Seattle Food Inspections

Technology Review

David Clancy, Sadra Fardhosseini, Greg Talpey and Harry Xie

Background

Starting 2012, jurisdictions across the country including King County have begun publishing health inspection scores using a standardized scoring system called LIVES. This open data allowed restaurant consumers to make informed decisions based on where they want to eat and motivated a lot of restaurant establishments to improve their inspection score in the hopes of attracting a bigger customer base.

Brief overview

1. **Trend Analysis**: What is the overall trend for the aggregate number of different risk violations from 2016 to 2019?
2. **Location Specific Contributions**: Where are the specific regions in Seattle that have the most high risk violations? How much do those regions contribute in terms of total high risk violations in Seattle?
3. **Interactive Map**: showing health inspection ratings in Seattle.
4. Then overlay these ratings with demographic information in the areas around restaurants.
5. Users will be able to explore **correlations** between the demographics and health inspection ratings to help make decisions about where to eat.

Dataset

1. Food Establishment Inspection Data - King County Open Data:

   https://data.kingcounty.gov/Health-Wellness/Food-Establishment-Inspection-Data/f29f-zza5

2. American Community Survey, 5-Year Estimates:

   https://data.census.gov/cedsci/table?q=S1903&table=S1903&tid=ACSST1Y2018.S1903&lastDisplayedRow=0

   https://data.census.gov/cedsci/table?q=S1201&table=S1201&tid=ACSST1Y2018.S1201&lastDisplayedRow=0

Seattle Restaurants Inspection

Select Year: 2014 | **Select Month:** Jan | **Select Day:** 01

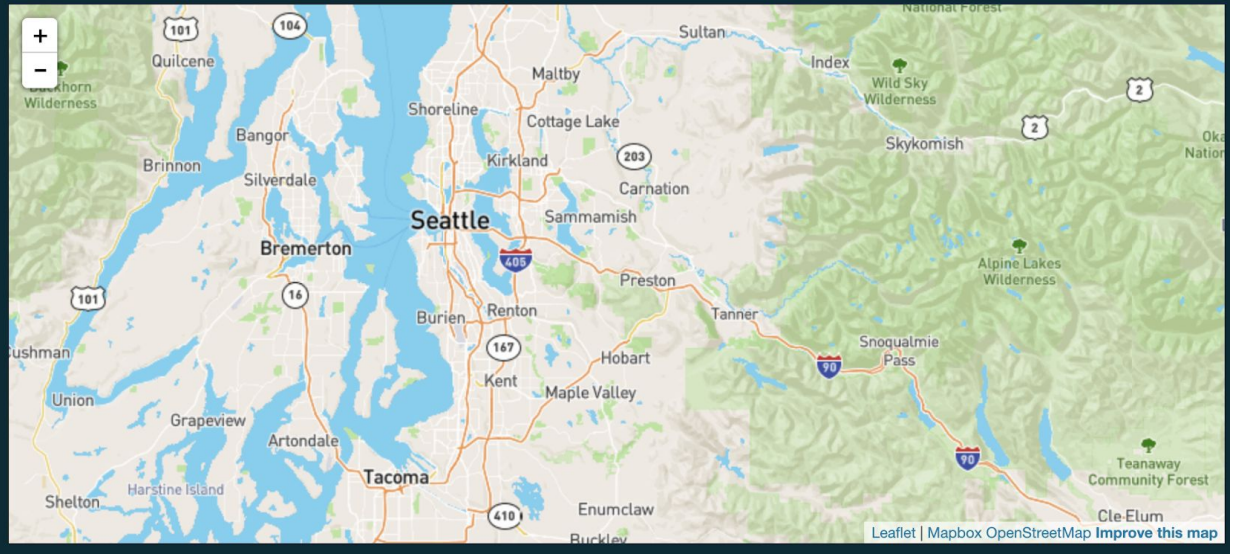
Violation Type: No Risk

☐ Display High Risk Violations Heatmap by Postcode

Plot Restaurants of HR Violations for Top 7 Postcode: 98105

SF Restaurant Location Map | Time Series: Number of Risk Violations | Bar Chart: Risk Violations by SF Postcodes

Bar Chart: Top 20 Restaurants Postcode



Use Case

BMC Public Health. 2014; 14: 571.
Published online 2014 Jun 7. doi: 10.1186/1471-2458-14-571
PMCID: PMC4057591
PMID: 24908104

Factors affecting food handling Practices among food handlers of Dangila town food and drink establishments, North West Ethiopia

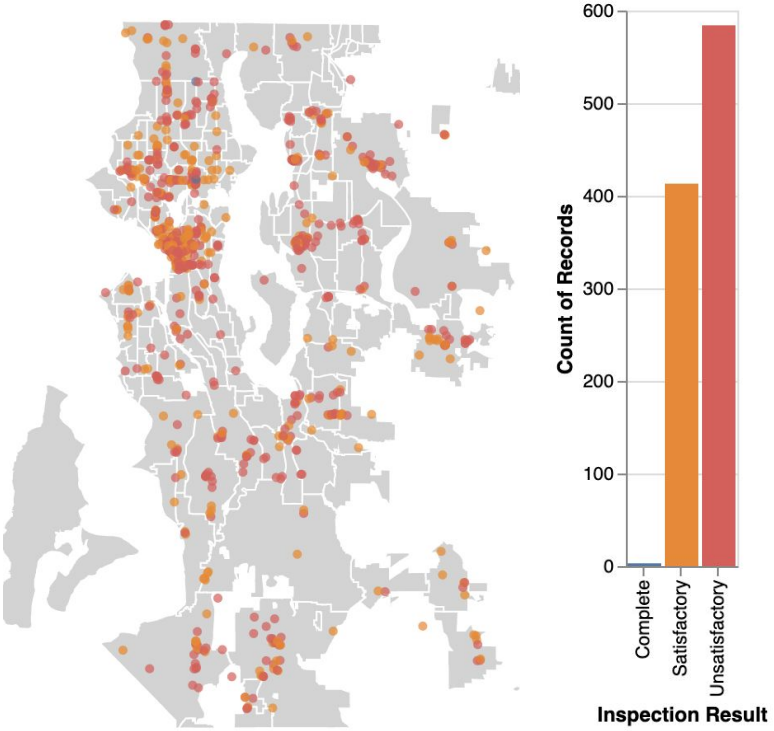
Ayehu Gashe Tessema,[1] Kassahun Alemu Gelaye,[1] and Daniel Haile Chercos[✉1]

At least in NW Ethiopia, certain demographic information of the workers is associated with food handling practices.

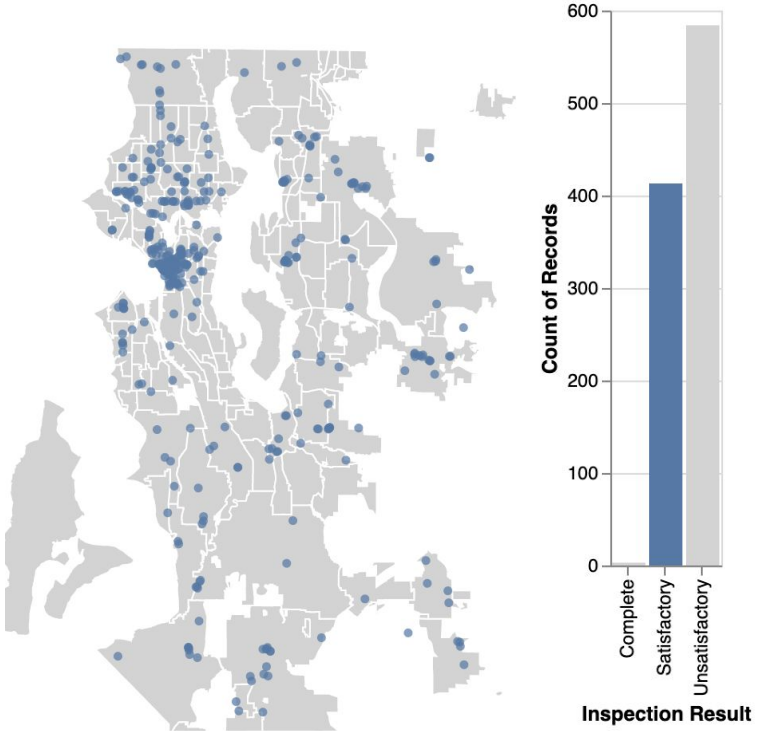
Users can see visually if such correlation between restaurant locations and some of the same demographic information in Seattle.

For example, seeing if satisfactory food inspection scores are associated with the marital status of the area surrounding the restaurant.

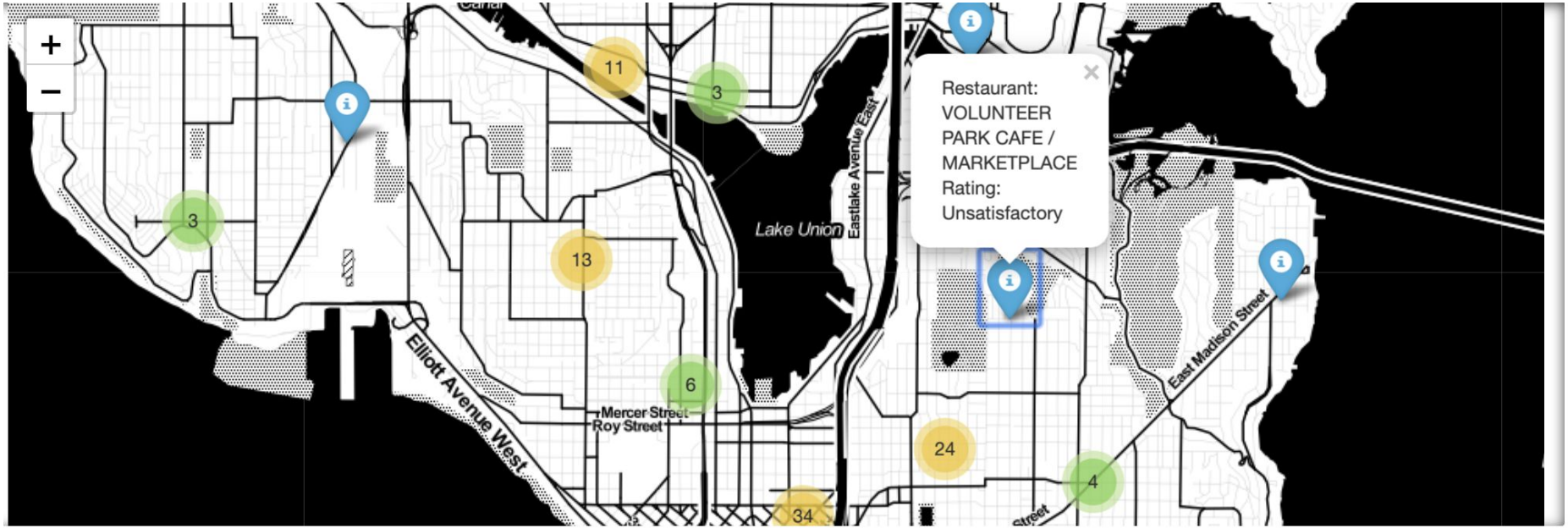
Libraries - Altair



Select a bar from the histogram



Libraries - Folium



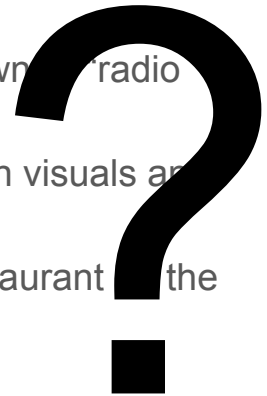
Altair v. Folium

Altair will enable:

- Filtering interaction from drop down or “radio button” menus
- Cross-filtering between correlation visuals and the map
- Seamless placement of each restaurant on the basemap, with interactive tooltips

The tradeoffs for taking advantage of Altair’s interaction capability:

- Map won’t zoom or pan
- May not be able to add streets



Folium will enable:

- Filtering interaction from drop down or “radio button” menus
- Typical “Google Maps” like features
- “Popup” information (like an Altair tooltip)

But, it would be harder to:

- Cross-filter with other visuals

Libraries - Pandas

Pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with “relational” or “labeled” data both easy and intuitive. We use Pandas for data analysis.

Libraries - Scikit-learn or Pysal

Scikit-learn is a simple and efficient tools for data mining and data analysis.

Pysal PySAL is an open source cross-platform library for geospatial data science with an emphasis on geospatial vector

We will **potentially** use these libraries to build a regression model predicting future food rates based on restaurants' location, region weather fluctuations and the high temperature.