R Code:

```
library(mapdata)
library(maps)
library(ggplot2)
library(readr)
library(dplyr)
library(tidyverse)
library(gganimate)

chicago_crime <- read.csv("Crimes2009to2019.csv")
mydata <- chicago_crime %>%
  select(Year, Primary.Type, Longitude, Latitude) %>%
  filter(Year %in% c(2010,2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019)) %>
%
  filter(!is.na(Longitude)) %>%
  filter(!is.na(Latitude)) %>%
  filter(Primary.Type %in% c("WEAPONS VIOLATION", "HOMICIDE"))


#illinois plot
states <- map_data("state")
chicago <- subset(states,region == "illinois")
chicago_plot <- ggplot(chicago, aes(long, lat, group = group)) +
  coord_fixed(1) +
  geom_polygon(color = 'black', fill = 'gray')
chicago_plot

#illinois by county plot
counties <- map_data("county")
illi_counties <- subset(counties, region = 'illinois')
chicago_plot + geom_polygon(data = illi_counties, fill = NA, color = 'white') +
  geom_polygon(fill = NA, color = 'black') + theme_bw() +
  scale_y_continuous(limits = c(36, 43))+
  scale_x_continuous(limits = c(-92.5, -86))

#cook county plot
s <- chicago_plot + geom_polygon(data = illi_counties, fill = NA, color = 'white')
+
  geom_polygon(fill = NA, color = 'black') +
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2)) +
  theme_bw()

#chicago crime plot 2009-2019
crimes_plot <-
  s +
geom_point(data= mydata, aes(Longitude, Latitude),pch=0, size=.5, color = 'blue',
inherit.aes = FALSE) +
          theme_bw()
crimes_plot

#WEAPONS VIOLATION plot 2009 - 2019
weapon_plot <-
  chicago_plot +
  geom_polygon(data = illi_counties, fill = NA, color = 'white') +
  geom_polygon(fill = NA, color = 'black')+
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2))+
  geom_point(data= filter(mydata, Primary.Type=="WEAPONS VIOLATION"),
aes(Longitude, Latitude),pch=0, size=.5, color = 'red', inherit.aes = FALSE) +
  theme_bw()
```

```
weapon_plot

#HOMICIDE plot 2009-2019
homicide_plot <-
  chicago_plot +
  geom_polygon(data = illi_counties, fill = NA, color = 'white') +
  geom_polygon(fill = NA, color = 'black')+
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2))+
  geom_point(data= filter(mydata, Primary.Type=="HOMICIDE"), aes(Longitude,
Latitude),pch=0, size=.5, color = 'purple', inherit.aes = FALSE) +
  theme_bw()
homicide_plot


#Animation crimes year-by-year
memory.size(max = FALSE)
memory.limit(size = 56000)
mydata1 <- mydata %>%
  filter(Year %in% c(2010,2011,2012,2013,2014))

crime_plot_by_year <-  chicago_plot +
  geom_polygon(data=illi_counties, fill=NA,color="white")+
  geom_polygon(fill = NA, color = 'black')+
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2))+
  geom_point(data= mydata, aes(Longitude, Latitude),pch=0, size=.5, color = 'blue',
inherit.aes = FALSE)+
  labs(title = "Chicago Crimes in: {closest_state}")+
  transition_states(Year, transition_length = 10, state_length = 20)+
  theme_bw()
crimes_plot_anim <- animate(crime_plot_by_year)
crimes_plot_anim
anim_save("Animation Crimes 2010-2019.gif")

#Animation WEAPONS VIOLATION year-by-year
weapon_plot_by_year <-  chicago_plot +
  geom_polygon(data=illi_counties, fill=NA,color="white")+
  geom_polygon(fill = NA, color = 'black')+
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2))+
  geom_point(data= filter(mydata, Primary.Type=="WEAPONS VIOLATION"),
aes(Longitude, Latitude),pch=0, size=.5, color = 'red', inherit.aes = FALSE)+
  labs(title = "Chicago Weapons Violation in: {closest_state}")+
  transition_states(Year, transition_length = 20, state_length = 30)+
  theme_bw()
weapon_plot_anim <- animate(weapon_plot_by_year)
weapon_plot_anim
anim_save("Chicago Weapons Violation 2010-2019.gif")
#Animation HOMICIDE year-by-year
homicide_plot_by_year <-  chicago_plot +
  geom_polygon(data=illi_counties, fill=NA,color="white")+
  geom_polygon(fill = NA, color = 'black')+
  coord_cartesian(xlim = c(xlim=-88.3, -87.5), ylim = c(41.4, 42.2))+
  geom_point(data= filter(mydata, Primary.Type=="HOMICIDE"), aes(Longitude,
Latitude),pch=0, size=.5, color = 'purple', inherit.aes = FALSE)+
  labs(title = "Homicide in: {closest_state}")+
  transition_states(Year, transition_length = 20, state_length = 30)+
  theme_bw()
homicide_plot_anim <- animate(homicide_plot_by_year)
homicide_plot_anim
anim_save("Chicago Homicide 2010-2019.gif")
```

```
#Heatmap for time
crime_df <-
  chicago_crime %>%
  filter(Year %in% c(2010,2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019)) %>
%
  filter(Primary.Type %in% c("WEAPONS VIOLATION", "HOMICIDE")) %>%
  mutate(
    DateTime = mdy_hms(as.character(Date)),
    Date_rw = as.Date.POSIXct(DateTime),
    Year = as.ordered(Year),
    Month = month(DateTime, label = T),
    DayofMonth = as.ordered(day(DateTime)),
    DayofWeek = wday(DateTime, label = T),
    Hour = as.ordered(hour(DateTime))
  )

heatmap_data <-
  crime_df %>%
  dplyr::select(DayofWeek, Hour) %>%
  mutate(Hour = factor(format(strptime(Hour, format='%H'), '%I%p'), levels =
format(strptime(0:23, format='%H'), '%I%p'))) %>%
  table()

heatmap(heatmap_data,
        Colv = NA,
        Rowv = NA,
        scale = "column")
#Heatmap with Chicago crime map
library(tmaptools)
library(sf)
library(leaflet)
library(lubridate)
library(rgdal)
library(raster)

mydummy <- chicago_crime %>%
  select(Year, Primary.Type, Longitude, Latitude, Community.Area) %>%
  filter(Year %in% c(2010,2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019)) %>
%
  filter(!is.na(Longitude)) %>%
  filter(!is.na(Latitude)) %>%
  filter(!is.na(Community.Area)) %>%
  filter(Primary.Type %in% c("WEAPONS VIOLATION", "HOMICIDE"))
mydummy$Crime.Type <- "Violent"


CommunityArea_data <-
  mydummy %>%
  group_by(Community.Area, Crime.Type) %>%
  na.omit() %>%
  summarize(Reported_incidents=n()) %>%
  spread(Crime.Type, Reported_incidents) %>%
  ungroup() %>%
  mutate(Violent_incident_rank = rank(-Violent))

chicago_shp <- st_read("geo_export_b5ad8297-56bf-4b79-b209-466c5a2d213b.shp")
chicago_shp$area_num_1 <- as.numeric(chicago_shp$area_num_1)
chicago_shp <- chicago_shp %>% left_join(CommunityArea_data, by = c("area_num_1" =
"Community.Area"))
```

```
violent_palette <- colorNumeric(palette = "Reds", domain = chicago_shp$Violent)

tooltips <- paste0(chicago_shp$area_num_1, " - Violent incidents: ",
formatC(chicago_shp$Violent, big.mark = ","),
                   " - Rank: ", chicago_shp$Violent_incident_rank)

chicago_map <- st_transform(chicago_shp, "+proj=longlat +datum=WGS84")

leaflet(chicago_map) %>%
  addProviderTiles("CartoDB.PositronNoLabels") %>%
  addPolygons(stroke = F,
              smoothFactor = 0.02,
              fillOpacity = 0.6,
              popup = tooltips,
              color = ~violent_palette(chicago_shp$Violent))
```

SAS CODE:

```
libname mydir "/home/u49456429/Project";

proc import out = mydir.chicrime
datafile="/home/u49456429/Project/Crimes2009to2019.csv" replace dbms=csv;
getnames=yes;
run;

data chicrime;
set mydir.chicrime;
WHere Year between 2010 and 2019;
run;

DATA mydata(DROP = ID Description "Location Description"n Arrest Domestic "X
Coordinate"n  "Y Coordinate"n "Case Number"n
Block IUCR Location Beat District Ward "Community Area"n "FBI Code"n "Updated
On"n);
SET chicrime;
datevar = datepart(date);
FORMAT datevar ddmmyy10.;
RUN;

proc contents data = mydata;

proc gchart data = mydata;
hbar "Primary Type"n;

proc freq data = mydata;
table "Location Description"n;

data dummy;
set mydata;
Where "Primary Type"n = "HOMICIDE" or "Primary Type"n = "WEAPONS VIOLATION";
month = datevar;
format month month.;
run;

proc sort data = dummy;
by "primary Type"n;
proc freq data = dummy;
table datevar/ out=very_dummy;
by "Primary Type"n;
run;

proc print data = very_dummy;

proc means data =

proc freq data = dummy;
table month/out=month_chart;

proc print data = month_chart;


proc gchart data = month_chart;
vbar month/ freq=COUNT discrete;
run;
```

```
data super_dummy;
set very_dummy;
monthyear = datevar;
Format monthyear monyy7.;
run;

proc means data = super_dummy;
var year;

proc freq data = super_dummy;
table monthyear*"Primary Type"n/out = testing2;
weight COUNT;
run;

proc transpose data = testing2 out = megamonth (DROP=_NAME_ _LABEL_);
by monthyear;
id "Primary Type"n;
var COUNT;
run;

Proc print data = megamonth;

data anova;
set megamonth;
year = year(monthyear);
crime = homicide + "WEAPONS VIOLATION"n;
RUN;

proc glm data = anova;
class year;
model crime = year;
means year;
run;

symbol1 color = black value = dot H=1 interpol=Join;
proc gplot data = anova;
plot crime*monthyear;

Symbol1 color = blue value=dot H=1 interpol=Join;
Symbol2 color = red value = dot H=1 interpol= Join;
proc gplot data = anova;
plot "WEAPONS VIOLATION"n*monthyear "HOMICIDE"n*monthyear/overlay legend;
run;

proc reg data = megamonth;
model "WEAPONS VIOLATION"n = monthyear;


proc esm data = megamonth out= nextyear print=forecasts;
id monthyear interval= month;
forecast _numeric_;
run;
title1 "Chicago Crime";
proc sgplot data=nextyear;
   series x=monthyear y="WEAPONS VIOLATION"n / markers
                           markerattrs=(symbol=circlefilled color=red)
                           lineattrs=(color=red);
   series x=monthyear y="HOMICIDE"n / markers
                           markerattrs=(symbol=circlefilled color=blue)
```

```
                          lineattrs=(color=blue);
   *xaxis values=('JAN2010'd to 'DEC2019'd by month);
   *yaxis values=(0 to 50 by 1) minor label='Websites';
run;

ods graphics on;
proc arima data = megamonth plots(only) = (forecast(forecast));
identify var="WEAPONS VIOLATION"n crosscorr=monthyear nlag=12;
estimate p = (3,5,10,14) q = 1 noint method=ml;
forecast lead=12 interval=month id = monthyear ;
title "Time Series for Weapons Violation";
run;
quit;

*auto correlated and the time series is not white noise
the time series is predictable
There is no difference between fitted model and data

ods graphics on;
proc arima data = megamonth plots(only) = (forecast(forecast));
identify var=HOMICIDE crosscorr=monthyear;

estimate P =1 Q = 1;
forecast lead=12 interval=month id = monthyear ;
title "Time Series for Homicide";
run;
quit;

data weapon;
set megamonth;
weapons = "WEAPONS VIOLATION"n;
run;


%let alpha=0.1;
/*Using PROC ARIMA*/
ods output "Autocorrelation Check for White Noise" = autocorr;
proc arima data=weapon;
identify var= weapons p=(0:7) q=(0:7) alpha=&alpha;
run;
quit;
ods output close;
proc print data = autocorr;
%macro checkacf;
data guicheck;
 format stddev comma12.8;
 set autocorr;
 stddev=1.96*stderr;
 if _n_=2 and correlation>stddev then do;
 call symput('weapons','weapons(1)');
 call symput('q','1:7');
 call symput('p','1:7');
 end;
 else if _n_=2 then do;
 call symput('weapons', 'weapons');
 call symput('q', '0:7');
 call symput('p', '0:7');
 end;
run;
```

```
%if %substr(&weapons,7,1)=1 %then %do;
ods output "Autocorrelation Check for White Noise" = autocorr1;
proc arima data=weapon;
identify var=&weapons minic p=(&p) q=(&q) alpha=&alpha;
run;
quit;
ods output close;
data _null_;
 set autocorr;
 if _n_=2 then call symput('stderr1', stderr);
run;
%let stddev1=%sysevalf(1.96*&stderr1);
data guicheck1;
 format stddev1 comma12.8;
 set autocorr1;
 stddev1=&stddev1;
 if _n_=2 and correlation>stddev1 then do;
 call symput("errormsg", "Failed");
 end;
 else call symput('errormsg', 'Passed');
run;
%if &errormsg=Failed %then %do;
%put ERROR: Weapons Unable to Pass Stationarity Test. DO NOT MODEL!!!!!!!! :-);
%return;
%end;
%end;
/*checking if original series is white noise or not*/
ods output "Autocorrelation Check for White Noise" = acfwn;
proc arima data=weapon;
identify var=&weapons minic p=(&p) q=(&q) alpha=&alpha;
run;
quit;
ods output close;
data _null_;
 set acfwn;
 if _n_=1 then call symput ('probchisq',probchisq);
run;
%if &probchisq<=&alpha %then %do;
%put ERROR: Series is White Noise. DO NOT MODEL!!!!!! :-);
%return;
%end;
/*ods output "ARMA(p+d,q) Tentative Order Selection Tests" = pdqtest;*/
/*proc arima data=megamonth;*/
/*identify var=&weapon minic esacf p=(&p) q=(&q) scan alpha=&alpha;*/
/*run;*/
/*quit;*/
/*ods output close;*/
ods output "Minimum Information Criterion"= minicdata;
proc arima data=weapon;
identify var=&weapons minic p=(&p) q=(&q) alpha=&alpha;
run;
quit;
ods output close;
data minicdata;
set minicdata;
/*want X_t based on at least 2 past observations*/
if rowname in('AR 0','AR 1') then delete;
run;
proc iml;
```

```
use minicdata;
read all into pqmat[rowname=rowname colname=ma_];
/*get number of rows and columns and calculate the minimum of the matrix.
This will be used in DO loop*/
nrow=nrow(pqmat);
ncol=ncol(pqmat);
act_min=min(pqmat);
/*display matrix, nrow, and ncol*/
print pqmat nrow ncol;
do i=1 to nrow;
do j=1 to ncol;
if pqmat[i,j]=act_min then call symput('i',left(char(i)));
if pqmat[i,j]=act_min then call symput('j',left(char(j)));
end;
end;
quit;
/*these steps are done to get the correct p and q*/
%let p1=%eval(&i+1);
%let q1=&j;
%put p1=&p1;
%put q1=&q1;
ods output "Outlier Details" = outliers;
proc arima data=weapon;
identify var=&weapons minic p=(&p) q=(&q) alpha=&alpha;
estimate p=&p1 q=&q1;
forecast alpha=&alpha lead=12 interval=month id=monthyear out=weapon1;
/*forecast for last 5 months of 2011*/
outlier maxnum=12 alpha=&alpha;
run;
quit;
ods output close;
%mend checkacf;
%checkacf;
/*calculate R^2 from linear regression*/
proc means data=weapon noprint;
 var weapons residual;
 output out=c css(weapons)=sst uss(residual)=sse;
 run;
 data c;
 set c;
 R_square = 1-(sse/sst);
 run;
 proc print data=c noobs;
 var R_square;
 run;
data weapon1;
set weapon1;
obs=_n_;
run;
proc sort data=outliers;
by obs;
run;
/*Using PROC REG (Time Trend Regression)*/
data ttrend;
set weapon1;
obs+1;
run;
proc reg data=ttrend;
model weapons = obs;
```

```
output out=weapon2 p=ptrend rstudent=rstud;
run;
quit;
data weapon2;
merge weapon2 outliers ttrend;
by obs;
run;
data weapon2;
set weapon2;
 if type ne '' then do;
 prob=weapons;
 probdate=monthyear;
 end;
format probdate monname.;
run;
%let bound=%sysevalf(100-100*&alpha);
%put &bound;
%let lbound=L&bound;
%let ubound=U&bound;
data final;
set weapon1(keep=monthyear weapons forecast &lbound &ubound rename=(weapons=actual
forecast=parima));
set weapon2(keep=monthyear ptrend prob probdate);
run;
data _null_;
set final end=last;
 where actual^=.;
if _n_=1 then do;
 call symputx('mon1',put(monthyear,monname.));
 call symputx('year1',put(monthyear,year.));
end;
if last then do;
 call symputx('mon2',put(monthyear,monname.));
 call symputx('year2',put(monthyear,year.));
end;
run;
%let dates=01%substr(&mon2,1,3)%substr(&year2,3,2);
goptions reset=all;
title "Weapons Violation from &mon1 &year1 to &mon2 &year2, for M3 Data";
symbol1 value=point pointlabel=(position=middle '#probdate' color=red
nodropcollisions) color=red;
symbol2 i=spline width=2 v=none c=black; /* for actual */
symbol3 i=spline width=2 v=none c=green; /* for parima forecast */
symbol4 i=spline width=2 v=none c=blue; /* for ptrend forecast */
symbol5 i=spline width=1 v=none c=grey l=20; /* &lbound for arima forecast,
dashed line */
symbol6 i=spline width=1 v=none c=grey l=20; /* &ubound for arima forecast,
dashed line */
axis1 label=('Year') minor=none;
axis2 label=(angle=90 'Weapons Violation');
legend1 across=1
cborder=black
position=(bottom inside right)
offset=(-2,0)
value=(tick=1 'Outlier' tick=2 'OBSERVED DATA' tick=3 'PROC ARIMA' tick=4 'TIME
TREND' tick=5 "&lbound" tick=6 "&ubound")
mode=protect label=none;
run;
proc gplot data=final;
```

```
format monthyear year4.;
plot prob*monthyear=1 actual*monthyear=2 parima*monthyear=3 ptrend*monthyear=4
&lbound*monthyear=5 &ubound*monthyear=6/ overlay noframe
href="&dates"d
chref=purple
vaxis=axis2
vminor=1
haxis=axis1
legend=legend1;
run;
quit;
%let alpha=.1;
goptions reset=all;
symbol1 width=1 v=dot c=black;
/*dm output 'clear'; */
dm odsresults 'clear';
/*histogram of residuals should be normally distributed*/
ods output "Tests For Normality"=tfn;
proc univariate data=weapon1 normal;
var residual;
qqplot residual / normal(mu=est sigma=est color=blue);
histogram / kernel;
run;
ods output close;
data tfn;
set tfn;
if pvalue>=&alpha then put "Normal residuals by " test;
if pvalue<&alpha then put "Nonnormal residuals by " test;
run;
```