Final Project Document

Jackyty

**Project code:**

1. Github url: <https://github.com/ytyi/spotcrime_project>
2. README.md contained
3. Packages Requirement: requests, flask, seaborn, sqlite3, json, bs4 (also figured out in README.md)

**Data sources:**

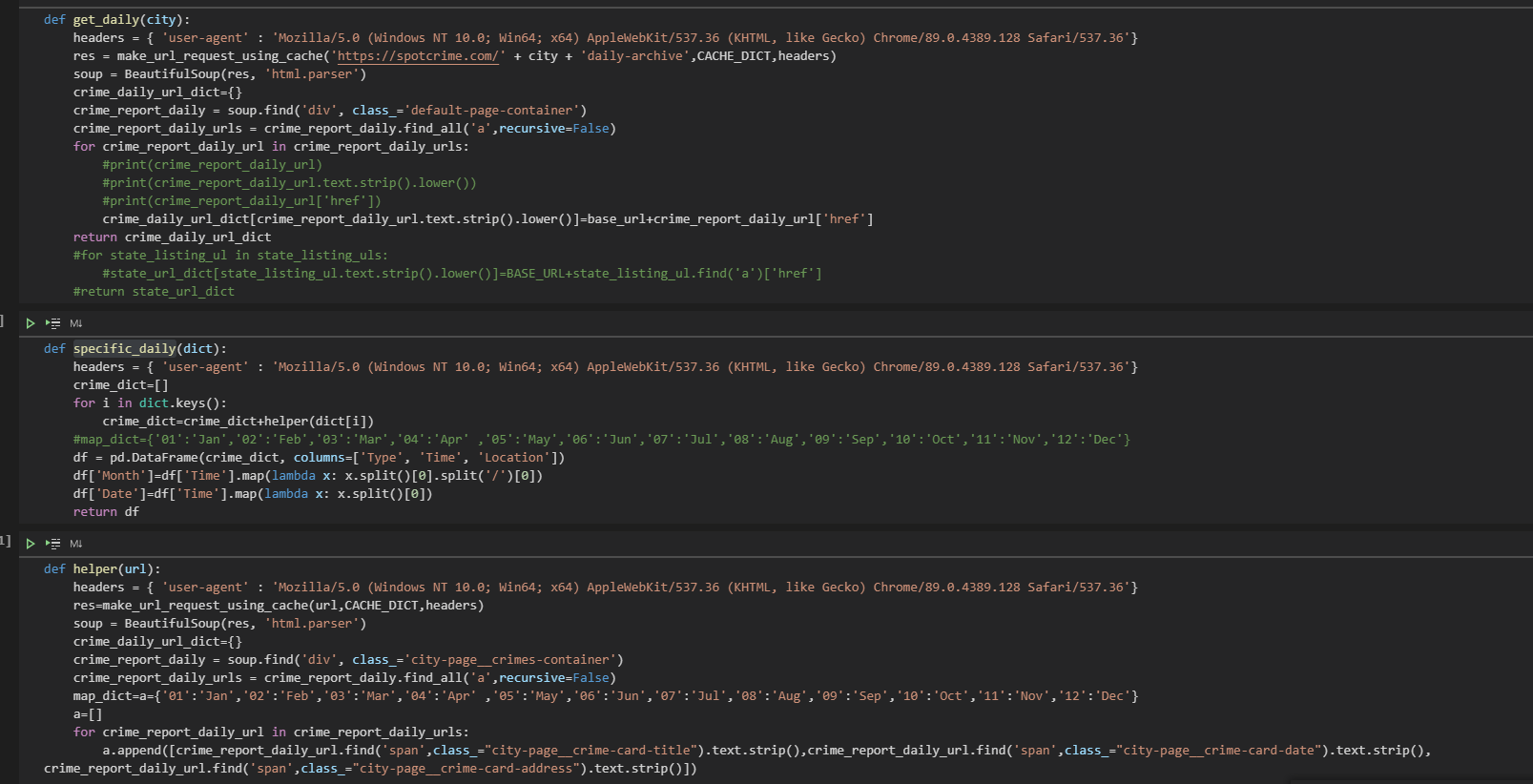
1. The data origin url: https://spotcrime.com/
2. The data document:

SpotCrime is a public facing crime map and crime alert service. With SpotCrime, it’s easier than ever to check crime anywhere in the United States and many other countries worldwide.Our goal is to provide the most accurate and timely crime information to the public.As the most visited crime mapping website, SpotCrime allows you to easily look up city and county crime data and reports -- all the way down to local neighborhood statistics.

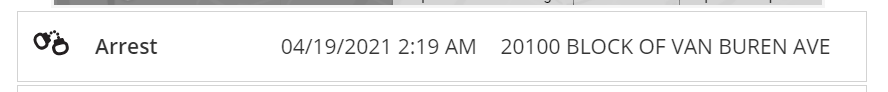
The content of document can be found at https://spotcrime.com/MI/Detroit/daily-archive/ etc

Daily archive records the detailed information of the crimes happened that day within the region given.

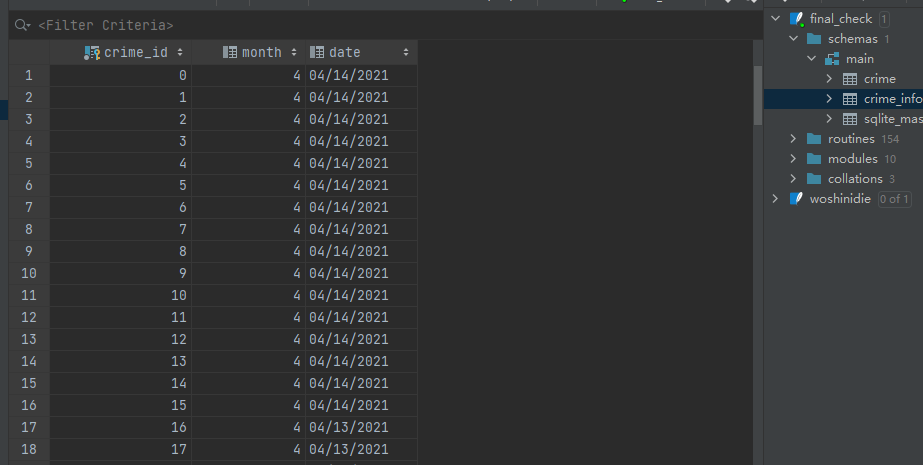
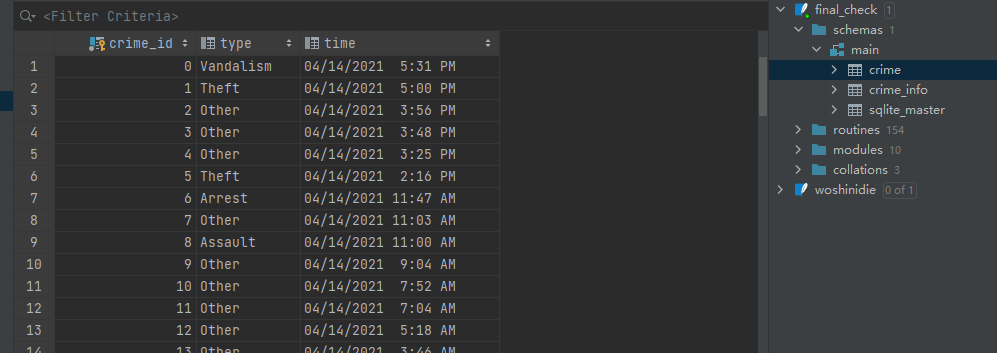
1. I access the data through crawling with login cookies needed and caching is applied for storing data crawled. Attached below is the part of codes for caching and crawling.



1. The summary of data:
2. For one city like Ann Arbor, there are 800 rows of records. Depending on the city numbers, the records number various. On spotcrime, there are also other options besides indexing by city. The total available record is estimated to be several millions records.
3. I collect data within 6 cities for analyze and the total record number is around 4000.
4. For one record, it contains crime type, time and location just as figure showing below



**Database:**

1. Database schema:
2. create\_crime = '''
3. create table crime
4. (
5. crime\_id int not null
6. constraint crime\_pk
7. primary key,
8. type text,
9. time text
10. );
12. '''
14. create\_crime\_info='''
15. create table crime\_info
16. (
17. crime\_id int not null
18. constraint crime\_info\_pk
19. primary key
20. constraint crime\_info\_crime\_crime\_id\_fk
21. references crime
22. on delete cascade,
23. month int,
24. date date
25. );
26. '''
27. Primary key is defined each table which is the crime\_id, the tables are connected by the foreign key crime\_id. The table design is in accordance with 1NF, 2NF, 3NF.
28. Screenshot of data in tables:

**Interaction and Presentation Plan:**

1. The program is primarily about city crime analysis and it provide users with 6 options in city selection. Detailed data visualization will be auto displayed after users choose the city they interested.
2. The interactive technology used: both flask and command line prompt are available. The default choice is Flask interaction.
3. User can select the city in which they are interested in the crime condition with a click on the button represent the city on the UI interface. Then the program will provide various data visualization, like pie chart showing the ratio of crime types, the bar chart showing the places with high crime occurrence, line chart showing the trend of crime cases, bar plot showing the occurrence of different crime types. The whole dataset will also be attached in case user want to do further analysis on the data itself