How did COVID-19 affect the crime rates in Chicago in 2020 compared to 2019

HaEun Choi

16/12/2020

Abstract

The purpose of this paper is to identify the differences in crime rates in Chicago between March 2019 and March 2020 using the dataset in Chicago Data Portal. It was found that the crime rates have significantly decreased after the COVID pandemic hit North America, compared to the previous year.

Introduction

Crimes-2019 dataset includes all the crime occurred in City of Chicago from 2001 to current year. This data is collected by the Chicago Police Department's CLEAR (Citizen Law Enforcement Analysis and Reporting) System. The main findings from this study is that some types of crime rate in 2020 has significantly decreased compared to the crime rate in 2019. In both years, the two most frequent crime types in 2019 are theft and battery. Another finding that was discovered was that there were high rates of crime committed in certain days of a week or month. These observations are interesting because by different types of crimes, the rates of that crime committed was different by the day.

Data

I examined the 2019 and 2020 crimes datasets provided by Chicago Police Department which includes the reported crime in Chicago. The data is available since 2001 and is constantly being updated. The crime data for 2019 and 2020 was downloaded as CSV files from the following locations: https://data.cityofchicago.org/Public-Safety/Crimes-2019/w98m-zvie https://data.cityofchicago.org/Public-Safety/Crimes-2020/qzdf-xmn8

I primarily focused on the crime type and the number of reported crimes in order to investigate the difference in numbers and categories of crimes. In order to properly categorize the types of crimes, I had to clean some types of crimes from the dataset. I also cleaned the categories that has low frequency and grouped it into "other offense". Additional variables were created to reflect the month, day of the week, and first of month indicator. I expect that seasonality exists in some types of crime and crime rates vary by day of the week. I removed months of January and February as there was no COVID in 2020 until 1-Mar-2020.

The combined dataset contains 376844 crime events in both years, 222433 in 2019 and 154411 in 2020. The following types of crime are most prevalent in Chicago: Theft (82603), battery (74322), criminal damage (42747), assault (31764).

```
## [1] 376844
```

```
##
##   2019   2020
## 222433 154411
```

Table 1. Total frequency of crime in Chicago in 2019-2020

```
d1[order(-d1$Frequency),]
```

```
##                            Crime Type Frequency
## 28                              THEFT     82603
## 3                             BATTERY     74322
## 6                      CRIMINAL DAMAGE     42747
## 2                             ASSAULT     31764
## 9                  DECEPTIVE PRACTICE     26666
## 21                     OTHER OFFENSE     22683
## 18                         NARCOTICS     16631
## 17                MOTOR VEHICLE THEFT     15271
## 4                             BURGLARY     14963
## 25                            ROBBERY     12763
## 29                 WEAPONS VIOLATION     12219
## 8                  CRIMINAL TRESPASS      8544
## 20        OFFENSE INVOLVING CHILDREN      3221
## 24            PUBLIC PEACE VIOLATION      2350
## 7            CRIMINAL SEXUAL ASSAULT      2173
## 26                       SEX OFFENSE      1769
## 13  INTERFERENCE WITH PUBLIC OFFICER      1718
## 11                           HOMICIDE      1113
## 1                              ARSON       813
## 22                     PROSTITUTION       786
## 27                         STALKING       332
## 5  CONCEALED CARRY LICENSE VIOLATION       304
## 16              LIQUOR LAW VIOLATION       302
## 14                      INTIMIDATION       266
## 15                      KIDNAPPING       241
## 10                        GAMBLING       159
## 19                       OBSCENITY        90
## 23                PUBLIC INDECENCY        18
## 12                HUMAN TRAFFICKING        13
```

Table 2. Daily crime rates for year 2019 and 2020 by crime type

```
table1[,3:4]
```

```
##                                  Daily rate 2019 Daily rate 2020
## ARSON                                       1.06            1.78
## ASSAULT                                    58.15           51.19
## BATTERY                                   139.77          115.66
## BURGLARY                                   26.78           24.80
## CONCEALED CARRY LICENSE VIOLATION           0.63            0.41
## CRIMINAL DAMAGE                            75.96           71.46
## CRIMINAL SEXUAL ASSAULT                     4.42            3.01
## CRIMINAL TRESPASS                          18.88           10.17
## DECEPTIVE PRACTICE                         51.30           40.21
## GAMBLING                                    0.45            0.08
```

```
## HOMICIDE                                  1.50           2.40
## HUMAN TRAFFICKING                         0.04           0.01
## INTERFERENCE WITH PUBLIC OFFICER          4.40           1.37
## INTIMIDATION                              0.44           0.49
## KIDNAPPING                                0.49           0.33
## LIQUOR LAW VIOLATION                      0.66           0.36
## MOTOR VEHICLE THEFT                      24.98          27.93
## NARCOTICS                                40.82          15.26
## OBSCENITY                                 0.17           0.14
## OFFENSE INVOLVING CHILDREN                6.37           4.66
## OTHER OFFENSE                            45.61          32.01
## PROSTITUTION                              1.90           0.76
## PUBLIC INDECENCY                          0.04           0.03
## PUBLIC PEACE VIOLATION                    4.34           3.75
## ROBBERY                                  22.41          21.64
## SEX OFFENSE                               3.65           2.40
## STALKING                                  0.60           0.54
## THEFT                                   175.39         106.24
## WEAPONS VIOLATION                        18.08          24.47
```

In the above table, I observed a decrease in the following four crime types: battery, theft, deceptive practice and narcotics. For example, narcotics daily rates decreased from 40.82 in 2019 to 15.26 in 2020, a decline of 63%.

Model

The primary analysis is performed to determine if daily crime rates have decreased significantly during the pandemic. Linear regression model was built for each type of crime with dependent variable being daily frequency of crime. Models include the following predictors: - Pandemic = binary indicator with 0 = year 2019 (pre-pandemic) and 1 = year 2020 (during pandemic) - Month = categorical variable with March being chosen as reference category - Day of the week = categorical variable with SUnday being selected as reference category - First day of the money = binary indicator with 0 = not first day of the month and 1 = first day of the month

Linear regression model equation can be written as following:

$$[DailyCrimeRate] = \beta_0 + \beta_1[Pandemic] + \beta_2[Apr] + ... + \beta_{10}[Dec] + \beta_{11}[Mon] + ... + \beta_{16}[Sat] + \beta_{17}[FirstDayOfMonth]$$

Data was reorganized so each row represents a particualar day and crime type. Each row has a count of how many crimes of this type were reported to police, with zero indicating no crime being reported that day. The sample below shows first 10 records in this re-organized file.

```
data1[1:10,]
```

```
##          Date                        PrimaryType Pandemic Month DayOfWeek
## 1  2019-03-01                              ARSON        0 March    Friday
## 2  2019-03-01                            ASSAULT        0 March    Friday
## 3  2019-03-01                            BATTERY        0 March    Friday
## 4  2019-03-01                           BURGLARY        0 March    Friday
## 5  2019-03-01 CONCEALED CARRY LICENSE VIOLATION        0 March    Friday
## 6  2019-03-01                     CRIMINAL DAMAGE        0 March    Friday
## 7  2019-03-01           CRIMINAL SEXUAL ASSAULT        0 March    Friday
```

```
## 8  2019-03-01                CRIMINAL TRESPASS        0 March    Friday
## 9  2019-03-01               DECEPTIVE PRACTICE        0 March    Friday
## 10 2019-03-01                         GAMBLING        0 March    Friday
##    FirstDayOfMonth NumOfEvents
## 1                1           1
## 2                1          47
## 3                1          97
## 4                1          23
## 5                1           1
## 6                1          63
## 7                1           5
## 8                1          12
## 9                1         105
## 10               1           0
```

Results

Four regression models were constructed for the following crime types: narcotics, theft, battery and deceptive practice.

NARCOTICS crime

```
summary(model1)
```

```
##
## Call:
## lm(formula = NumOfEvents ~ Pandemic + Month + DayOfWeek + FirstDayOfMonth,
##     data = data1[which(data1$PrimaryType == "NARCOTICS"), ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -27.2647  -4.7158  -0.4507   4.5005  28.9227
##
## Coefficients:
##                    Estimate Std. Error t value             Pr(>|t|)
## (Intercept)         43.3614     1.3097  33.107 < 0.0000000000000002 ***
## Pandemic           -26.1058     0.6686 -39.047 < 0.0000000000000002 ***
## MonthApril          -9.9996     1.4202  -7.041   0.0000000000055841 ***
## MonthMay            -7.7569     1.4084  -5.508   0.0000000554054719 ***
## MonthJune           -8.2648     1.4198  -5.821   0.0000000098230831 ***
## MonthJuly           -6.5011     1.4087  -4.615   0.0000048722449878 ***
## MonthAugust         -2.8425     1.4083  -2.018             0.044018 *
## MonthSeptember      -5.7462     1.4199  -4.047   0.0000591822928668 ***
## MonthOctober        -5.7856     1.4087  -4.107   0.0000460217277144 ***
## MonthNovember       -7.7053     1.4198  -5.427   0.0000000852479574 ***
## MonthDecember      -12.2975     1.7568  -7.000   0.0000000000073157 ***
## DayOfWeekMonday      2.6889     1.2101   2.222             0.026675 *
## DayOfWeekTuesday     4.0661     1.2145   3.348             0.000868 ***
## DayOfWeekWednesday   4.9807     1.2181   4.089   0.0000496211037259 ***
## DayOfWeekThursday    3.2064     1.2186   2.631             0.008743 **
```

```
## DayOfWeekFriday      9.4954     1.2140   7.822   0.0000000000000258 ***
## DayOfWeekSaturday    3.0974     1.2144   2.551             0.011019 *
## FirstDayOfMonth      2.1238     1.8306   1.160             0.246472
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.839 on 563 degrees of freedom
## Multiple R-squared:  0.7503, Adjusted R-squared:  0.7428
## F-statistic: 99.51 on 17 and 563 DF,  p-value: < 0.00000000000000022
```

Model overall is statistically significant ($p < .001$) with R-square being 0.7503. The model explains 75.03% of variability in daily narcotics crime rates. During the COVID pandemic, daily narcotics crime has decreased significantly by 26.1058 events ($p < .001$), controlling for other factors. The lowest daily rates are reported in December (12.2975 lower than March). Friday appears to have significantly higher daily rates compared to Sunday, 9.4954 higher.

THEFT crime

```
summary(model2)
```

```
##
## Call:
## lm(formula = NumOfEvents ~ Pandemic + Month + DayOfWeek + FirstDayOfMonth,
##     data = data1[which(data1$PrimaryType == "THEFT"), ])
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -86.64 -12.57  -0.66  10.84  73.48
##
## Coefficients:
##                    Estimate Std. Error t value             Pr(>|t|)
## (Intercept)         147.231      3.229  45.601 < 0.0000000000000002 ***
## Pandemic            -68.739      1.648 -41.707 < 0.0000000000000002 ***
## MonthApril          -10.230      3.501  -2.922              0.00362 **
## MonthMay              4.426      3.472   1.275              0.20292
## MonthJune            16.108      3.500   4.602    0.000005169657036 ***
## MonthJuly            25.825      3.473   7.437    0.000000000000387 ***
## MonthAugust          32.555      3.472   9.378 < 0.0000000000000002 ***
## MonthSeptember       21.219      3.500   6.062    0.000000002465564 ***
## MonthOctober         11.430      3.473   3.291              0.00106 **
## MonthNovember        -2.395      3.500  -0.684              0.49405
## MonthDecember        12.846      4.331   2.966              0.00314 **
## DayOfWeekMonday      18.065      2.983   6.056    0.000000002552292 ***
## DayOfWeekTuesday     15.570      2.994   5.201    0.000000278278099 ***
## DayOfWeekWednesday   18.569      3.003   6.184    0.000000001201331 ***
## DayOfWeekThursday    15.931      3.004   5.303    0.000000163895343 ***
## DayOfWeekFriday      26.730      2.993   8.932 < 0.0000000000000002 ***
## DayOfWeekSaturday    16.227      2.994   5.420    0.000000088347917 ***
## FirstDayOfMonth      14.631      4.513   3.242              0.00126 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 19.33 on 563 degrees of freedom
## Multiple R-squared:  0.7965, Adjusted R-squared:  0.7903
## F-statistic: 129.6 on 17 and 563 DF,  p-value: < 0.00000000000000022
```

Model overall is statistically significant ($p < .001$) with R-square being 0.7965. The model explains 79.65% of variability in daily theft crime rates. During the COVID pandemic, daily theft crime has decreased significantly by 68.739 events ($p < .001$), controlling for other factors. The lowest daily rates are reported in April (10.230 lower than March), while the highest are in August (32.555 higher than March). Friday appears to have significantly higher daily rates compared to Sunday, 26.730 higher. First day of the month is associated with 14.631 more thefts compared to other days of the month.

BATTERY crime

```
summary(model3)
```

```
##
## Call:
## lm(formula = NumOfEvents ~ Pandemic + Month + DayOfWeek + FirstDayOfMonth,
##     data = data1[which(data1$PrimaryType == "BATTERY"), ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -47.568 -12.455   0.297  11.119  63.188
##
## Coefficients:
##                    Estimate Std. Error t value             Pr(>|t|)
## (Intercept)         160.322      2.947  54.400 < 0.0000000000000002 ***
## Pandemic            -25.919      1.504 -17.229 < 0.0000000000000002 ***
## MonthApril          -10.324      3.196  -3.231             0.001307 **
## MonthMay             10.911      3.169   3.443             0.000618 ***
## MonthJune            21.980      3.195   6.880   0.0000000000159579 ***
## MonthJuly            24.191      3.170   7.632   0.0000000000000996 ***
## MonthAugust          16.165      3.169   5.101   0.0000004621431728 ***
## MonthSeptember        7.279      3.195   2.278             0.023095 *
## MonthOctober         -4.383      3.170  -1.383             0.167305
## MonthNovember       -11.852      3.195  -3.710             0.000228 ***
## MonthDecember       -12.117      3.953  -3.065             0.002279 **
## DayOfWeekMonday     -31.502      2.723 -11.570 < 0.0000000000000002 ***
## DayOfWeekTuesday    -37.859      2.733 -13.854 < 0.0000000000000002 ***
## DayOfWeekWednesday  -37.122      2.741 -13.544 < 0.0000000000000002 ***
## DayOfWeekThursday   -34.824      2.742 -12.700 < 0.0000000000000002 ***
## DayOfWeekFriday     -27.896      2.732 -10.213 < 0.0000000000000002 ***
## DayOfWeekSaturday    -8.237      2.733  -3.014             0.002692 **
## FirstDayOfMonth       1.776      4.119   0.431             0.666533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 17.64 on 563 degrees of freedom
## Multiple R-squared:  0.6271, Adjusted R-squared:  0.6159
## F-statistic:  55.7 on 17 and 563 DF,  p-value: < 0.00000000000000022
```

Model overall is statistically significant ($p < .001$) with R-square being 0.6271. The model explains 62.71% of variability in daily battery crime rates. During the COVID pandemic, daily battery crime has decreased significantly by 25.919 events ($p < .001$), controlling for other factors. The lowest daily rates are reported in December (12.117 lower than March), while the highest are in July (24.191 higher than March). Sunday has the highest daily rates of battery crime compared to all other days of the week, the lost rate are in Tuesday and Wednesday.

DECEPTIVE PRACTICE

```
summary(model4)
```

```
##
## Call:
## lm(formula = NumOfEvents ~ Pandemic + Month + DayOfWeek + FirstDayOfMonth,
##     data = data1[which(data1$PrimaryType == "DECEPTIVE PRACTICE"),
##         ])
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -34.175  -6.555  -0.532   5.921  57.881
##
## Coefficients:
##                     Estimate Std. Error t value             Pr(>|t|)
## (Intercept)         34.74605    1.87838  18.498 < 0.0000000000000002 ***
## Pandemic           -11.29961    0.95885 -11.785 < 0.0000000000000002 ***
## MonthApril          -0.78095    2.03682  -0.383               0.7016
## MonthMay            -1.58267    2.01990  -0.784               0.4336
## MonthJune           -0.21346    2.03621  -0.105               0.9165
## MonthJuly            3.92034    2.02029   1.940               0.0528 .
## MonthAugust          0.84557    2.01970   0.419               0.6756
## MonthSeptember       0.02959    2.03641   0.015               0.9884
## MonthOctober        11.67277    2.02029   5.778        0.0000000125 ***
## MonthNovember        9.09538    2.03621   4.467        0.0000095985 ***
## MonthDecember        0.23851    2.51955   0.095               0.9246
## DayOfWeekMonday     17.86657    1.73546  10.295 < 0.0000000000000002 ***
## DayOfWeekTuesday    17.53037    1.74175  10.065 < 0.0000000000000002 ***
## DayOfWeekWednesday  15.88951    1.74688   9.096 < 0.0000000000000002 ***
## DayOfWeekThursday   16.63276    1.74771   9.517 < 0.0000000000000002 ***
## DayOfWeekFriday     19.00006    1.74102  10.913 < 0.0000000000000002 ***
## DayOfWeekSaturday    4.47259    1.74169   2.568               0.0105 *
## FirstDayOfMonth     31.51421    2.62541  12.004 < 0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.24 on 563 degrees of freedom
## Multiple R-squared:  0.5143, Adjusted R-squared:  0.4997
## F-statistic: 35.07 on 17 and 563 DF,  p-value: < 0.00000000000000022
```

Model overall is statistically significant ($p < .001$) with R-square being 0.5143. The model explains 51.43% of variability in daily deceptive practice crime rates. During the COVID pandemic, daily deceptive practice crime has decreased significantly by 11.29961 events ($p < .001$), controlling for other factors. The highest

daily rates are reported in October (11.67277 higher than March). Friday has the highest daily rates of deceptive practice compared to Sunday (19.00006 higher). First day of the month is associated with 31.51421 higher number of deceptive practice crimes.