**E:** the original data was obtained from Kaggle. It is about LA restaurants inspection to see whether they violate rules and the case follow up.

**T:** Two csv datasets was found about the LA restaurant inspection. The following columns were extracted from the tables. These data can be used to analysis the status of the restaurant violence. The rest of data was cleaned out. The header was renamed with a different format. Replace the index with “SERIAL\_NUMBER”

There are two tables: table 1: Inspection\_transformed; table 2: Violation\_transformed.

Table 1 "facility\_name", "facility\_id", "owner\_id", "service\_description", "program\_status", "serial\_numer"

Table 2: "serial\_number", "facility\_name", "violation\_code", "violation\_status", "grade"

**Load the data:**

* use Pgadmin to create a database named “ETL\_project”.
* Create tables: ”inspection” and “violation”
* Create headers for the tables
* Create database connection
* Load dataframes into database
* Inner join the two tables on “serials\_number”
* Use pd.read\_sql\_query to check the data to make sure the data was upload successfully