COLLEGE TOWN CRIME

DS 2002 Final Project

JP Meyer, Lily Rademacher, Owen Himmel, Owen Shafer, Cara Seely

DATA SELECTION

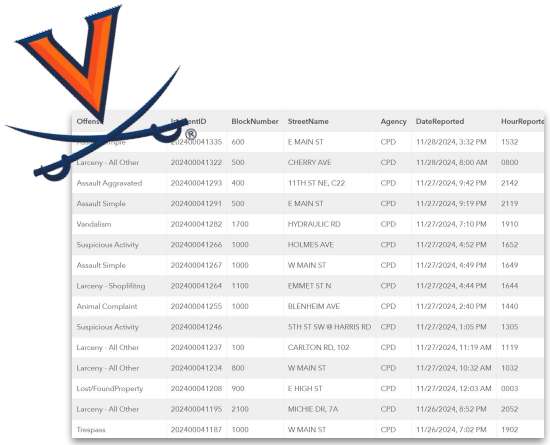
We have chosen to explore datasets representing crime data from Charlottesville, VA and Ann Arbor, MI

RATIONALE

- Compare relative safety of prominent college towns
- Safety is a vital choice for school selection
- Hope to find insights on incidence in crime and types of crime students may face

TAKEAWAYS

- Alignment between research objectives and datasets is crucial
- Important to analyze underlying raw data and data sources
- Datasets rarely tell the full story



Offense	IncidentID	BlockNumber	StreetName	Agency	DateReported	HourReported
Theft From Motor Vehicle	202400041335	600	E MAIN ST	CPD	11/28/2024, 3:32 PM	1532
Larceny - All Other	202400041322	500	CHERRY AVE	CPD	11/28/2024, 8:00 AM	0800
Assault Aggravated	202400041293	400	11TH ST NE, C22	CPD	11/27/2024, 9:42 PM	2142
Assault Simple	202400041291	500	E MAIN ST	CPD	11/27/2024, 9:19 PM	2119
Vandalism	202400041282	1700	HYDRAULIC RD	CPD	11/27/2024, 7:10 PM	1910
Suspicious Activity	202400041266	1000	HOLMES AVE	CPD	11/27/2024, 4:52 PM	1652
Assault Simple	202400041267	1000	W MAIN ST	CPD	11/27/2024, 4:49 PM	1649
Larceny - Shoplifting	202400041264	1100	EMMET ST N	CPD	11/27/2024, 4:44 PM	1644
Animal Complaint	202400041255	1000	BLENHEIM AVE	CPD	11/27/2024, 2:40 PM	1440
Suspicious Activity	202400041246		5TH ST SW @ HARRIS RD	CPD	11/27/2024, 1:05 PM	1305
Larceny - All Other	202400041237	100	CARLTON RD, 102	CPD	11/27/2024, 11:19 AM	1119
Larceny - All Other	202400041234	800	W MAIN ST	CPD	11/27/2024, 10:32 AM	1032
Lost/FoundProperty	202400041208	900	E HIGH ST	CPD	11/27/2024, 12:03 AM	0003
Larceny - All Other	202400041195	2100	MICHIE DR, 7A	CPD	11/26/2024, 8:52 PM	2052
Trespass	202400041187	1000	W MAIN ST	CPD	11/26/2024, 7:02 PM	1902

VS

CASE_NUM	CRIME_DESC	CHARGE_CODE	CHARGE_DESC
230001504	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL
230001517	DESTRUCTION/DAMAGE/VANDALISM OF PROPERTY		2902 DAMAGE TO PROPERTY - PRIVAT
230001583	FALSE PRETENSE/SWINDLE/CONFIDENCE GAME		2699 FRAUD (OTHER)
230001585	THEFT FROM BUILDING		2308 LARCENY - FROM BUILDING (INCLUDES
230001456	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL DOMESTIC AND
230001425	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL DOMESTIC AND POLICE OFFICER
230001520	FALSE PRETENSE/SWINDLE/CONFIDENCE GAME		2699 FRAUD (OTHER)
230001543	FALSE PRETENSE/SWINDLE/CONFIDENCE GAME		2699 FRAUD (OTHER)
230001577	MOTOR VEHICLE THEFT		2404 VEHICLE THEFT UDAA (REPORTED BY YOUR JURISDICTION)
230001547	SHOPLIFTING		3078 RETAIL FRAUD, THEFT 3RD DEGREE
230001611	FAMILY OFFENSE NONVIOLENT		3806 NEGLECT CHILD
230001611	ALL OTHER		5013 CONDITIONAL RELEASE VIOLATION
230001541	THEFT FROM MOTOR VEHICLE		2305 LARCENY - PERSONAL PROPERTY FROM VEHICLE - LFA
230001541	DESTRUCTION/DAMAGE/VANDALISM OF PROPERTY		2996 DAMAGE PROPERTY - MDOP- THROWING STONE, ETC, AT TRAIN OR MOTI
230001509	ALL OTHER LARCENY		2399 LARCENY (OTHER)
230001546	ALL OTHER LARCENY		2399 LARCENY (OTHER)
230001818	FALSE PRETENSE/SWINDLE/CONFIDENCE GAME		2602 FRAUD - SWINDLE
230001687	ALL OTHER LARCENY		2399 LARCENY (OTHER)
230001687	IDENTITY THEFT FRAUD		2609 IDENTITY THEFT
230001717	IDENTITY THEFT FRAUD		2609 IDENTITY THEFT
230001790	WEAPON LAW VIOLATIONS		5289 WEAPONS CONCEALED (OTHER)
230001692	SHOPLIFTING		3074 RETAIL FRAUD THEFT 2ND DEGREE
230001660	THEFT FROM BUILDING		2308 LARCENY - FROM BUILDING (INCLUDES LIBRARY, OFFICE USED BY PUBLI
230001876	SHOPLIFTING		3073 RETAIL FRAUD THEFT 1ST DEGREE
230001889	DISORDERLY CONDUCT		5311 DISORDERLY CONDUCT
230001948	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL DOMESTIC AND POLICE OFFICER
230001676	THEFT OF MOTOR VEHICLE PARTS/ACCESSORIES		2304 LARCENY - PARTS & ACCESSORIES FROM VEHICLE - LFA
230001937	DISORDERLY CONDUCT		5311 DISORDERLY CONDUCT
230001875	SHOPLIFTING		3073 RETAIL FRAUD THEFT 1ST DEGREE
230001856	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL DOMESTIC AND POLICE OFFICER
230002009	SIMPLE ASSAULT		1313 ASSAULT/ BATTERY/SIMPLE (INCL DOMESTIC AND POLICE OFFICER

ETL Pipeline

Extraction

RecordID	Offense	IncidentID	BlockNumb	StreetNam	Agency	DateReport	HourRepor
1	Assist Citizen - Menta	2.02E+11	900	ELLIOTT AV	CPD	2024/10/0	2024
2	Larceny - Of Veh Parts	2.02E+11	1600	6TH ST SE,	CPD	2024/10/0	1738
3	Hit and Run	2.02E+11	900	EMMET ST I	CPD	2024/10/0	1711
4	Disorderly Conduct	2.02E+11	900	ST CLAIR A	CPD	2024/10/0	1659
5	Lost/FoundProperty	2.02E+11	700	RUGBY RD	CPD	2024/10/0	1423
6	Crisis Assessment	2.02E+11	1300	LONG ST	CPD	2024/10/0	1100
7	Trespass	2.02E+11	500	PRESTON A	CPD	2024/10/0	1007
8	Assault Simple	2.02E+11	900	ROCK CREE	CPD	2024/10/0	1002
9	Hit and Run	2.02E+11		CHERRY AV	CPD	2024/10/0	2116
10	Assist Citizen - Menta	2.02E+11	200	CARLTON R	CPD	2024/10/0	2100
11	Trespass	2.02E+11	1200	W MAIN ST	CPD	2024/09/3	1835
12	Larceny - Shoplifitng	2.02E+11	200	E MAIN ST	CPD	2024/09/3	1739
13	Assist Citizen - Menta	2.02E+11	200	DOUGLAS A	CPD	2024/09/3	1635
14	Larceny - Of Veh Parts	2.02E+11	500	RUGBY RD	CPD	2024/09/3	1547
15	Larceny - Shoplifitng	2.02E+11	1200	EMMET ST I	CPD	2024/09/3	1546
16	Larceny - Shoplifitng	2.02E+11	1100	5TH ST SW	CPD	2024/09/3	1533
17	Assist Citizen - Menta	2.02E+11	200	RIDGE ST	CPD	2024/09/3	1358
18	Larceny - Theft from B	2.02E+11	400	17TH ST NV	CPD	2024/09/3	1340
19	Hit and Run	2.02E+11	100	MINOR COI	CPD	2024/09/3	1317
20	Animal Complaint	2.02E+11	800	MONTROSE	CPD	2024/09/3	1112
21	Vandalism	2.02E+11	300	E HIGH ST	CPD	2024/09/3	1040
22	Vandalism	2.02E+11	1900	EMMET ST I	CPD	2024/09/3	819
23	Larceny - From Motor	2.02E+11	1200	KING ST	CPD	2024/09/3	733
24	Assist Citizen - Menta	2.02E+11	600	E MARKET S	CPD	2024/09/3	443
25	Domestic Disturbanc	2.02E+11	700	MADISON A	CPD	2024/09/3	10
26	Assist Agency - Backu	2.02E+11	800	HARDY DR,	CPD	2024/09/3	2047
27	Disorderly Conduct	2.02E+11	800	FRANKLIN S	CPD	2024/09/3	2032

Steps:
Download data from local government.

pd.read_csv...

Transformation

	RecordID	Offense	IncidentID	BlockNumber	StreetName	Agency	DateReported	HourReported
0	1	Assist Citizen - Mental/TDO/ECO	202400034093	900.0	ELLIOTT AVE, A	CPD	2024/10/02 00:24:02+00	2024
1	2	Larceny - Of Veh Parts/Access	202400034075	1600.0	6TH ST SE, 28	CPD	2024/10/01 21:38:47+00	1738
2	3	Hit and Run	202400034071	900.0	EMMET ST N	CPD	2024/10/01 21:11:45+00	1711
3	4	Disorderly Conduct	202400034068	900.0	ST CLAIR AVE, A	CPD	2024/10/01 20:59:17+00	1659
4	5	Lost/FoundProperty	202400034047	700.0	RUGBY RD	CPD	2024/10/01 18:23:15+00	1423
5	6	Crisis Assessment	202400034030	1300.0	LONG ST	CPD	2024/10/01 15:00:33+00	1100
6	7	Trespass	202400034026	500.0	PRESTON AVE	CPD	2024/10/01 14:07:55+00	1007
7	8	Assault Simple	202400034025	900.0	ROCK CREEK RD	CPD	2024/10/01 14:02:57+00	1002
8	9	Hit and Run	202400033979	NaN	CHERRY AVE / 5TH ST SW	CPD	2024/10/01 01:16:59+00	2116
9	10	Assist Citizen - Mental/TDO/ECO	202400033977	200.0	CARLTON RD	CPD	2024/10/01 01:00:58+00	2100

Steps:
Manipulate data frame for desired analysis

Ex: pd.to_datetime...

Loading

Steps:
Connect to MongoDB, create Crime_2023 db and collections for each city

Pymongo

```
client = MongoClient(uri,
server_api=ServerApi('1'))
```

Cloud Storage

Google Cloud

- Used google.cloud storage library to interact with Google Cloud storage.
- Created an uploading function for quicker storage of new files:
- `upload_to_google(bucket_name, source_file_name, destination_blob_name)`

Advantages and Uses

- MongoDB allows for flexibility in data needed when sourcing for different locations.
- Takes advantage other Google Cloud Services

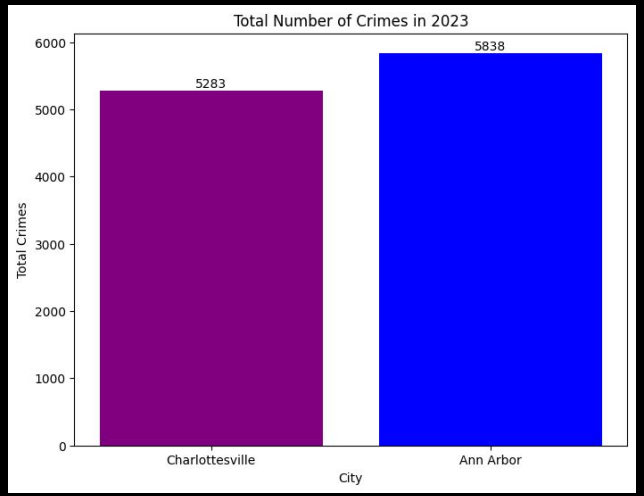
Credentials Management and Access Control:

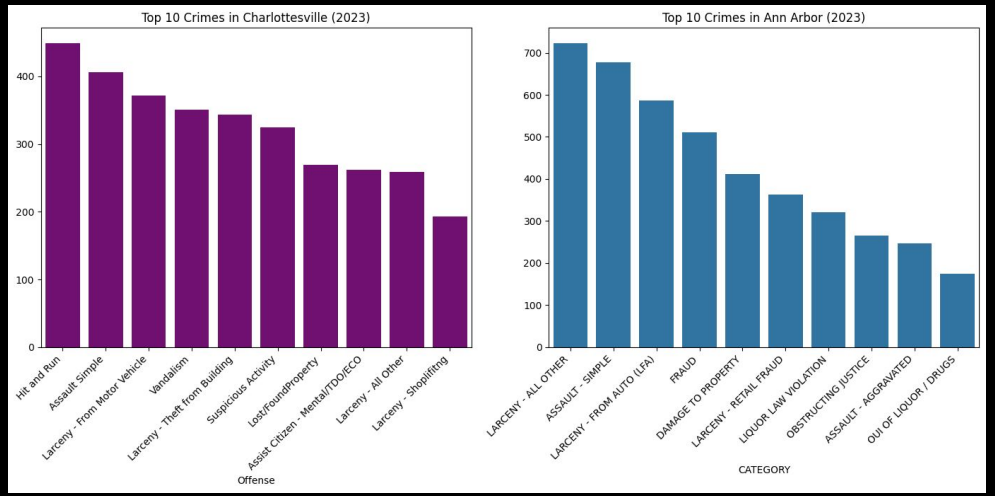
- Access is controlled by possession of JSON file associated with Cloud account
- The key/URI file is stored locally and accessed through an environmental variable so it is not exposed when uploaded.

```
with open("mongoaccess.json") as f:
    config = json.load(f)
    uri = config["MONGO_URI"]

# Create a new client and connect to the server
client = MongoClient(uri, server_api=ServerApi('1'))
```

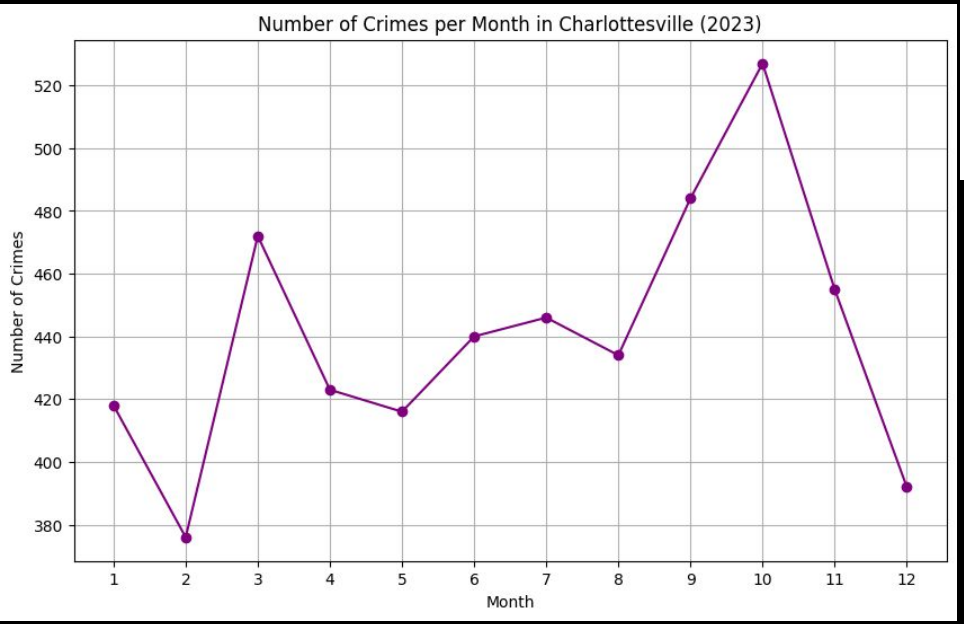
Visualizations & Stats: Total Crime

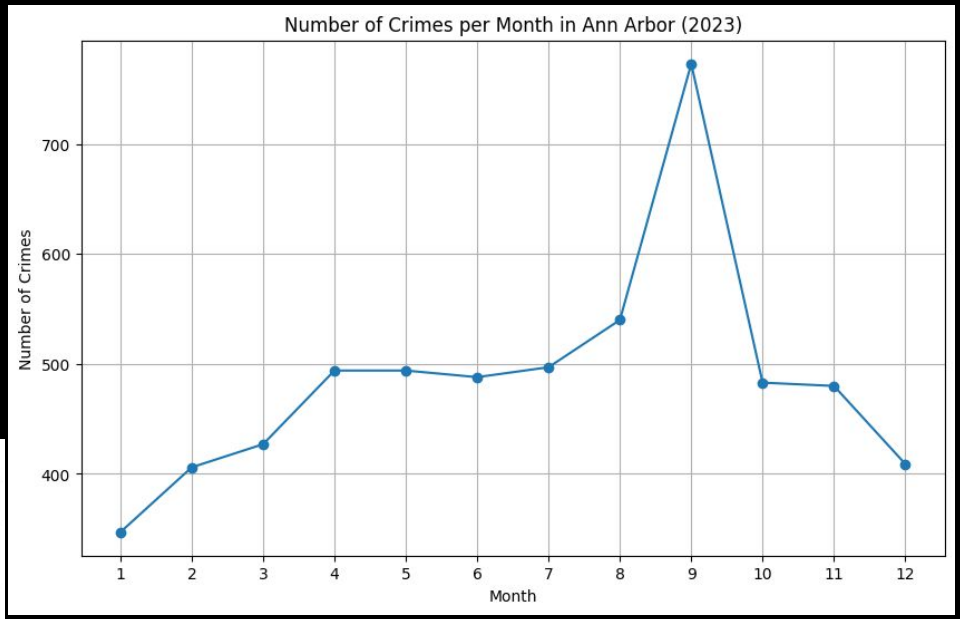




- In 2023, there were 555 more reports of crime in Ann Arbor than Charlottesville
- Ann Arbor's crime is approximately 10.5% higher than Charlottesville's in terms of reported incidents, but Ann Arbor is also more than double the size of Charlottesville
    - Scaling by population, Ann Arbor has fewer crimes per capita (1 crime per 20 people) than Charlottesville (1 crime per 9 people)
- The most predominant crimes were consistent across Charlottesville and Ann Arbor (assault and larceny)

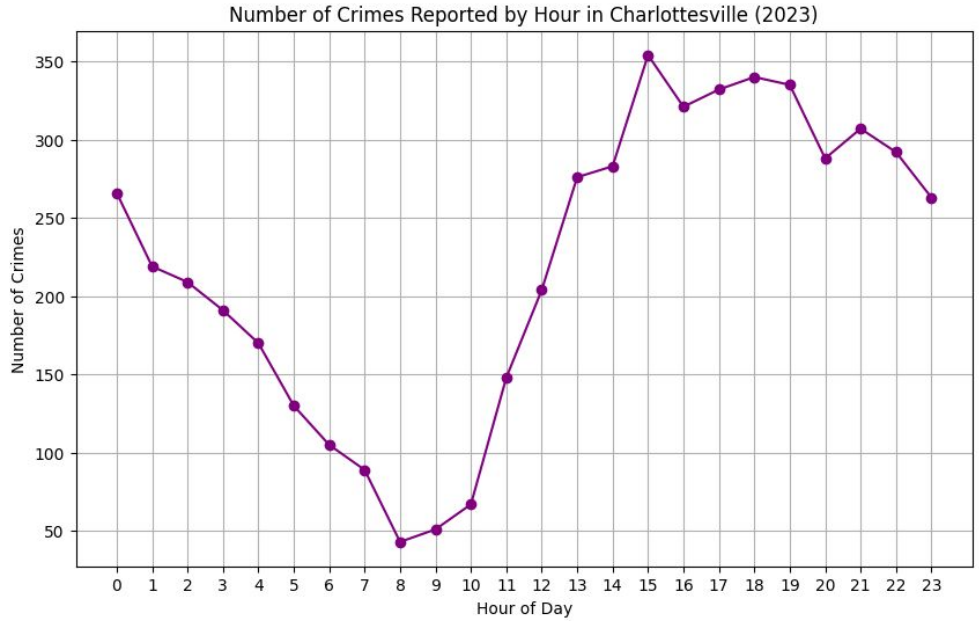
Visualizations & Stats: Monthly Frequencies

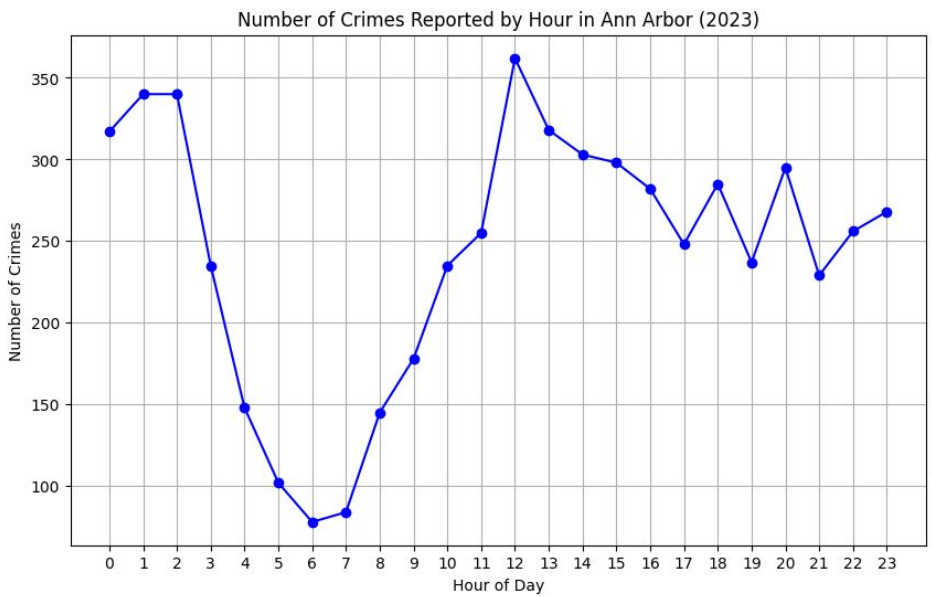




Both Charlottesville and Ann Arbor experience their greatest spikes in crime towards the end of the year (Charlottesville in October and Ann Harbor in September. In Charlottesville, However, crime peaks in March and July as well. Ann Arbor's crime steadily increases between January and April, followed by relatively steady crime rates through the month of July and August.

Visualizations & Stats: Hourly Frequencies





These two towns show similarities in the number of total crimes reported by time of day in 2023. Crime peaks early afternoon (Charlottesville around 3pm and Ann Arbor around noon) with lulls in the early morning. Most crime is reported in the early evening and night.

Visualizations & Stats: Monthly Frequencies

Crime by whether or not students are present

	Average number of assault crimes during months with no students present	Average number of assault crimes during months *with* students present
Charlottesville	43	49
Ann Arbor	78	76

In Charlottesville, assault crimes are most likely to occur when UVA students are mostly present. In Ann Arbor, assault crimes are less likely to occur when Michigan students are mostly present. For the purpose of this data manipulation, students are assumed to not be present during January, May, June, July and August.

Visualizations & Stats: Heat Maps

BlockNumber	StreetName
800.0	RIDGE ST

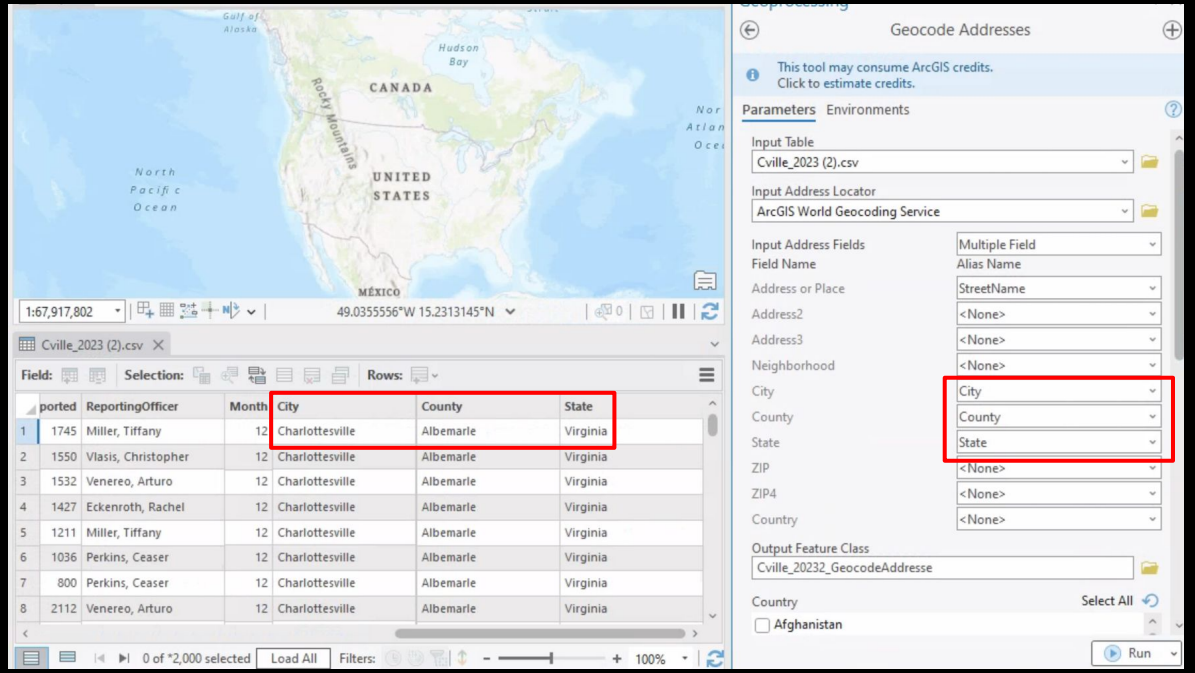
VS

POINT_X	POINT_Y
-83.749555	42.282469

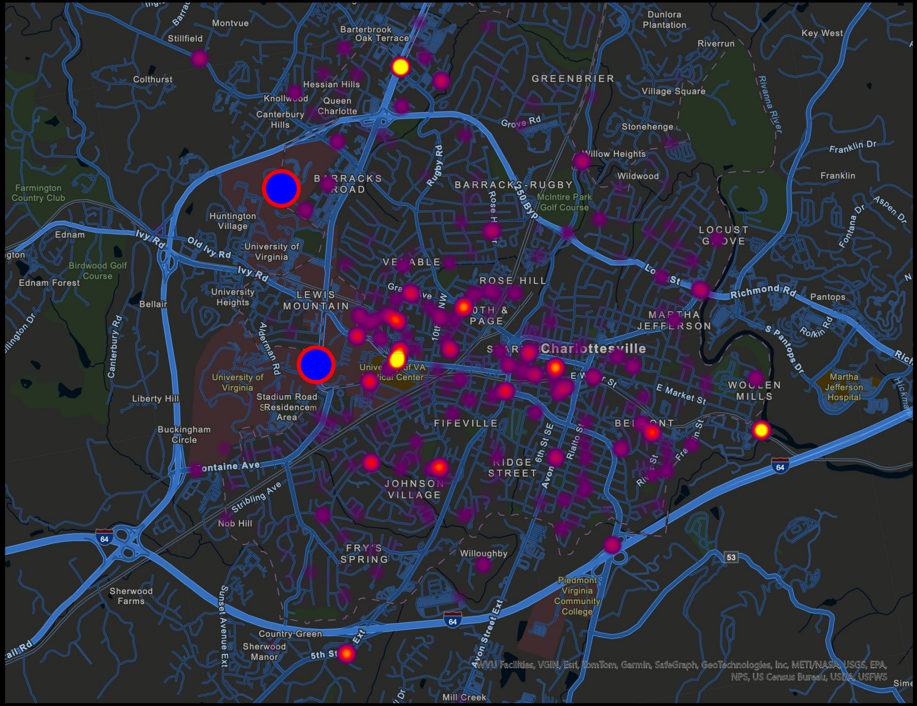
```
#add city column to cville_2023 for increased accuracy in Arc GIS Geocoding
Cville_2023['City'] = 'Charlottesville'
Cville_2023['County'] = 'Albemarle'
Cville_2023['State'] = 'Virginia'
```

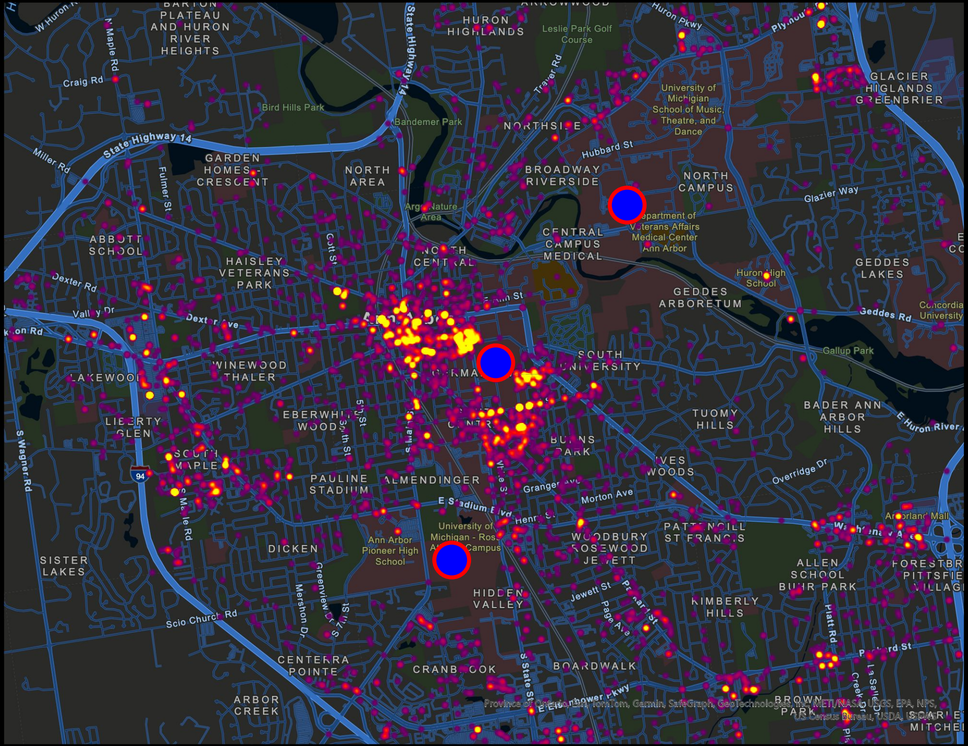
Once the tabular data was processed and saved as a .csv file, it can be exported into software like ArcGIS Pro for geospatial analysis. The data for Charlottesville and Ann Arbor was spatially different (coordinates vs addresses), so we adjusted our approach for each data set.

- XYTabletoPoint function for Ann Arbor
- Geocode Address function for Charlottesville data
  - Added city, county and State columns for accuracy



Visualizations & Stats: Heat Maps





In both Charlottesville and Ann Arbor, the town's reported crime is most intense in close proximity to the campuses of UVA and U-M, respectively. Both town's universities represent a relatively central location on the maps where crime is concentrated. However, we do see some decrease in crime reported on and immediately surrounding each campus (red areas with blue dots). Crime reports are less frequent and more spread out with greater distance from the center.

Key Findings/Insights

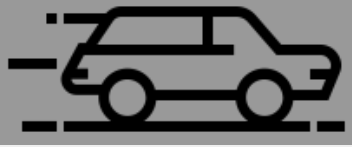
Which is a more safe college town to live in: Charlottesville or Ann Arbor? → Well, it depends...

Total Crime

More crime was reported in 2023 in Ann Arbor (5838 reports) than in Charlottesville (5283). Assault crimes are reported more on average in Ann Arbor. Charlottesville had more crime per capita in 2023 than Ann Arbor.

Crime Type

Charlottesville's most frequently reported crime in 2023 was hit and runs, whereas Ann Arbor's were larceny crimes.



Frequency

Both towns greatest peaks in crime occurred towards the end of the year. In Charlottesville, assault crimes during months where students are present are more likely to occur, whereas in Ann Arbor, these crimes are less likely to occur. In both towns, 2023 crime reports occurred less in the early morning and more in the afternoons and night.

Crime Location

The majority of reported crimes in 2023 occurred surrounding the town's university.

THANK YOU