1) **To separate useful data**
   select * except(latitude, longitude, location, location_description, x_coordinate, y_coordinate, description) from `73.raw_acd` where council_district_code is not null and clearance_status is not null
2) **Make binary column for clearance_status**
   update `73.u_acd` set status_int=0 where clearance_status like 'Not%'
3) **To simplify the large variety of type of crime (for one-hot encoding)**
   update `73.u_acd` set primary_type='Theft' where primary_type like '%Theft%'
   Note: Same query for all the distinct primary_type attributes
4) **One-hot encode primary_type column for regression**
   update `73.u_acd` set theft_int=1 where primary_type='Theft'
   Note: Same query for all the distinct primary_type attributes
5) **Get data for regression**
   select * except(unique_key, address, clearance_date, zipcode, clearance_status, district, census_tract, primary_type) from `73.u_acd`
6) **Regression using BigQuery ML**
   CREATE OR REPLACE MODEL `73.acdmodel` OPTIONS(
   model_type='logistic_reg', input_label_cols=['status_int']) AS
   SELECT * except(year), status_int as si FROM `73.ml_acd` where year!=2016
7) **Prediction results**
   SELECT * FROM ML.PREDICT(MODEL `73.acdmodel`, (select *, status_int as si from `73.ml_acd`))