

## PART 5 – Report (50points)

### **1. The goals for your project including what APIs/websites you planned to work with and what data you planned to gather (5 points)**

- Goals: Our goals for the project were to look at New York's different Boroughs and see differences in statistics occurring across all the boroughs. Initially we intended to look at tree count distributions across boroughs and see how demographics lined up with equitable green spreads. Our APIs were going to be NYC OpenData's APIs on health statistics (asthma counts), demographics(black pop.), and tree counts per borough.
  - The goal switched through the project when we realized that the API's we chose wouldn't work because of the base url, so we switched out focus to connecting different NYC data

### **2. The goals that were achieved including what APIs/websites you actually worked with and what data you did gather (5 points)**

- Our end goals focused more on boroughs in NYC and also to see what statistics may not have correlation to demographic or borough. We looked at demographics and how that could correlate with yelp reviews and car safety in collisions. The APIs we ended up using changed around but we ended up landing at NYC OpenData's APIs for Collision data (borough, people injured, and collision IDs) and demographics (black pop. per borough), a Yelp API that had ratings across boroughs, and a weather API that reports weather averages per borough (temp, humidity, wind speeds).

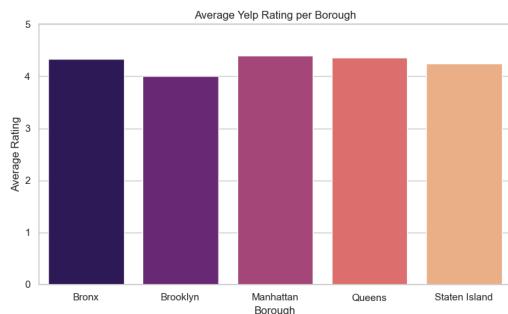
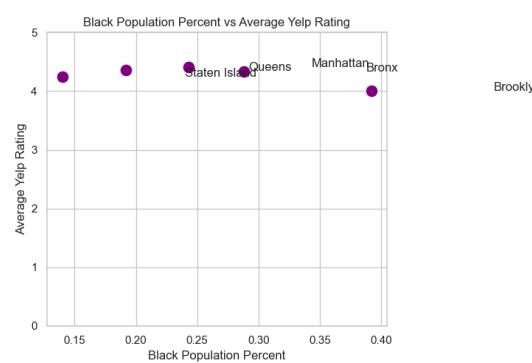
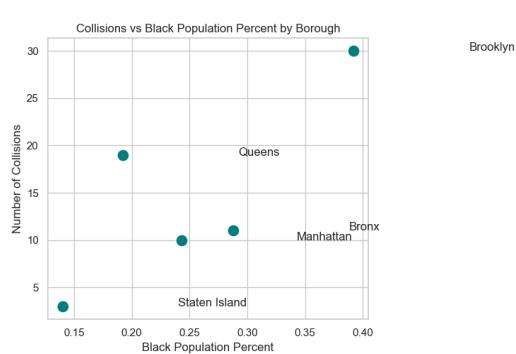
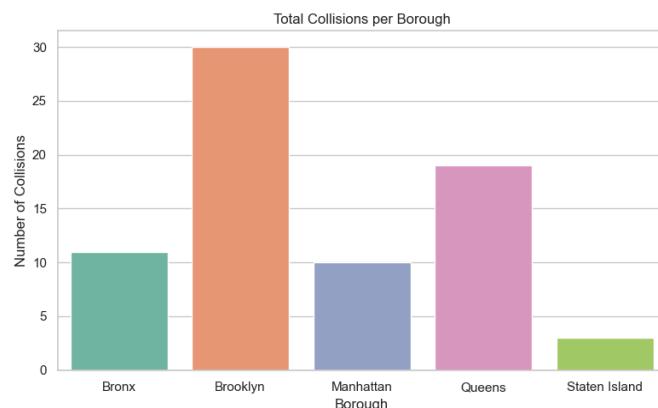
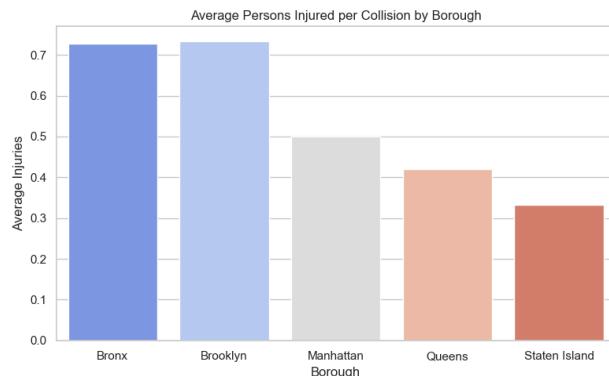
### **3. The problems that you faced (5points)**

- We faced many major problems, including rewriting most of the project in the final few days after the review since we had all of our data coming from different APIs all under NYC OpenData and we were told we had to have them all be through different APIs with different base URLs.
- Another challenge faced included figuring out how to write code in different files and how to integrate and connect them to one another because we had not done much of that in class.
- Working the Database wasn't the easiest section too and navigating joins and duplicate string data was confusing for our group and took some help from AI debuggers.

4. The calculations from the data in the database (i.e. a screen shot) (5points)

```
Ξ calculations.txt ×  
Ξ calculations.txt  
1 === Collisions per Borough ===  
2 Bronx: 11  
3 Brooklyn: 30  
4 Manhattan: 10  
5 Queens: 19  
6 Staten Island: 3  
7  
8 === Average Persons Injured per Borough ===  
9 Bronx: 0.73  
10 Brooklyn: 0.73  
11 Manhattan: 0.50  
12 Queens: 0.42  
13 Staten Island: 0.33  
14  
15 === Black Population Percent by Borough ===  
16 Manhattan: 0.24%  
17 Brooklyn: 0.39%  
18 Queens: 0.19%  
19 Bronx: 0.29%  
20 Staten Island: 0.14%  
21  
22 === Average Yelp Rating per Borough ===  
23 Bronx: 4.34  
24 Brooklyn: 4.00  
25 Manhattan: 4.40  
26 Queens: 4.36  
27 Staten Island: 4.25  
28
```

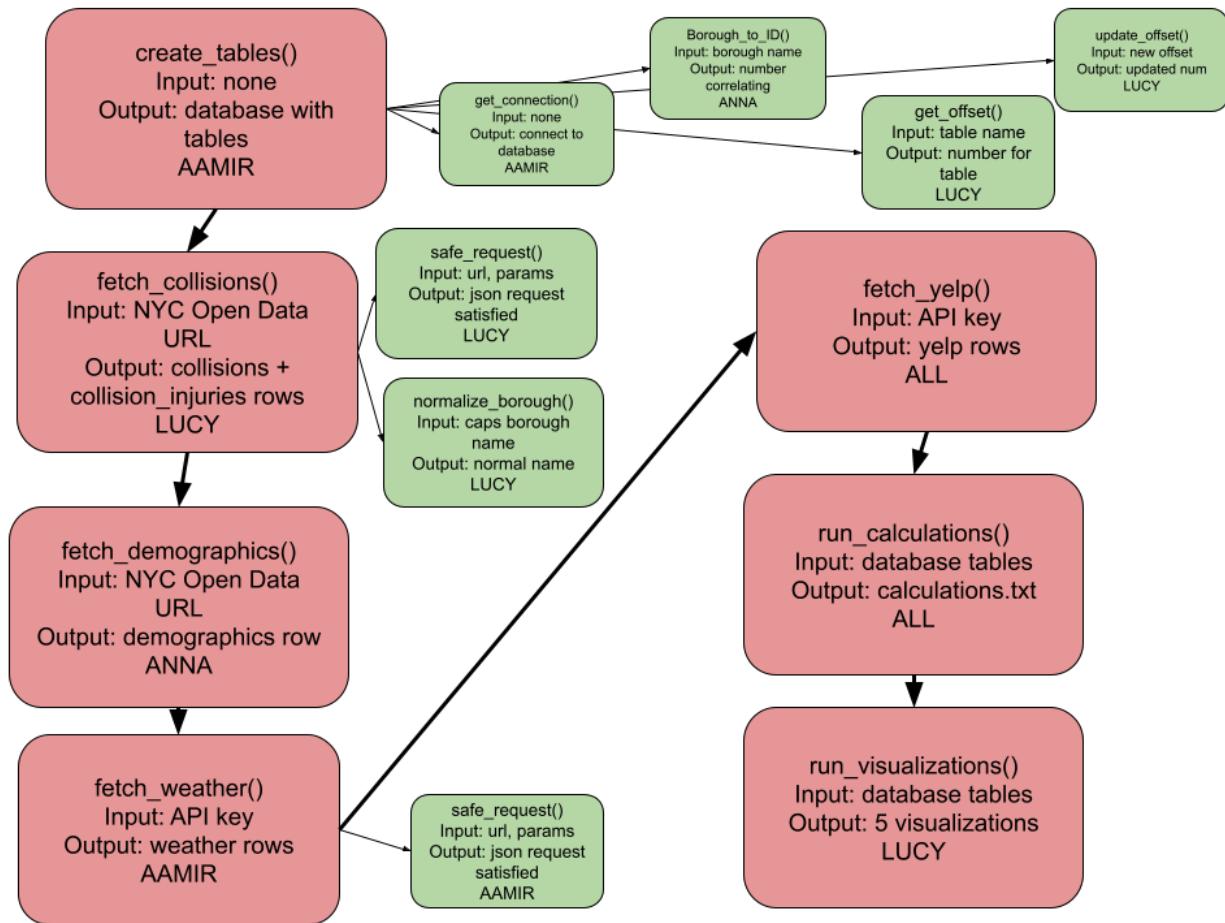
## 5.The visualization that you created (i.e. screen shot or image file)(5points)



## 6. Instructions for running your code (5points)

1. Run db\_utils.py to create the database
2. Run fetch\_collisions.py, fetch\_weather.py, fetch\_demographics.py, and fetch\_yelp.py multiple times since each inserts 25 rows
3. Run calculations.py to get calculations.txt
4. Run visualizations.py to get visualizations

## 7. An updated function diagram with the names of each function, the input, and output and who was responsible for that function(10 points)



## 8. Resources

Date	Issue Description	Location of Resource	Result (Did it solve the issue?)
12/5	Database structure confusion	SQLite schema setup (lec/dis slides)	Yes, correctly linked all tables
12/5	API requests and key errors	API examples in homeworks and lecture	Yes, Data fetched reliable
12/8	Borough name mismatches in tables	ChatGPT data-normalization guidance	Yes, Removed 'unknown borough' errors
12/11	Locked database errors	ChatGPT SQLite connection-handling help	Yes, made sure we were writing to database correctly
12/11	Disorganized fetch scripts	ChatGPT python project structure examples	Yes, we were able to work out consistent non-breaking scripts
12/12	Collisions script not inserting rows	ChatGPT guidance in debugging conversation	Yes, identified normalization and logic issues
12/12	"unknown borough" being printed repeatedly	ChatGPT debugging instructions, told us to rewrite borough normalization function	Yes, fixed borough parsing by mapping all variants properly
12/12	SQLite error: database is locked when updating collision offset	ChatGPT explanation and debugged update_offset()	Yes, solved by performing offset update inside the same DB connection
12/12	Confusion between weather fetch and collisions fetch since they were new APIs	ChatGPT clarified	Yes, clarified which script was causing the issue
12/12	Incorrectly written script for offsets and collisions	ChatGPT showed what was wrong in the code	Yes, collisions offset now increments correctly

12/12	Struggled to separate calculations and visualizations into two files	ChatGPT provided explanation of how to do it	Yes, both files created in correct format
12/12	Needed help generating calculations.txt output format	ChatGPT explanation + file-writing code	Yes — created file output logic