Region Comparisons

Laura.w

2/25/2020

```
library(readxl)

library(readr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

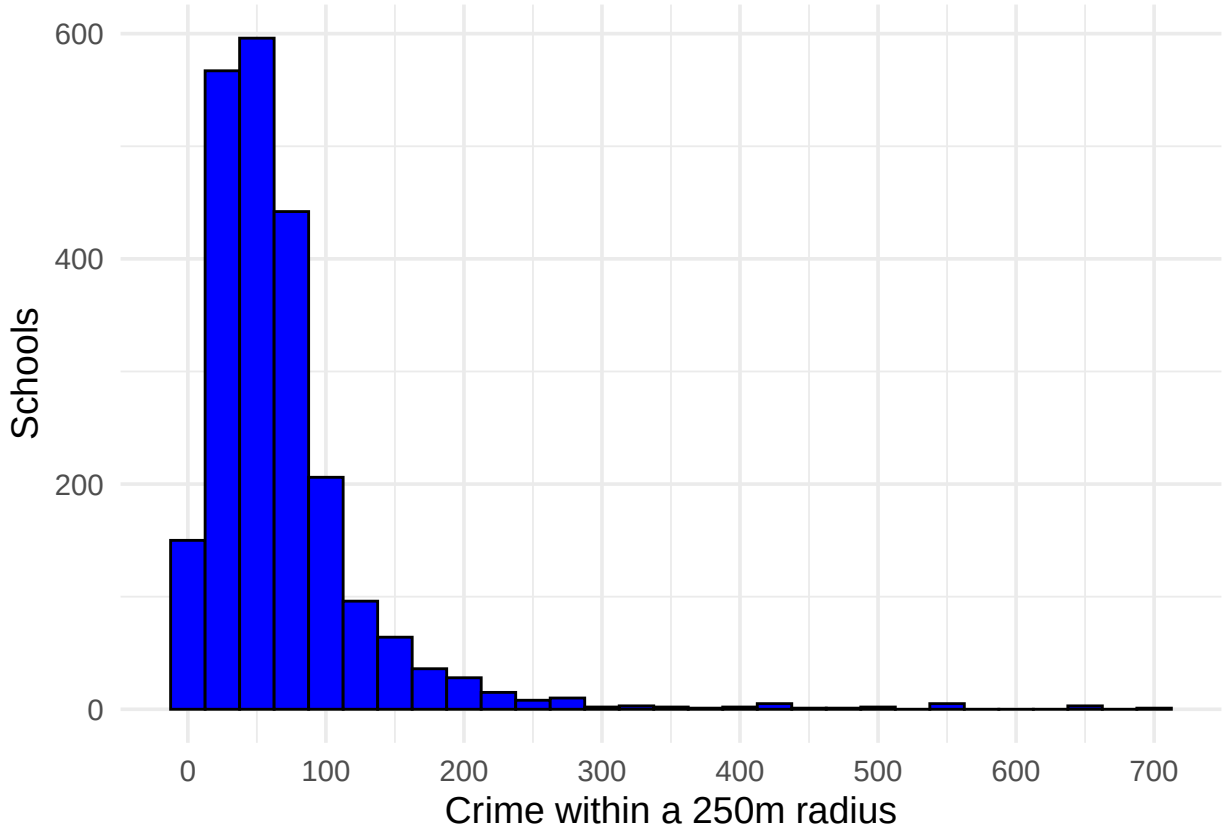
```
## -- Attaching packages ---------------------------------------------------------------- tidyverse 1.3.
```

```
## v ggplot2 3.2.1     v purrr   0.3.3
## v tibble  2.1.3     v stringr 1.4.0
## v tidyr   1.0.2     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------ tidyverse_conflicts()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
data_250 <- read_xlsx('C://Users//LW//Box//Mexico City 2020//data//buffer_250_crime1_TableToExcel.xlsx')

data_150 <- read_xlsx('C://Users//LW//Box//Mexico City 2020//data//buffer_150_crime2.xlsx')

data_100 <- read_xlsx('C://Users//LW//Box//Mexico City 2020//data//buffer_100_crime2_TableToExcel.xlsx')
```

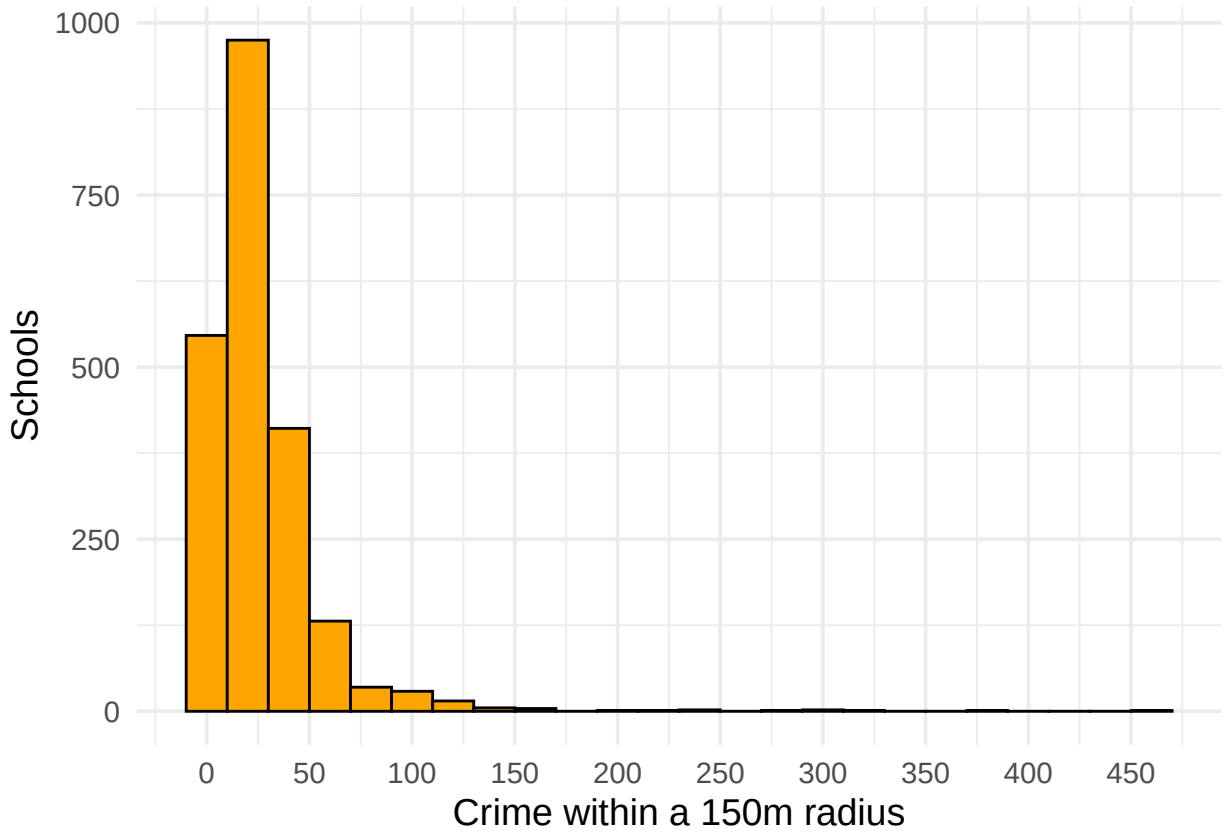
```
library(ggplot2)

ggplot(data_250) +
    geom_histogram(aes(x = data_250$Join_Count),
                   binwidth = 25, fill = "blue", color = "black") + theme_minimal(14) + scale_x_continuous
```




```
summary(data_250$Join_Count)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   31.00   55.00   67.63   82.00  691.00
```

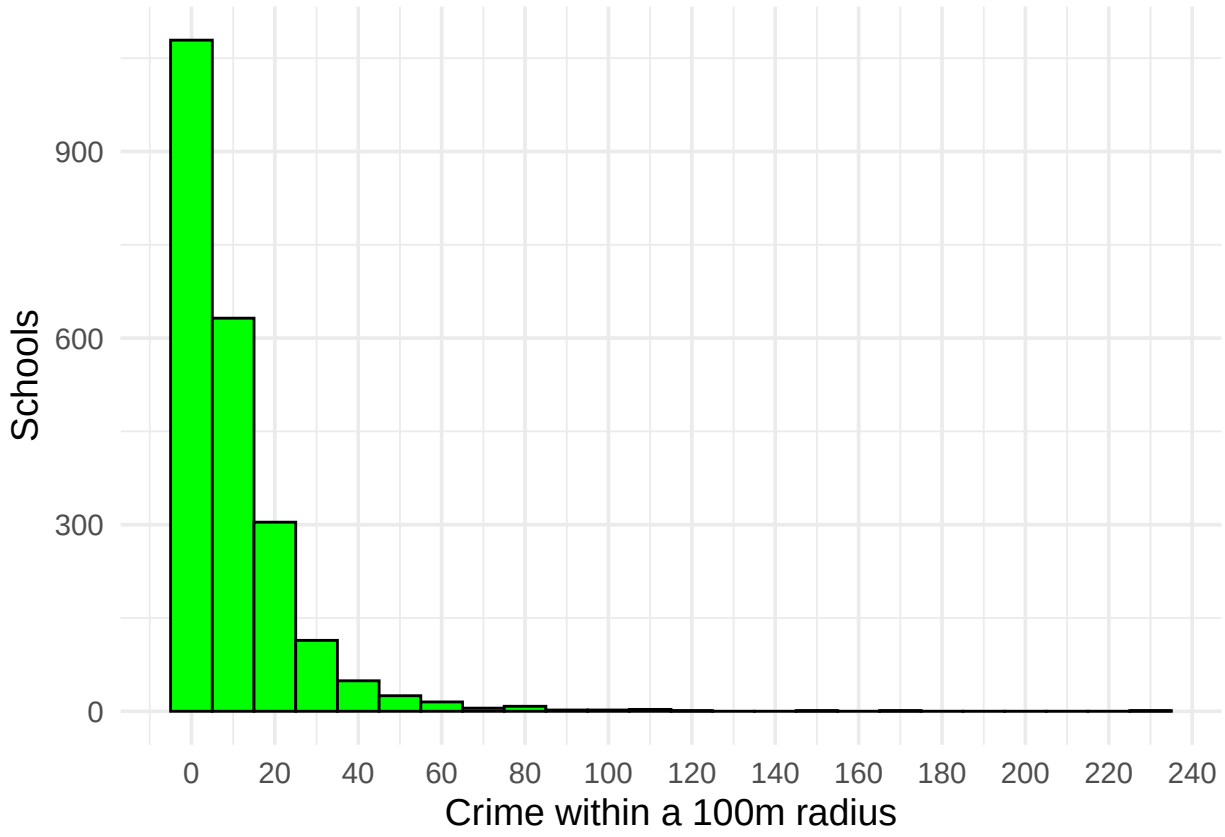
```
ggplot(data_150) +
    geom_histogram(aes(x = data_150$Join_Count),
                   binwidth = 20, fill = "orange", color = "black") + theme_minimal(14) + scale_x_continuou
```




```
summary(data_150$Join_Count)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   10.00   18.00   25.73   34.00  468.00
```

```
ggplot(data_100) +
    geom_histogram(aes(x = data_100$Join_Count),
                   binwidth = 10, fill = "green", color = "black") + theme_minimal(14) + scale_x_continuous
```




```
summary(data_100$Join_Count)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00    1.00    8.00   11.28   15.00  232.00
```

The overall number of crime lowers as the radius around the school lowers. For example, in a radius of 100m, the maximum number of crimes commited was 232 with an average of 11.28 crimes in the region overall.

For a radius of 150m, the maximum number of crimes commited was 468 with an average of 26 crimes in the region overall.

Lastly, for a radius of 250m, the maximum number of crimes commited was 691 with an average of 68 crimes throughout the region.

Next steps: Locate regions (schools) with top crime counts and review from there.

Top 10%

```
n <- 10
nomdelito <- data.frame(data_250)

topdies <- data_250[data_250$Join_Count > quantile(data_250$Join_Count, prob=1-n/100),]



topdies <- topdies[order(-topdies$Join_Count, topdies$nombre),]

topdies[c(4,2,5,1,3,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26)]
```

```
## # A tibble: 222 x 26
```

```
##    nombre Join_Count domicilio_   FID TARGET_FID coordenada latitud
##    <chr>       <dbl> <chr>      <dbl>      <dbl> <chr>        <dbl>
##  1 PRIMA~        691 DOMICILIO~   194        194 -99.14937~    19.4
##  2 ESCUE~        652 DOMICILIO~   849        849 -99.14645~    19.4
##  3 PRIMA~        643 DOMICILIO~   158        158 -99.14647~    19.4
##  4 <NA>          643 <NA>          NA         NA <NA>          NA
##  5 PRIMA~        553 DOMICILIO~  1279       1279 -99.12126~    19.4
##  6 PREES~        547 DOMICILIO~  1321       1321 -99.15434~    19.4
##  7 ESCUE~        546 DOMICILIO~  1572       1572 -99.12108~    19.4
##  8 PRIMA~        543 DOMICILIO~    74         74 -99.15446~    19.4
##  9 TELES~        540 DOMICILIO~  2160       2160 -99.16588~    19.4
## 10 PRIMA~        504 DOMICILIO~  1865       1865 -99.14596~    19.4
## # ... with 212 more rows, and 19 more variables: longitud <dbl>,
## #   domicilio <chr>, BUFF_DIST <dbl>, ORIG_FID <dbl>, ao_hechos <chr>,
## #   mes_hechos <chr>, fecha_hech <chr>, delito <chr>, categoria_ <chr>,
## #   fiscalia <chr>, agencia <chr>, unidad_inv <chr>, alcaldia_h <chr>,
## #   colonia_he <chr>, ao_inicio <chr>, mes_inicio <chr>, fecha_inic <chr>,
## #   calle_hech <chr>, calle_he_1 <chr>
```