Chicago Domestic Crimes During Pandemic

Samuel, Natalie, & Stacy

INTRODUCTION AND DATA

-introduce your general research question and your data (where it came from, how it was collected, what are the cases, what are the variables, etc.).

METHODOLOGY

-variables used to address your research question -useful visualizations or summary statistics -introduce and justify the statistical method(s) that you believe will be useful in answering your research question.

```
domvio<-domvio %>%
  mutate(MONTH = as.integer(substring(DATE..OF.OCCURRENCE, 0, 2))) %>%
  mutate(DAY = as.integer(substring(DATE..OF.OCCURRENCE, 4,5))) %>%
  mutate(YEAR = as.integer(substring(DATE..OF.OCCURRENCE, 7,10)))

#here's what the data looks like
glimpse(domvio)
```

```
## Rows: 231,002
## Columns: 20
## $ CASE.                  <chr> "JD163753", "JD212847", "JC497784", "JC459410...
## $ DATE..OF.OCCURRENCE    <chr> "02/24/2020 08:15:00 PM", "04/10/2020 10:56:0...
## $ BLOCK                  <chr> "031XX W LEXINGTON ST", "005XX W 103RD ST", "...
## $ IUCR                   <chr> "1153", "0560", "0860", "0560", "0810", "0820...
## $ PRIMARY.DESCRIPTION    <chr> "DECEPTIVE PRACTICE", "ASSAULT", "THEFT", "AS...
## $ SECONDARY.DESCRIPTION <chr> "FINANCIAL IDENTITY THEFT OVER $ 300", "SIMPL...
## $ LOCATION.DESCRIPTION   <chr> "", "RESIDENCE", "DEPARTMENT STORE", "SIDEWAL...
## $ ARREST                 <chr> "N", "N", "N", "N", "N", "N", "N", "N", "N", ...
## $ DOMESTIC               <chr> "N", "N", "N", "N", "N", "N", "N", "N", "N", ...
## $ BEAT                   <int> 1134, 2232, 1924, 122, 123, 2433, 312, 914, 3...
## $ WARD                   <int> 24, 9, 44, 4, 25, 48, 20, 11, 5, 26, 27, 37, ...
## $ FBI.CD                 <chr> "11", "08A", "06", "08A", "06", "06", "08A", ...
## $ X.COORDINATE           <int> NA, 1174583, NA, NA, NA, NA, 1180030, 1171590...
## $ Y.COORDINATE           <int> NA, 1836593, NA, NA, NA, NA, 1862317, 1887793...
## $ LATITUDE               <dbl> NA, 41.70700, NA, NA, NA, NA, 41.77747, 41.84...
## $ LONGITUDE              <dbl> NA, -87.63629, NA, NA, NA, NA, -87.61556, -87...
## $ LOCATION               <chr> "", "(41.707000821, -87.636288063)", "", "", ...
## $ MONTH                  <int> 2, 4, 11, 10, 5, 12, 5, 5, 4, 5, 4, 5, 5, 5, ...
## $ DAY                    <int> 24, 10, 3, 4, 24, 5, 7, 3, 28, 7, 25, 7, 7, 5...
## $ YEAR                   <int> 2020, 2020, 2019, 2019, 2020, 2019, 2020, 202...
```

```
domvio_mut<-domvio
```

```
#create an indicator for being before covid entirely
#this indicator is "isbeforecovid" 2019/7/10-2020/1/24
domvio_mut<-domvio_mut %>%
  mutate(isbeforecovid = case_when(
    YEAR == 2019 ~ 1,
    YEAR == 2020 & MONTH == 1 & DAY <=24 ~ 1,
    YEAR == 2020 & MONTH == 1 & DAY > 24 ~ 0,
    YEAR == 2020 & MONTH > 1 ~ 0,
    ))
```

```
#create an indicator for being before lockdown (tho covid is in the USA)
#this indicator is "isprelockdown" 2020/1/25-2020/3/21
domvio_mut<-domvio_mut %>%
  mutate(isprelockdown = case_when(
    YEAR == 2019 ~ 0,
    YEAR == 2020 & MONTH == 1 & DAY <=24 ~ 0,
    YEAR == 2020 & MONTH == 1 & DAY > 24 ~ 1,
    YEAR == 2020 & MONTH == 2  ~ 1,
    YEAR == 2020 & MONTH == 3 & DAY < 22 ~ 1,
    YEAR == 2020 & MONTH == 3 & DAY >= 22 ~ 0,
    YEAR == 2020 & MONTH > 3 ~ 0
    ))
```

```
#create an indicator for being in the lockdown
#this indicator is "islockdown" 2020/3/22-2020/4/30
domvio_mut<-domvio_mut %>%
  mutate(islockdown = case_when(
    YEAR == 2019 ~ 0,
    YEAR == 2020 & MONTH < 3 ~ 0,
    YEAR == 2020 & MONTH == 3 & DAY < 22 ~ 0,
    YEAR == 2020 & MONTH == 3 & DAY >= 22 ~ 1,
    YEAR == 2020 & MONTH == 4 ~ 1,
    YEAR == 2020 & MONTH >= 5 ~ 0
    ))
```

```
#create an indicator for phase 2 in the recovery
#this indicator is "isphase2" 2020/5/1-2020/6/2
domvio_mut<-domvio_mut %>%
  mutate(isphase2 = case_when(
    YEAR == 2019 ~ 0,
    YEAR == 2020 & MONTH < 5 ~ 0,
    YEAR == 2020 & MONTH == 5 ~ 1,
    YEAR == 2020 & MONTH == 6 & DAY < 3 ~ 1,
    YEAR == 2020 & MONTH == 6 & DAY >= 3 ~ 0,
    YEAR == 2020 & MONTH >= 7 ~ 0
    ))
```

```
#create an indicator for phase 3 in the recovery
#this indicator is "isphase3" 2020/6/3-2020/6/25
domvio_mut<-domvio_mut %>%
  mutate(isphase3 = case_when(
    YEAR == 2019 ~ 0,
    YEAR == 2020 & MONTH < 6 ~ 0,
    YEAR == 2020 & MONTH == 6 & DAY < 3 ~ 0,
    YEAR == 2020 & MONTH == 6 & DAY >= 3 & DAY <= 25 ~ 1,
```

```
    YEAR == 2020 & MONTH == 6 & DAY > 25 ~ 0,
    YEAR == 2020 & MONTH >= 7 ~ 0
    ))
```

```
#create an indicator for phase 4 in the recovery
#this indicator is "isphase4" 2020/6/26-2020/7/8 (last day on dataset)
domvio_mut<-domvio_mut %>%
  mutate(isphase4 = case_when(
    YEAR == 2019 ~ 0,
    YEAR == 2020 & MONTH < 6 ~ 0,
    YEAR == 2020 & MONTH == 6 & DAY < 26 ~ 0,
    YEAR == 2020 & MONTH == 6 & DAY >= 26 ~ 1,
    YEAR == 2020 & MONTH >= 7 ~ 1
    ))
```

```
#find which labels have to do with domestic violence
domvio_cat<-domvio_mut %>%
  group_by(SECONDARY.DESCRIPTION) %>%
  summarise(n=n())
domvio_cat
```

```
## # A tibble: 430 x 2
##    SECONDARY.DESCRIPTION               n
##    <chr>                           <int>
##  1 $500 AND UNDER                  20888
##  2 ABUSE / NEGLECT - CARE FACILITY     5
##  3 ABUSE/NEGLECT: CARE FACILITY       10
##  4 AGG CRIM SEX ABUSE FAM MEMBER      78
##  5 AGG CRIMINAL SEXUAL ABUSE          63
##  6 AGG PO HANDS ETC SERIOUS INJ        8
##  7 AGG PO HANDS NO/MIN INJURY        593
##  8 AGG PRO EMP HANDS SERIOUS INJ      16
##  9 AGG PRO.EMP: HANDGUN               26
## 10 AGG PRO.EMP: OTHER DANG WEAPON     85
## # ... with 420 more rows
```

```
#create an indicator for cases of dom violence
#this indicator is "isdomviolence"
domvio_mut<-domvio_mut %>%
  mutate(isdomviolence = case_when(
    SECONDARY.DESCRIPTION == "VIOLATION OF BAIL BOND - DOMESTIC VIOLENCE" |
    SECONDARY.DESCRIPTION == "DOMESTIC BATTERY SIMPLE" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: OTHER DANGEROUS WEAPON" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: OTHER DANG WEAPON" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: KNIFE/CUTTING INST" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: KNIFE / CUTTING INSTSTRUMENT" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: HANDS/FIST/FEET SERIOUS INJURY" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY: HANDGUN" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY - OTHER FIREARM" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY - OTHER DANGEROUS WEAPON" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY - KNIFE / CUTTING INSTRUMENT" |
    SECONDARY.DESCRIPTION == "AGGRAVATED DOMESTIC BATTERY - HANDGUN" |
    SECONDARY.DESCRIPTION == "AGG. DOMESTIC BATTERY - HANDS, FISTS, FEET, SERIOUS INJURY" ~ 1,
    SECONDARY.DESCRIPTION != "VIOLATION OF BAIL BOND - DOMESTIC VIOLENCE" &
    SECONDARY.DESCRIPTION != "DOMESTIC BATTERY SIMPLE" &
```

```
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: OTHER DANGEROUS WEAPON" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: OTHER DANG WEAPON" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: KNIFE/CUTTING INST" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: KNIFE / CUTTING INSTSTRUMENT" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: HANDS/FIST/FEET SERIOUS INJURY" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY: HANDGUN" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY - OTHER FIREARM" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY - OTHER DANGEROUS WEAPON" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY - KNIFE / CUTTING INSTRUMENT" &
    SECONDARY.DESCRIPTION != "AGGRAVATED DOMESTIC BATTERY - HANDGUN" &
    SECONDARY.DESCRIPTION != "AGG. DOMESTIC BATTERY - HANDS, FISTS, FEET, SERIOUS INJURY" ~ 0
    ))
```

```
#if you are wondering about how the actual data looks like there it is

#number of observations during in 2020 pre covid=140697
domvio_mut %>%
  count (isbeforecovid)
```

```
##   isbeforecovid      n
## 1             0  90305
## 2             1 140697
```

```
#number of observations during prelockdown=34619
domvio_mut %>%
  count (isprelockdown)
```

```
##   isprelockdown      n
## 1             0 196383
## 2             1  34619
```

```
#number of observations during lockdown=16842
domvio_mut %>%
  count (islockdown)
```

```
##   islockdown      n
## 1          0 214160
## 2          1  16842
```

```
#number of observations during p2=18578
domvio_mut %>%
  count (isphase2)
```

```
##   isphase2      n
## 1        0 212424
## 2        1  18578
```

```
#number of observations during p3=12722
domvio_mut %>%
  count (isphase3)
```

```
##   isphase3      n
## 1        0 218280
## 2        1  12722
```

```
#number of observations during p4=7544
domvio_mut %>%
  count(isphase4)
```

```
##   isphase4      n
## 1        0 223458
## 2        1   7544
```

```
#number of cases of domestic violence=24897
domvio_mut %>%
  count(isdomviolence)
```

```
##   isdomviolence      n
## 1             0 206105
## 2             1  24897
```

```
#see number of cases of domestic violence on a given day
domvio_mut %>%
  filter(YEAR==2019) %>%
  filter(MONTH==7) %>%
  filter(DAY==9) %>%
  count(isdomviolence)
```

```
## [1] isdomviolence n
## <0 rows> (or 0-length row.names)
```

```
#observations of domestic violence in a given month
domvio_mut %>%
  filter(YEAR==2019) %>%
  filter(MONTH==07) %>%
  count(isdomviolence)
```

```
##   isdomviolence     n
## 1             0 15502
## 2             1  1842
```

```
domvio_mut<-domvio_mut %>%
  mutate(chisquare_indicators= case_when(isbeforecovid == 1 ~ "pre",
                                         isprelockdown == 1 ~ "pre",
                                         islockdown == 1 ~ "lockdown",
                                         isphase2 == 1 ~ "lockdown",
                                         isphase3 == 1 ~ "post",
                                         isphase4 == 1 ~ "post"
                                         ) )
```

RESULTS

-Showcase how -Provide the main results from your analysis

```
table(domvio_mut$isdomviolence, domvio_mut$chisquare_indicators)
```

Chi-Square Test

```
##
##     lockdown   post    pre
##   0    31109  17599 157397
##   1     4311   2667  17919
```

```
chisq.test(table(domvio_mut$isdomviolence, domvio_mut$chisquare_indicators))
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(domvio_mut$isdomviolence, domvio_mut$chisquare_indicators)
## X-squared = 247.63, df = 2, p-value < 2.2e-16
```

Since the data (Table 1) satisfies the independent sampling assumption and is large enough (i.e. each cell $>$ 10), we will be performing a chi-square test at the $\alpha = 0.05$ significance level. We test the two hypotheses below:

H0 : The frequency of domestic violence cases in Chicago is unrelated to the phases of the pandemic. H1 : The frequency of domestic violence cases in Chicago is related to the phases of the pandemic.

Under the null hypothesis, our test statistic has a chi-square distribution with 2 degrees of freedom. We performed the test and obtained a chi-square value of 247.63, which corresponds to a p-value of < 0.001. Thus, at an $\alpha = 0.05$ significance level, we reject the null hypothesis; there is sufficient evidence to suggest that the frequency of domestic violence cases in Chicago is related to the phases of the pandemic.

Step Down 2 Proportion Z-Tests -with Bonferroni correction

Regression Analysis

DISCUSSION

-summary of what you have learned about your research question along -statistical arguments supporting your conclusions -critique your own methods and provide suggestions for improving your analysis (Issues pertaining to the reliability and validity of your data and appropriateness of the statistical analysis) -what you would do differently -what you would do next if you were going to continue work on the project