

## Food Inspections Report - An end-to-end project

Saipavan Katta

Aim: To take data from two datasets, clean, transform, and load into staging tables, to MySQL DB. Then perform dimensional modeling and create dimensions, facts tables in MySQL DB. Load data into dimensions and fact tables in MySQL DB from the staging table. Connect PowerBI and Tableau to MySQL DB to create visualizations and address business requirements.

### Datasets Overview:

#### Chicago Dataset:

This dataset has been derived from inspections of restaurants and other food establishments in Chicago from January 1, 2010, to the present. Inspections are performed by staff from the Chicago Department of Public Health's Food Protection Program using a standardized procedure. The results of the inspection are inputted into a database, then reviewed and approved by a State of Illinois Licensed Environmental Health Practitioner (LEHP).

#### Dallas Dataset:

This dataset conveys essential information about various establishments, including their names, physical addresses, inspection dates, overall inspection scores, and specific point deductions assigned for individual violations. It provides a comprehensive overview of the inspection process, highlighting key details such as the establishment's identity, location, inspection timeline, overall assessment, and the specific areas where point deductions occurred due to violations. The dataset serves as a valuable resource for understanding and analyzing the performance and compliance of different establishments during inspections.

### Project steps:

#### Step 1: Understanding the Dataset

- Analyze the Schema and Content of the provided dataset
- Identified potential data quality issues, outliers, or anomalies

#### Step 2: Data Profiling in Alteryx

- Imported both the datasets separately into Alteryx
- Utilized Alteryx tools for data profiling and perform transformations
- Studied the datatype of each column, their type (single valued or multi valued)

#### Step 3: Devising a common schema for the final table

- Based on the data available in both datasets and the business requirements, a schema devised

for the final table

#### Step 4: Data transformation in Alteryx

- Perform data transformations to clean the data, separate multivalued columns into separate ones, format the columns as required, clean, address the null values, datatypes and add audit columns
- Load the data into staging tables, present in MySQL DB

#### Step 5: Created a dimensional model using ER Studio

- Design fact table, dimensions based on the business requirements
- Generate physical model, DDL script and run it in MySQL DB

#### Step 6: Load the data into dimensional models using Talend

- Loaded data in the dimensional model with Talend, into MySQL DB
- Added audit columns

#### Step 7: Connect to BI tools

- Established connection to MySQL DB in Tableau C PowerBI
- Created dashboards based on business requirements
- Publish the dashboards to cloud

#### Data Profiling:

Chicago Dataset: 280,858 records

Name	Data Type	Missing Values (%)	Unique Values	Observations	Min Value	Max Value
AKA Name	V_WString	0.865864083	31385	This field records the commonly called names of the restaurants, it has an higher number of missing values	EL	MAE'S EARLY CHILDHOOD DEVELOPMENT AND THERAPEUTIC DAY CARE CENTER, INCORPORATED
Longitude	V_WString	0.347271691	18412	Records the longitude co-ordinates of the restaurants, it has null data of 34%.	-87.66946465	-87.62616828
Address	V_WString	0	19874	Captures complete address info of the restaurants.		1400 S JEAN BAPTISTE POINTE DUSABLE LAKESHORE DRIVE
State	V_WString	0.022439094	7	Captures the state names info of the restaurants.	IL	IL
DBA Name	V_WString	0	32990	Records the registered names of the business, it has no null values	N	MAE'S EARLY CHILDHOOD DEVELOPMENT AND THERAPEUTIC DAY CARE CENTER, INCORPORATED
Inspection ID	V_WString	0	280760	Unique identifier exclusive of Chicago dataset, helped in uniquely identifying records	88698	2603151
License #	V_WString	0.00641117	45806	Alternate key exclusive of Chicago dataset assigned by Dept.of Business Affairs and Consumer Protection.	0	1617528

Latitude	V_WString	0.347271691	18412	Records the latitude co-ordinates of the restaurants. Helpful in performing geographic visualizations	41.78613073	41.88310863
City	V_WString	0.055207294	86	Captures the city names of the restaurants located.	CH	BANNOCKBURNDEERFIELD
Violations	V_WString	27.59189343	201864		63. REMOVAL OF SUSPENSION SIGN	2. FACILITIES TO MAINTAIN PROPER TEMPERATURE - Comments: All food establishments that display, prepare, or store potentially hazardous food shall have adequate refrigerated food storage facilities.....
Risk	V_WString	0.029562616	5	Restaurants are classified as per the risk levels from 1 to 3 with 1 representing highest risk.	All	Risk 2 (Medium)
Location	V_WString	0.347271691	18412	This column represents the combined co-ordinates of latitude and longitude of the establishments	(41.73536766705, -87.638879040659)	(41.883108627266935, -87.62794877080378)
Inspection Type	V_WString	0.000356176	111	Categorizes various inspection types such as routine canvass and complaint responses.	SFP	LICENSE TASK FORCE / NOT -FOR-PROFIT CLUB
Inspection Date	V_WString	0	3738	Column representing the dates of each inspection	9/19/2024	9/19/2024
Facility Type	V_WString	1.841074227	520	Column that represents the type of facilities that are present in the respective establishments.	bar	Mobile Frozen Dessert Dispenser_non Motorized.
Zip	V_WString	0.01460322	127	Column that captures the zipcodes of the establishments.	60656	60656
Results	V_WString	0	7	Column that captures the outcomes of the inspections whether the establishments passed, passed with conditions, or failed.	Pass	Business Not Located

Dallas Dataset: 78,984 records

Name	Observations	Data_Type	Percent Missing	Unique Values	Min Value	Max Value
Inspection Date	Date when the inspection happened	V_WString	47.11444452	2309	10/30/2019	10/30/2019
Inspection Month		V_WString	47.11444452	90	Oct-19	Oct-19
Inspection Score		V_WString	47.11444452	59	0	100
Inspection Type	Code indicating the inspection type, such as	V_WString	47.11444452	4	Routine	Complaint

	Routine, Follow-up, Complaint, Temporary and Mobile.					
Inspection Year	Year part of the Inspection date	V_WString	47.11444452	9	FY2020	FY2020
Lat Long Location	Denotes a location point on a longitude line (perpendicular to the equator) and latitude line (parallel to the equator)	V_WString	47.11444452	7824	907 ELM ST	"4351 DALLAS FORT WORTH TPKE #110
Restaurant Name	Name of the Restaurant	V_WString	0.014062814	21085	KFC	(32.768615587000056, 96.82313404199994)"
Street Address	Street Address of the restaurant	V_WString	47.11444452	7824	605 ELM ST	50326 THE TRADITION-PRESTONWOOD ASSISTED LIVING AND MEMORY-4TH FL
Street Direction	Direction of the Street (N,W,E,S)	V_WString	82.53465479	5	W	7932 S GREAT TRINITY FOREST WAY #101A
Street Name	Name of the street where the resturant is there	V_WString	47.11444452	843	Tl	W
Street Number	Street number for the address of the facility.	V_WString	47.11444452	3444	0	MOUNTAIN CREEK PKWY #BU23
Street Type	Street type for the address of the facility. For example, AVE, LN, ST, etc.	V_WString	48.23545168	20	RD	10295
Street Unit	Unit number or apartment number for the address of the facility.	V_WString	81.13105203	991	3	BLVD
Violation Description - 1	Description for the type of violation.	V_WString	51.51543561	624	*23 No Water	#3000
Violation Description - 10	Description for the type of violation.	V_WString	93.12462332	413	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 11	Description for the type of violation.	V_WString	95.38873636	378	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 12	Description for the type of violation.	V_WString	97.1385522	330	*23 No Water	*21 Person in charge. The person in charge shall ensure that §228.34 (1-15) Relating to Duties of PI
Violation Description - 13	Description for the type of violation.	V_WString	98.18254872	290	*23 No Water	*21 Person in charge. The person in charge shall ensure that §228.34 (1-15) Relating to Duties of PI
Violation Description - 14	Description for the type of violation.	V_WString	98.72162325	277	*23 No Water	*21 Person in charge. The person in charge shall ensure that §228.34 (1-15) Relating to Duties of PI
Violation Description - 15	Description for the type of violation.	V_WString	99.09931025	238	*45 First Aid	*09 Food protected cross contamination by preparing each type

						of food at different times or in separate
Violation Description - 16	Description for the type of violation.	V_WString	99.40065626	192	*45 First Aid	*39 In-use utensils, between-use storage. During pauses in food preparation or dispensing, food prep
Violation Description - 17	Description for the type of violation.	V_WString	99.62967923	177	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 18	Description for the type of violation.	V_WString	99.76294114	137	*45 First Aid	*39 In-use utensils, between-use storage. During pauses in food preparation or dispensing, food prep
Violation Description - 19	Description for the type of violation.	V_WString	99.86071118	95	*47 Other Violations	*21 Person in charge. The person in charge shall ensure that §228.34 (1-15) Relating to Duties of PI
Violation Description - 2	Description for the type of violation.	V_WString	56.57001272	612	*23 No Water	*09 Food protected from cross contamination separating types of raw animal food storage, preparation, holding
Violation Description - 20	Description for the type of violation.	V_WString	99.91160517	71	*47 Other Violations	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 21	Description for the type of violation.	V_WString	99.95044532	52	*34 Pest Control	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 22	Description for the type of violation.	V_WString	99.97790129	26	*23 No Water	*09 Cooked ready to eat - Food protected from cross contamination by separating, storage, preparation
Violation Description - 23	Description for the type of violation.	V_WString	99.98794616	18	*47 OTHER VIOLATIONS	*09 Food protected from cross contamination separating types of raw animal food storage, preparation, holding
Violation Description - 24	Description for the type of violation.	V_WString	99.99196411	11	*45 Drying Mops-air dry	*21 A establishment shall have written procedures for employees to follow when responding to vomit in
Violation Description - 25	Description for the type of violation.	V_WString	99.99598205	7	*47 Other Violations	*46 Water, Plumbing, and Waste Plumbing Systems-good repair
Violation Description - 3	Description for the type of violation.	V_WString	62.55742316	615	*23 No Water	*29 Cold/hot hold unit thermometer easily viewable

Violation Description - 4	Description for the type of violation.	V_WString	68.50599344	588	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 5	Description for the type of violation.	V_WString	74.01392888	563	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 6	Description for the type of violation.	V_WString	78.97810219	554	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 7	Description for the type of violation.	V_WString	83.37306636	498	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 8	Description for the type of violation.	V_WString	87.16801714	482	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Description - 9	Description for the type of violation.	V_WString	90.39978571	422	*23 No Water	*05 Reheat from 41°F to 165°F within 2 hours and/or Commercially processed RTE foods reheated from 4
Violation Detail - 1	Detailed description of the violation.	V_WString	51.88642604	561	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	*21 Person in charge. The person in charge shall ensure that §228.34 (1-15) Relating to Duties of PI
Violation Detail - 10	Detailed description of the violation.	V_WString	93.23243822	371	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies.....
Violation Detail - 11	Detailed description of the violation.	V_WString	95.45637179	342	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies.....
Violation Detail - 12	Detailed description of the violation.	V_WString	97.19547311	295	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies that do .....

Violation Detail - 13	Detailed description of the violation.	V_WString	98.2193799	259	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.71 Food. Cooking. (a) Raw animal foods. (1) Except as specified in paragraphs (2) - (4) of this subsection, raw animal foods such as eggs, fish, meat, poultry, and foods containing these raw animal foods.....
Violation Detail - 14	Detailed description of the violation.	V_WString	98.7417129 8	253	228.120 Equipment, Utensils, and Linens. Laundering, Frequency, Specifications. (d) Wet wiping cloths shall be laundered daily.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 15	Detailed description of the violation.	V_WString	99.1160517	216	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 16	Detailed description of the violation.	V_WString	99.4120404 5	178	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original

						container is opened in a food establishment.....
Violation Detail - 17	Detailed description of the violation.	V_WString	99.6377151 3	167	228.75 Food. Time and temperature control. (a) Frozen food. Stored frozen foods shall be maintained frozen.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 18	Detailed description of the violation.	V_WString	99.7682984	130	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 19	Detailed description of the violation.	V_WString	99.8627201 5	90	228.75 Food. Time and temperature control. (a) Frozen food. Stored frozen foods shall be maintained frozen.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 2	Detailed description of the violation.	V_WString	56.9651108 3	550	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking.

					aid kit shall be provided.	(2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 20	Detailed description of the violation.	V_WString	99.91562312	68	228.61 Food. Condition safe, unadulterated, and honestly presented. Food shall be safe, unadulterated, and, as specified in §228.78(b) of this title, honestly presented.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies that do not meet the definition of a public water system as defined by 30 TAC §290.38(66) or that are not regulated by the Texas Commission on Environmental Quality (TCEQ) shall comply with the requirements of this subchapter ....
Violation Detail - 21	Detailed description of the violation.	V_WString	99.95379361	49	228.176 Physical Facilities. Toilets and urinals. (b) Toilet tissue, availability. A supply of toilet tissue shall be available at each toilet.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 22	Detailed description of the violation.	V_WString	99.97991027	25	228.224 REQUIREMENTS APPLICABLE TO CERTAIN ESTABLISHMENTS Outfitter Operations. (i) Equipment. All equipment and utensils intended for food contact shall be approved for food use.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall

						be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 23	Detailed description of the violation.	V_WString	99.98794616	18	228.173 Physical Facilities. Floors, walls, and ceilings. (e) Floor covering, mats and duckboards. Mats and duckboards shall be designed to be removable and easily cleanable.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 24	Detailed description of the violation.	V_WString	99.99196411	11	228.149 Water, Plumbing, and Waste. Plumbing, operation and maintenance. (e) System maintained in good repair. A plumbing system shall be: (2) maintained in good repair.	228.75 Food. Time and temperature control. (g) Ready-to-eat, TCS food, date marking. (2) Except as specified in paragraphs (5) - (7) of this subsection, refrigerated, ready-to-eat, time/temperature controlled for safety food prepared and packaged by a food processing plant shall be clearly marked, at the time the original container is opened in a food establishment.....
Violation Detail - 25	Detailed description of the violation.	V_WString	99.99665171	6	228.175 Physical Facilities. Handwashing Sinks. (c) Hand drying provision. Each handwashing lavatory or group of adjacent lavatories shall be provided with: (1) individual, disposable towels;	228.106 Equipment, Utensils, and Linens. Functionality of equipment. (a) Ventilation hood systems, drip prevention. Exhaust ventilation hood systems in food preparation and warewashing areas including components such as hoods, fans, guards, and ducting shall be designed to prevent grease or condensation from draining or dripping onto food, equipment, utensils, linens, and single-service and single-use articles.

Violation Detail - 3	Detailed description of the violation.	V_WString	62.91100248	550	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.106 Equipment, Utensils, and Linens. Functionality of equipment. (l) Temperature measuring devices. (2) Except as specified in paragraph (3) of this subsection, cold or hot holding equipment used for time/temperature control for safety (TCS) food shall be designed to include and shall be equipped with at least one integral or permanently affixed temperature measuring device that is located to allow easy viewing of the device's temperature display.
Violation Detail - 4	Detailed description of the violation.	V_WString	68.84082234	531	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.35 Management and Personnel. Responsibilities and Reporting Symptoms and Diagnosis. (a) Responsibility of Permit Holder, Person in Charge, and Conditional Employees. The permit holder shall require food employees and conditional employees to report to the person in charge, information about their health and activities.....
Violation Detail - 5	Detailed description of the violation.	V_WString	74.33469497	504	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.35 Management and Personnel. Responsibilities and Reporting Symptoms and Diagnosis. (a) Responsibility of Permit Holder, Person in Charge, and Conditional Employees. The permit holder shall require food employees and conditional employees to report to the person in charge, information about their health and activities.....
Violation Detail - 6	Detailed description of the violation.	V_WString	79.24194736	492	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies that do not meet the definition

						of a public water system as defined by 30 TAC §290.38(66) or that are not regulated by the Texas Commission on Environmental Quality (TCEQ) shall comply with the requirements of this subchapter (2) Water pressure....
Violation Detail - 7	Detailed description of the violation.	V_WString	83.5960624 1	450	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.35 Management and Personnel. Responsibilities and Reporting Symptoms and Diagnosis (a) Responsibility of Permit Holder, Person in Charge, and Conditional Employees. The permit holder shall require food employees and conditional employees to report to the person in charge, information about their health and activities as they relate to diseases that are transmissible.....
Violation Detail - 8	Detailed description of the violation.	V_WString	87.3441371 5	430	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.35 Management and Personnel. Responsibilities and Reporting Symptoms and Diagnosis (a) Responsibility of Permit Holder, Person in Charge, and Conditional Employees. The permit holder shall require food employees and conditional employees to report to the person in charge, information about their health and activities as they relate to diseases that are transmissible.....
Violation Detail - 9	Detailed description of the violation.	V_WString	90.5451014 5	385	228.210 Poisonous or Toxic Materials. First Aid Supplies, Availability. A first aid kit shall be provided.	228.35 Management and Personnel. Responsibilities and Reporting Symptoms and Diagnosis (a) Responsibility of Permit Holder, Person in Charge, and Conditional Employees. The permit holder shall require food employees and conditional employees to report to the person in charge, information about their health and

						activities as they relate to diseases that are transmissible.....
Violation Memo - 1	Additional notes for the violation.	V_WString	62.4047411 8	44343	.	228.271 PRIVATE WATER SYSTEMS Water Supply and Pressure Food Service Establishments having water supplies that do not meet the definition of a public water system as defined by 30 TAC §290.38(66) or that are not regulated by the Texas Commission on Environmental Quality (TCEQ) shall comply with the requirements of this subchapter.....
Violation Memo - 10	Additional notes for the violation.	V_WString	94.2402732 2	7072	PIC	
Violation Memo - 11	Additional notes for the violation.	V_WString	96.1313868 6	4836	BAR	WARNING! ALSO ALL COOLERS, DEEP FRYERS, GRILL, BAR , BAR SINKS ICE MACHINE ,
Violation Memo - 12	Additional notes for the violation.	V_WString	97.6441438 4	3047	l	
Violation Memo - 13	Additional notes for the violation.	V_WString	98.4959485 7	2042	WIC	DUE TO ITEM E ABOVE. Apply for new FOOD ESTABLISHMENT permit, due to change of ownership, from 320 E. Jefferson Blvd, Room 118 within 24 hours. Please bring sales tax permit, articles of INC., certificate of occupancy and application fees for FOOD ESTABLISHMENT permit. OBSERVED DIFFERENT OWNER'S NAME BETWEEN SALE TAX PERMIT AND FOOD PERMIT. **PLEASE FOLLOW INSTRUCTION FROM BUILDING INSPECTOR REGARDING ALL STRUCTURE C EQUIPMENT REQUIREMENT. PLANS C PERMITTING OFFICE NUMBER: 214/948-4480
Violation Memo - 14	Additional notes for the violation.	V_WString	98.9426103 3	1445	Raid	1) RESTOCK YOUR FIRST AID KIT 2) Heimlich maneuver sign SEC. 17-11.2. ADDITIONAL

						REQUIREMENTS. (a) General. All food establishments that provide dining areas shall post the Heimlich maneuver sign in a place conspicuous to employees and customers. (b) Specifications: The sign shall: (1) be no smaller than 11 inches wide by 17 inches long;...
Violation Memo - 15	Additional notes for the violation.	V_WString	99.2640460 7	1004	NONE	SEC. 17-11.2. ADDITIONAL REQUIREMENTS. (a) General. All food establishments that provide dining areas shall post the Heimlich maneuver sign in a place conspicuous to employees and customers. (b) Specifications: The sign shall: (1) be no smaller than 11 inches wide by 17 inches long; (2) be printed in English ...
Violation Memo - 16	Additional notes for the violation.	V_WString	99.5011049 4	690	TILE	SEC. 17-11.2. ADDITIONAL REQUIREMENTS. (a) General. All food establishments that provide dining areas shall post the Heimlich maneuver sign in a place conspicuous to employees and customers. (b) Specifications: The sign shall: (1) be no smaller than 11 inches wide by 17 inches long; (2) be printed in English and Spanish...
Violation Memo - 17	Additional notes for the violation.	V_WString	99.6986539 9	428	70	ceiling with grease and soils accumulations for grill outside vent hood must be clean ceiling and readjust equipment inside vent hood area to avoid cross contamination floors behind equipment and all floors with blood, food debris and litter must be clean to avoid cross contamination walls with spills, grease accumulations and soils deposits must be

						clean to avoid cross contamination
Violation Memo - 18	Additional notes for the violation.	V_WString	99.8044599 2	284	90 F	CLEAN INTERIOR OF COMPARTMENT SINK EXTERIOR OF REACH IN COOLERS MUST BE CLEANED C SANITIZED (INCLUDING EXTERIOR OF BLUE CONTAINERS) EXTERIOR OF COOKING EQUIPMENT MUST BE CLEANED CSANITIZED CLEAN MICROWAVE OVEN/ SMALL OVEN
Violation Memo - 19	Additional notes for the violation.	V_WString	99.8868278 3	165	CLEAN ALL	obtain r fsm certificate from City of Dallas within 45 days Bring learn 2 serv certificate, DL or ID and money (cash, check, money order no credit or debit cards), \$30.00 per year. 7901 Goforth RD Dallas TX 75238 214 670 8083
Violation Memo - 2	Additional notes for the violation.	V_WString	66.6764883 1	39400	d	cooler units in all areas including frame and gaskets doors dirty with mold , soils and grease deposits must be clean to avoid cross contamination WI cooler dirty with soils and grease accumulations must be clean to avoid cross contamination
Violation Memo - 20	Additional notes for the violation.	V_WString	99.9249983 3	112	EMPLOYEES	3 compartment sink faucet leaking.....
Violation Memo - 21	Additional notes for the violation.	V_WString	99.9578115 6	63	salsa	using recycle non approved containers must be use approved commercial containers to avoid chemical contamination use of thank you bags as food containers must be have use approved commercial food containers to avoid chemical contamination
Violation Memo - 22	Additional notes for the violation.	V_WString	99.9792406 1	32	label 3 comp sink	equipment and single use utensils contaminated by trash container located inside shelving area must be remove equipment and discard single use utensils contaminated by the

						trash receptacle inside shelving area
Violation Memo - 23	Additional notes for the violation.	V_WString	99.98861582	18	PROVIDE	light shields removed in back storage area replace missing vent a hood filters food employees must focus on one task at a time food staff recommend have cleaning bucket label mop sink recommend a small cart to store dirty dishes provide lids for bulk ingredients bins or cover them recommend using food temp logs strong sewer smell observed inthe back storage area
Violation Memo - 24	Additional notes for the violation.	V_WString	99.99196411	13	Bulk items	cooking equipment dirty with soils and grease accumulations must be clean to avoid cross contamination
Violation Memo - 25	Additional notes for the violation.	V_WString	99.99598205	7	CLEAN ALL	reseal hand sink to wall near 3 com sink clean the floor throughout the kitchen area cover floor drain in grill area seal opening above the bag-n box check interior of walk in cooler
Violation Memo - 3	Additional notes for the violation.	V_WString	71.21141097	34368	p	Dirty dust on fans and vents above food prep areas and counters. dirty food containers, outside RIC doors, stoves, ovens, walls, vents, vent filters. Ice on freezer floor.
Violation Memo - 4	Additional notes for the violation.	V_WString	75.62110761	29198	a	SINK AT 0 PPM. CONTAINER IS EMPTIED . CONTAINER IS OUT.
Violation Memo - 5	Additional notes for the violation.	V_WString	79.72677962	24516	33	ICE SCOOP STORAGE BIN
Violation Memo - 6	Additional notes for the violation.	V_WString	83.46413982	20032	T	
Violation Memo - 7	Additional notes for the violation.	V_WString	86.73206991	16147	LA	
Violation Memo - 8	Additional notes for the violation.	V_WString	89.6571352	12549	CO	
Violation Memo - 9	Additional notes for the violation.	V_WString	92.140896	9605	ALL	
Violation Points - 1	The amount of points assigned to this violation.	V_WString	51.51543561	5	2	2
Violation Points - 10	The amount of points assigned to this violation.	V_WString	93.12462332	4	3	3
Violation Points - 11	The amount of points assigned to this violation.	V_WString	95.38873636	4	1	1
Violation Points - 12	The amount of points assigned to this violation.	V_WString	97.1385522	4	1	1

Violation Points - 13	The amount of points assigned to this violation.	V_WString	98.18254872	5	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 14	The amount of points assigned to this violation.	V_WString	98.72162325	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 15	The amount of points assigned to this violation.	V_WString	99.09931025	4	<a href="#">2</a>	<a href="#">2</a>
Violation Points - 16	The amount of points assigned to this violation.	V_WString	99.40065626	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 17	The amount of points assigned to this violation.	V_WString	99.62967923	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 18	The amount of points assigned to this violation.	V_WString	99.76294114	4	<a href="#">2</a>	<a href="#">2</a>
Violation Points - 19	The amount of points assigned to this violation.	V_WString	99.86071118	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 2	The amount of points assigned to this violation.	V_WString	56.57001272	5	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 20	The amount of points assigned to this violation.	V_WString	99.91160517	4	<a href="#">2</a>	<a href="#">2</a>
Violation Points - 21	The amount of points assigned to this violation.	V_WString	99.95044532	4	<a href="#">3</a>	<a href="#">3</a>
Violation Points - 22	The amount of points assigned to this violation.	V_WString	99.97790129	4	<a href="#">2</a>	<a href="#">2</a>
Violation Points - 23	The amount of points assigned to this violation.	V_WString	99.98794616	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 24	The amount of points assigned to this violation.	V_WString	99.99196411	3	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 25	The amount of points assigned to this violation.	V_WString	99.99598205	4	<a href="#">3</a>	<a href="#">3</a>
Violation Points - 3	The amount of points assigned to this violation.	V_WString	62.55742316	5	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 4	The amount of points assigned to this violation.	V_WString	68.50599344	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 5	The amount of points assigned to this violation.	V_WString	74.01392888	5	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 6	The amount of points assigned to this violation.	V_WString	78.97810219	4	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 7	The amount of points assigned to this violation.	V_WString	83.37306636	5	<a href="#">1</a>	<a href="#">1</a>
Violation Points - 8	The amount of points assigned to this violation.	V_WString	87.16801714	4	<a href="#">2</a>	<a href="#">2</a>
Violation Points - 9	The amount of points assigned to this violation.	V_WString	90.39978571	5	<a href="#">3</a>	<a href="#">3</a>
Zip Code	Zip Code of the establishment	V_WString	47.11444452	157	<a href="#">75237</a>	75238-2632

## Staging table Plan:

- Initially, the data in each of the datasets will be cleaned and transformed into a preferred format
- They will be stored in separate stage tables (with no data being dropped)
- In a separate pipeline, both the staging tables will be combined and put to a final staging table.
- This final staging table will be used to load data into the facts and dimensions.

## Items followed in all stage tables:

- The column names are converted to not having spaces. Eg. Risk Category is renamed as Risk\_Category.
- The datatypes of the columns are set to a preferred one. Eg. Inspection date is converted to date datatype from string
- All the numeric columns with nulls are kept being -9999 (considering unavailable), after checking if they can be calculated from any field.
- All the text columns with nulls are kept being NA, after checking if they can be calculated from any other field.
- All leading, trailing white spaces, tabs have been removed
- All the text columns have been made UPPERCASE for uniformity
- Audit columns have been added
- No data has been dropped when loading the data to initial staging table. However, for the final staging table, only the required data has been chosen.

## Column wise key transformation decisions

### 1. Dallas:

Name of the Column	Transformation done
Inspection Date	Changed String to Date format
Inspection Month	Kept as is
Inspection Score	Calculated based on total violation score
Inspection Type	Kept as is
Inspection Year	Kept as is
Lat Long Location	Split to Latitude, longitude
Restaurant Name	Kept as is
Street Address	Kept as is
Street Direction	Kept as is
Street Name	Kept as is
Street Number	Kept as is
Street Type	Kept as is
Street Unit	Kept as is
All Violation Description Columns	The starting number is considered as Violation category ID and the trailing text

	as violation category. They are put into two separate columns
All Violation Detail columns, Violation memo columns	These columns are combined and put into a column named as Violation others
All Violation points columns	These points are considered to be the points for the violation category ID
Zip Code	The first 5 digits are considered
InspectionID	Kept as is
InspectionID_Created	This is a column that is created based on auto increment for all records in the initial phase that starts with "DA"
City	Dallas is assigned as default
State	TX is assigned as default
Facility Type	Eatery DA is assigned as default
Sum_Violation_Point	This is a sum of all violation points, unique to a inspection ID
Risk Category	This is a calculated column based on Inspection score
Inspection_Result	This is a calculated column based on Inspection score
FileName	FileName
User_Name	User_Name
Load_Date	Load_date

Important Notables - The nature of inspection\_score column is different in dallas than that of Chicago. In dallas, as per metadata, the inspection score calculated by deducting the sum of all the violation points from 100. Therefore, a facility having more violations would lead to a very less inspection score and the facility with very less violations would have a high violation score.

Logic for the calculation of Inspection\_Result, Risk\_Category:

Risk\_Category:

```

IF [Inspection_Score] >= 85 THEN "Low"
ELSEIF [Inspection_Score] >= 70 AND [Inspection_Score] <= 84 THEN "Medium"
ELSE "High"
ENDIF

```

This is different from what was provided in assignment, since the inspection score scale for both Dallas and Chicago dataset are not the same.

Inspection\_Result:

```

IF [Inspection_Score] >= 85 THEN "Pass"
ELSEIF [Inspection_Score] >= 70 AND [Inspection_Score] <= 84 THEN "Pass with Warning"
ELSE "Fail"
ENDIF

```

This is different from what was provided in assignment, since the inspection score scale for both Dallas and Chicago dataset are not the same.

## 2. Chicago

Name of the column	Transformation done
AKA Name	Kept as is
Longitude	Kept as is
Address	Kept as is
State	Kept as is
DBA Name	Kept as is
Inspection ID	Kept as is
License #	0 padding was done in the start if the length of license is <7
Latitude	Kept as is
City	Kept as is
Violations	Based on the separator " ", the violations were separated. In each violation, the starting number was considered as the violation category ID, the following text as violation category and the trailing text as violation comments
Risk	This column was split into Risk Level and Risk Category
Location	Kept as is
Inspection Type	Kept as is
Inspection Date	Converted string to date
Facility Type	Kept as is
Zip	Considered the first 5 digits
Results	Kept as is
Violation Score	Calculated based on provided logic
Result_Score	Calculated based on provided logic
Risk_Score	Calculated based on provided logic
Total_Violation_Score	Calculated based on provided logic
Inspection_Result	Calculated based on provided logic
FileName	FileName
User_Name	User_Name
Load_Date	Load_Date

### Logic for Calculated columns:

if Violations is null, score always 0, if any violation then  $2 * \text{count of violations}$

if Result is Failed it should be 70

else if Result is Pass w/conditions it should be 40

else if Result is Pass it should be 0

else if Result is any Other it should be 0

if Risk is High: 30,

Low: 10,

Medium : 20

Any Other: 0

Total Violation Score = Summation of all mentioned above. If Total Violation Score > 100 then  
Total Violation Score=100

### Calculating Inspection-Result

Total Violation Score Range

60 above Fail

below 60 and above 30: Pass with Warning,

30 and below: Pass

any other value: Other

### 3. Combining Chicago, Dallas - Final Stage table:

Only the required columns from each of the staging tables of datasets have been chosen. Some columns have been renamed to perform a union operation between both tables.

From Chicago staging table:

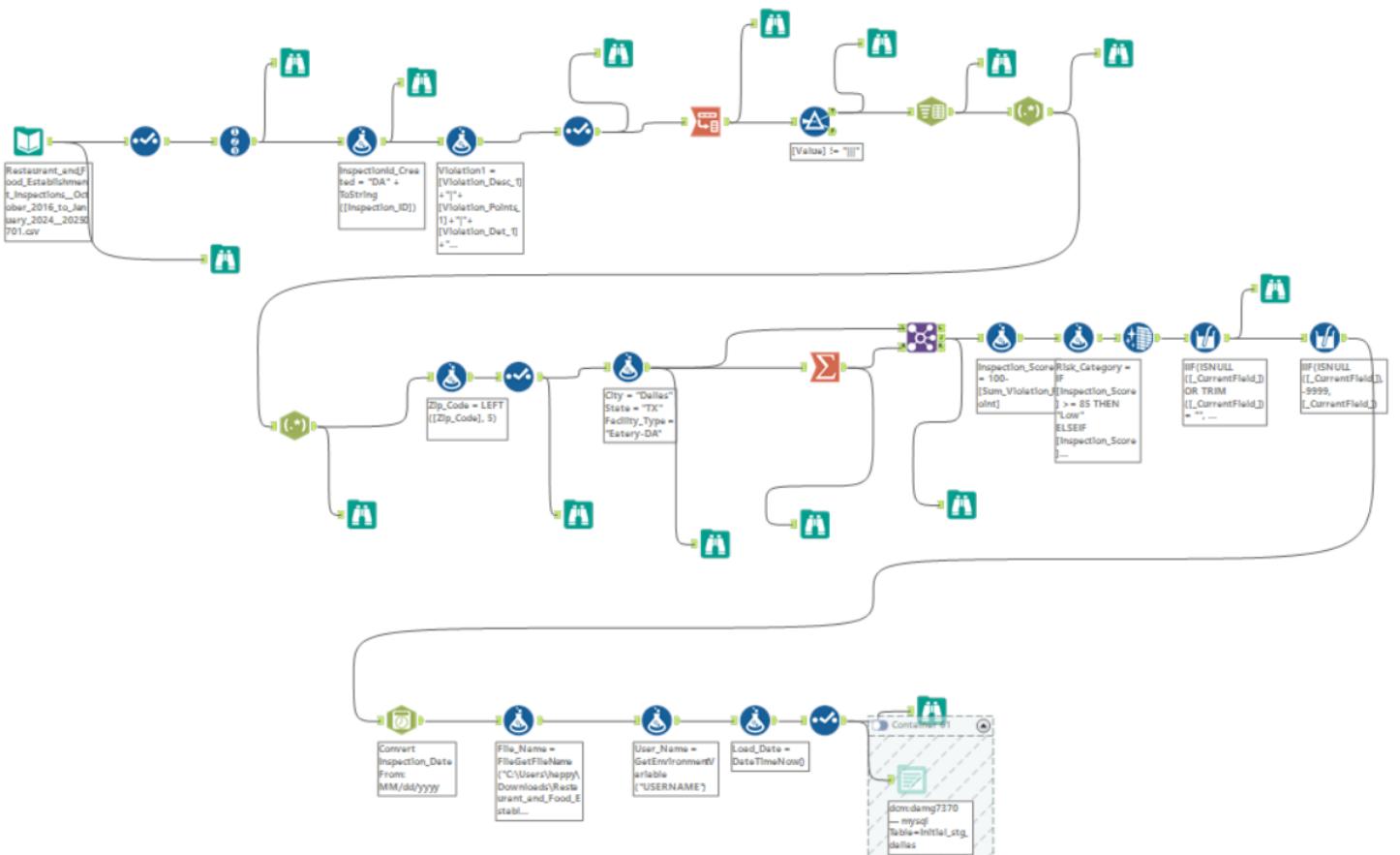
	<input type="checkbox"/> Column	Type	Size	Rename	I
>	<input checked="" type="checkbox"/> Inspection_ID	V_WString	▼ 167...		
	<input checked="" type="checkbox"/> DBA_Name	V_String	▼ 167...	Name	
	<input type="checkbox"/> AKA_Name	V_String	▼ 79		
	<input type="checkbox"/> License_no	String	▼ 7		
	<input checked="" type="checkbox"/> Facility_Type	V_String	▼ 47		
	<input type="checkbox"/> Risk_Level	FixedDecimal	▼ 21.0		
	<input checked="" type="checkbox"/> Risk_Category	String	▼ 6		
	<input checked="" type="checkbox"/> Address	V_String	▼ 51		
	<input checked="" type="checkbox"/> City	V_String	▼ 20		
	<input checked="" type="checkbox"/> State	String	▼ 2		
	<input checked="" type="checkbox"/> Zip	Int64	▼ 8	Zip_Code	
	<input checked="" type="checkbox"/> Inspection_Date	String	▼ 10		
	<input checked="" type="checkbox"/> Inspection_Type	V_String	▼ 41		
	<input type="checkbox"/> Results	V_String	▼ 20		
	<input checked="" type="checkbox"/> Latitude	Double	▼ 8		
	<input checked="" type="checkbox"/> Longitude	Double	▼ 8		
	<input type="checkbox"/> Location	String	▼ 40		
	<input checked="" type="checkbox"/> Violation_Category_ID	Int64	▼ 8	Violation_Category_Id	
	<input checked="" type="checkbox"/> Violation_Category	V_String	▼ 134		
	<input type="checkbox"/> Violation_Comments	V_String	▼ 5422		
	<input checked="" type="checkbox"/> Violation_Score	Int64	▼ 8		
	<input type="checkbox"/> Result_Score	FixedDecimal	▼ 21.0		
	<input type="checkbox"/> Risk_Score	FixedDecimal	▼ 21.0		
	<input checked="" type="checkbox"/> Total_Violation_Score	Int64	▼ 8		
	<input checked="" type="checkbox"/> Inspection_Result	V_String	▼ 17		
	<input type="checkbox"/> DateTime_Out	Date	▼ 10		
	<input type="checkbox"/> File_Name	String	▼ 25		
	<input type="checkbox"/> User_Name	String	▼ 5		
	<input type="checkbox"/> Load_Date	DateTime	▼ 19		

From Dallas Staging Table:

	Column	Type	Size	Rename
>	Inspection_ID	FixedDecimal	21.0	
<input checked="" type="checkbox"/>	InspectionId_Created	String	9	Inspection_ID
<input checked="" type="checkbox"/>	Restaurant_Name	V_String	65	Name
<input checked="" type="checkbox"/>	Inspection_Type	String	9	
<input checked="" type="checkbox"/>	Inspection_Date	V_WString	20	
<input checked="" type="checkbox"/>	Inspection_Score	Int64	8	Total_Violation_Score
	Street_Number	Int32	4	
	Street_Name	V_String	25	
	Street_Direction	String	2	
	Street_Type	String	4	
	Street_Unit	String	5	
<input checked="" type="checkbox"/>	Street_Address	V_String	37	Address
<input checked="" type="checkbox"/>	City	String	6	
<input checked="" type="checkbox"/>	State	String	2	
<input checked="" type="checkbox"/>	Zip_Code	Int64	8	
	Inspection_Month	String	8	
	Inspection_Year	String	6	
<input checked="" type="checkbox"/>	Latitude	Double	8	
<input checked="" type="checkbox"/>	Longitude	Double	8	
<input checked="" type="checkbox"/>	Violation_Category_Id	Int64	8	
<input checked="" type="checkbox"/>	Violation_Category	V_WString	98	
	Violation_Others	V_WString	2757	
	Violation_Point	FixedDecimal	21.0	
<input checked="" type="checkbox"/>	Facility_Type	String	9	
<input checked="" type="checkbox"/>	Sum_Violation_Point	Int64	8	Violation_Score
<input checked="" type="checkbox"/>	Risk_Category	String	6	
<input checked="" type="checkbox"/>	Inspection_Result	V_String	17	

## Alteryx Profiling C Transformation workflow

Stg\_Dallas Table loading



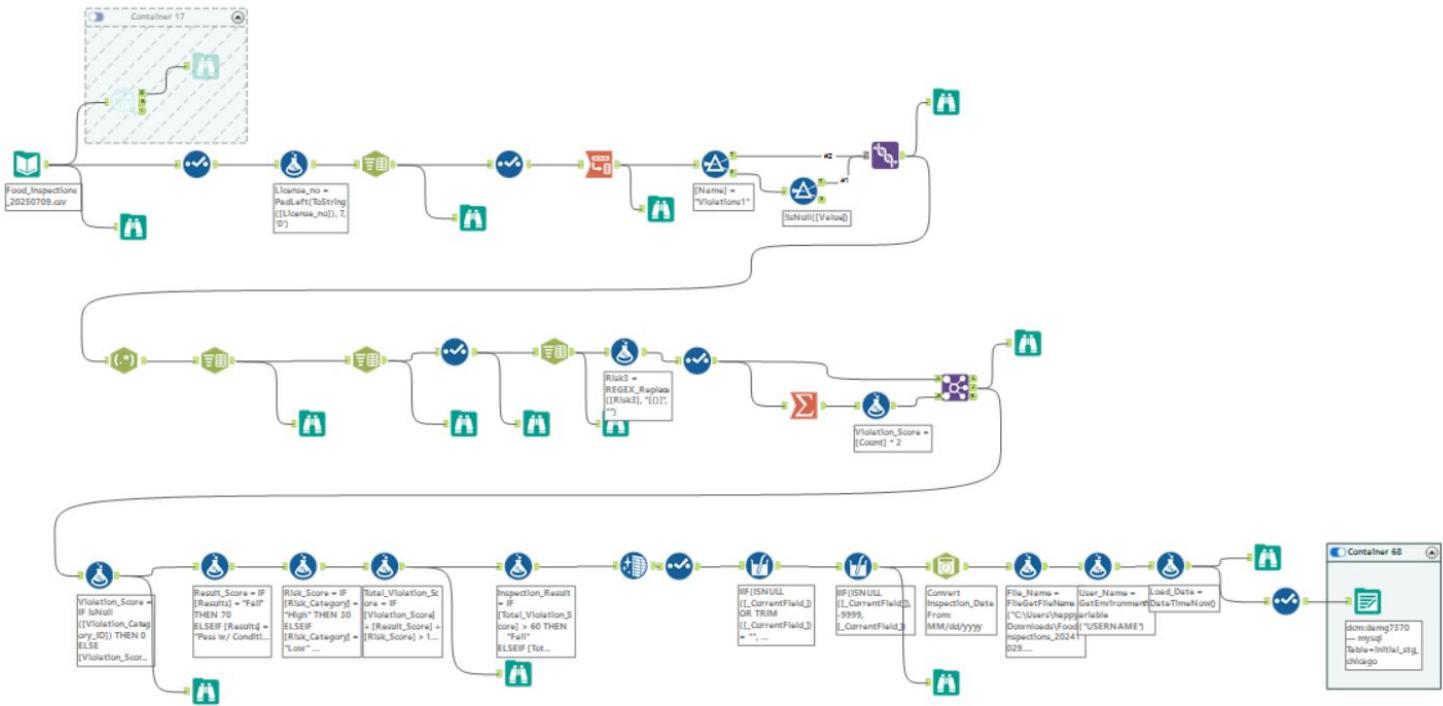
Records Loaded: 398408

DDL Script:

```
CREATE TABLE `initial_stg_dallas` (
  `Inspection_ID` bigint DEFAULT NULL,
  `InspectionId_Created` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Restaurant_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Inspection_Type` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Inspection_Date` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Inspection_Score` bigint DEFAULT NULL,
  `Street_Number` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Street_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Street_Direction` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Street_Type` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Street_Unit` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Street_Address` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `City` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `State` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Zip_Code` bigint DEFAULT NULL,
```

```
`Inspection_Month` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Inspection_Year` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Latitude` double DEFAULT NULL,  
 `Longitude` double DEFAULT NULL,  
 `Violation_Category_Id` bigint DEFAULT NULL,  
 `Violation_Category` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Violation_Others` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Violation_Point` bigint DEFAULT NULL,  
 `Facility_Type` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Sum_Violation_Point` bigint DEFAULT NULL,  
 `Risk_Category` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Inspection_Result` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `DateTime_Out` datetime DEFAULT NULL,  
 `File_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `User_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,  
 `Load_Date` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci;
```

Stg\_Chicago Loading



Records Loaded: 998578

DDL Script:

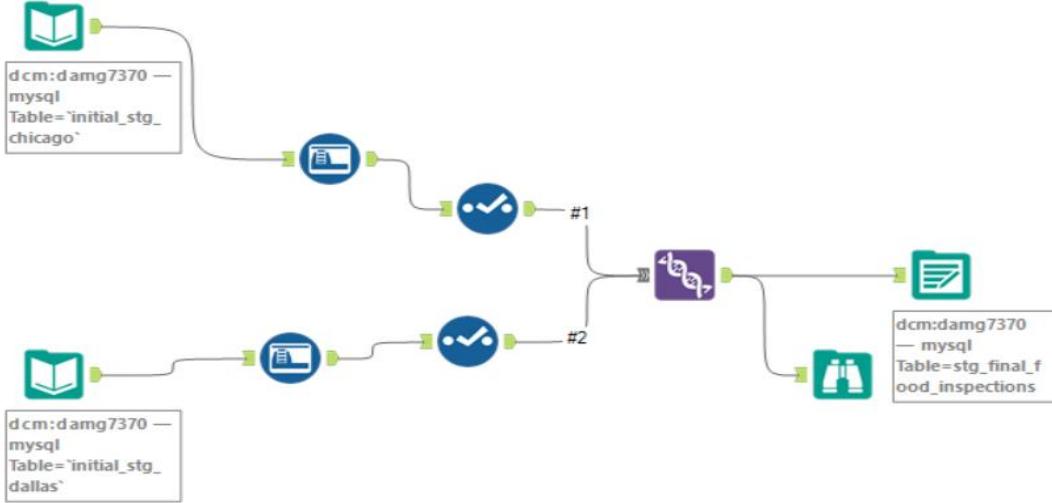
```

CREATE TABLE `initial_stg_chicago` (
  `Inspection_ID` bigint DEFAULT NULL,
  `DBA_Name` varchar(100) DEFAULT NULL,
  `AKA_Name` text,
  `License_no` text,
  `Facility_Type` text,
  `Risk_Level` bigint DEFAULT NULL,
  `Risk_Category` text,
  `Address` text,
  `City` text,
  `State` text,
  `Zip` bigint DEFAULT NULL,
  `Inspection_Date` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `Inspection_Type` text,
  `Results` text,
  `Latitude` double DEFAULT NULL,
  `Longitude` double DEFAULT NULL,
  `Location` text,
  `Violation_Category_ID` bigint DEFAULT NULL,
  `Violation_Category` text,
  `Violation_Comments` text,
  `Violation_Score` bigint DEFAULT NULL,
  `Result_Score` bigint DEFAULT NULL,
  `Risk_Score` bigint DEFAULT NULL,
  `Total_Violation_Score` bigint DEFAULT NULL,
  `Inspection_Result` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `DateTime_Out` datetime DEFAULT NULL,
  `File_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  `User_Name` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci,
  ...
)

```

```
`Load_Date` text CHARACTER SET utf8mb3 COLLATE utf8mb3_general_ci  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci;
```

## Final\_staging\_Table loading



Records Loaded: 1396986

DDL Script:

```
create or replace TABLE FOOD_INSPECTIONS.DADABI_MIDTERM.STG_FINAL_FOOD_INSPECTIONS (
    "Inspection_ID" VARCHAR(16777216),
    "Name" VARCHAR(16777216),
    "Facility_Type" VARCHAR(47),
    "Risk_Category" VARCHAR(6),
    "Address" VARCHAR(51),
    "City" VARCHAR(20),
    "State" VARCHAR(2),
    "Zip_Code" NUMBER(20,0),
    "Inspection_Date" VARCHAR(20),
    "Inspection_Type" VARCHAR(41),
    "Latitude" FLOAT,
    "Longitude" FLOAT,
    "Violation_Category_Id" NUMBER(20,0),
```

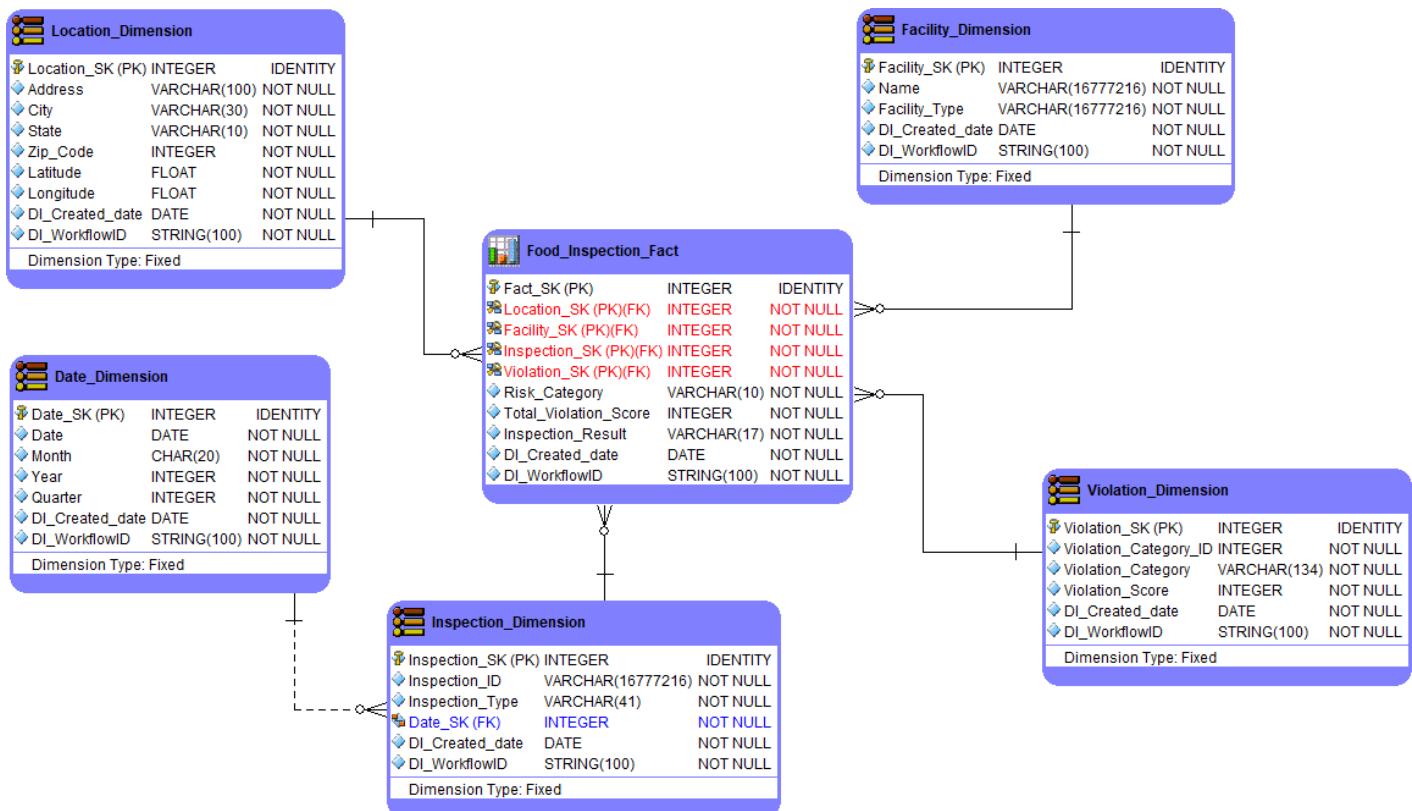
```

    "Violation_Category" VARCHAR(134),
    "Violation_Score" NUMBER(20,0),
    "Total_Violation_Score" NUMBER(20,0),
    "Inspection_Result" VARCHAR(17)

);

```

### Dimensional modeling:



- All dimensions and the fact contain audit columns - DI Created date and DI Workflow ID and also Surrogate Keys
- The fact table contains numbers, and Risk\_Category, Inspection\_Result, and all SKs
- All the datatypes, null characteristics, keys have been displayed in the image.

### DDL Script generated:

--

-- ER/Studio Data Architect SQL Code Generation

-- Project : Food\_Inspection.DM1

--  
-- Date Created : Sunday, November 03, 2024 00:16:24

-- Target DBMS : MySQL DB

--  
--

-- TABLE: Date\_Dimension

--

-- TABLE: Date\_Dimension

CREATE TABLE Date\_Dimension (

Date\_SK INT AUTO\_INCREMENT PRIMARY KEY,

Date DATE NOT NULL,

Month VARCHAR(20) NOT NULL,

Year INT NOT NULL,

Quarter INT NOT NULL,

DI\_Created\_date DATE NOT NULL,

DI\_WorkflowID VARCHAR(100) NOT NULL

) ENGINE=InnoDB;

-- TABLE: Facility\_Dimension

CREATE TABLE Facility\_Dimension (

Facility\_SK INT AUTO\_INCREMENT PRIMARY KEY,

Name VARCHAR(255) NOT NULL,

Facility\_Type VARCHAR(255) NOT NULL,

DI\_Created\_date DATE NOT NULL,

DI\_WorkflowID VARCHAR(100) NOT NULL

) ENGINE=InnoDB;

-- TABLE: Location\_Dimension

```
CREATE TABLE Location_Dimension (
    Location_SK INT AUTO_INCREMENT PRIMARY KEY,
    Address VARCHAR(100) NOT NULL,
    City VARCHAR(30) NOT NULL,
    State VARCHAR(10) NOT NULL,
    Zip_Code INT NOT NULL,
    Latitude DOUBLE NOT NULL,
    Longitude DOUBLE NOT NULL,
    DI_Created_date DATE NOT NULL,
    DI_WorkflowID VARCHAR(100) NOT NULL
) ENGINE=InnoDB;
```

-- TABLE: Violation\_Dimension

```
CREATE TABLE Violation_Dimension (
    Violation_SK INT AUTO_INCREMENT PRIMARY KEY,
    Violation_Category_ID INT NOT NULL,
    Violation_Category VARCHAR(134) NOT NULL,
    Violation_Score INT NOT NULL,
    DI_Created_date DATE NOT NULL,
    DI_WorkflowID VARCHAR(100) NOT NULL
) ENGINE=InnoDB;
```

-- TABLE: Inspection\_Dimension

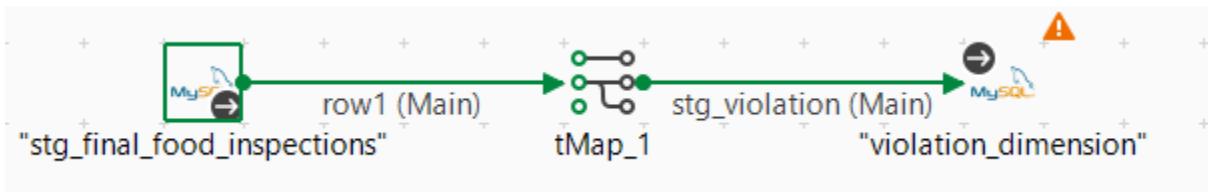
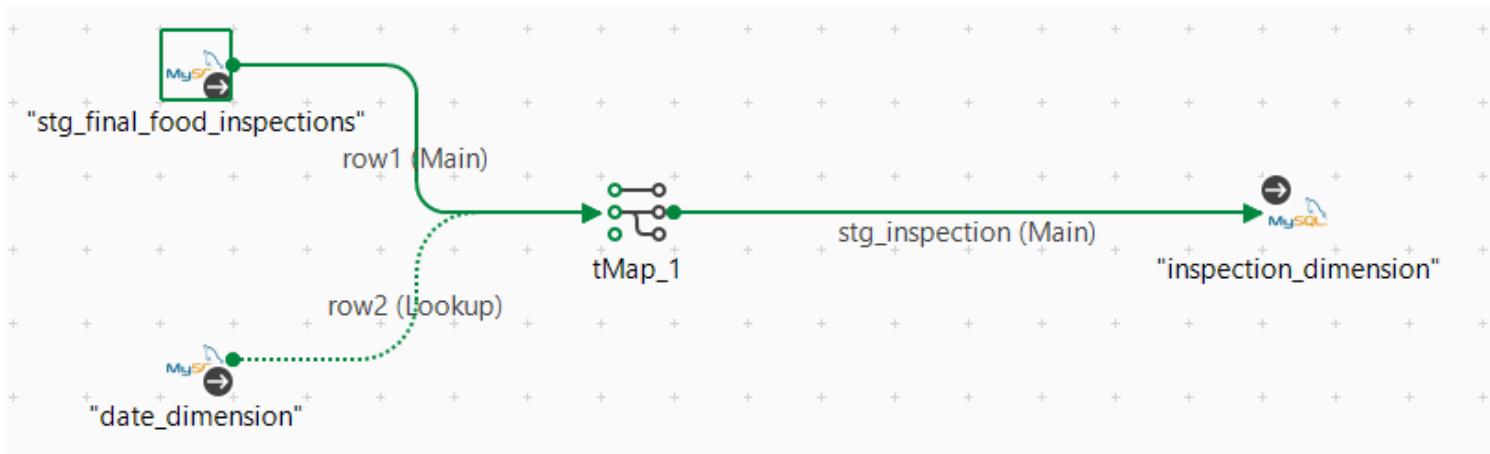
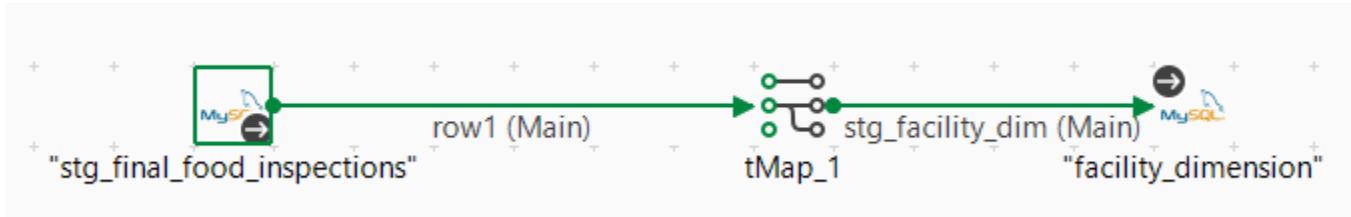
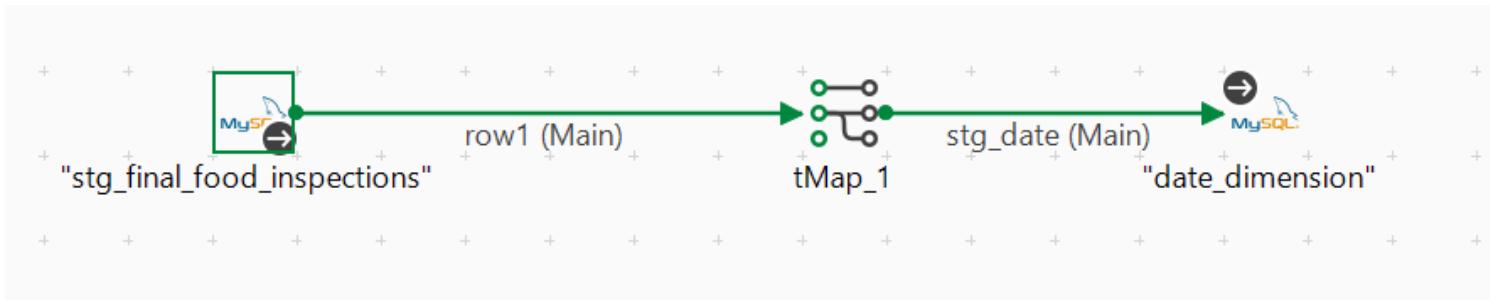
```
CREATE TABLE Inspection_Dimension (
    Inspection_SK INT AUTO_INCREMENT PRIMARY KEY,
    Inspection_ID VARCHAR(255) NOT NULL,
```

```
Inspection_Type VARCHAR(41) NOT NULL,  
Date_SK INT NOT NULL,  
DI_Created_date DATE NOT NULL,  
DI_WorkflowID VARCHAR(100) NOT NULL,  
FOREIGN KEY (Date_SK) REFERENCES Date_Dimension(Date_SK)  
) ENGINE=InnoDB;
```

-- TABLE: Food\_Inspection\_Fact

```
CREATE TABLE Food_Inspection_Fact (  
Fact_SK INT AUTO_INCREMENT,  
Location_SK INT NOT NULL,  
Facility_SK INT NOT NULL,  
Inspection_SK INT NOT NULL,  
Violation_SK INT NOT NULL,  
Risk_Category VARCHAR(10) NOT NULL,  
Total_Violation_Score INT NOT NULL,  
Inspection_Result VARCHAR(17) NOT NULL,  
DI_Created_date DATE NOT NULL,  
DI_WorkflowID VARCHAR(100) NOT NULL,  
PRIMARY KEY (Fact_SK, Location_SK, Facility_SK, Inspection_SK, Violation_SK),  
FOREIGN KEY (Location_SK) REFERENCES Location_Dimension(Location_SK),  
FOREIGN KEY (Facility_SK) REFERENCES Facility_Dimension(Facility_SK),  
FOREIGN KEY (Inspection_SK) REFERENCES Inspection_Dimension(Inspection_SK),  
FOREIGN KEY (Violation_SK) REFERENCES Violation_Dimension(Violation_SK)  
) ENGINE=InnoDB;
```

Loading of data into dimensions and fact table from the final staging table using talend:



Records count in final staging table and the fact table:

```
1 •   select count(*) from stg_final_food_inspections;
```

Result Grid		Filter Rows:	Export:	Wrap Cell C
count(*)				
▶ 1438943				

```
1 •   select count(*) from food_inspection_fact;
```

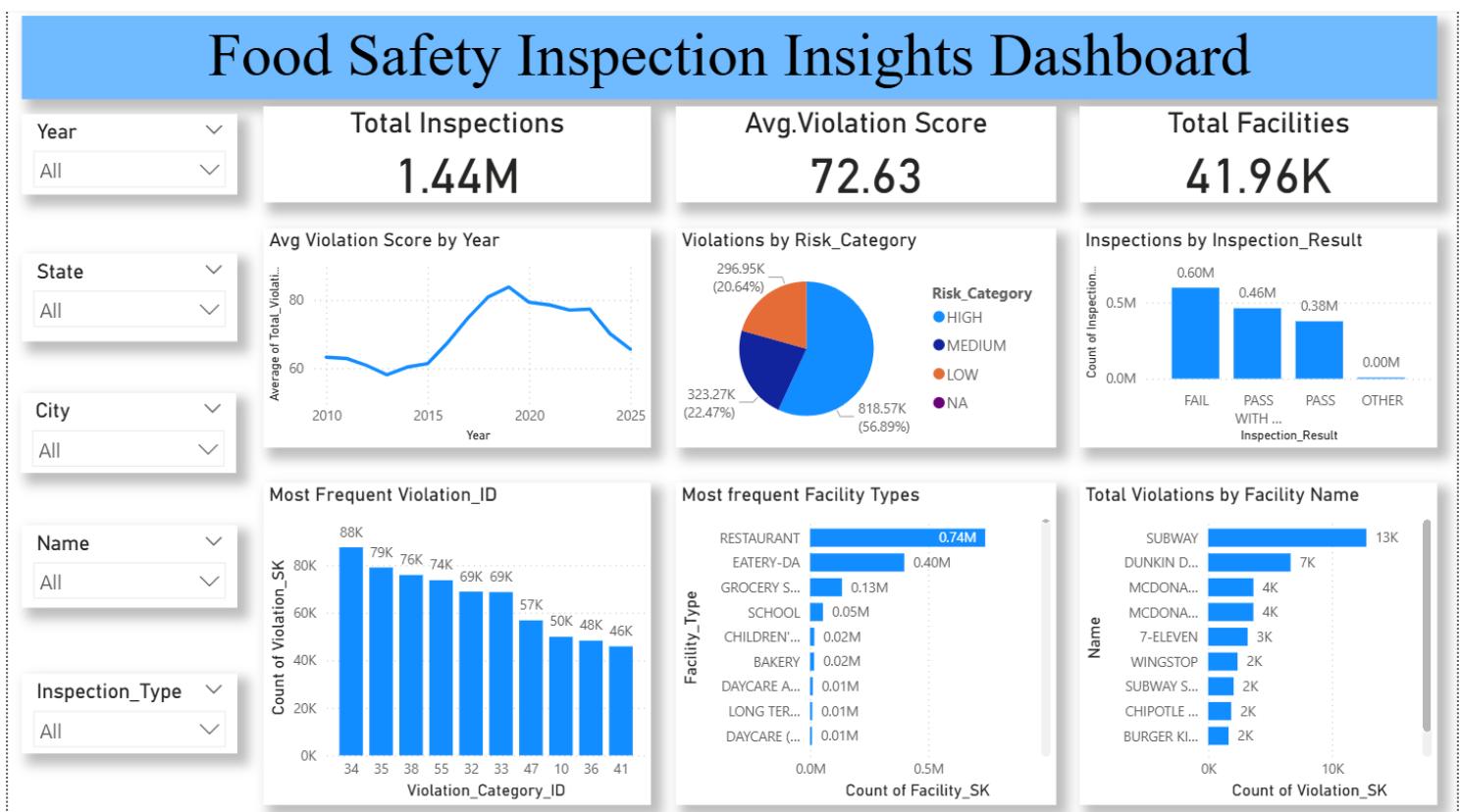
Result Grid		Filter Rows:	Export:	Wrap Cell C
count(*)				
▶ 1438943				

Both have the same number of records - since dimensional modeling is designed in that way

## Dashboards:

- Both Tableau and Power BI have been connected to the MySQL DB
- They have been designed to look similar
- The design idea was to keep the filters on the top, numbers on the right and the visualizations in the center
- Overall theme was to keep it simple and easy to use
- Some visualizations have extra detail in the tool tip
- The visualizations meet all the business requirements provided
- Dashboards in tableau and PowerBI have been uploaded to cloud

## Power BI



# Food Safety Inspection Insights Dashboard

Year

All

Total Inspections

1.44M

Avg.Violation Score

72.63

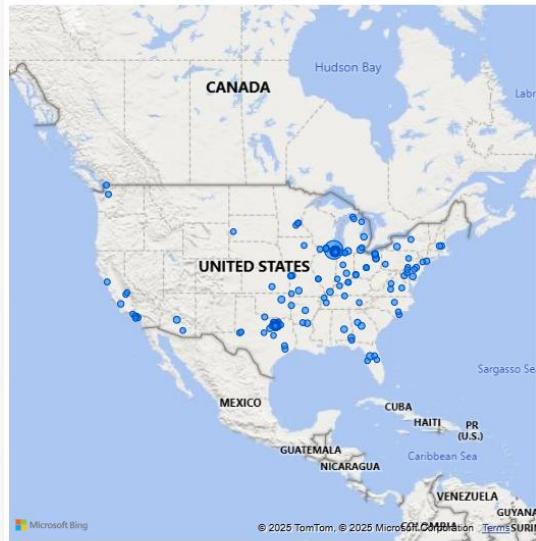
Total Facilities

41.96K

State

All

Total Violation Score by Latitude and Longitude



City

All

Name

All

Inspection\_Type

All

Violations by Violation\_Category









