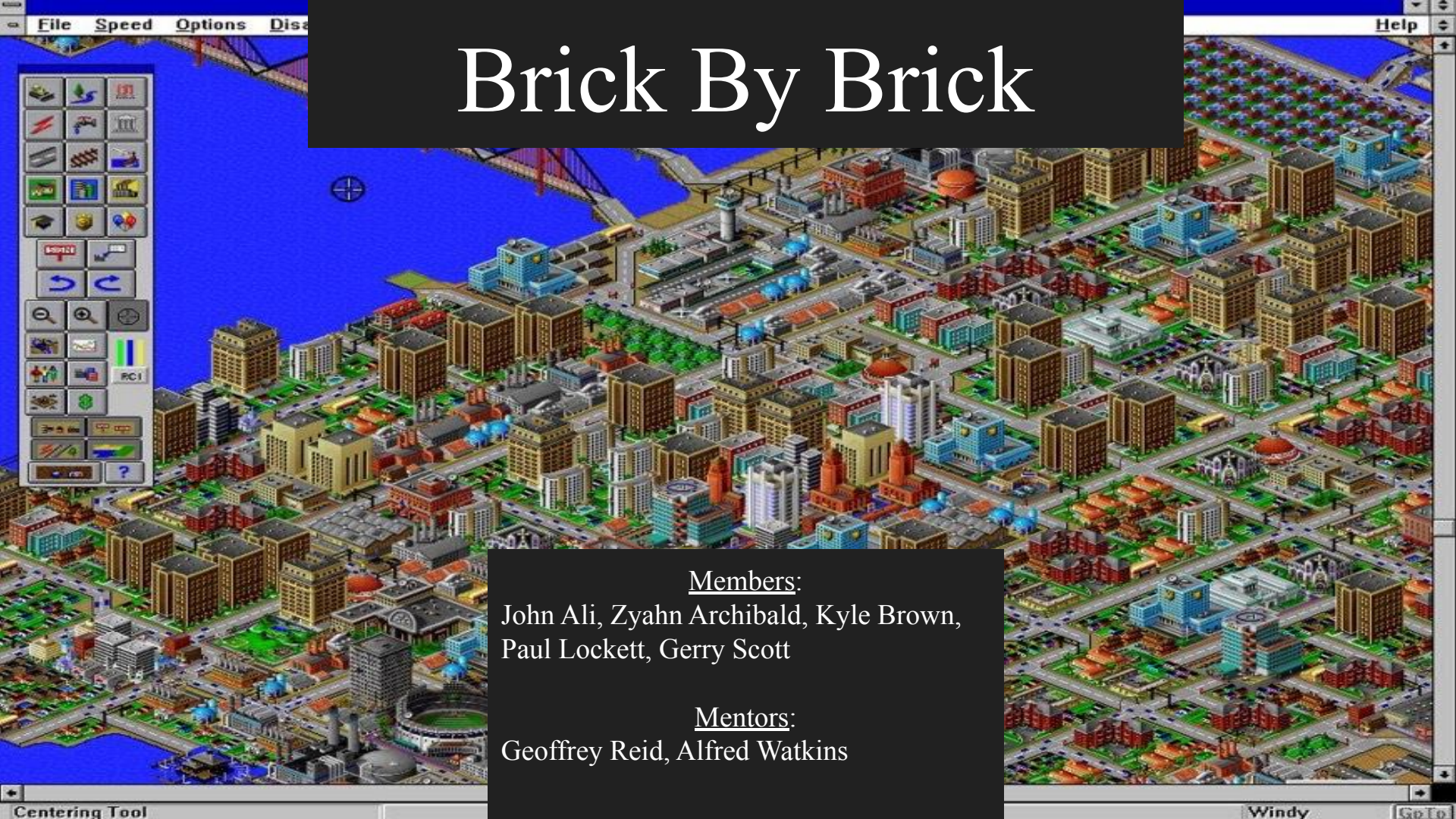
Brick By Brick



Members:
John Ali, Zyahn Archibald, Kyle Brown, Paul Lockett, Gerry Scott

Mentors:
Geoffrey Reid, Alfred Watkins

Overall Goal

- Final Goal:
  - Wrap our functions into some type of GUI such as a web page.

- Initial Goal:
  - Create a system which was composed of a scoring system that analyzed housing data and crime data reported in St. Louis between 2019-2020 and map and explain any correlations.

Implementation

- Google Collab
- Jupyter Notebook
- CSV Files and Data Collection
- Python Functions
- Host on GPC

Issues

- Initial errors in functions
- Issues correctly reading in data from CSV files.

4065 lines (4065 sloc) | 593 KB

We can't make this file beautiful and searchable because it's too large.

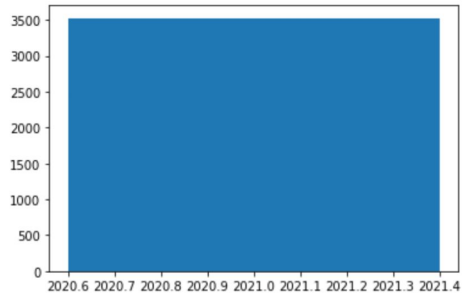
```
1  Complaint,CodedMonth,DateOccur,FlagCrime,FlagUnfounded,FlagAdminis
2  19-015802,2019-04,01/01/2003 12:00,Y, , ,1, ,115400,6,STLG BY DECE
3  19-013866,2019-04,01/01/2014 00:01,Y, , ,1, ,171200,5,SEX OFFNS-ST
4  19-015518,2019-04,01/01/2019 08:00,Y, , ,1, , 91124,3,"ASSAULT, CH
5  19-016861,2019-04,01/03/2019 13:30,Y, , ,1, ,263899,4,OBSTRUCT GOV
6  19-000824,2019-04,01/05/2019 23:00, , , ,-1, , 67601,4,"LARCENY-FR
7  19-015203,2019-04,01/07/2019 10:35,Y, , ,1, , 51322,3,BURGLARY-RES
8  19-015231,2019-04,01/10/2019 09:00,Y, , ,1, , 51212,5,BURGLARY-RES
9  19-001659,2019-04,01/10/2019 17:40,Y, , ,1, , 41016,4,ASSLT-AGGRAV
10 19-017901,2019-04,01/12/2018 13:01,Y, , ,1, , 51322,3,BURGLARY-RES
11 19-016631,2019-04,01/15/2013 10:56,Y, , ,1, ,115400,1,STLG BY DECE
12 19-019026,2019-04,01/15/2019 08:00,Y, , ,1, , 65701,5,LARCENY-MTR
13 19-014671,2019-04,01/15/2019 12:00,Y, , ,1, , 69702,3,LARCENY-ALL
14 19-002257,2019-04,01/15/2019 12:05, , , ,-1, ,265321,3,LEAVING SCE
```

Demo

Function 1

```
import pandas as pd
import matplotlib.pyplot as plt

year = 2019
month = 0
months = ['January', 'February', 'March', 'April', 'May', 'June', 'July', 'August', 'September', 'October', 'Nov
url_prefix = 'https://raw.githubusercontent.com/genericlastname/hackhpc-urban-renewal/master/'
crimes = []

while year <= 2020:
  while month < 12:
    crimes.append(pd.read_csv("{}{}{}.CSV".format(url_prefix, months[month], year), encoding="latin-1"))
    month += 1
  year += 1
plt.bar(year, len(crimes[0]))
```

```
<BarContainer object of 1 artists>
```

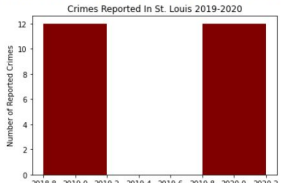



```
[7] import pandas as pd
```

Function 2

```
import matplotlib.pyplot as plt

year = [2019, 2020]
month = 0
months = ['January', 'February', 'March', 'April', 'May', 'June', 'July', 'August', 'September','October', 'November', 'December']
url_prefix = 'https://raw.githubusercontent.com/genericlastname/hackhpc-urban-renewal/master/'
tempCrimesarr1 = []
tempCrimesarr2 = []
crimesByYear = []

for x in year:
  for m in months:
    if x < 2020 and x >= 2019:
      tempCrimesarr1.append(pd.read_csv("{}{}{}.CSV".format(url_prefix, m, x),encoding="latin-1"))
    elif x == 2020:
      tempCrimesarr2.append(pd.read_csv("{}{}{}.CSV".format(url_prefix, m, x),encoding="latin-1"))

month = 0

crimesByYear = [tempCrimesarr1, tempCrimesarr2]
print("This is for 2019\n", crimesByYear[0])
print("This is for 2020\n", crimesByYear[1])
plt.bar(year, [len(crimesByYear[0]), len(crimesByYear[1])], color ='maroon', width = .4)
plt.xlabel("Year")
plt.ylabel("Number of Reported Crimes")
plt.title("Crimes Reported In St. Louis 2019-2020")
```

```
0     20-045219    2020-10     1/1/2020 0:01  ...     CANAAN  896321.1  1051430.0
1     20-045219    2020-10     1/1/2020 0:01  ...     CANAAN  896321.1  1051430.0
2     20-045219    2020-10     1/1/2020 0:01  ...     CANAAN  896321.1  1051430.0
3     20-045219    2020-10     1/1/2020 0:01  ...     CANAAN  896321.1  1051430.0
4     20-045518    2020-10    1/1/2020 12:00  ...        LEE  894389.3  1035481.0
...         ...        ...              ...  ...        ...       ...        ...
3688  20-049535    2020-10  10/31/2020 23:11  ...  WEST PINE  892547.4  1021595.0
3689  20-045341    2020-10  12/23/2019 16:00  ...        7TH  908113.4  1017396.0
3690  20-045889    2020-10  12/29/2019 14:17  ...     SEMPLE  886514.0  1034042.0
3691  20-045890    2020-10  12/29/2019 15:47  ...     SEMPLE  886514.0  1034042.0
3692  20-045892    2020-10  12/29/2019 17:13  ...     SEMPLE  886514.0  1034042.0

[3693 rows x 20 columns],       Complaint CodedMonth         DateOccur  ...   CADStreet    XCoord     Ycoord
0     20-032651    2020-11     1/1/1966 0:01  ...    BANCROFT  884351.0  1003449.0
1     20-033147    2020-11     1/1/2017 0:01  ...  WASHINGTON  889183.6  1025751.0
2     20-052101    2020-11     1/1/2019 0:01  ...    SULLIVAN  906581.0  1026620.0
3     20-052101    2020-11     1/1/2019 0:01  ...    SULLIVAN  906581.0  1026620.0
4     20-052101    2020-11     1/1/2019 0:01  ...    SULLIVAN  906581.0  1026620.0
...         ...        ...              ...  ...         ...       ...        ...
3838  20-038188    2020-11     12/5/2019 0:01  ...       KLEIN  906619.3  1028372.0
3839  19-061882    2020-11  12/11/2019 16:59  ...   JEFFERSON  899242.4  1009280.0
3840  20-053663    2020-11   12/17/2019 8:00  ...  WASHINGTON  904856.5  1019655.0
3841  20-054337    2020-11   12/20/2019 8:00  ...     CATALAN  886791.9   985739.5
3842  19-064562    2020-11  12/29/2019 12:41  ...    PLEASANT  902201.8  1030604.0

[3843 rows x 20 columns],       Complaint CodedMonth         DateOccur  ... CADStreet    XCoord     YCoord
0     21-000043    2020-12     1/1/2011 2:00  ...    WATSON       0.0        0.0
1     20-057309    2020-12     1/1/2019 8:00  ...    MARKET       0.0        0.0
2     20-054345    2020-12     1/1/2019 12:00  ...      FAIR  899784.3  1034793.0
3     20-055184    2020-12     1/1/2020 0:01  ...     GRAND       0.0        0.0
4     20-054871    2020-12     1/7/2020 20:00  ...       NaN  889532.9  1029862.0
...         ...        ...              ...  ...       ...       ...        ...
3314  20-058746    2020-12  12/31/2020 22:20  ...    SPRUCE       0.0        0.0
3315  20-058749    2020-12  12/31/2020 22:50  ...   BENNETT       0.0        0.0
3316  21-000004    2020-12  12/31/2020 23:00  ...       NaN       0.0        0.0
3317  21-000001    2020-12  12/31/2020 23:15  ...       NaN       0.0        0.0
3318  21-000010    2020-12  12/31/2020 23:50  ...     GRAND       0.0        0.0

[3319 rows x 20 columns]]
Text(0.5, 1.0, 'Crimes Reported In St. Louis 2019-2020')
```



Function 3 and 4

Can we correlate the black and white home ownership with

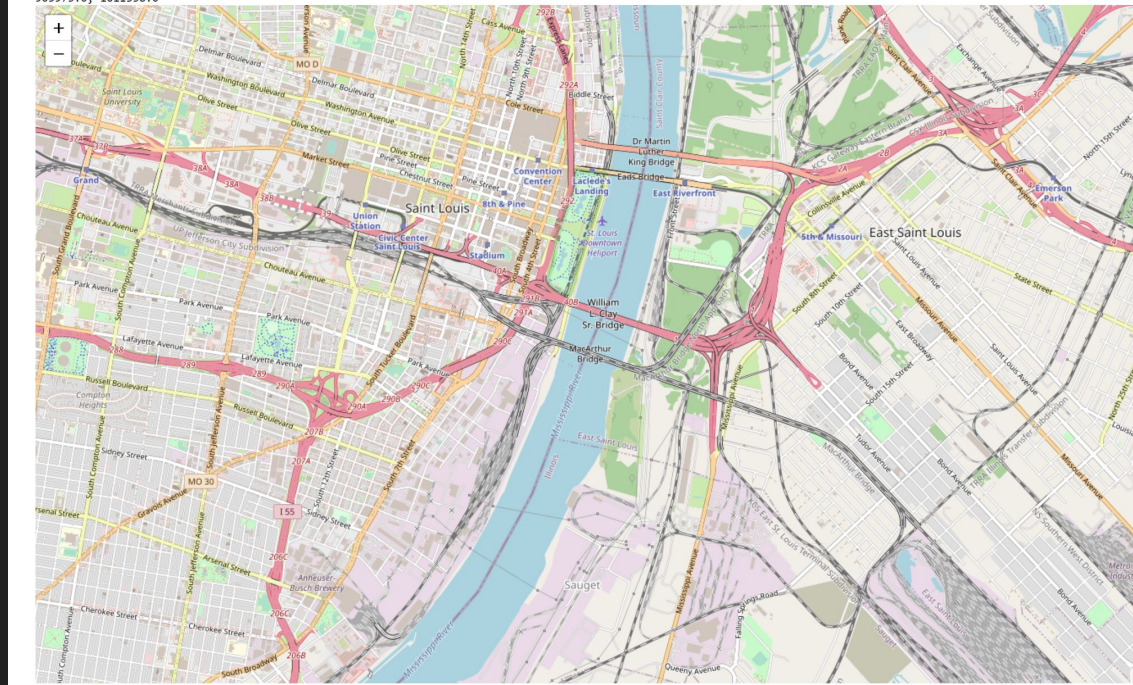
```
import pandas as pd

data = pd.read_csv('https://pastebin.com/raw/G5A17VVV')

print(data)
```

```
   regionid  ... Implied Housing Gap
0    395121  ...           -0.707739

[1 rows x 11 columns]
```

```
import folium
import pandas as pd

stlMap = folium.Map(
    location=[38.627003, -90.199402],
    zoom_start=12,
    max_zoom=18)

months = ['February', 'March', 'April', 'May', 'June', 'July', 'August', 'September','October', 'November', 'December']
url_prefix = 'https://raw.githubusercontent.com/genericlastname/hackhpc-urban-renewal/master/'
crimes_2020 = pd.read_csv(url_prefix + "January2020.CSV", encoding='latin-1')

for month in months:
  curr = pd.read_csv(url_prefix + month + "2020.CSV", encoding="latin-1", skiprows=1)
  crimes_2020.append(curr)

for index, row in crimes_2020.iterrows():
  print('{}, {}'.format(row['XCoord'], row['YCoord']))
  folium.Marker([row['XCoord'], row['YCoord']],
                icon=folium.Icon(color='red')
                ).add_to(stlMap)


stlMap
```

Map From Function 4

```
889128.1, 1032112.0
903668.9, 1010436.0
895104.3, 1002015.0
907524.7, 1017960.0
889757.8, 996082.2
882832.8, 1035589.0
908031.7, 1020149.0
905973.6, 1011338.0
```



Questions?