

Project Proposal

Weiling Li

12/14/2019

1. Project

1.1 Topics

Do restaurants exhibit a geospatial clustering by genre? For example, Chinese cuisine clustered in allston and chinatown. Italian in North End, Vietnamese in Dorchester. Does the rating of restaurants also clustered geospatially? When conditioned in one attribute(genre or rating), do this two clusterings have a strong interaction?(Maybe the restaurant in it's genre cluster are likely to associated with higher ratings or not?)

How does the findings in Boston differs with the findings in San Francisco and why?

1.2 Data Overview

Data used in this project comes from yelp, specifically from yelp **api**.

Because of the limit of yelp's free api and the scope of this project, the following cuisine genre are observed:

- Chinese
- Korean
- Japanese
- Mediterranean
- Italian
- Greek
- Spanish

As described above, the City being compared are: Boston and San Francisco. The original yelp challenge dataset does not include these two cities, so api was inevitably used to analyse the problem.

1.3 Deliverables

The final deliverables, as discussed in class, are:

1. Preliminary Document(this one)
2. A Reproducible Presentation
3. A Shiny App