

# Notebook

February 14, 2020



Use the `head` command on your four files again. This time, describe at least one potential problem with the data you see. Consider issues with missing values and bad data.

Please write your answer in the markdown cell below. You may create new cells below your answer to run code, but **please never add cells between a question cell and the answer cell below it.**

Latitude and longitude coordinates are unable for some places, showing up as -9999, -9999 instead.



---

## 1 6: Compute Something Interesting

Play with the data and try to compute something interesting about the data. Please try to use at least one of groupby, pivot, or merge (or all of the above).

Please show your work in the cell below and describe in words what you found in the same cell. This question will be graded leniently but good solutions may be used to create future homework problems.

**Please have both your code and your explanation in the same one cell below. Any work in any other cell will not be graded.**

```
In [310]: high_risk_restaurants = vio.merge(ins2vio).merge(ins).merge(bus)
          high_risk_restaurants = high_risk_restaurants.loc[high_risk_restaurants['risk_category'] == "High Risk"]

          print("This table gives all restaurants with a risk category of High Risk, along with any information available")
```

This table gives all restaurants with a risk category of High Risk, along with any information available