Anpassung der Daten, um sie nutzbar zu machen. Es wurde dafür ein Python Skript geschrieben, welches die Yelp Daten mit den Inspektionsdaten verknüpft hat. Dafür wurden allen eindeutig zuordenbaren Restaurants in dem Inspektionsdatensatz eines der meistbewerteten Businesses aus Yelp mit der Kategorie Restaurant zugeordnet und in einer mapping Tabelle dargestellt.

fakingScript.py

Danach müssen die Daten bereinigt werden. Lokale ohne eindeutigen INSPECTION DATE (hier 01/01/1900) müssen entfernt werden, sowie die den zugehörigen

Zwei SQL-Tabellen benötigt

1. Yelp Daten
2. Inspection results

Yelp Daten: Object ID = Primärschlüssel

Business Id, Name, Address, Postal Code, Borough, latitude, longitude, Phone, Stars, Reviews, Cuisine Description

Inspection results: Inspection (Object) ID = Primärschlüssel, KAMI (Business ID) = Fremdschlüssel

Inspection Date, Action, Violation Code, Violation Description, Critical Flag, Score, Grade, Inspection Type, Community Board, Council District, Census Tract, BIN, BBL, NTA