Detroit Police Department Carjacking Offenses Data Analysis

Final Project

Prepared For: CST8390_010 BI & Data Analytics

Prepared By: Abdullah Zeki Ilgun, Elaiza Rivera

April 3, 2021

Frame Question: Under what circumstances the carjacking happens?

Table of Contents

Introduction

This study will analyze the DPD (Detroit Police Department) Carjacking Offenses confirmed by Victims. Retrieved from Detroit's Open Data Portal DPD Carjacking Offenses Victims Confirmed | Detroit's Open Data Portal (detroitmi.gov). The data was based on the city's 2015 – 2017 profile. With this study we will examine what are the common factors that affect these crimes such as time, location, certain day and whether a police officer is present around area, or the proximity of the officer's position.

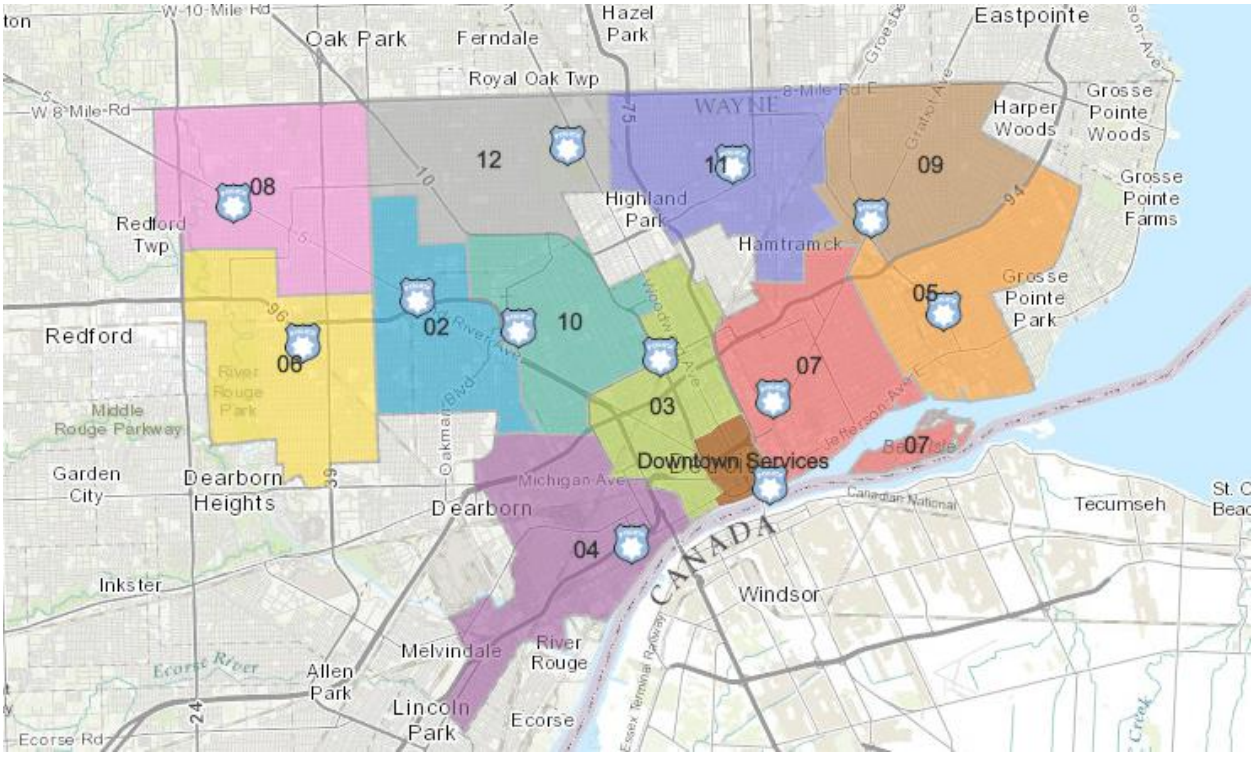
Carjacking usually subsets a form of other crime, specifically robbery. In some articles like in abc Detroit [1] news from 2019, an armed man was charged of multiple crime-related activities including multiple number of carjacking offenses of armed robbery in which attacks mostly happened late at night until early morning. Though there are incidents that happened late afternoon as well. Another article of the same city from 2020 [2] the criminal approached a man who was sitting on his car, knocked on his window by the firearm, robbed his possession and took the car, the incident happens approximately 5:04 am.

Why we chose this data set?

The movie called "Minority Report" was one of the reasons of why analysing this data set caught our interest. Synopsis of the movie is all about preventing crimes based on the predictions, which have almost perfect accuracy, of a technology developed in the future. Even though he current accuracy of those kind of prediction systems in real life is yet to be perfect, the idea is interesting and very useful as well as the massive growing in the amount of data is leading AI and the foresees to get closer to perfect accuracy and to a utopic new world.

In this study we will investigate the common factors as to when and where attackers pounce most commonly and how often is the occurrence of the attack in a day and in what location. We found interest on this dataset as to this sort of criminality can happen anywhere, this may share awareness and may reduce the number of carjacking cases in every city. We want to learn if there are scheme that we can predict of, by the behaviour of the carjacking offender.

Prior to the analysis, we predict that the most incidents happen in the 2nd, 9th and 10th precincts, as they seem to be the most dangerous ones, even though the number of stolen cars has been decreased over the years. (https://www.mlive.com/news/2016/02/where_are_detroits_most_danger_1.html) Also, we think that most of the incidents should be happening on Fridays where the citizens are chilling or having Friday night fun. Finally, obviously the nighttime is mostly preferred by the criminals, though we do not have an exact time interval foresee.



Detroit Precinct Map

Data Collection

This dataset was collected from the city of Detroit's Open Data Portal. These data were gathered from the and confirmed by the victims occurred between 2015 and 2017. These data were used by the Detroit Police Department for their strategic planning and crime analysis, thus this gathering information regarding these datasets are still ongoing and is not final. The dataset contains 18 attributes and 1,210 records in overall.

The data set at first was downloaded as a spreadsheet in CSV format. In that sheet, there were some differences from the records' overview on the website that would prevent doing a proper analysis, which are the date value was in scientific format that cannot be converted to a proper format, and 75% of the time value was missing, and without the time value, the purpose of the analysis would not be fulfilled.

FID	X	Y	DAY	DATE	TIME	REPORT_N	PRECINCT	ADDRESS	CITY	STATE	Scout_Car	Zip_Codes	Detroit_Ne	Hardest_H	Counties	City_Coun	ObjectId
	-83.0479	42.33292		1.48E+12	########	1.61E+09	1	15TH/POP	DETROIT	MI	4	20014	142		1345	5	1
	-83.0479	42.33292		1.48E+12	########	1.22E+09	8	192XX SCH	DETROIT	MI	4	20014	142		1345	5	2
	-83.0479	42.33292		1.47E+12	########	1.61E+09	8	GREENVIE	DETROIT	MI	4	20014	142		1345	5	3
	-83.0479	42.33292		1.48E+12	########	1.61E+09	6	LIVERNOI	DETROIT	MI	4	20014	142		1345	5	4
	-83.0479	42.33292		1.48E+12	########	1.61E+09	11	204XX MA	DETROIT	MI	4	20014	142		1345	5	5
	-83.0479	42.33292		1.47E+12	########	1.61E+09	8	194XX PLY	DETROIT	MI	4	20014	142		1345	5	6
	-83.0479	42.33292		1.47E+12	########	1.61E+09	11	81XX EMIL	DETROIT	MI	4	20014	142		1345	5	7
	-83.0479	42.33292		1.45E+12	########	1.6E+09	12	ASBURY PI	DETROIT	MI	4	20014	142		1345	5	8
	-83.0479	42.33292		1.47E+12	########	1.61E+09	10	WARREN/	DETROIT	MI	4	20014	142		1345	5	9
	-83.0479	42.33292		1.47E+12	########	1.61E+09	5	XX MILE/M	DETROIT	MI	4	20014	142		1345	5	10
	-83.0479	42.33292		1.46E+12	########	1.6E+09	3	15XX W. G	DETROIT	MI	4	20014	142		1345	5	11
	-83.0479	42.33292		1.47E+12	########	1.61E+09	7	88XX E JEF	DETROIT	MI	4	20014	142		1345	5	12
	-83.0479	42.33292		1.48E+12	########	1.61E+09	5	105XX JEFI	DETROIT	MI	4	20014	142		1345	5	13
	-83.0479	42.33292		1.47E+12	########	1.61E+09	2	LYON/DEA	DETROIT	MI	4	20014	142		1345	5	14
	-83.0479	42.33292		1.47E+12	########	1.61E+09	2	151XX GRE	DETROIT	MI	4	20014	142		1345	5	15
	-83.0479	42.33292		1.47E+12	########	1.61E+09	8	SCHOOLCF	DETROIT	MI	4	20014	142		1345	5	16
	-83.0479	42.33292		1.47E+12	########	1.61E+09	10	KENTUCKY	DETROIT	MI	4	20014	142		1345	5	17
	-83.0479	42.33292		1.46E+12	########	1.6E+09	10	INTERVALI	DETROIT	MI	4	20014	142		1345	5	18
	-83.0479	42.33292		1.47E+12	########	1.61E+09	12	134XX W E	DETROIT	MI	4	20014	142		1345	5	19
	-83.0479	42.33292		1.47E+12	########	1.61E+09	2	145XX MC	DETROIT	MI	4	20014	142		1345	5	20
	-83.0479	42.33292		1.48E+12	########	1.61E+09	12	165XX LIT	DETROIT	MI	4	20014	142		1345	5	21
	-83.0479	42.33292		1.45E+12	########	1.6E+09	5	46XX BEN	DETROIT	MI	4	20014	142		1345	5	22
	-83.0479	42.33292		1.46E+12	########	1.6E+09	7	GRATIOT/I	DETROIT	MI	4	20014	142		1345	5	23
	-83.0479	42.33292		1.48E+12	########	1.61E+09	10	33XX JOY I	DETROIT	MI	4	20014	142		1345	5	24
	-83.0479	42.33292		1.45E+12	########	1.6E+09	6	182XX JOY	DETROIT	MI	4	20014	142		1345	5	25
	-83.0479	42.33292		1.46E+12	########	1.6E+09	9	120XX HAF	DETROIT	MI	4	20014	142		1345	5	26
	-83.0479	42.33292		1.47E+12	########	1.61E+09	11	FENELON/	DETROIT	MI	4	20014	142		1345	5	27
	-83.0479	42.33292		1.48E+12	########	1E+09	11	191XX HAS	DETROIT	MI	4	20014	142		1345	5	28

First version of the data set

Therefore, after trying the other formats, it was seen the JSON format had the proper time values. It was downloaded and converted to CSV format using an online source (http://convertcsv.com/json-to-csv.htm).

Below are the lists of attributes that were collected from the victims. As it cannot be found proper definitions of the attributes online, the definitions were based on the literal meaning of the word and through the researchers' understanding:

Attribute	**Definition**
FID	Force Investigation Division, the number of the division responsible for the crime.
X and Y	Respectively longitude and latitude. The exact coordinate on the map in decimal format that the crime happened.

Day	Day of the week when the crime happened.
Date	It is assumed that this attribute represents the date when the crime was reported rather than when it happened, because the values conflict with the values of the time attribute.
Time	Shows the hour and minute when the crime happened.
Report No	The report number of the crime.
Precinct	Number of the area which the certain police is responsible for.
Address	The address of the crime, including street and the door number with incomplete detail.
City	As the data set is all about Detroit incidents, all values point to Detroit.
State	The state where the city of Detroit belongs – Michigan (MI)
Scout Car	The amount of scout cars presents near the area at the time when the incident happened.
Zip Codes	Zip code of the area.
Detroit Neighborhoods	The area code of the neighborhood.
Hardest Hit Fund	The degree of which the victim was affected mostly.
Counties	The region, same as city and state attributes, nearly all the counties point to 1345.
City Council Districts	The city council district of the area which the incident happened.
Object ID	The unique id for the report.

Preprocessing

A proper analysis could not be done on the set before preprocessing for certain reasons:

- Irrelevant attributes – are attributes that we think won't affect much or relates to the result we want to come up, which are not about the conditions of the incident.
- Redundant attributes - which represents the information that are already given by another attribute, but presented in different ways for analysis i.e., map coordinates and actual address.
- Missing values - which an algorithm cannot simply work without handling them.
- Outliers - attributes that has too big or small values, though which the dataset does not contain some values are having conflict with the values of another attribute.
- Data formats - which are needed to be in the same format for running a certain algorithm.

Irrelevant Attributes

Report number, FID and Object ID attributes concern the police department rather than the analysis.

Hardest hit fund has nothing to do with the analysis because the purpose of the analysis is to predict any incident that might happen regardless the effect of the incident.

It is assumed that the date value is about the date when the report was taken rather than the exact date of the incident as it conflicts with the values in the time value.

Redundant Attributes

City, state and county values are all about Detroit. There are some missing data in the city and state attributes, and some of them are checked to make sure if they are in Detroit, Michigan as well, and thus they are in Detroit too. There is very little amount of data in the counties which are different than 1345, and that can be ignored since they are in Detroit as well.

X and Y values already represent the exact position where the incident happened, so there is no need for the address. The address data is also hard to implement to an analysis as it cannot be converted to a nominal nor a numerical format. However, some values in the attribute are used for detecting and handling the wrong data given in the X and Y coordinates. Its values are also used for making sure all the records are concerning Detroit.

As there were a lot of wrong data in city council district value, and also as there is no need for it when there are X and Y coordinates, it is removed as well.

Zip code value is redundant and also gives wrong data about the code. It can be seen on the Internet that the zip codes in the Detroit are totally different than the given values.

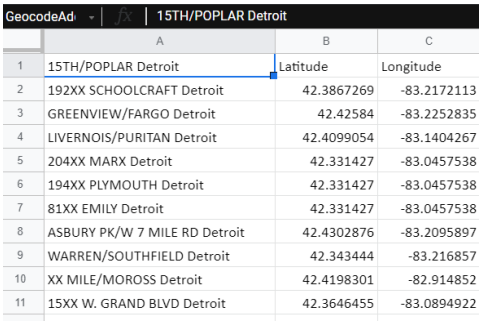
Remaining Attributes: X, Y, Day, Time, Precinct, Scout Car

Outliers

There were significant number of outliers in the data set, which would have given wrong results if the analysis had been made with these.

It can be clearly seen that roughly 25% of the X and Y data were having the same values even though the addresses were different, which proves that they were filled randomly. If they were simply removed, that amount of data loss would not have been a good practice, so it is decided to find the correct coordinates for the addresses.

X	Y	DAY	DATE	TIME	REPORT_N	PRECINCT	ADDRESS	CITY
-83.0479	42.33292		1.48E+12	########	1.61E+09	1	15TH/POP	DETROIT
-83.0479	42.33292		1.48E+12	########	1.22E+09	8	192XX SCH	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	8	GREENVIE	DETROIT
-83.0479	42.33292		1.48E+12	########	1.61E+09	6	LIVERNOIS	DETROIT
-83.0479	42.33292		1.48E+12	########	1.61E+09	11	204XX MA	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	8	194XX PLY	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	11	81XX EMIL	DETROIT
-83.0479	42.33292		1.45E+12	########	1.6E+09	12	ASBURY PI	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	10	WARREN/	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	5	XX MILE/N	DETROIT
-83.0479	42.33292		1.46E+12	########	1.6E+09	3	15XX W. G	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	7	88XX E JEF	DETROIT
-83.0479	42.33292		1.48E+12	########	1.61E+09	5	105XX JEF	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	2	LYON/DEA	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	2	151XX GRE	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	8	SCHOOLCF	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	10	KENTUCKY	DETROIT
-83.0479	42.33292		1.46E+12	########	1.6E+09	10	INTERVALI	DETROIT
-83.0479	42.33292		1.47E+12	########	1.61E+09	12	134XX W 8	DETROIT



GeocodeAd | fx | 15TH/POPLAR Detroit

	A	B	C
1	15TH/POPLAR Detroit	Latitude	Longitude
2	192XX SCHOOLCRAFT Detroit	42.3867269	-83.2172113
3	GREENVIEW/FARGO Detroit	42.42584	-83.2252835
4	LIVERNOIS/PURITAN Detroit	42.4099054	-83.1404267
5	204XX MARX Detroit	42.331427	-83.0457538
6	194XX PLYMOUTH Detroit	42.331427	-83.0457538
7	81XX EMILY Detroit	42.331427	-83.0457538
8	ASBURY PK/W 7 MILE RD Detroit	42.4302876	-83.2095897
9	WARREN/SOUTHFIELD Detroit	42.343444	-83.216857
10	XX MILE/MOROSS Detroit	42.4198301	-82.914852
11	15XX W. GRAND BLVD Detroit	42.3646455	-83.0894922

Filling them all manually would not be a good practice as well, and after some search on how to do that automatically, it is decided to use Google's Geocode extension for its sheets. All the wrong address values are copied into Google Sheets and latitude/longitude (Y/X) values were automatically generated by Geocode.

However, still some of the values were not correct, as the addresses were not totally clear (The door numbers are incomplete and some of the digits are replaced by 'X') Geocode has automatically perceived these unclear values as central Detroit. So, these several addresses were looked up from Google Maps, and manually filled.

Other than that, by sorting in Excel, it could be seen that some of the values were significantly bigger than the others, and outside the Detroit. These few records were removed from the set as the focus of the analysis is Detroit.

The other issue was the scout car and city council districts were having the same values as well for these repetitive X and Y records. It is made sure whether they are the mode values of the attribute, and after making sure, no further operation has been done on these records.

Data Formats

Two of the attributes' data were in another format than decimal, day, and time.

To convert the day to a numeric format, all the days are replaced by the equivalent number of the day's order in the week starting from Monday. To illustrate, monday-1, tuesday-2 and so on.

Converting the time was not that simple. These steps below were followed to do the conversion:

1. Firstly, time data should have been converted to 24-hour format. However, not all the time values were in a clean form to do the conversion automatically. For example, AM and PM indicators were stick to the number. By using Excel's features, all the values were cleaned, and became in the same form.
2. All the values were converted into 24-hour format by the text formula [=TEXT(A1, "HH:MM:SS")]

3. Hour, minute and second were split up to different columns by the delimiter ':', and the column showing second was removed as all the data were '0'.
4. Finally, by using a sum formula, which is multiplying hour by 60, and adding the minutes to it, all the data were converted to a decimal form.

Missing Values

There were not much missing values for most of the attributes that were going to be used in the analysis. Few of them are checked by filtering in Excel, and simply put the mode values of the attribute. There were 25% missing values in the day attribute, so the mode of the days was calculated, and these missing values are filled with the mode, which is 7 (Sunday)

Analysis

In order to extract more results from the set, two different methods were decided to be used. Decision Tree and Association Rule Mining.

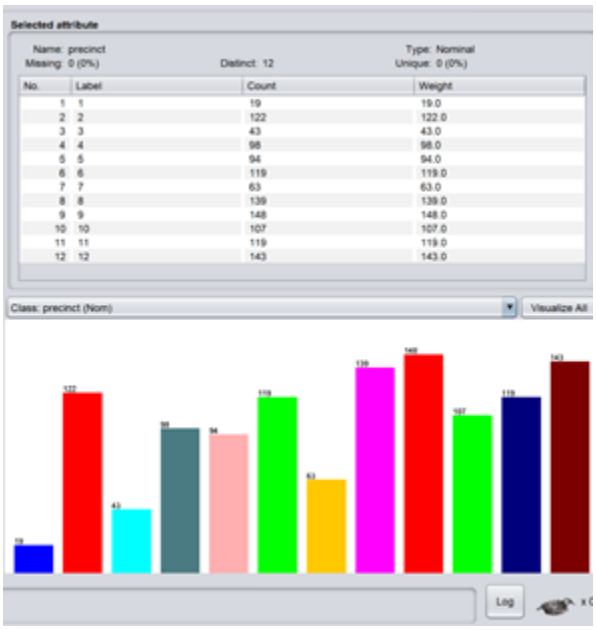
Decision Tree

Decision tree is selected as the main method to use in the analysis. It is expected that this method will help to figure out the possible chances that the incident will most likely to happen. By definition, decision trees help research operations and decision analysis, and it helps to identify a strategy to reach the goal.

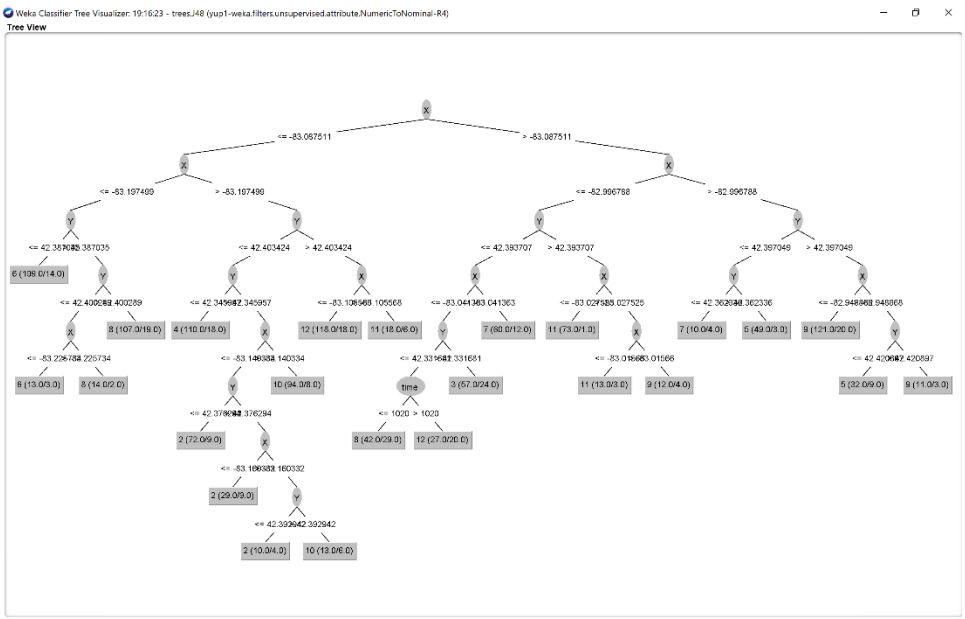
Two sets were created for analysis to compare:

First set

Attributes were X and Y coordinates, time with precinct as a class category. On this set it was to consider only the time value and location. Since time does not contain is the only attribute that is concrete, and it can be said that it is the actual. Location is important to determine on what certain time an area was usually an incident happen. With this, it can be determined the behaviour of a possible attack. Below is the table and graph to visualize how many carjacking incidents had happened on different precinct. It can as well be read the most and least affected area.



Couple of analysis were run, and there the two of trees that were generated. Basing on the coordinates, north and west part of the city is hardly hit after 5:00 pm. It can be seen on this tree that the precinct 8 had more attack when clock strikes at 5:00 PM onwards. But we want to see more details.



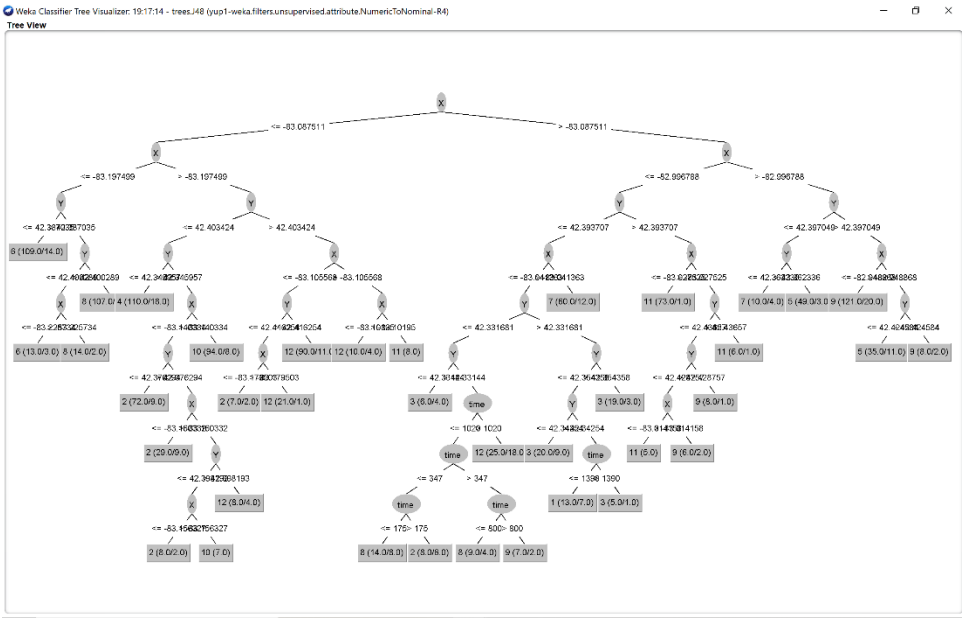
First tree

The first tree has more correctly classified instances compare to the second tree, yet we will still take into consideration the second tree to further see the time when incident happens. Below is the summary of the first tree and its confusion matrix. Our got 73.8499% on our correctly classified instance and 26.1501% incorrectly classified.

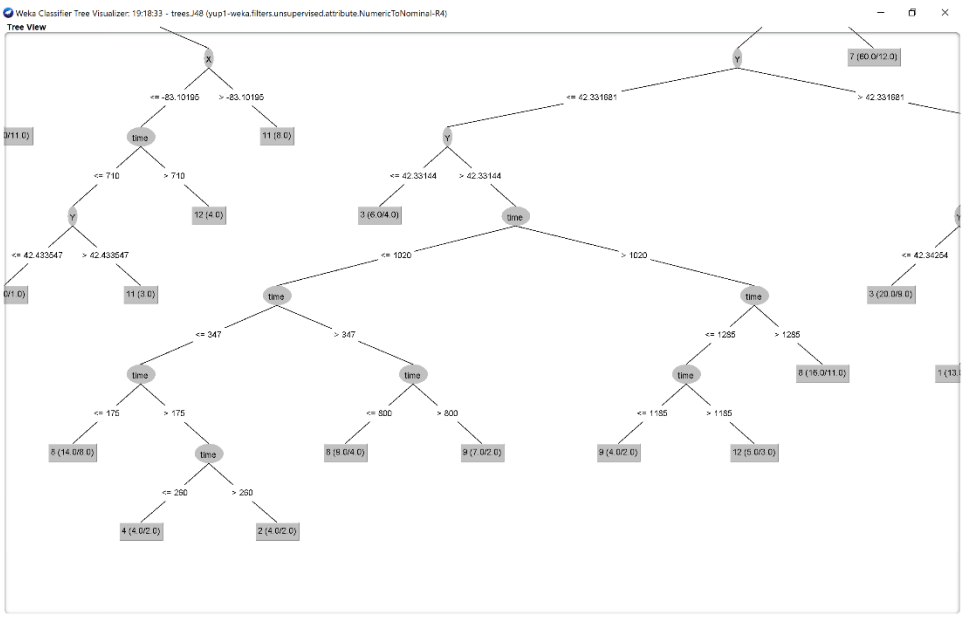
```
Number of Leaves  :     24

Size of the tree :      47


Time taken to build model: 0.01 seconds

=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances         305               73.8499 %
Incorrectly Classified Instances       108               26.1501 %
Kappa statistic                          0.7105
Mean absolute error                      0.0638
Root mean squared error                  0.1916
Relative absolute error                 42.2901 %
Root relative squared error             69.6538 %
Total Number of Instances              413


 === Confusion Matrix ===

   a  b  c  d  e  f  g  h  i  j  k  l   <-- classified as
   0  1  1  1  0  0  1  0  0  1  0  0 |  a = 1
   0 33  0  5  0  1  0  1  2  2  0  1 |  b = 2
   0  0 12  1  0  0  4  0  0  1  0  0 |  c = 3
   0  0  0 27  0  0  0  0  2  0  0  1 |  d = 4
   0  0  0  0 21  0  7  0  8  0  0  2 |  e = 5
   0  2  0  0  0 29  1  4  0  0  0  2 |  f = 6
   0  0  1  0  0  0 22  0  2  0  0  0 |  g = 7
   0  0  0  0  0  3  0 34  3  0  0  4 |  h = 8
   0  0  0  0 10  0  0  1 44  0  1  1 |  i = 9
   0  2  2  0  0  3  0  0  0 32  0  0 |  j = 10
   0  0  1  0  3  0  1  0  2  0 19  2 |  k = 11
   0  2  0  0  0  0  0  2  2  2  6 32 |  l = 12
```

Second tree was generated with minNumObj set to 5 to compare if coordinate values will be changed and to see more of the time value. After comparing with the first tree, the second tree is able to expand to us the time after 5:00 PM, yet coordinates are similar to the first one. It can be

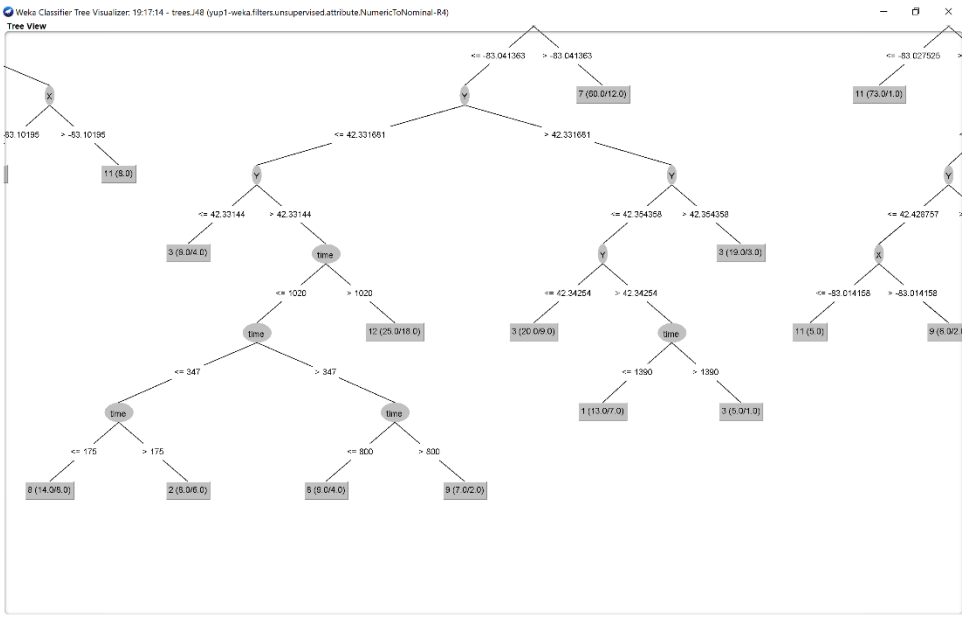
said that on the same areas and precinct 8, 2 and 9 located on the northwest part of the cities are the areas where mostly attack happens later at night.



This tree had the minNumObj reduced to 3 to see more of the time value. When we will look more the time, on the third diagram, it may be seen that later than 7:00pm more attacks happen.



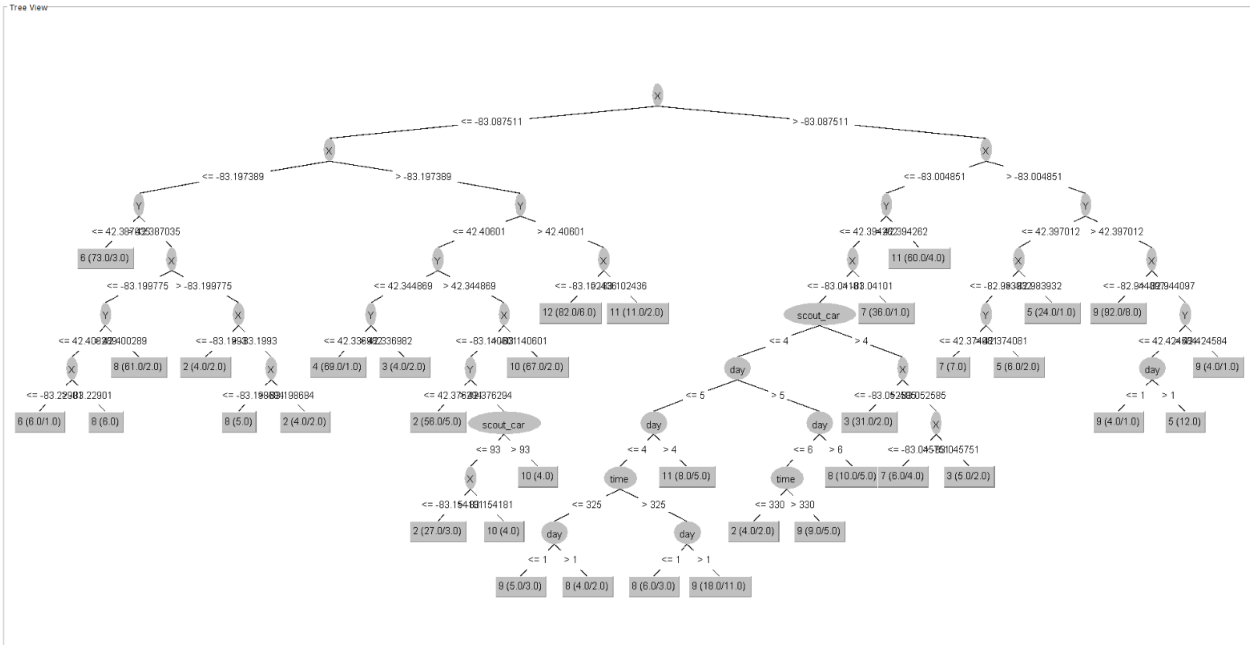
This was trimmed from the second tree to see the value of time. It can be seen that if the location is not further west attacks are more prevalent when time approaches 11:00 PM



A different set of data were run, to consider other factors. The time of the most attacked area was gotten, but we also want to see which day is usually the attack and if the presence of scout cars helps reduce the incident to happen.

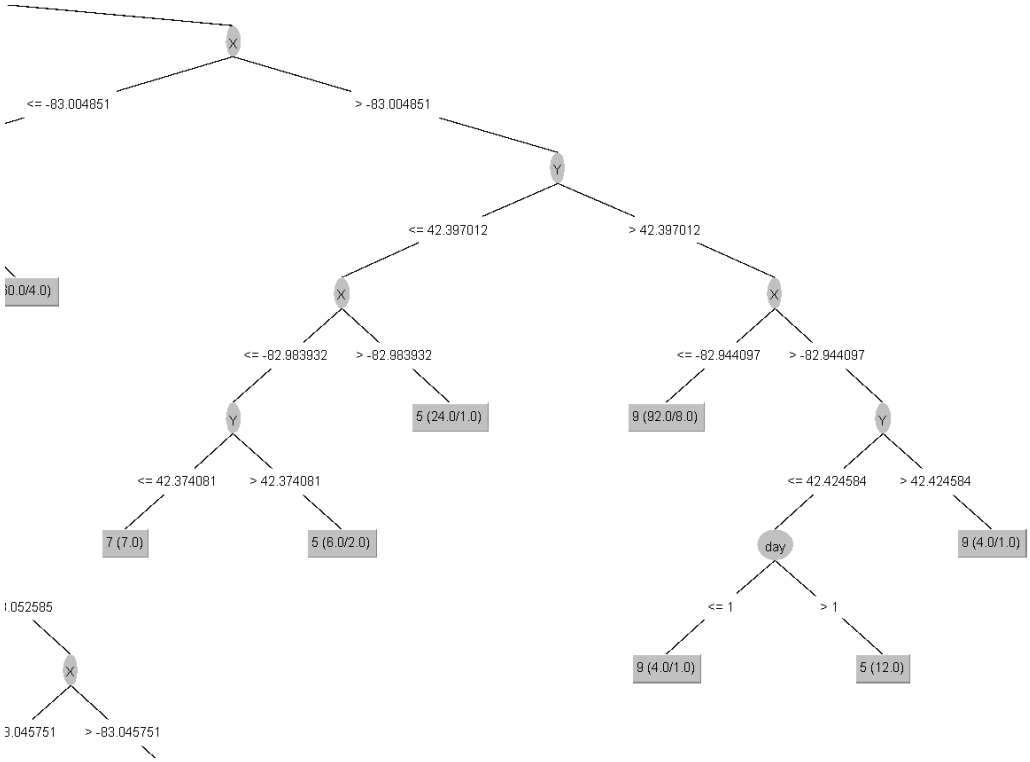
Second Set

In this set, it was aimed to consider all the attributed to make the prediction, but with the less amount of data. Roughly 400 of the records were removed due to their repetitive data in day, scout car and city council district columns. Therefore, the attributes included in this set are X, Y, day, time, precinct, scout car and city council districts.

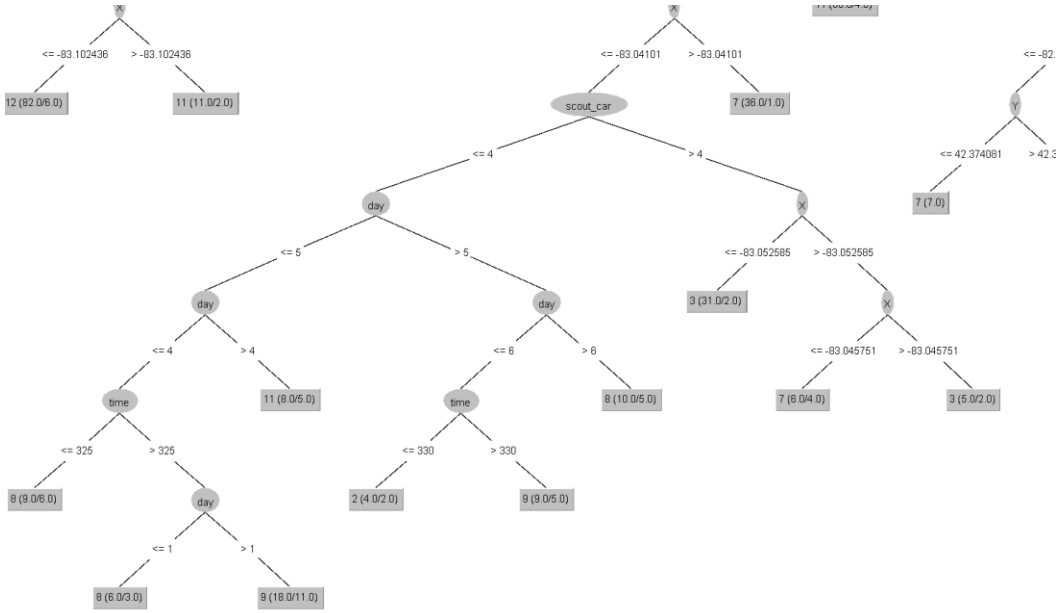


Second Tree

From the tree, the coordinates, and especially latitude is also the most significant factor of the set.



When we look at here, it is seen that south of the city is more likely for a crime to happen. At the most northwest side of the city has the most car jackings.



Above capture of the tree also shows the other factors affecting an incident to happen. As it is expected, there is a reversed ratio between the amount of scout cars and the car jacks.

Also, in overall, the weekends include less incidents. On the weekend, early morning till 5.30 is the most dangerous time for the crime. It is same for the weekdays as well, so that shows the criminals mostly prefer early morning for a car jack.

Association Rule Mining

This method is decided to be used as a supporter method, which would help to get more specific results for critical decisions.

Firstly, to run this algorithm, all attributes should be nominal. To achieve that, X and Y attributes were excluded from the set, as it would have been impossible to categorize them meaningfully. Then, the other attributes were categorized via Excel functions, by the methods explained below.

Precinct: Even though it is already categorized by the values 1,2,3,4,5,6,7,8,9,10,11,12, when the algorithm was run, no proper association was gotten, so it was decided to generalize it more by splitting them to “East” (1,3,5,7,9,11), and “West” (2,4,6,8,10,12).

Day: It is also categorized naturally, and only think done on that is to convert them into their word formats again. (Monday, Tuesday etc.)

Scout Cars: Its values were categorized as “Low” (0-40), “Medium” (40-80), and “High” (Above 80).

Time: That conversion was the more complicated one as they should have been more categories, and a conversion. In order to convert the decimal values to their hour format again, a simple Java program was written and used for it to be make the process more convenient.

```java
import java.util.Scanner;

public class Yay {
    public static void main(String args[]) {

        while(true) {
            Scanner hey = new Scanner(System.in);
            System.out.print("Enter the num: ");
            int num = hey.nextInt();

            int min = num % 60;
            int hour = num / 60;

            String convertedNum = hour + ":" + min;
            System.out.println(convertedNum);
        }

    }
}
```

```
C:\Users\karka\Desktop\New folder (2)>java Yay
Enter the num: 1200
20:0
Enter the num: 1440
24:0
Enter the num: 300
5:0
Enter the num: 720
12:0
Enter the num:
```

Finally, it is decided to label the hours as; 8pm-5am is night, 5am to 12pm is morning, 12pm to 5pm is noon, and 5pm to 8pm is evening. They are converted by the Excel formula;

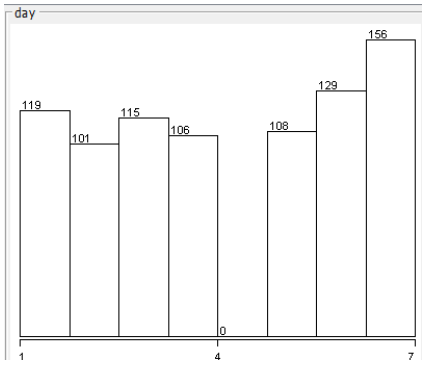
=IF(OR(AND(D3>1200, D3<=1440), D3<=300), "Night", IF(AND(D3>300, D3<=720), "Morning", IF(AND(D3>720, D3<=1020), "Noon", IF(AND(D3>1020, D3<=1200), "Evening", "Null"))))

Afterwards, the file is loaded to Weka, and the Apriori algorithm was run on the set with minimum support set to 1%.

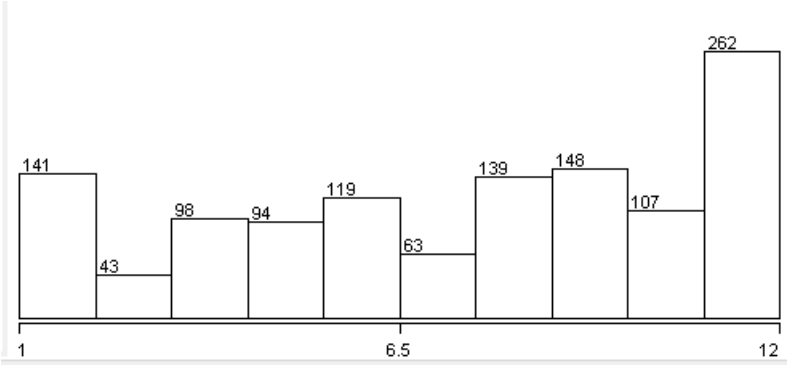
```
Best rules found:

 1. Time=Morning Day=Wednesday 16 ==> Precinct=West 15    <conf:(0.94)> lift:(1.53) lev:(0.01) [5] conv:(3.11)
 2. Scout_Car_Amount=Low Time=Morning Day=Wednesday 12 ==> Precinct=West 11    <conf:(0.92)> lift:(1.5) lev:(0) [3] conv:(2.34)
 3. Time=Evening Day=Thursday 11 ==> Precinct=West 10    <conf:(0.91)> lift:(1.49) lev:(0) [3] conv:(2.14)
 4. Scout_Car_Amount=Low Time=Evening Day=Saturday 11 ==> Precinct=West 10    <conf:(0.91)> lift:(1.49) lev:(0) [3] conv:(2.14)
 5. Time=Noon Day=Tuesday 10 ==> Scout_Car_Amount=Low 9    <conf:(0.9)> lift:(1.58) lev:(0) [3] conv:(2.15)
```

Despite that there were not many rules found, yet couple of meaningful results were gotten. The results are concerning the west side of the city, and about the different times and days that will be explained in the results.

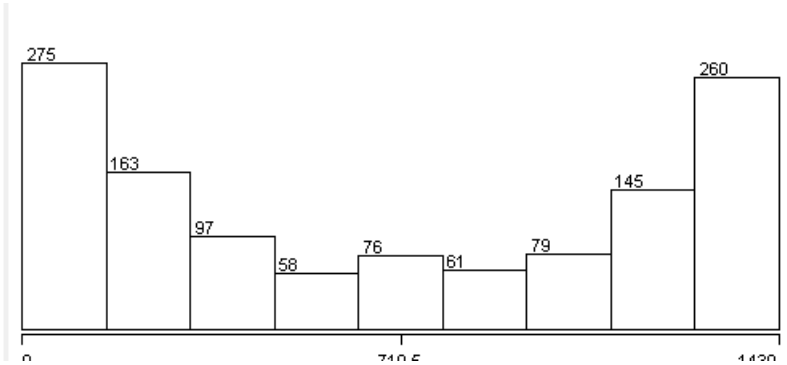
Important Statistics from the Analysis



Sunday has the most crime rate.



Precinct 12 has the most incident rate, whereas 2, 6, 9 are having the least.



Between 9.20 pm and 2.40 am mostly preferred by the car jackers.

Results and Summary

With the two analysis with different attributes, both are resulted with same area location to which carjacking usually happen that is the north/west part of the city. In which if we want to see what is may attack carjackers on this part of Detroit city, on this article [3] in the north part of the city were people who are sports stars, entrepreneurs and automotive execs whose last name aren't Ford. This area of the city were big and roomy houses were also located; you can actually say that is a rich neighborhood. On the west part of the city were the aging baby boomers, technology professionals and lots of scientist, doctors, professors and students. This area also is surrounded with big developments which have downtowns with shops, restaurants, and condos. This might be the factors as to why so far this is the burglar's favorite spot.

We first mentioned and notice that when the time starts at 5:00 PM onwards, incidents happen as to the first analysis result. And will continue to rise after 11:00 PM. Yet our second analysis resulted and is telling us that early mornings usually until 5:30 am, when the usual activities happen regardless of if it is the weekend or weekdays. This may be because where streets are quitter or having less traffic. Though if there are more than 4 scout cars present in the area, it helps reduce the chance were on this unwanted event.

Looking at the bottom leaf of our last tree presented if it is weekdays and around 5:41 AM with less than four scout cars present. Precinct 8 and 9 had more incidents on record. We can conclude that mostly attackers strike on weekdays compares to weekend and are usually at dawn or early mornings but if more scout cars are actively present incidents are reduced.

Illustration of Wealthy Detroit Neighbourhood

From the association mining, we have seen that, again, on the west side of the city, Wednesday morning, Thursday evenings and Tuesday noon times are specifically popular among the criminals. We could not find by the researching that whether it is a coincidence, or there are specific reasons for them to prefer these times.

Conclusion & Recommendations

By doing this analysis, we aimed to make a meaningful prediction and give Detroit Police Department a recommendation of how to prevent these incidents in the future.

Prior to the analysis, we predicted that the most incident should have been happening in the most dangerous precincts; however, in fact criminals most likely to prefer the wealthy neighbourhoods due to the fact that they might see these areas as gold digging areas. Also, it was no surprise that they prefer the night time for the crime, yet their day preference, which is Friday, is different than our prediction.

It is hard to predict the exact point where the crime will happen, but by this analysis, it could be seen which areas and when they are more likely for an incident to happen. It could be also said, naturally and by looking at the results, that more scout cars means less crime rate. Thus, it is recommended for the Police Department to locate more scout cars on the high crime rate areas.

- The police department should more focus on the northwest part of Detroit in general, and especially to 12th Precinct.
- Scout cars should be mostly active on Sundays regarding the day.
- Scout cars should also be active between 5pm and 5.30 am, and especially between 9.20 pm and 2.40 am the hours to be most careful.
- For the west side; noon time, between 12pm and 5pm, on Tuesdays, evening, between 5pm and 8pm, on Thursdays, and morning, between 5am and 12pm, time on Wednesdays should require specific attention. More scout cars would be a good idea as well, but most importantly searching the reason why do the criminals choose these hours would be more helpful to find and solve the core problem of the crime. It might be coincidental, yet beneficial to make an investigation.

References

[1] ABD news 7, "Detroit man charged in crime spree involving carjacking, armed robberies and assaults," Scrippps Media, Detroit, 2019.

[2] Department of Justice U.S. Attorney's Office Eastern District of Michigan, "Detroit Man Charged with Carjacking and Firearms Offenses Carjacked a retired police officer and shot up a gas station," U.S. Department of Justice, Detroit, 2020.

[3] Great American Country, "5 Great Neighborhoods in Detroit," Detroit, 2021.