**Create data instruments and procedures:**

**7. Data Dictionary:**

|  |  |  |
| --- | --- | --- |
| Variable | Data Type | Tips for accurate measurement |
| Region Name | text | region in the U.S. |
| Restaurant Name | text |  |
| Health Score | integer | min: 0 max: 100 |

**8. Sampling Procedure:**

* Opening the website of Yelp(<https://www.yelp.com/>)
* Searching the restaurant in one region
* Accessing the restaurant main page and find the Health Score
* Gathering the first 100 restaurant by python scraping from San Francisco and Washington, DC (Yelp places the items randomly)

**9. Cite a journal article:**

* Restaurant Inspection Scores and Foodborne Disease (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3323064/>)

**10.Formulate Hypothesis:**

* Hypothesis: The health condition of Washington, DC might be better than San Francisco.
* Reason: Washington, DC is the capital city of the U.S. It should have the best city management system in the world including the management system of health condition of its public building

 