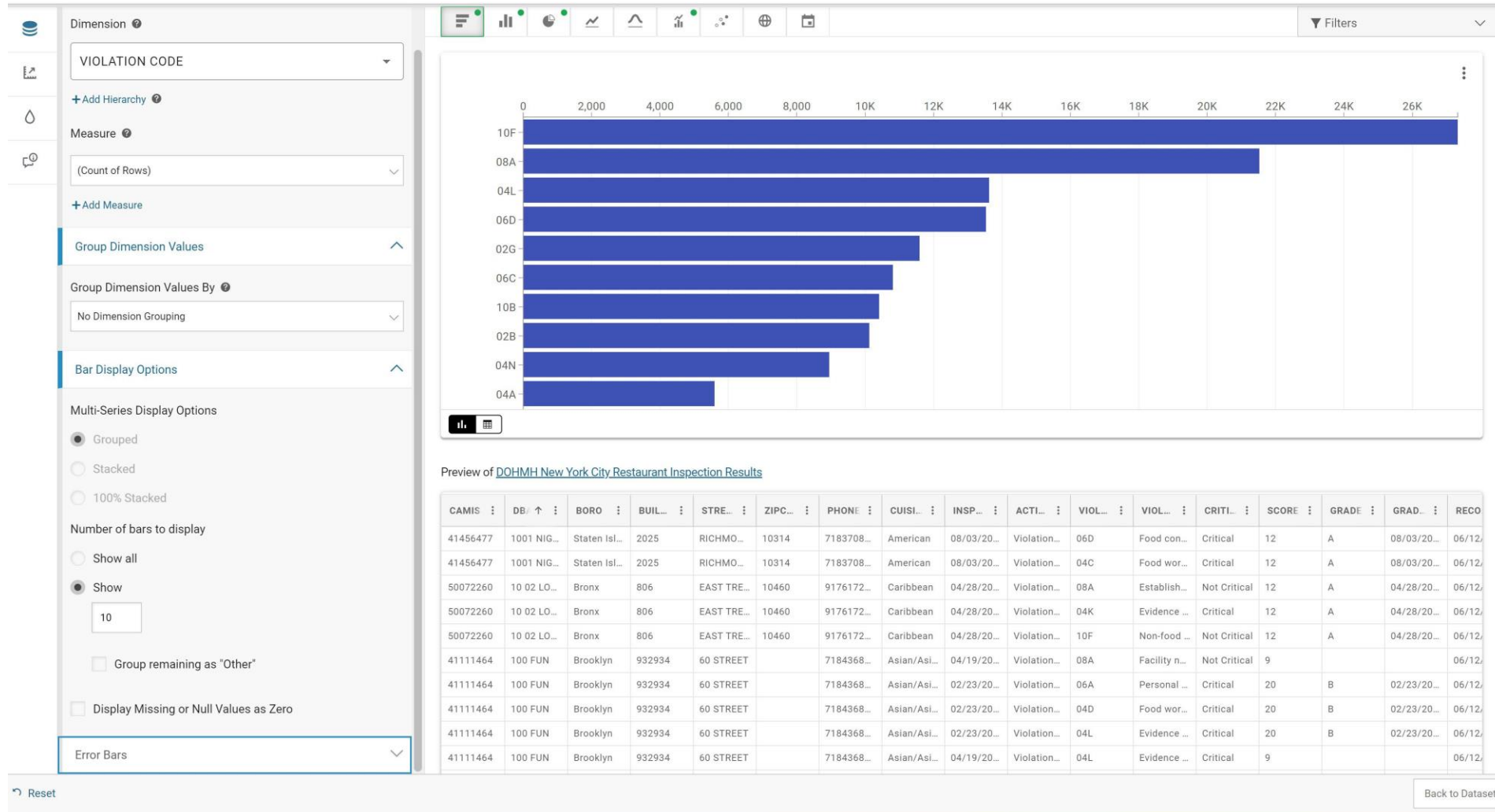
Less



NYC Food Inspection DATA – Columns



NYC Food Inspection DATA – Columns



NYC Food Inspection – Business questions to answer

- Examine food inspection results by –
 - Inspection type
 - Inspection result, Score, Grade
 - Actions
 - Facility type
 - Violations (Codes, descriptions)
 - Business inspected
 - DBA (Doing Business As), AKA (Also Know As), License
 - Location: Borough, Address, other
 - Trends over time

NYC Food Inspection – Business questions to answer

- Data Visualizations examples
 - Restaurants Grades – Map
 - Longitude, Latitude, Grade, Zip Code to create a map
 - Street and Grade for filtering the results
 - Distribution of Cuisine Type – Heatmap
 - Cuisine Description, Number of Records
 - The Distribution of Critical Flags by Neighborhood – Bar Chart
 - Boro, Number of records, Critical flag
 - Number of Inspections by Neighborhood from 2015 to 2020 – Line Chart
 - Year, Boro, Number of Records
 - Grades Distribution by Neighborhood – Horizontal Bar Chart
 - Grade, Boro, Number of Records

NYC Food Inspection – Data Preparation

- Load Stage & Perform Data Profiling
- Inspection Results Data to Staging Schema: Workflow
- Inspection Results Data to Staging Schema: Results
 - Time to load
 - Row counts

NYC Food Inspection – Creating dimensional data model

- Food Inspections
 - Determine Dimensions & Facts
 - Data Model

NYC Food Inspection – Load to Integration Schema (Dimensional Model)

- Load Facts & Dimensions
 - Number of tables count
 - Number of row counts in each table

NYC Food Inspection – data visualizations

Power BI and Tableau

- How many food inspections over time
- What have been the inspection results over time
 - Pass vs fail
 - Grades, scores, or score ranges
 - Number, severity and types of violations
- Food establishments inspected:
 - Top ten most inspected (over the last 2 full years)
 - Top ten with worst results (over the last 2 full years)
- Inspections and inspection results by food establishment attributes. Gifferent data sets have different attributes available, such as:
 - Food establishment business type
 - Cuisine
 - NYC boroughs
 - Too be determined (TBD) – you provide possibilities based on the dataset
- Map food inspections (if possible)