

NYC Food Inspection

Submission by Yamini Manral

NUID 002703881

Part 1

- 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)

DOHMH New York City Restaurant Inspection Results

Health

View Data

Visualize

Export

API

...

Updated
June 18, 2023

Data Provided by
Department of Health
(DOHMH)

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](#)

Exporting
data as
TSV

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:

Workflow for staging Table

Alteryx Designer x64 - nyc_food_inspection.yxmd*

File Edit View Options Help

Search for tools, help, and resources

★ Favorites Recommended In/Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface

Browse Input Data Output Data Text Input Data Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Workflow - Configuration Canvas Workflow Runtime Events Meta Info

Canvas Options

Layout Direction: Horizontal

Annotations: Show

Connection Progress: Show Only When Run

Start Here.yxmd x nyc_food_inspection.yxmd* x +

DOHMH_New_Yor
k_City_Restaurant
Inspection_Results
.tsv

Convert GRADE
DATE From:
MM/dd/yyyy

Convert RECORD
DATE From:
MM/dd/yyyy

Convert
INSPECTION
DATE From:
MM/dd/yyyy

DI_CreatedDate =
DateTimeNow()
DI_WorkflowFile
Name =
[Engine.Workflow
FileName]

aka:nyc_food_insp
ection
Query=stg_nyc_fo
od_inspection

Results - Workflow - Messages

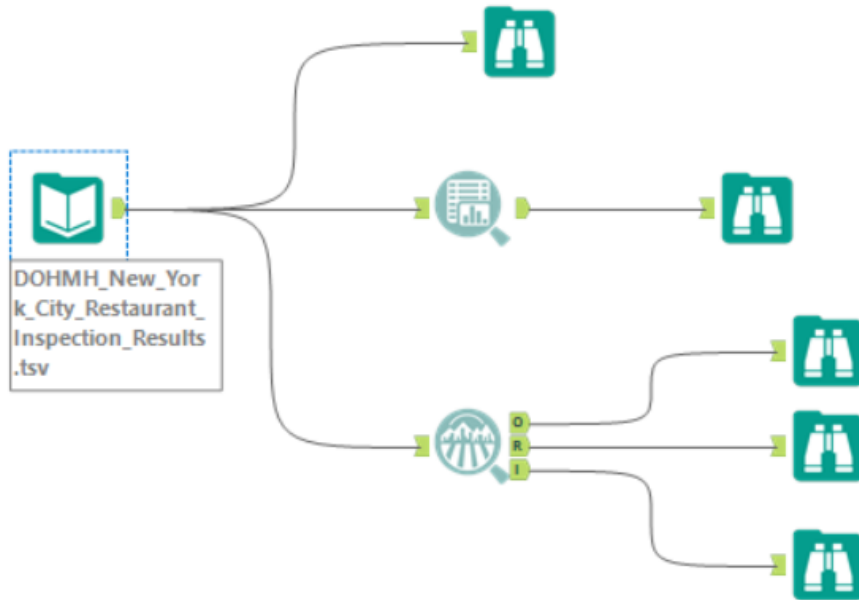
All 0 Errors 11 Conv Errors 0 Warnings 32 Messages 2 Files

Last Run Configuration

Auto Field (2)	The FieldType of "BIN" changed to: Int32
Auto Field (2)	The FieldType of "BBL" changed to: Int64
Auto Field (2)	The FieldType of "NTA" changed to: String(4)
Auto Field (2)	The FieldType of "Location Point1" changed to: Bool
Output Data (10)	Alias translated to odbc:DRIVER={ODBC Driver 17 for SQL Server};UID=damg7370;PWD=__EncPwd1__;DATABASE=nyc_food_inspection;SERVER=localhost stg_nyc_food_inspection
Output Data (10)	ODBC Driver version: 03.80
Browse (9)	198,329 records
Output Data (10)	198329 records were written to aka:nyc_food_inspection (stg_nyc_food_inspection)
Designer x64	Finished running nyc_food_inspection.yxmd in 45.7 seconds with 11 field conversion errors

dbo.stg_nyc_food_inspection
Columns
RecordID (int, null)
CAMIS (int, null)
DBA (nvarchar(95), null)
BORO (varchar(13), null)
BUILDING (varchar(10), null)
STREET (varchar(40), null)
ZIPCODE (char(5), null)
PHONE (char(12), null)
CUISINE DESCRIPTION (varchar(30), null)
INSPECTION DATE (char(10), null)
ACTION (varchar(130), null)
VIOLATION CODE (char(5), null)
VIOLATION DESCRIPTION (nvarchar(254), null)
CRITICAL FLAG (char(14), null)
SCORE (tinyint, null)
GRADE (char(1), null)
GRADE DATE (char(10), null)
RECORD DATE (char(10), null)
INSPECTION TYPE (varchar(254), null)
Latitude (float, null)
Longitude (float, null)
Community Board (smallint, null)
Council District (char(2), null)
Census Tract (char(6), null)
BIN (int, null)
BBL (bigint, null)
NTA (char(4), null)
Location Point1 (bit, null)
GradeDate_new (date, null)
RecordDate_new (date, null)
InspectionDate_new (date, null)
DI_CreatedDate (nvarchar(max), null)
DI_WorkflowFilename (nvarchar(max), null)

Stage table creation



Workflow for Data profiling

Results - Workflow - Messages

All	0 Errors	11 Conv Errors	0 Warnings	5 Messages	6 Files	Last Run	Configuration
?	Input Data (1)	There were 23053 values that were truncated (1st 10 shown)					
	Input Data (1)	198329 records were read from "C:\Users\Yamini Manral\OneDrive\Desktop\Summer Semester\DAMG7370 DADABI\Assignments - DADABI\A5\NYC\DOHMH New York City Restaurant Inspection R					
	Browse (2)	198,329 records					
	Designer x64	The Designerx64 reported: Beginning to compact waiting packets to reduce memory usage					
	Browse (8)	567 records					
	Browse (6)	1 records					
	Browse (5)	27 records					
	Browse (7)	1 records					
!	Designer x64	Finished running chicago_data_profiling.yxmd in 7.4 seconds with 11 field conversion errors					

Part 2

- 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

Dimensional Model

