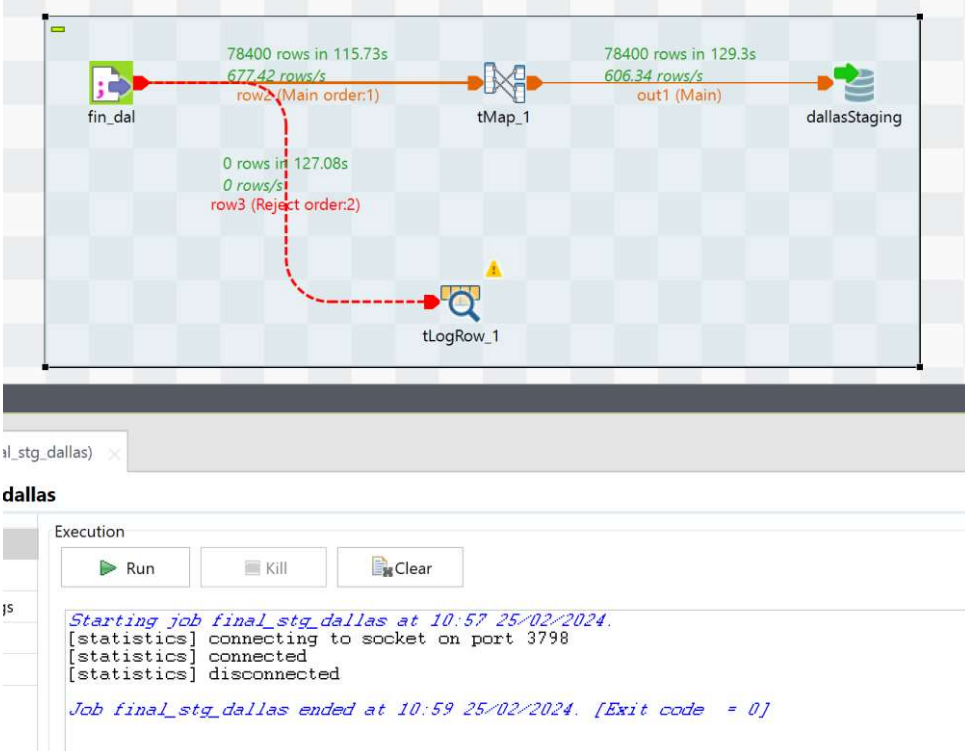
Midterm Talend Execution

Stagging Loads

Job Name – final_stg_chicago

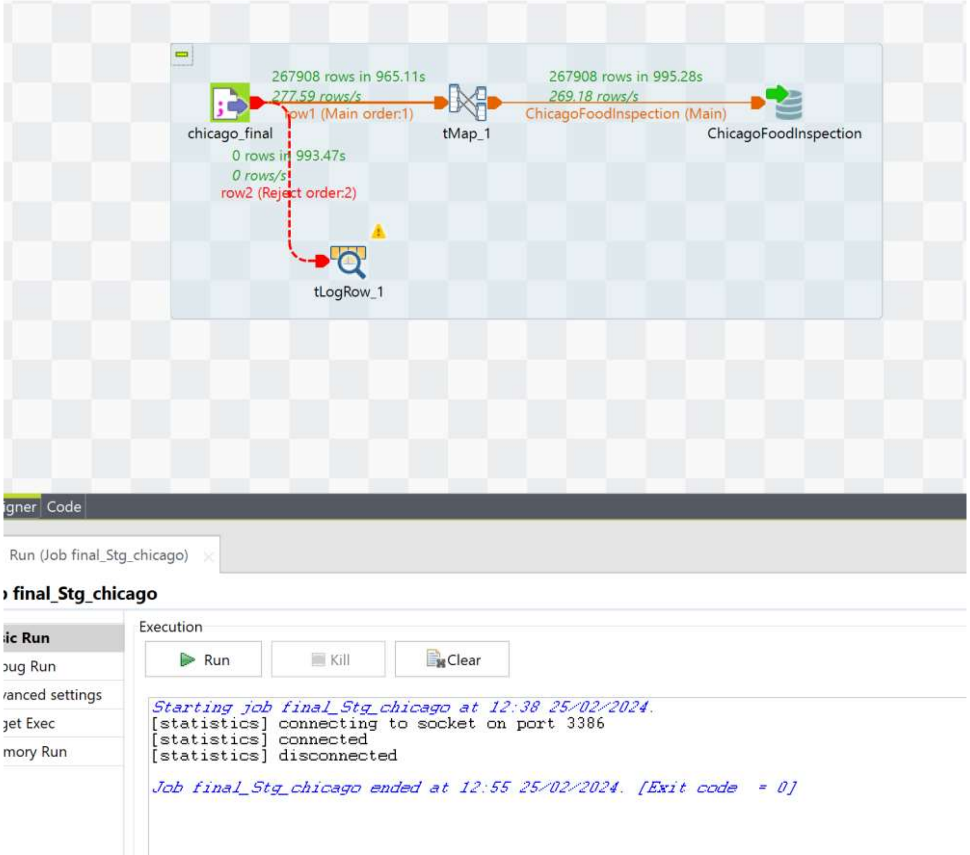
1. stg_dallas_inspection



```
Starting job final_stg_dallas at 10:57 25/02/2024.
[statistics] connecting to socket on port 3798
[statistics] connected
[statistics] disconnected

Job final_stg_dallas ended at 10:59 25/02/2024. [Exit code  = 0]
```

Job Name – final_stg_chicago

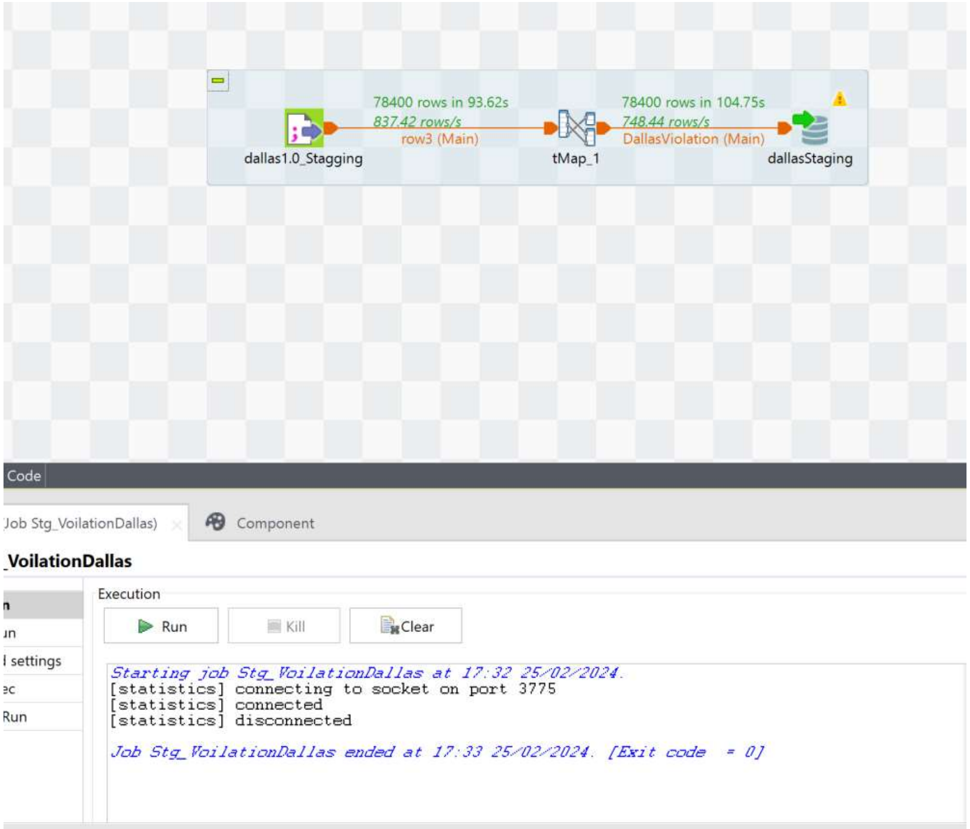
2. Stg_chicagofoodinspection



```
Starting job final_Stg_chicago at 12:38 25/02/2024.
[statistics] connecting to socket on port 3386
[statistics] connected
[statistics] disconnected

Job final_Stg_chicago ended at 12:55 25/02/2024. [Exit code  = 0]
```

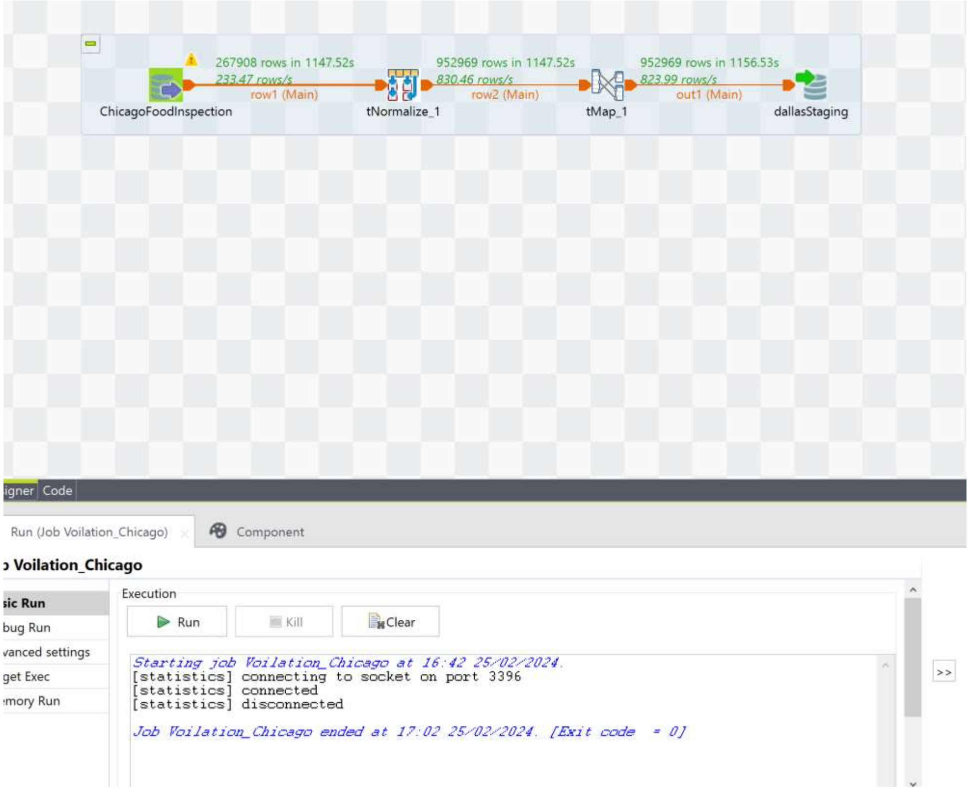
Job Name – stg_violation_Dallas

3. Stg_VoilationDallas



```
Starting job Stg_VoilationDallas at 17:32 25/02/2024.
[statistics] connecting to socket on port 3775
[statistics] connected
[statistics] disconnected

Job Stg_VoilationDallas ended at 17:33 25/02/2024. [Exit code  = 0]
```

4. Stg_VoilationChicago



› Voilation_Chicago

```
Starting job Voilation_Chicago at 16:42 25/02/2024.
[statistics] connecting to socket on port 3396
[statistics] connected
[statistics] disconnected

Job Voilation_Chicago ended at 17:02 25/02/2024. [Exit code  = 0]
```

Integration layer

Dimension and Fact Tables

"DimAddress"

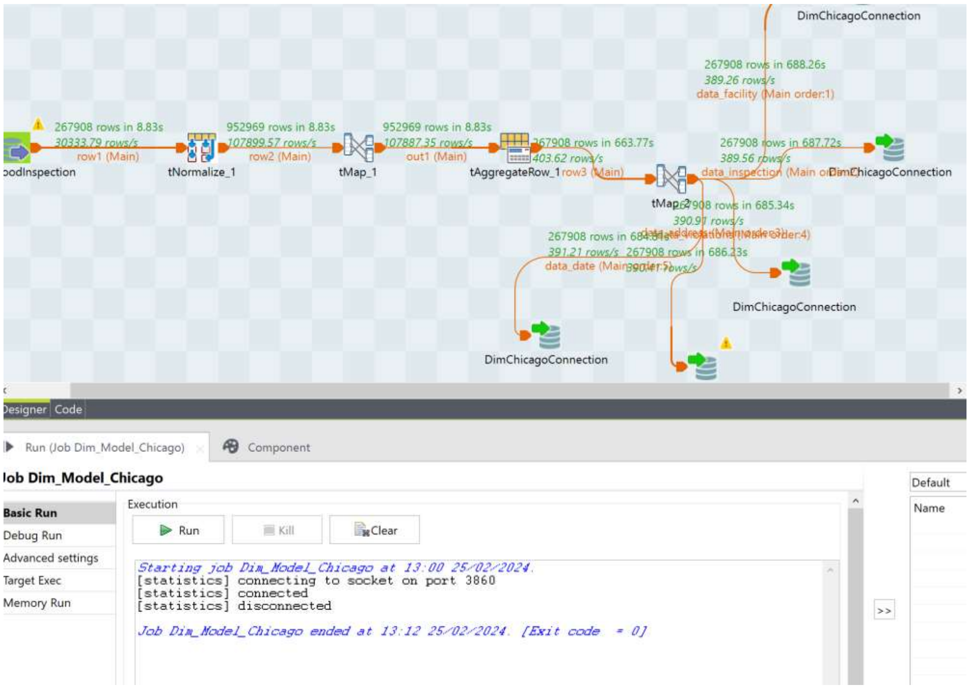
"FactInspection"

"DimFacility"

"DimDate"

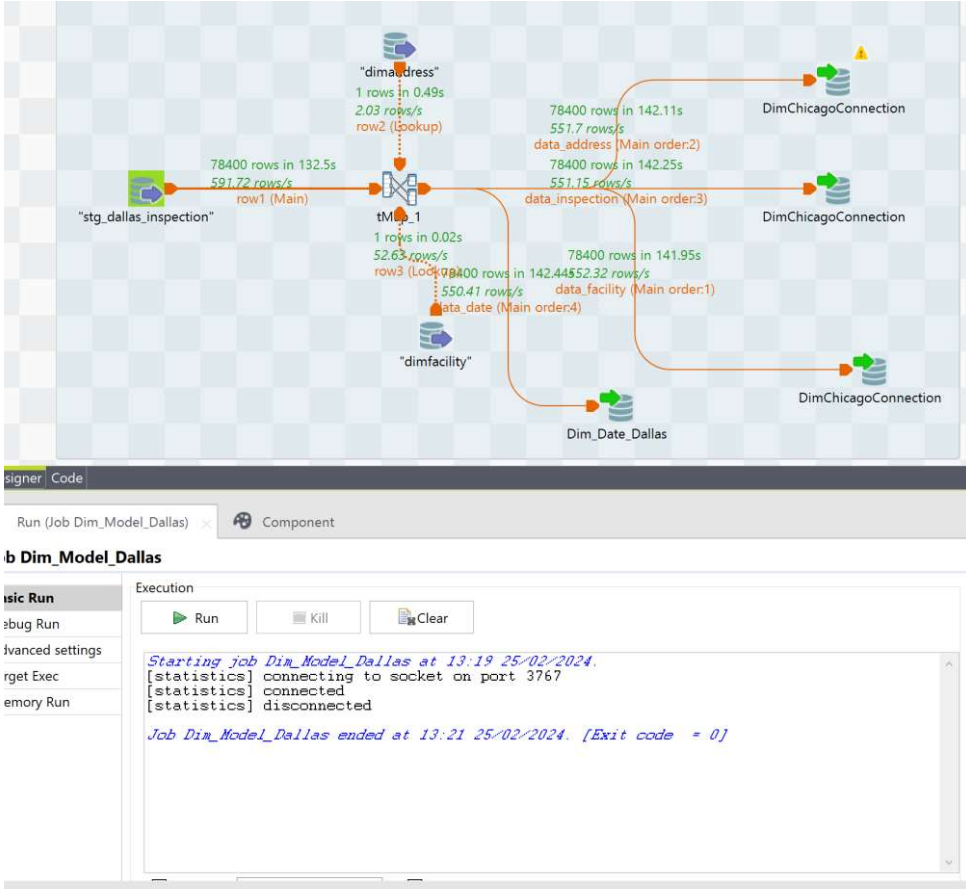
"stg_VoilationDallas"

"stg_VoilationChicago"

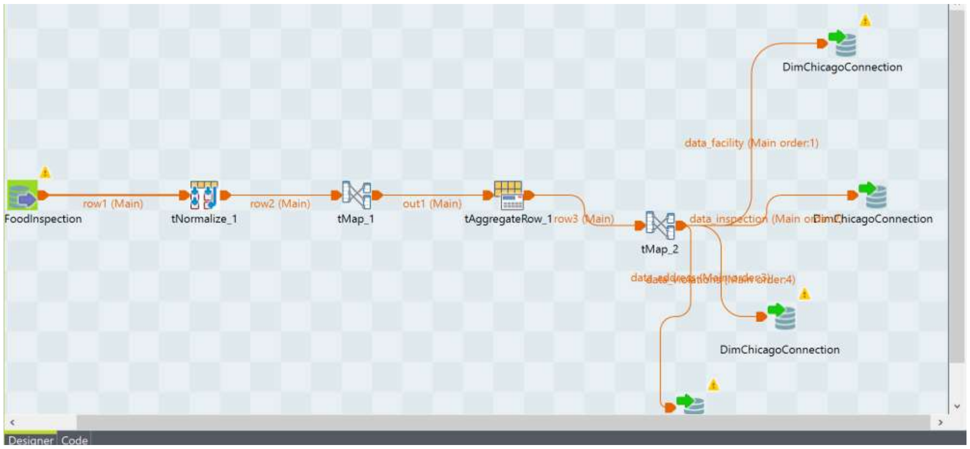


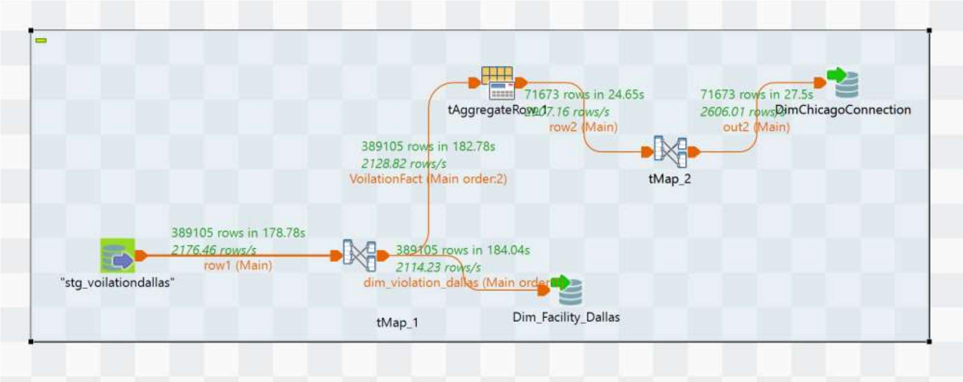
Job Dim_Model_Chicago

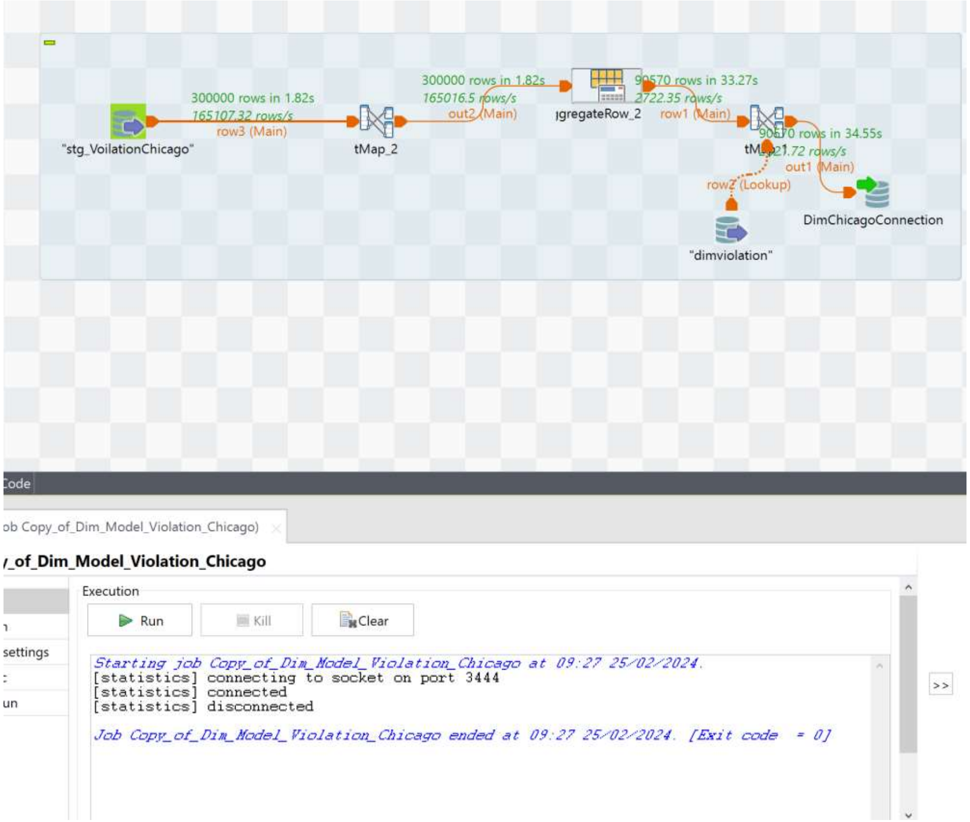
```
Starting job Dim_Model_Chicago at 13:00 25/02/2024.
[statistics] connecting to socket on port 3860
[statistics] connected
[statistics] disconnected

Job Dim_Model_Chicago ended at 13:12 25/02/2024. [Exit code  = 0]
```



```
Starting job Dim_Model_Dallas at 13:19 25/02/2024.
[statistics] connecting to socket on port 3767
[statistics] connected
[statistics] disconnected

Job Dim_Model_Dallas ended at 13:21 25/02/2024. [Exit code  = 0]
```





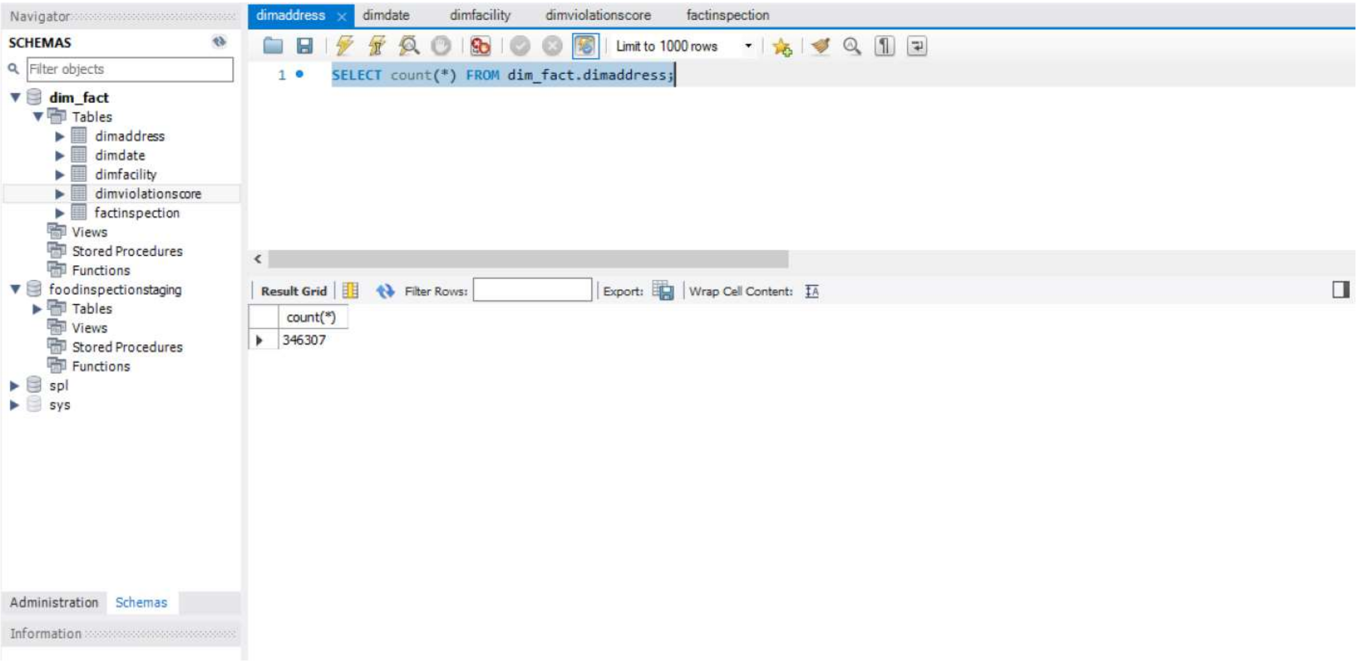


```
Starting job Copy_of_Dim_Model_Violation_Chicago at 09:27 25/02/2024.
[statistics] connecting to socket on port 3444
[statistics] connected
[statistics] disconnected

Job Copy_of_Dim_Model_Violation_Chicago ended at 09:27 25/02/2024. [Exit code  = 0]
```

COUNTS OF EACH TABLE:

(I) DIMENSIONS

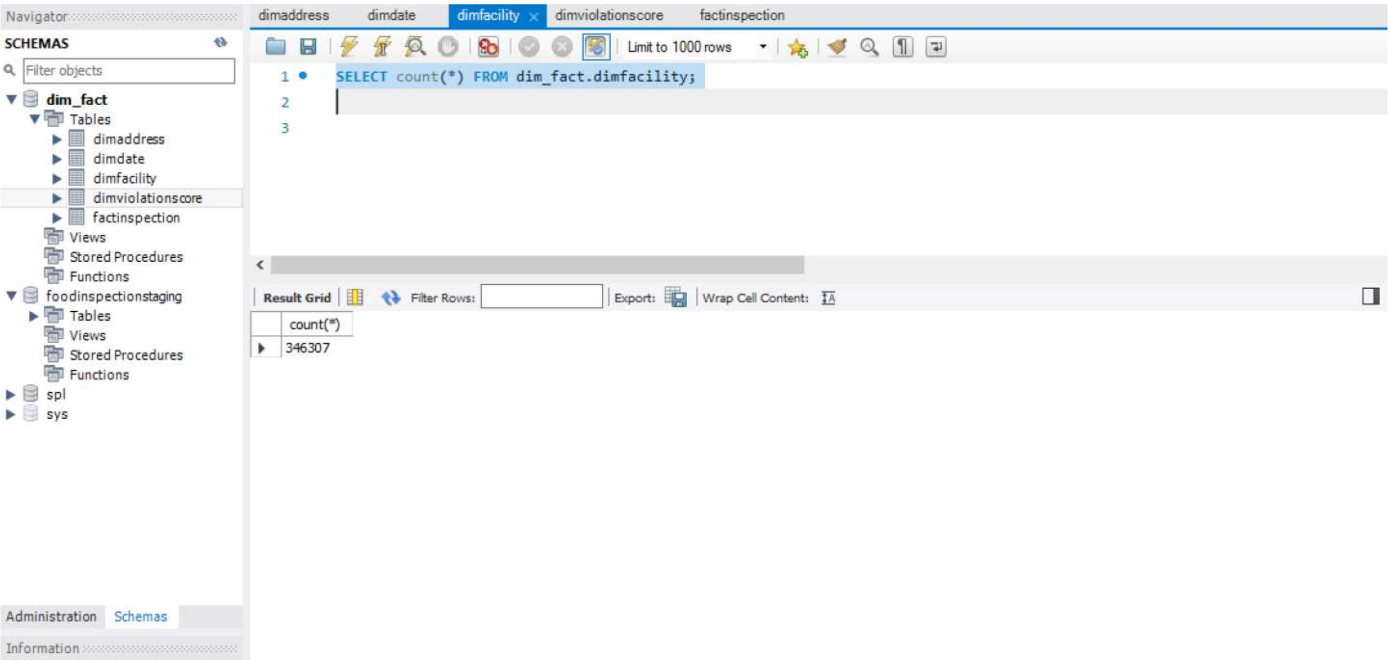
1. DIMADDRESS



```
SELECT count(*) FROM dim_fact.dimaddress;
```

count(*)
346307

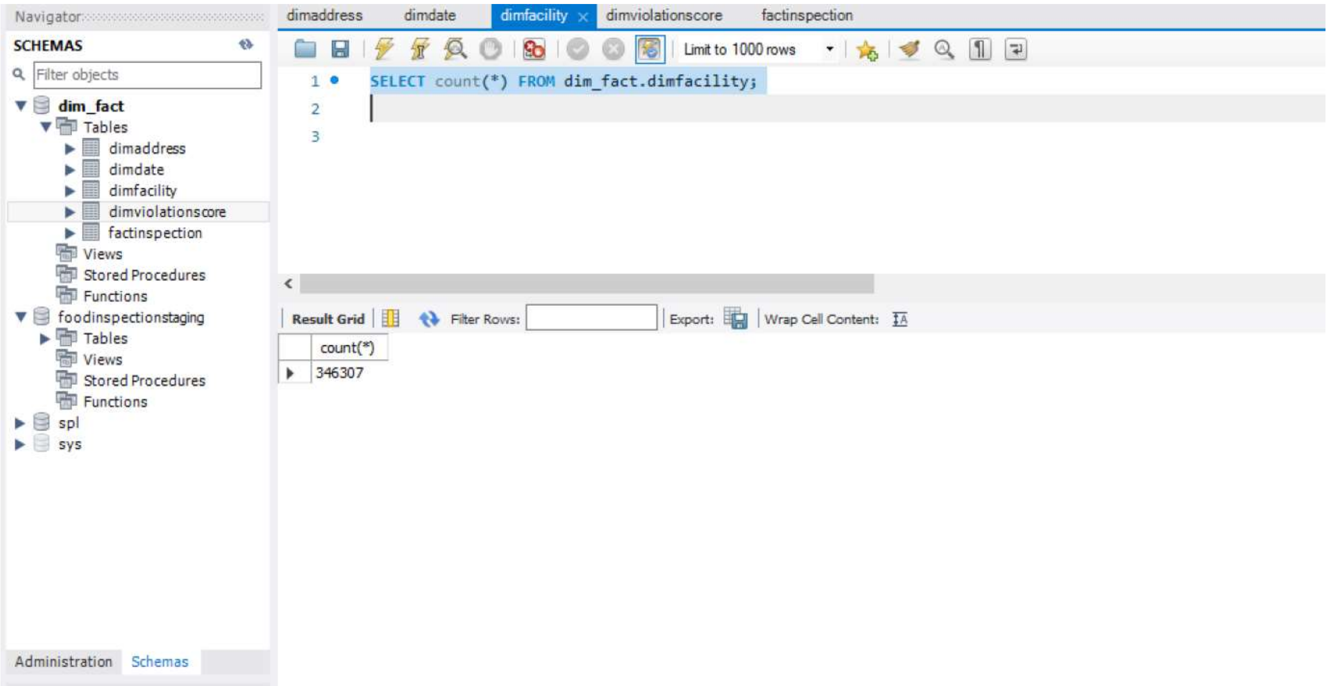
2. DIMDATE



```
SELECT count(*) FROM dim_fact.dimfacility;
```

count(*)
346307

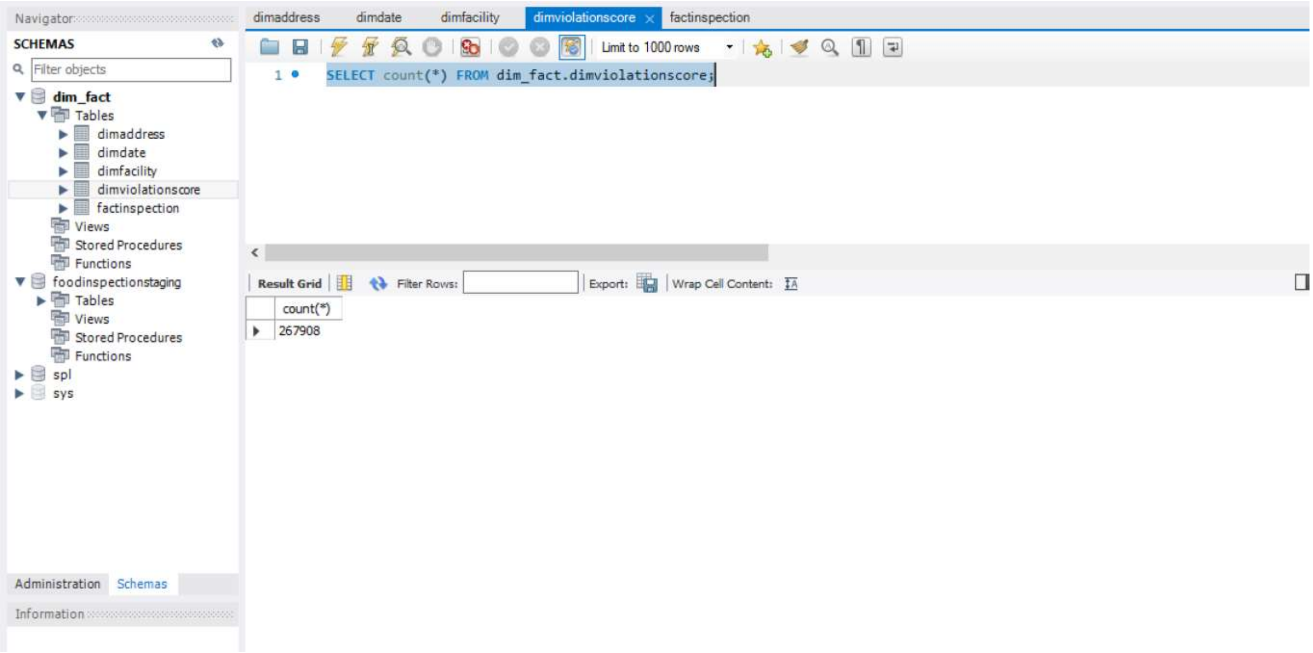
3. DIMFACILITY



```
SELECT count(*) FROM dim_fact.dimfacility;
```

```
count(*)
346307
```

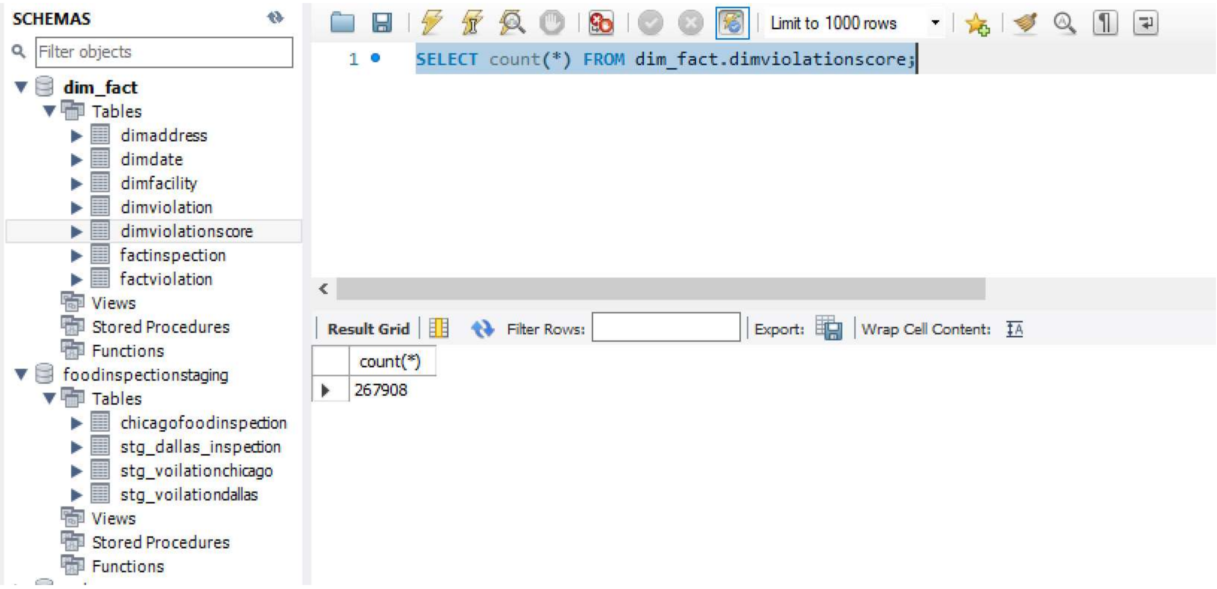
4. DIMVIOLATIONSCORE



```
SELECT count(*) FROM dim_fact.dimviolationscore;
```

```
count(*)
267908
```

5. DIMVIOLATION

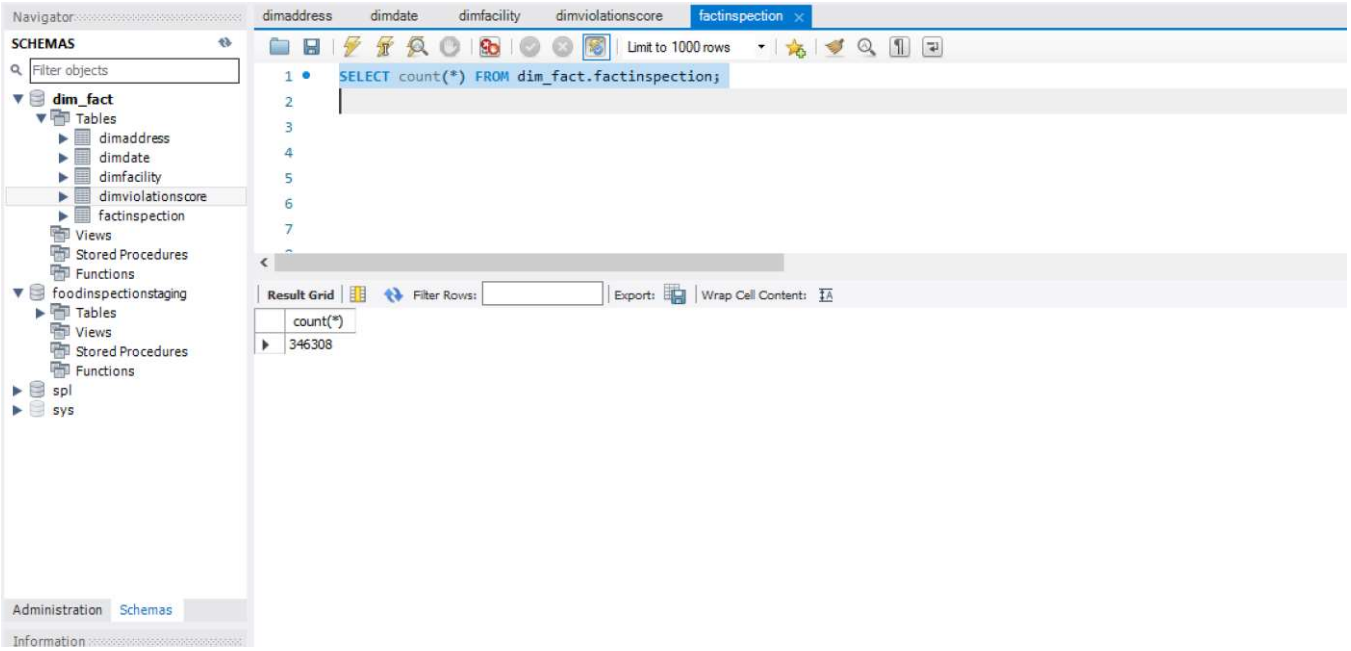


```
SELECT count(*) FROM dim_fact.dimviolationscore;
```

count(*)
267908

(II) FACTS

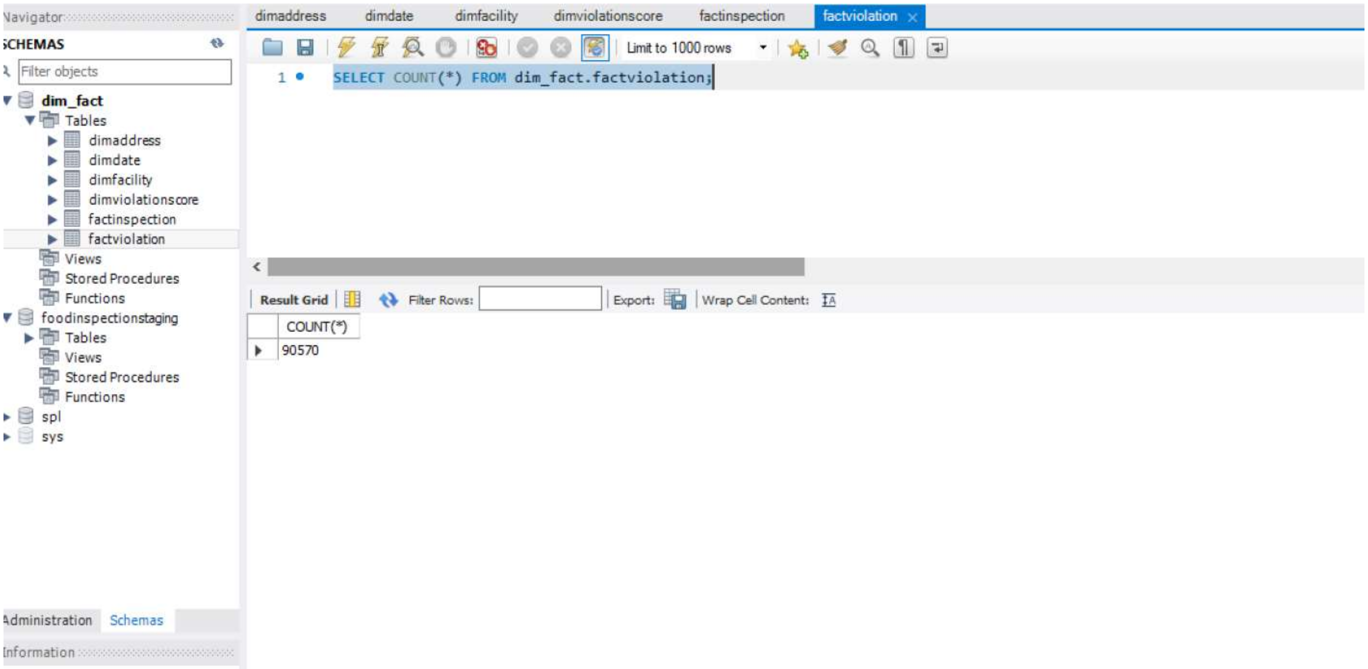
1. FACTINSPECTION



```
SELECT count(*) FROM dim_fact.factinspection;
```

```
count(*)
346308
```

2. FACTVIOLATION



```
SELECT COUNT(*) FROM dim_fact.factviolation;
```

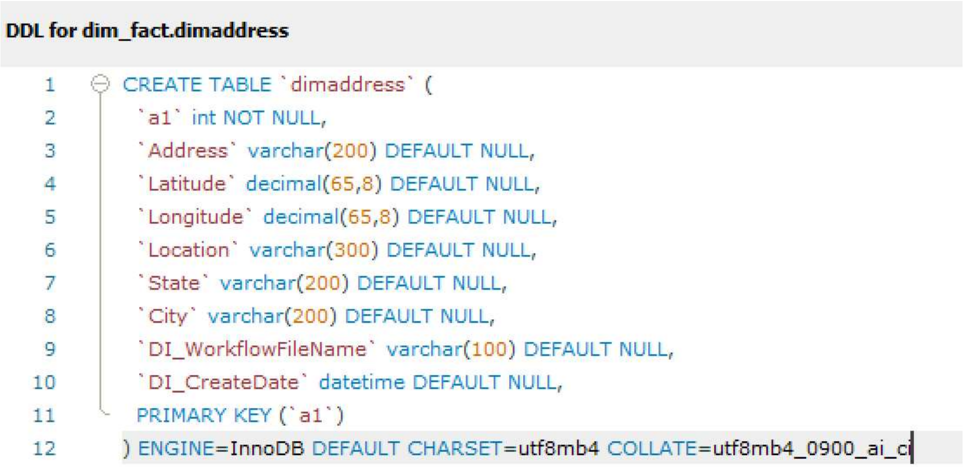
```
COUNT(*)
90570
```

DDL SCRIPTS:

(I) DIMENSIONS

1. DIMENSION ADDRESS

```
CREATE TABLE `dimaddress` (
 `a1` int NOT NULL,
 `Address` varchar(200) DEFAULT NULL,
 `Latitude` decimal(65,8) DEFAULT NULL,
 `Longitude` decimal(65,8) DEFAULT NULL,
 `Location` varchar(300) DEFAULT NULL,
 `State` varchar(200) DEFAULT NULL,
 `City` varchar(200) DEFAULT NULL,
 `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
 `DI_CreateDate` datetime DEFAULT NULL,
 PRIMARY KEY (`a1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

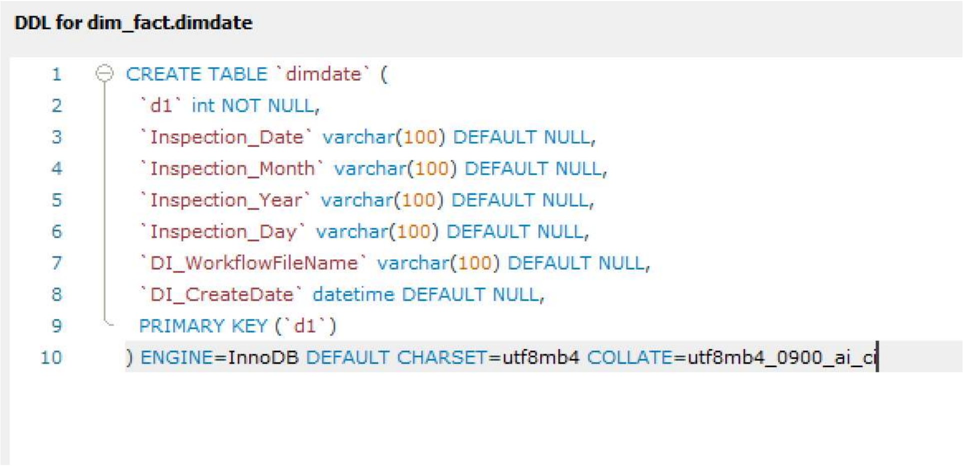


DDL for dim_fact.dimaddress

```
CREATE TABLE `dimaddress` (
  `a1` int NOT NULL,
  `Address` varchar(200) DEFAULT NULL,
  `Latitude` decimal(65,8) DEFAULT NULL,
  `Longitude` decimal(65,8) DEFAULT NULL,
  `Location` varchar(300) DEFAULT NULL,
  `State` varchar(200) DEFAULT NULL,
  `City` varchar(200) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`a1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

2. DIMENSION DATE

```
CREATE TABLE `dimdate` (
  `d1` int NOT NULL,
  `Inspection_Date` varchar(100) DEFAULT NULL,
  `Inspection_Month` varchar(100) DEFAULT NULL,
  `Inspection_Year` varchar(100) DEFAULT NULL,
  `Inspection_Day` varchar(100) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`d1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

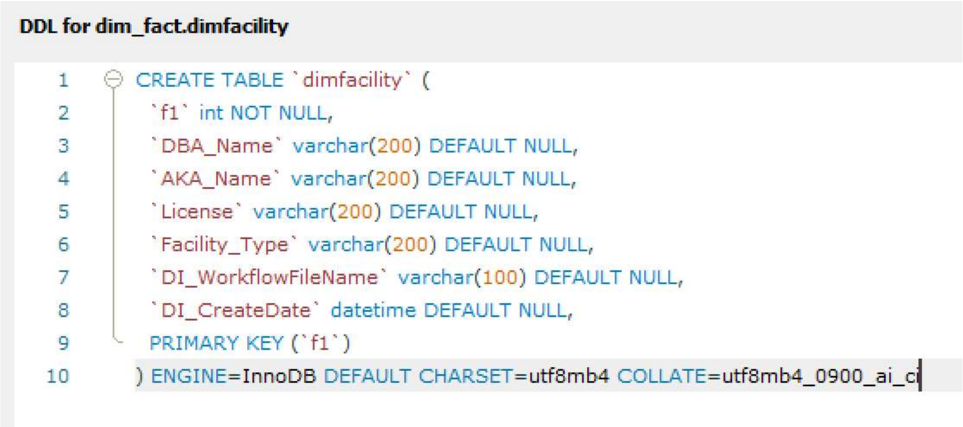


DDL for dim_fact.dimdate

```
CREATE TABLE `dimdate` (
  `d1` int NOT NULL,
  `Inspection_Date` varchar(100) DEFAULT NULL,
  `Inspection_Month` varchar(100) DEFAULT NULL,
  `Inspection_Year` varchar(100) DEFAULT NULL,
  `Inspection_Day` varchar(100) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`d1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

3. DIMENSION FACILITY

```
CREATE TABLE `dimfacility` (
  `f1` int NOT NULL,
  `DBA_Name` varchar(200) DEFAULT NULL,
  `AKA_Name` varchar(200) DEFAULT NULL,
  `License` varchar(200) DEFAULT NULL,
  `Facility_Type` varchar(200) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`f1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```



DDL for dim_fact.dimfacility

```
CREATE TABLE `dimfacility` (
  `f1` int NOT NULL,
  `DBA_Name` varchar(200) DEFAULT NULL,
  `AKA_Name` varchar(200) DEFAULT NULL,
  `License` varchar(200) DEFAULT NULL,
  `Facility_Type` varchar(200) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`f1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

4. DIMENSION VIOLATION SCORE

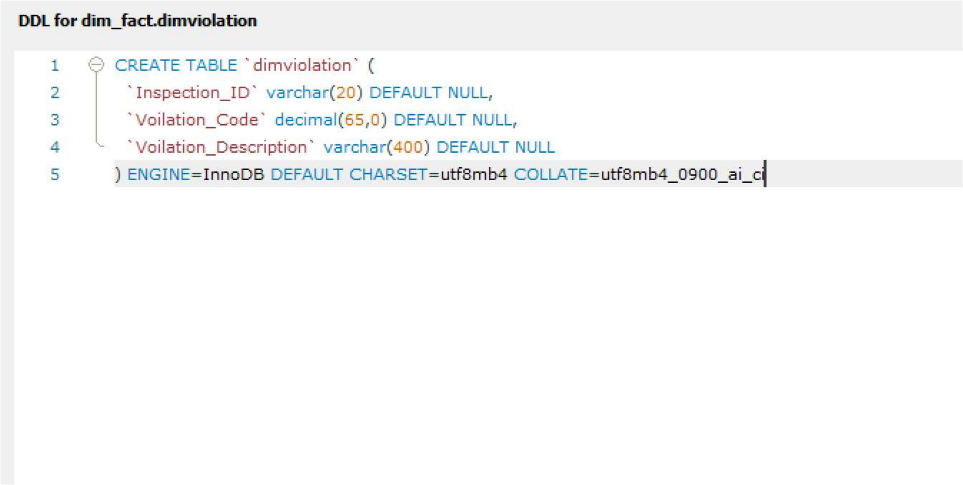
```
CREATE TABLE `dimviolationscore` (
  `null` int NOT NULL,
  `Total_violation_score` int DEFAULT NULL,
  `DI_WorkflowFileName` varchar(50) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`null`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

DDL for dim_fact.dimviolationscore

```
CREATE TABLE `dimviolationscore` (
  `null` int NOT NULL,
  `Total_violation_score` int DEFAULT NULL,
  `DI_WorkflowFileName` varchar(50) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`null`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

5. DIMENSION VIOLATION

```
CREATE TABLE `dimviolation` (
  `Inspection_ID` varchar(20) DEFAULT NULL,
  `Voilation_Code` decimal(65,0) DEFAULT NULL,
  `Voilation_Description` varchar(400) DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```



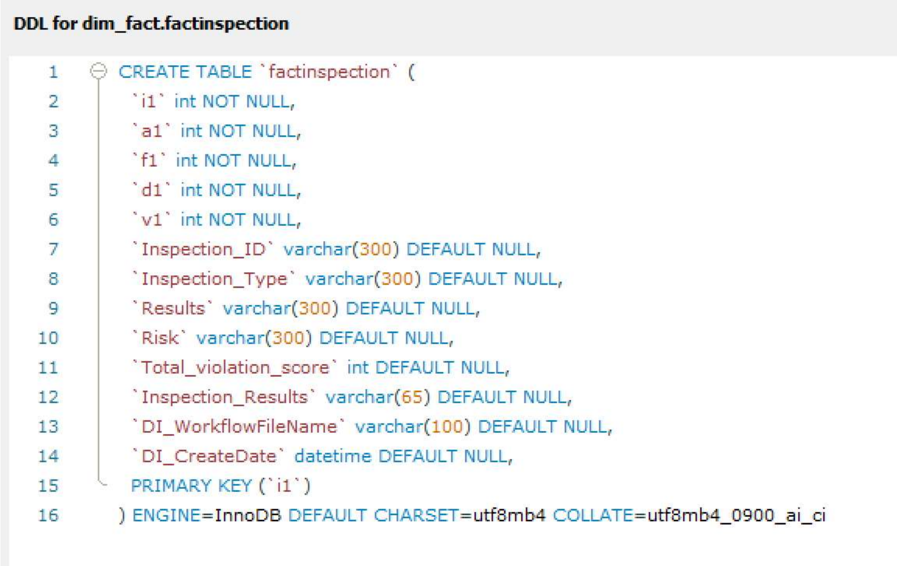
DDL for dim_fact.dimviolation

```
CREATE TABLE `dimviolation` (
  `Inspection_ID` varchar(20) DEFAULT NULL,
  `Voilation_Code` decimal(65,0) DEFAULT NULL,
  `Voilation_Description` varchar(400) DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

(II) FACTS

1. FACT INSPECTION

```
CREATE TABLE `factinspection` (
  `i1` int NOT NULL,
  `a1` int NOT NULL,
  `f1` int NOT NULL,
  `d1` int NOT NULL,
  `v1` int NOT NULL,
  `Inspection_ID` varchar(300) DEFAULT NULL,
  `Inspection_Type` varchar(300) DEFAULT NULL,
  `Results` varchar(300) DEFAULT NULL,
  `Risk` varchar(300) DEFAULT NULL,
  `Total_violation_score` int DEFAULT NULL,
  `Inspection_Results` varchar(65) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`i1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```



DDL for dim_fact.factinspection

```
CREATE TABLE `factinspection` (
  `i1` int NOT NULL,
  `a1` int NOT NULL,
  `f1` int NOT NULL,
  `d1` int NOT NULL,
  `v1` int NOT NULL,
  `Inspection_ID` varchar(300) DEFAULT NULL,
  `Inspection_Type` varchar(300) DEFAULT NULL,
  `Results` varchar(300) DEFAULT NULL,
  `Risk` varchar(300) DEFAULT NULL,
  `Total_violation_score` int DEFAULT NULL,
  `Inspection_Results` varchar(65) DEFAULT NULL,
  `DI_WorkflowFileName` varchar(100) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`i1`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

2. FACT VIOLATION

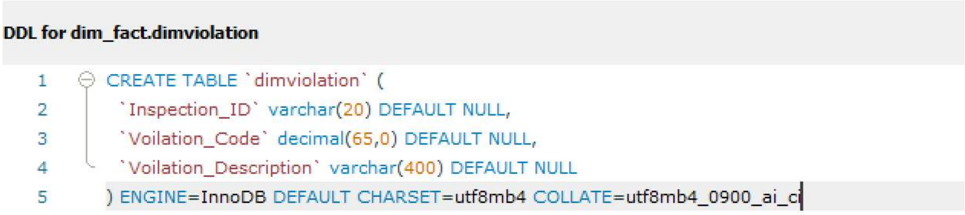
CREATE TABLE `factviolation` (

`v1` int NOT NULL,

`Inspection_ID` varchar(20) DEFAULT NULL,

`Voilations_Code` decimal(65,0) DEFAULT NULL

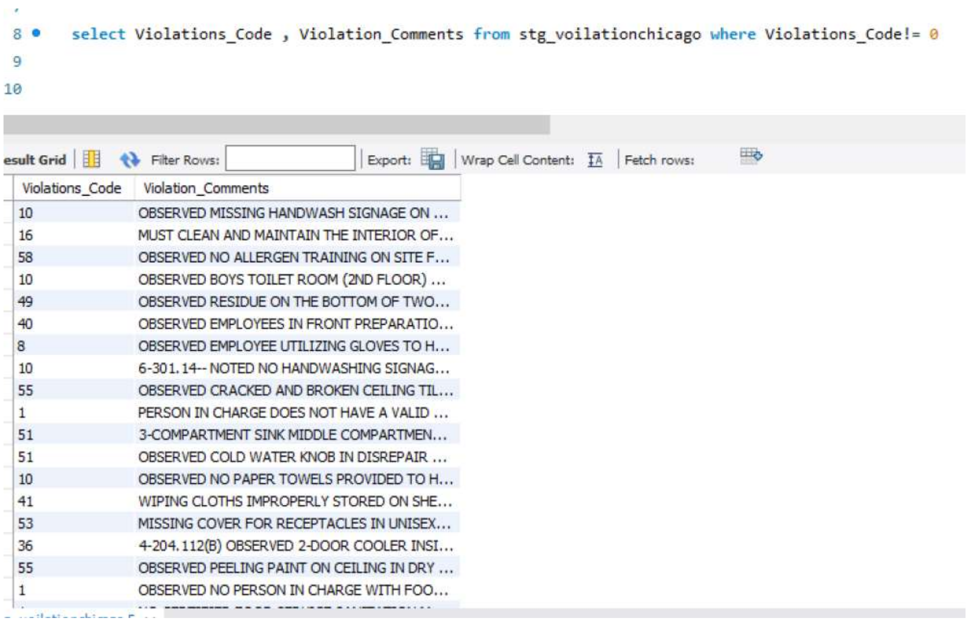
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci



DDL for dim_fact.dimviolation

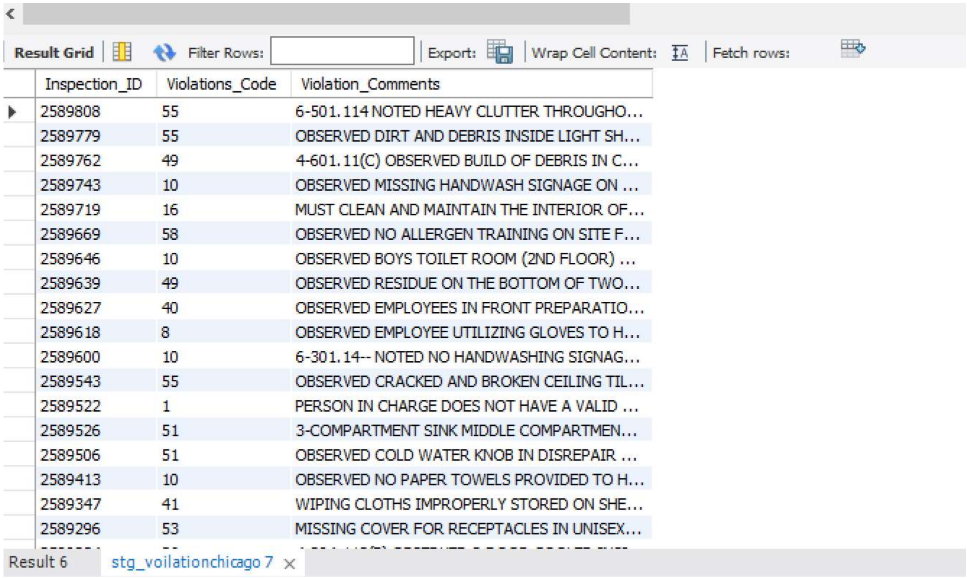
```
CREATE TABLE `dimviolation` (
  `Inspection_ID` varchar(20) DEFAULT NULL,
  `Voilation_Code` decimal(65,0) DEFAULT NULL,
  `Voilation_Description` varchar(400) DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

Showing Violation_Comments given by inspector at inspection



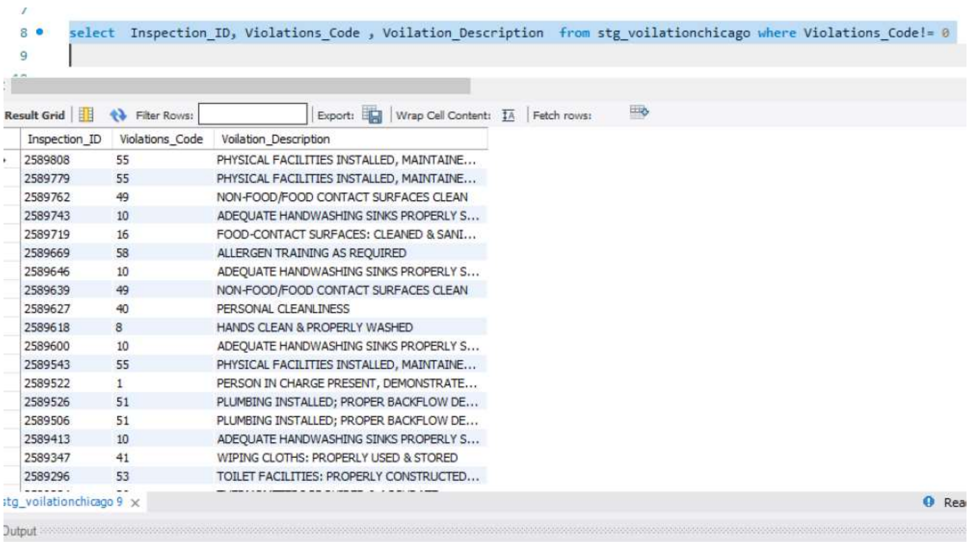
```
select Violations_Code , Violation_Comments from stg_voilationchicago where Violations_Code!= 0
```

Violations_Code	Violation_Comments
10	OBSERVED MISSING HANDWASH SIGNAGE ON ...
16	MUST CLEAN AND MAINTAIN THE INTERIOR OF...
58	OBSERVED NO ALLERGEN TRAINING ON SITE F...
10	OBSERVED BOYS TOILET ROOM (2ND FLOOR) ...
49	OBSERVED RESIDUE ON THE BOTTOM OF TWO...
40	OBSERVED EMPLOYEES IN FRONT PREPARATIO...
8	OBSERVED EMPLOYEE UTILIZING GLOVES TO H...
10	6-301.14-- NOTED NO HANDWASHING SIGNAG...
55	OBSERVED CRACKED AND BROKEN CEILING TIL...
1	PERSON IN CHARGE DOES NOT HAVE A VALID ...
51	3-COMPARTMENT SINK MIDDLE COMPARTMEN...
51	OBSERVED COLD WATER KNOB IN DISREPAIR ...
10	OBSERVED NO PAPER TOWELS PROVIDED TO H...
41	WIPING CLOTHS IMPROPERLY STORED ON SHE...
53	MISSING COVER FOR RECEPTACLES IN UNISEX...
36	4-204.112(B) OBSERVED 2-DOOR COOLER INSI...
55	OBSERVED PEELING PAINT ON CEILING IN DRY ...
1	OBSERVED NO PERSON IN CHARGE WITH FOO...



Inspection_ID	Violations_Code	Violation_Comments
2589808	55	6-501.114 NOTED HEAVY CLUTTER THROUGHO...
2589779	55	OBSERVED DIRT AND DEBRIS INSIDE LIGHT SH...
2589762	49	4-601.11(C) OBSERVED BUILD OF DEBRIS IN C...
2589743	10	OBSERVED MISSING HANDWASH SIGNAGE ON ...
2589719	16	MUST CLEAN AND MAINTAIN THE INTERIOR OF...
2589669	58	OBSERVED NO ALLERGEN TRAINING ON SITE F...
2589646	10	OBSERVED BOYS TOILET ROOM (2ND FLOOR) ...
2589639	49	OBSERVED RESIDUE ON THE BOTTOM OF TWO...
2589627	40	OBSERVED EMPLOYEES IN FRONT PREPARATIO...
2589618	8	OBSERVED EMPLOYEE UTILIZING GLOVES TO H...
2589600	10	6-301.14-- NOTED NO HANDWASHING SIGNAG...
2589543	55	OBSERVED CRACKED AND BROKEN CEILING TIL...
2589522	1	PERSON IN CHARGE DOES NOT HAVE A VALID ...
2589526	51	3-COMPARTMENT SINK MIDDLE COMPARTMEN...
2589506	51	OBSERVED COLD WATER KNOB IN DISREPAIR ...
2589413	10	OBSERVED NO PAPER TOWELS PROVIDED TO H...
2589347	41	WIPING CLOTHS IMPROPERLY STORED ON SHE...
2589296	53	MISSING COVER FOR RECEPTACLES IN UNISEX...

Violation Description



```
select  Inspection_ID, Violations_Code , Voilation_Description  from stg_voilationchicago where Violations_Code!= 0
```

Inspection_ID	Violations_Code	Voilation_Description
2589808	55	PHYSICAL FACILITIES INSTALLED, MAINTAINE...
2589779	55	PHYSICAL FACILITIES INSTALLED, MAINTAINE...
2589762	49	NON-FOOD/FOOD CONTACT SURFACES CLEAN
2589743	10	ADEQUATE HANDWASHING SINKS PROPERLY S...
2589719	16	FOOD-CONTACT SURFACES: CLEANED & SANI...
2589669	58	ALLERGEN TRAINING AS REQUIRED
2589646	10	ADEQUATE HANDWASHING SINKS PROPERLY S...
2589639	49	NON-FOOD/FOOD CONTACT SURFACES CLEAN
2589627	40	PERSONAL CLEANLINESS
2589618	8	HANDS CLEAN & PROPERLY WASHED
2589600	10	ADEQUATE HANDWASHING SINKS PROPERLY S...
2589543	55	PHYSICAL FACILITIES INSTALLED, MAINTAINE...
2589522	1	PERSON IN CHARGE PRESENT, DEMONSTRATE...
2589526	51	PLUMBING INSTALLED; PROPER BACKFLOW DE...
2589506	51	PLUMBING INSTALLED; PROPER BACKFLOW DE...
2589413	10	ADEQUATE HANDWASHING SINKS PROPERLY S...
2589347	41	WIPING CLOTHS: PROPERLY USED & STORED
2589296	53	TOILET FACILITIES: PROPERLY CONSTRUCTED...

SQL queries to validate your BI visualizations

Viz: CNT(inspection_id) VS year

Query:

```
SELECT YEAR(Inspection_Date) AS Year, COUNT(Inspection_Id) AS Inspection_Count
FROM your_table_name
GROUP BY YEAR(Inspection_Date)
ORDER BY YEAR(Inspection_Date);
```

Viz: CNT(inspection_results) VS year

Query:

```
SELECT YEAR(inspection_date) AS inspection_year,
    inspection_result,
    COUNT(*) AS result_count
FROM fact_inspection
GROUP BY YEAR(inspection_date), inspection_result
ORDER BY inspection_year, inspection_result;
```

Viz:Inspection Results VS location

Query:

```
SELECT fi.Inspection_Result,
    da.Longitude,
    da.Latitude
FROM factinspection fi
JOIN dimaddress da ON fi.A1 = da.A1;
```

Viz:Inspection Id VS Facility_Type

Query:

```
SELECT
  df.Facility_Type,
  COUNT(fi.Inspection_ID) AS inspection_count
FROM
  factinspection fi
JOIN
  dimfacility df ON fi.f1 = df.f1
GROUP BY
  df.Facility_Type;
```

Viz: TVS VS DBA Name

Query:

```
SELECT
  df.DBA_Name AS Restaurant_Name,
  SUM(fi.Total_violation_score) AS TVS
FROM
  factinspection fi
JOIN
  dimfacility df ON fi.f1 = df.f1
GROUP BY
  df.DBA_Name;
```

Viz: Risk VS Location

Query:

```
SELECT
   da.Latitude,
   da.Longitude,
   fi.Risk
FROM
   factinspection fi
JOIN
   dimaddress da ON fi.a1 = da.a1;
```

Viz:inspection_type VS Location

Query:

```
SELECT
   da.Latitude,
   da.Longitude,
   fi.Inspection_Type
FROM
   factinspection fi
JOIN
   dimaddress da ON fi.a1 = da.a1;
```

Viz: TVS VS Facility_Type

Query:

```
SELECT
  df.Facility_Type,
  SUM(fi.Total_violation_score) AS TVS
FROM
  factinspection fi
JOIN
  dimfacility df ON fi.f1 = df.f1
GROUP BY
  df.Facility_Type;
```

Viz: Inspection_Result VS License

Query:

```
SELECT
  f.License,
  fi.Inspection_Results,
  COUNT(fi.Inspection_Results) AS result_count
FROM
  factinspection fi
JOIN
  dimfacility f ON fi.f1 = f.f1
GROUP BY
  f.License, fi.Inspection_Results;
```

Viz: Risk count by AKA name

Query:

```
SELECT
  f.License,
  fi.Inspection_Results,
  COUNT(fi.Inspection_Results) AS result_count
FROM
  factinspection fi
JOIN
  dimfacility f ON fi.f1 = f.f1
GROUP BY
  f.License, fi.Inspection_Results;
```

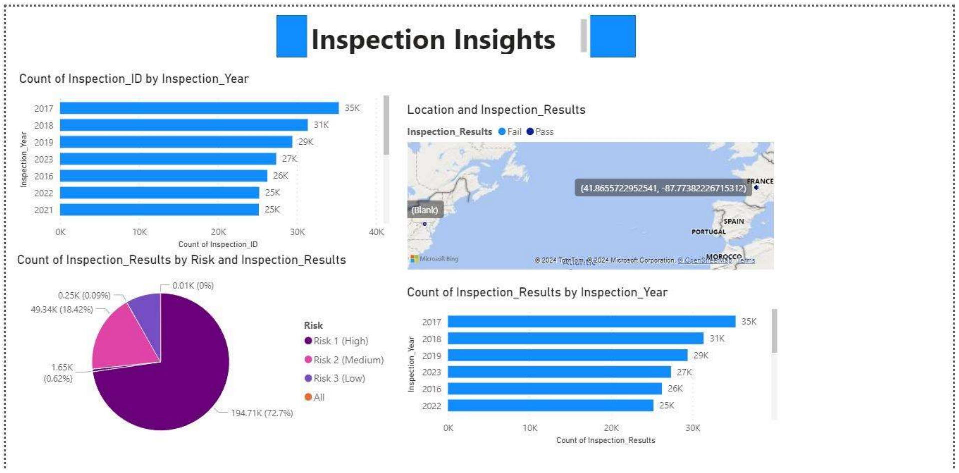
Viz: count of Inspection_Result VS Risk

Query:

```
SELECT
  fi.Risk,
  fi.Inspection_Results,
  COUNT(fi.Inspection_Results) AS result_count
FROM
  factinspection fi
GROUP BY
  fi.Risk, fi.Inspection_Results;
  factinspection fi
JOIN
  dimaddress da ON fi.a1 = da.a1;
```

BI Visualizations

Power BI:



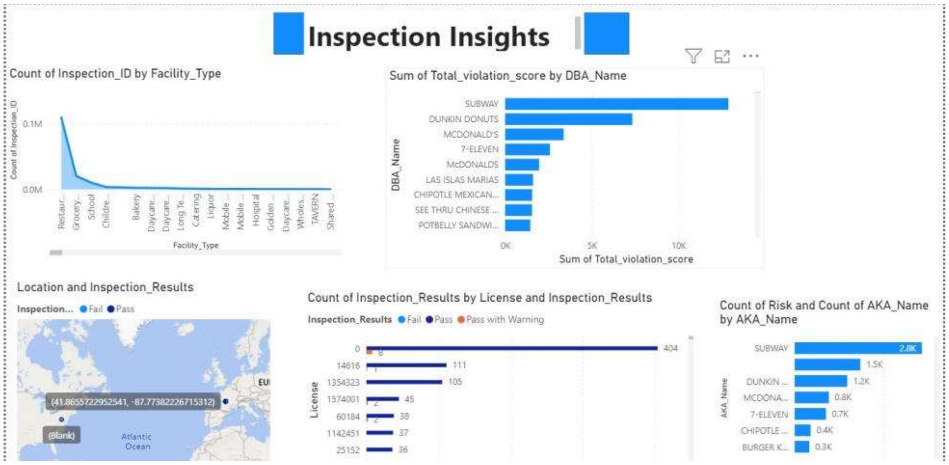
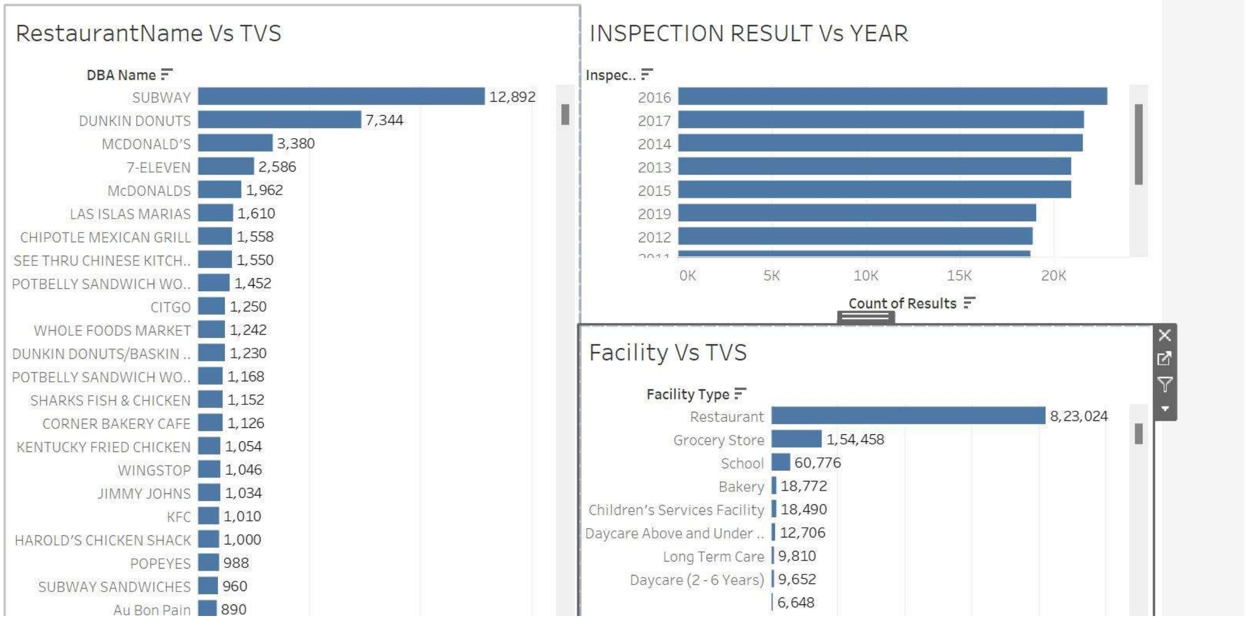


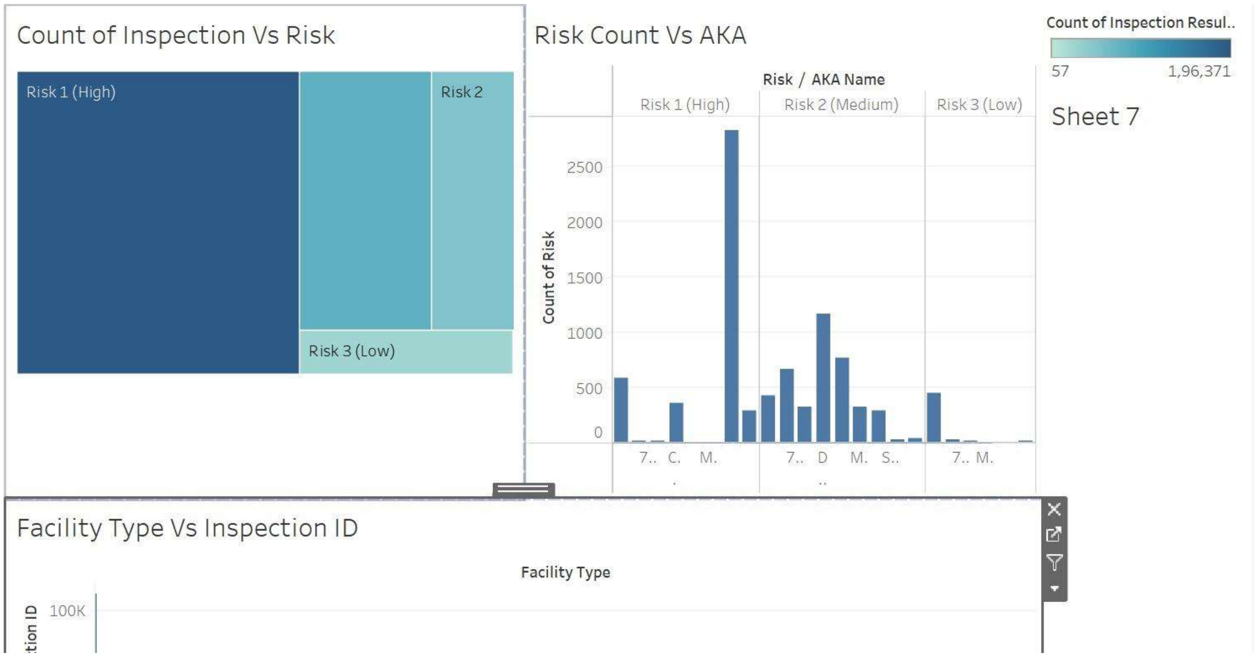
Tableau:



RestaurantName Vs TVS

INSPECTION RESULT Vs YEAR

Facility Vs TVS



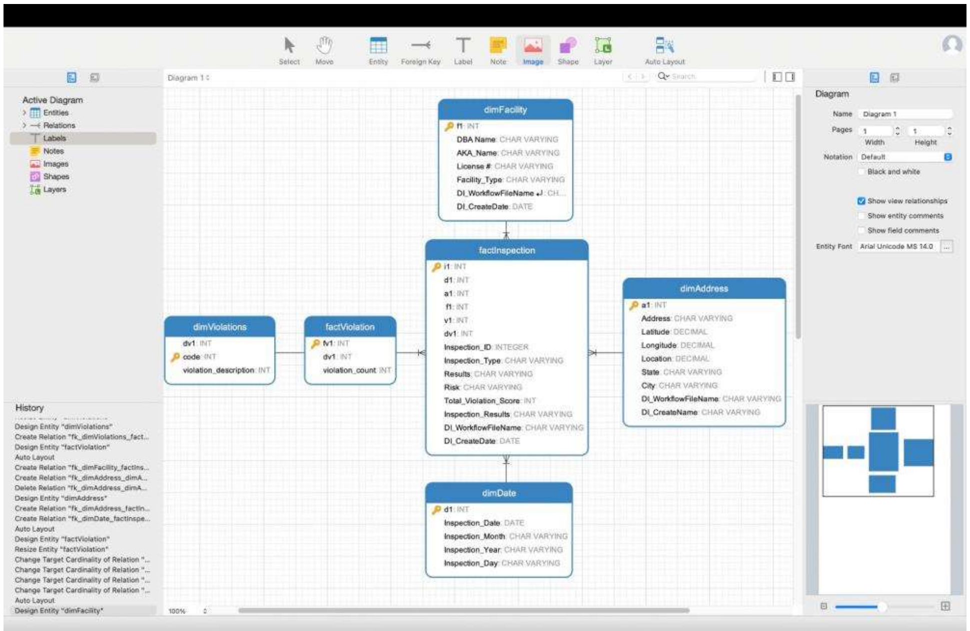
Count of Inspection Vs Risk

Risk Count Vs AKA

Sheet 7

Facility Type Vs Inspection ID

Dimensional Model:



dimFacility

- f1: INT
- DBA Name: CHAR VARYING
- AKA_Name: CHAR VARYING
- License #: CHAR VARYING
- Facility_Type: CHAR VARYING
- DI_WorkflowFileName ↵: CH...
- DI_CreateDate: DATE

factInspection

- i1: INT
- d1: INT
- a1: INT
- f1: INT
- v1: INT
- dv1: INT
- Inspection_ID: INTEGER
- Inspection_Type: CHAR VARYING
- Results: CHAR VARYING
- Risk: CHAR VARYING
- Total_Violation_Score: INT
- Inspection_Results: CHAR VARYING
- DI_WorkflowFileName: CHAR VARYING
- DI_CreateDate: DATE

dimViolations

- dv1: INT
- code: INT
- violation_description: INT

factViolation

- fv1: INT
- dv1: INT
- violation_count: INT

dimAddress

- a1: INT
- Address: CHAR VARYING
- Latitude: DECIMAL
- Longitude: DECIMAL
- Location: DECIMAL
- State: CHAR VARYING
- City: CHAR VARYING
- DI_WorkflowFileName: CHAR VARYING
- DI_CreateName: CHAR VARYING

dimDate

- d1: INT
- Inspection_Date: DATE
- Inspection_Month: CHAR VARYING
- Inspection_Year: CHAR VARYING
- Inspection_Day: CHAR VARYING

History

```
Design Entity "dimViolations"
Create Relation "fk_dimViolations_fact...
Design Entity "factViolation"
Auto Layout
Create Relation "fk_dimFacility_factIns...
Create Relation "fk_dimAddress_dimA...
Delete Relation "fk_dimAddress_dimA...
Design Entity "dimAddress"
Create Relation "fk_dimAddress_factIn...
Create Relation "fk_dimDate_factInspe...
Auto Layout
Design Entity "factViolation"
Resize Entity "factViolation"
Change Target Cardinality of Relation "...
Change Target Cardinality of Relation "...
Change Target Cardinality of Relation "...
Change Target Cardinality of Relation "...
Auto Layout
Design Entity "dimFacility"
```

Diagram

Name: Diagram 1

Pages: 1 (Width) 1 (Height)

Notation: Default

□ Black and white

☑ Show view relationships

□ Show entity comments

□ Show field comments

Entity Font: Arial Unicode MS 14.0