Lab 5

Amrutha Rao and Madison Goodwin

Research question: How do the types and severity levels of reported crimes in Cincinnati vary by neighborhood, and are there areas with higher rates of violent crimes compared to non-violent crimes?

Data Source 1: [PDI\_\_Police\_Data\_Initiative\_\_CPD\_Shootings.csv](file:///C:\DSC%20200\PDI__Police_Data_Initiative__CPD_Shootings.csv)

The database shows confirmed shootings in Cincinnati. The data is sourced from data.cincinatti-oh.gov and the source of the data is the Cincinnati Police Department. The data was created through the City’s computer-aided dispatch (CAD) system and is stored in the City’s Records Management System (RMS). This continually evolving dataset currently accounts for 6500 confirmed shooting incidents, but the data is updated daily. The data is provided by the city of Cincinnati, and I contacted the dataset owner to confirm its reliability. The database is stored as a CSV file as CSV files are efficient for storing and transferring large datasets. CSV files are also easy to work with as their simple structure is easily understandable. CSV files also maintain data integrity as there is less risk of data corruption. CSV files are text-based and therefore more suited for GitHub. This makes it easier for the group to work on the assignment together. This dataset passes the “smell test”. The author is a veritable source to be contacted for answers to questions. As previously mentioned, the data is regularly updated and checked for errors. The data comes with information as to how it was acquired. Given my overall knowledge of the topic, the data does seem plausible. I’ve used this website in the past so I’m familiar with its credibility. The data set can easily be opened in a simple document viewer, we used Microsoft Excel, without crashing the computer. The data appears to be properly labeled and organized, so you can easily pluck out each piece of information. The data is easily stored and moved if you need to use more than one laptop or computer while working on it. The data is real-time and accessible via an API, meaning you can get the data you need by requesting it live.

Data Source 2: [PDI\_\_Police\_Data\_Initiative\_\_Crime\_Incidents.csv](file:///C:\DSC%20200\PDI__Police_Data_Initiative__Crime_Incidents.csv)

The database shows reported crime incidents in Cincinnati. The data is sourced from data.cincinatti-oh.gov and the source of the data is the Cincinnati Police Department. The data was created through the City’s computer-aided dispatch (CAD) system and is stored in the City’s Records Management System (RMS). The crime incidents are the records of reported crimes, collated by an agency for management. This continually evolving dataset currently accounts for 515270 reported crime incidents, but the data is updated daily. The data is provided by the city of Cincinnati, and I contacted the dataset owner to confirm its reliability. The database is stored as a CSV file as CSV files are efficient for storing and transferring large datasets. CSV files are also easy to work with as their simple structure is easily understandable. CSV files also maintain data integrity as there is less risk of data corruption. CSV files are text-based and therefore more suited for GitHub. This makes it easier for the group to work on the assignment together. This dataset passes the “smell test”. The author is a veritable source to be contacted for answers to questions. As previously mentioned, the data is regularly updated and checked for errors. The data comes with information as to how it was acquired. Given my overall knowledge of the topic, the data does seem plausible. I’ve used this website in the past so I’m familiar with its credibility. The data set can easily be opened in a simple document viewer, we used Microsoft Excel, without crashing the computer. The data appears to be properly labeled and organized, so you can easily pluck out each piece of information. The data is easily stored and moved if you need to use more than one laptop or computer while working on it. The data is real-time and accessible via an API, meaning you can get the data you need by requesting it live.

Data Source 3: [PDI\_\_Police\_Data\_Initiative\_\_Police\_Calls\_for\_Service\_\_CAD\_.csv](file:///C:\DSC%20200\PDI__Police_Data_Initiative__Police_Calls_for_Service__CAD_.csv)

The database shows all Cincinnati Police Department Calls for Service in Cincinnati. The data is sourced from data.cincinatti-oh.gov and the source of the data is the Cincinnati Police Department. The data was created through the City’s computer-aided dispatch (CAD) system and is stored in the City’s Records Management System (RMS). This continually evolving dataset currently accounts for 1048575 calls for service, but the data is updated daily. The data is provided by the city of Cincinnati, and I contacted the dataset owner to confirm its reliability. The database is stored as a CSV file as CSV files are efficient for storing and transferring large datasets. CSV files are also easy to work with as their simple structure is easily understandable. CSV files also maintain data integrity as there is less risk of data corruption. CSV files are text-based and therefore more suited for GitHub. This makes it easier for the group to work on the assignment together. This dataset passes the “smell test”. The author is a veritable source to be contacted for answers to questions. As previously mentioned, the data is regularly updated and checked for errors. The data comes with information as to how it was acquired. Given my overall knowledge of the topic, the data does seem plausible. I’ve used this website in the past so I’m familiar with its credibility. The data set can easily be opened in a simple document viewer, we used Microsoft Excel, without crashing the computer. The data appears to be properly labeled and organized, so you can easily pluck out each piece of information. The data is easily stored and moved if you need to use more than one laptop or computer while working on it. The data is real-time and accessible via an API, meaning you can get the data you need by requesting it live.

ASK: For Part 2 of Lab 5, we used the same website and csv files for all the data sets, so is it okay for each paragraph to look the same other than a few specific details.