Nicholas Fry
November 24, 2020
Coursera Applied Data Science Capstone

Targeting Family Counseling Services Across Kansas City

1 Introduction

1.1 Background

As the United States confronts an ongoing coronavirus, more people are adapting to new social habitation standards. With families in closer proximity during isolation, quarantine, job lay-off, there has been a rise in domestic violence (Maclen, 2020). Kansas City is no exception to the rise in domestic violence seen across the world (Kraske, 2020). As the pandemic persists more families are falling into cycles of domestic violence, often with no other choice than to lockdown with their abusers.

1.2 Problem

This rise in domestic violence requires the attention of family counseling services, battered females' shelters, and other organizations that can intervene, prevent, and deescalate tensions between family members. The goal of this data analysis project is to identify where the abuse takes place and bring attention to the zip codes that most need the services these organizations provide. To examine the need for services, each zip code across Kansas City will undergo a domestic violence increase analysis, k-means clustering to bin the neighborhoods, and identification of the clusters with the highest need.

2 Developing Data and Cleaning

2.1 Data Sources

Kansas City maintains an open data portal. This open data portal contains a repository of crime data (Mesa, 2020). An csv is available for download. The crime data table contains a column that indicates if the offense was a domestic violence incident. It also contains information on the area where the crime took place. Isolating only the domestic violence incidents with complete zip code information there are 10,141 occurrences in 2020, for Kansas City – shown in Table 1.

1 The head() of the long table of domestic violence occurrences in Kansas City for 2020. DVFlag 'Y' indicates that, "yes" the crime was domestic violence.

	Reported_Date	Zip Code	DVFlag
0	2/5/2020	66206	Y
1	2/5/2020	66206	Y
2	5/18/2020	66110	Y
3	5/18/2020	66110	Y
4	5/18/2020	66106	Y

This table, however, lacks latitude and longitude for the domestic violence occurrences. Location data for all of Kansas City's zip codes is available in CSV format (*US Zip Code Latitude and Longitude*, n.d.). The joined tables appear in Table 2.

2 DVFlag here indicates a domestic violence incident in Kansas City, MO. This table represents the first 5 rows in the 2020 Kansas City domestic violence crimes list, made in the data preparation phase.

	Unnamed: 0	Reported_Date	Zip Code	DVFlag	Latitude	Longitude
0	2089	1/12/2020	64153	Y	39.281602	-94.71439
1	3309	2/24/2020	64153	Y	39.281602	-94.71439
2	3498	3/3/2020	64153	Y	39.281602	-94.71439
3	3930	1/31/2020	64153	Y	39.281602	-94.71439
4	4527	1/24/2020	64153	Y	39.281602	-94.71439

2.2   Data Cleaning

The original extent of the 2020 domestic violence crimes spreadsheet for Kansas City was over 10,100 rows long. This also contained incorrect zip codes, NaN values, among others. I was able to prepare the spreadsheet in Python by dropping NaN values. When it came to connecting with a geocoding API, none of my calls were returned. I reduced the scope from individual addresses to zip codes. There were 47 unique zip codes that I could connect a latitude and longitude to, however, no calls were returned from the Google API or ArcGIS. I opted instead to pull another csv file from OpenDataSoft.com (*US Zip Code Latitude and Longitude*, n.d.). The result was Table 2. This contains 10,098 rows of domestic violence crimes with errant zip codes removed.

2.3   Data Summary

This data is prepared for analysis in part 2 of this capstone project. In the next phase I intend to determine the areas with higher levels of domestic violence taking place. This information could be related to other externalities, such as the topology or composition of the neighborhood and the services found within it.

3   Methodology

3.1   Exploratory Data Analysis

I took the table from **Error! Reference source not found.** and created a list count of domestic violence per zip code. This list was merged to a data frame of the 47 zip codes across the Kansas City area. The first 5 rows appear in Table 3.

3 Counts per zip code join to the data frame from Figure 2.

	Zip	City	State	Latitude	Longitude	DVCount
0	64101	Kansas City	MO	39.103037	-94.60066	5
1	64102	Kansas City	MO	39.093854	-94.60411	10
2	64105	Kansas City	MO	39.104204	-94.58868	47
3	64106	Kansas City	MO	39.103704	-94.57311	203
4	64108	Kansas City	MO	39.084554	-94.58296	192

I explored the domestic counts for each of the zip codes. The average count of domestic violence per zip code was 215. The maximum count per zip code was 1235 incidents, indicating a clear discrepancy of domestic violence incidents by location. Numeric descriptions appear in Table 4.

4 Numeric descriptions of the domestic violence counts across all zip codes.

	DVCount
count	47.000000
mean	214.851064
std	254.549472
min	2.000000
25%	32.000000
50%	152.000000
75%	288.000000
max	1235.000000

After joining this information and sorting the statistics, I sought a method of targeting the zip codes that report the highest incidents of domestic violence. I sorted the data by highest to lowest.

After sorting the highest offending zip codes, I ran a machine learning k-means clustering code. That produces 8 clusters from the 47 original zip codes. These clusters were placed on a folium map. I refined the clusters to the highest offending areas. After refinement I again placed those clusters on a map. Finding the geographic areas to target for domestic violence offenses, I explored the existing women's shelters. Exploring the location of existing women's shelters helped me determine where the concentration of needs versus the concentration of services was.

4   Results

4.1   The top ten and bottom ten zip codes for domestic violence in Kansas City

5 The top ten and bottom ten zip codes for domestic violence offenses in 2020. The total count of incidents for the year-to-date appears in the second column.

Zip	DVCount	Zip	DVCount
64130	1235	64161	24
64127	817	64147	21
64134	774	64113	11
64132	694	64102	10
64131	568	64139	9
64128	484	64146	7
64109	372	64101	5
64111	341	64149	4
64124	317	64121	2
64110	297	64163	2

4.2   A bar graph of the highest to lowest domestic violence offenses in Kansas City
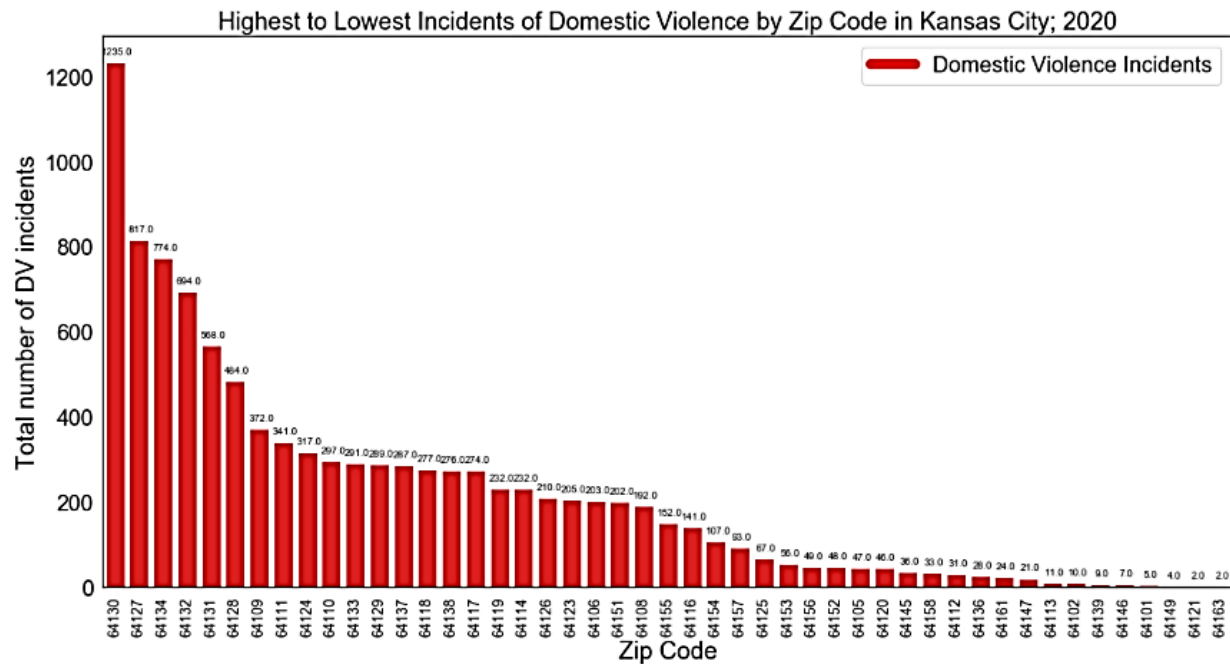

Figure 1 A bar graph of all domestic violence incidents through mid-November in 2020, Kansas City metro area. The bars are sorted by zip code.

4.3 K-means Clustering and Mapping

The first 5 rows of k-means clustering results appear in Table 6.

6 After k-means clustering, each row is assigned a cluster value of 0-7.

	Zip	Latitude	Longitude	DVCount	Cluster
0	64101	39.103037	-94.60066	5	0
1	64102	39.093854	-94.60411	10	0
2	64105	39.104204	-94.58868	47	0
3	64106	39.103704	-94.57311	203	3
4	64108	39.084554	-94.58296	192	3

The k-means cluster results are projected to a folium map to explore domestic violence clusters geographically (Figure 2).
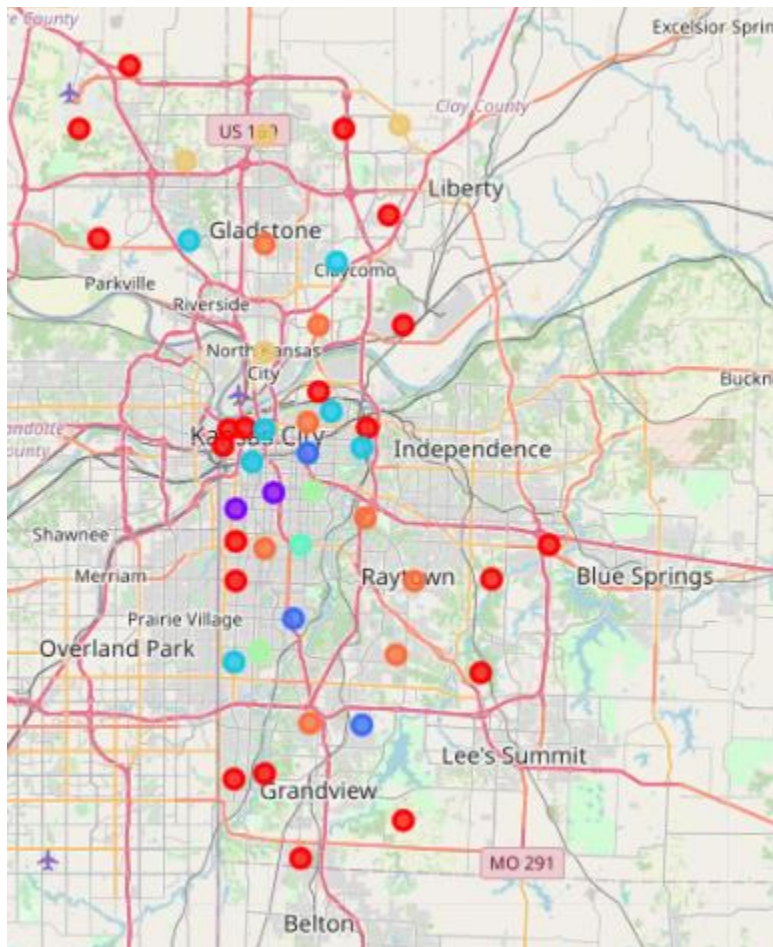

Figure 2 K-means clustering outcomes on a map. The zip codes were sorted into 8 clusters.

The outcomes of the k-means clustering largely follows the magnitude of domestic violence counts by area. The top three clusters are isolated and projected to a map in Figure 3. These high-incident areas largely center around Raytown.
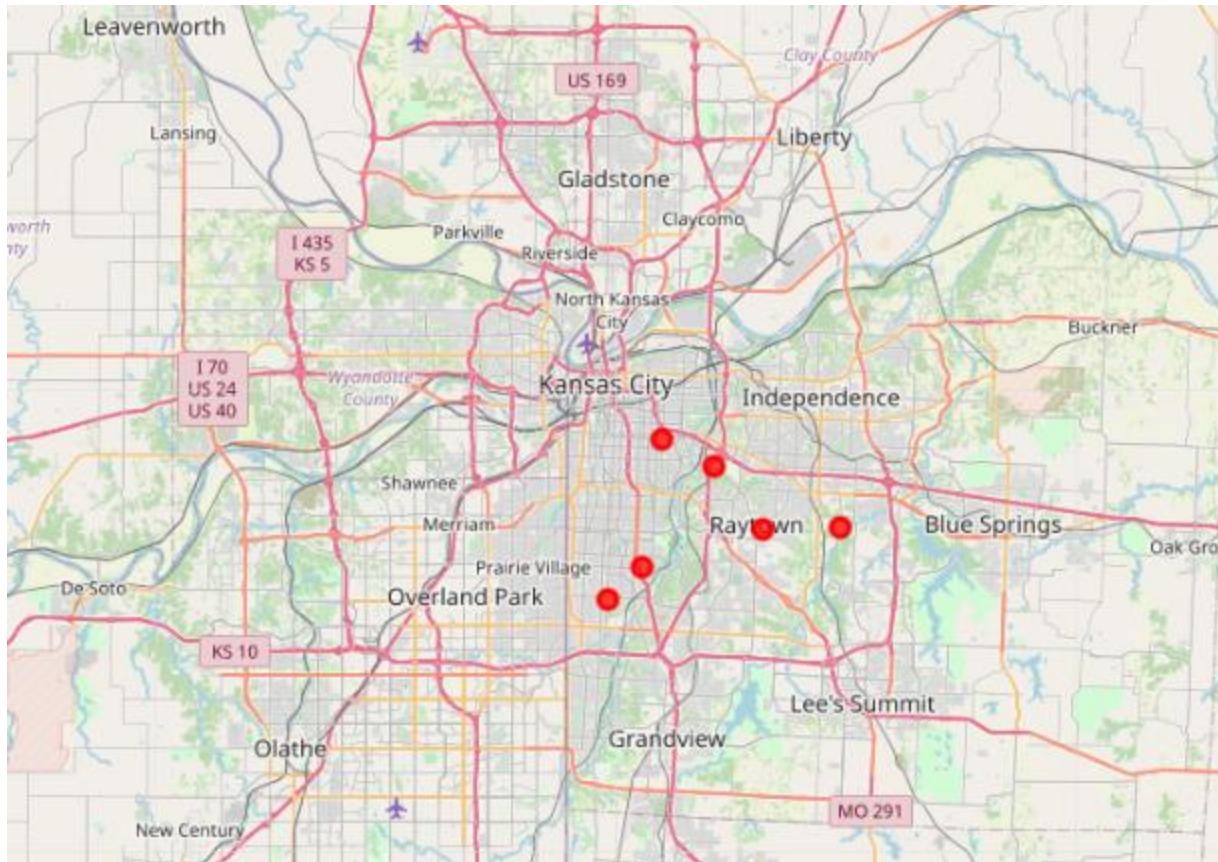
Figure 3 Top three clusters of domestic violence incidents in 2020, Kansas City metro area.

A list of women's shelters was compiled. It appears in Table 7.

7 List of women's shelters in the Kansas City metro area.

	Name	Address	ZipCode
0	Friends of Yates	1418 Garfield Avenue Kansas City, KS	66104
1	Synergy Services	400 East 6th Street Parkville, MO	64152
2	Safehome Inc.	PO Box 4563 Overland Park, KS	66204
3	Mother's Refuge	3721 Delridge Rd Independence, MO	64055
4	Shelter KC	2611 E. 11th Street Kansas City, MO	64127
5	El Centro, Inc.	650 Minnesota Ave. Kansas City, KS	66101
6	Metropolitan Organization to Counter Sexual As...	3100 Broadway, Suite 400 Kansas City, MO	64111

These addresses were joined to latitude and longitude coordinates, appearing in Table 8.

8 List of women's shelters with locations.

	Name	Address	ZipCode	Latitude	Longitude
0	Friends of Yates	1418 Garfield Avenue Kansas City, KS	66104	39.125983	-94.645101
1	Synergy Services	400 East 6th Street Parkville, MO	64152	39.191914	-94.678679
2	Safehome Inc.	PO Box 4563 Overland Park, KS	66204	38.985800	-94.671610
3	Mother's Refuge	3721 Delridge Rd Independence, MO	64055	39.054243	-94.456902
4	Shelter KC	2611 E. 11th Street Kansas City, MO	64127	39.100266	-94.551020
5	El Centro, Inc.	650 Minnesota Ave. Kansas City, KS	66101	39.115727	-94.626596
6	Metropolitan Organization to Counter Sexual As...	3100 Broadway, Suite 400 Kansas City, MO	64111	39.070635	-94.590631

These locations are plotted on a map to see if they overlap with the highest incidents of domestic violence, in Figure 4.
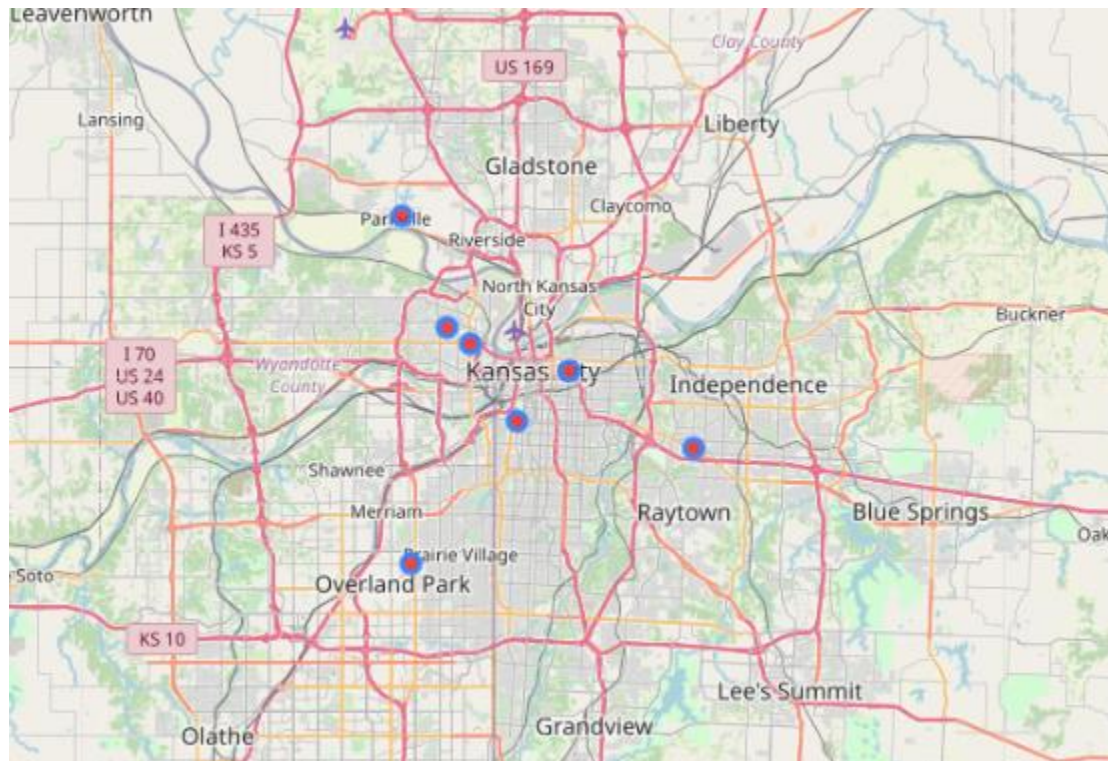

Figure 4 The location of women's shelters in the Kansas City metro area.

Mothers Refuge and Safehome Inc. are the only women's shelters in the most affected areas by domestic violence. It is apparent that the areas with the highest rates of domestic violence do not have the number of social service organizations that low-incidence areas have.

5   Conclusion

The number of domestic violence incidents across the Kansas City metro area was looked up. This was plotted on a bar graph. Highest incidents of domestic violence were clustered using the k-means machine learning algorithm. After isolating the high incident clusters, the number of women's shelter services for those areas were explored. It is apparent that the zip codes suffering from the highest rates of domestic violence also have the fewest social services. The Kansas City metro area would be wise to concentrate more domestic violence social services on the Raytown area and surrounding neighborhoods.

Kraske, S. (2020, November). *Coronavirus Keeps Victims Of Domestic Violence In Lockdown*

With Their Abusers. KCUR 89.3 - NPR in Kansas City. Local News, Entertainment and

Podcasts. https://www.kcur.org/podcast/up-to-date/2020-11-02/coronavirus-keeps-

victims-of-domestic-violence-in-lockdown-with-their-abusers

Nicholas Fry
November 24, 2020
Coursera Applied Data Science Capstone

Maclen, S. (2020, May 9). *Why the increase in domestic violence during COVID-19?*

Psychology Today. https://www.psychologytoday.com/blog/making-sense-

chaos/202005/why-the-increase-in-domestic-violence-during-covid-19

Mesa, K. (2020, November 17). *KCPD Crime Data 2020*. https://data.kcmo.org/Crime/KCPD-

Crime-Data-2020/vsgj-uufz

US Zip Code Latitude and Longitude. (n.d.). Retrieved November 21, 2020, from

https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/