crimenes_250_50m

Laura.w

2/20/2020

```
library(readxl)
library(readr)

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

data_250
```

```
## # A tibble: 2,246 x 26
##      FID Join_Count TARGET_FID nombre domicilio_ coordenada latitud
##    <dbl>      <dbl>      <dbl> <chr>  <chr>      <chr>        <dbl>
##  1     0         51          0 INTER~ DOMICILIO~ -99.10185~    19.4
##  2     1         45          1 PRIMA~ DOMICILIO~ -99.10427~    19.3
##  3     2         74          2 PREES~ DOMICILIO~ -99.23743~    19.3
##  4     3         48          3 PREES~ DOMICILIO~ -99.03688~    19.4
##  5     4          7          4 PRIMA~ DOMICILIO~ -99.11412~    19.4
##  6     5         71          5 PRIMA~ DOMICILIO~ -99.04822~    19.4
##  7     6         72          6 ESCUE~ DOMICILIO~ -99.02922~    19.3
##  8     7          5          7 SECUN~ DOMICILIO~ -99.06236~    19.3
##  9     8         60          8 PREES~ DOMICILIO~ -99.09861~    19.4
## 10     9         69          9 PRIMA~ DOMICILIO~ -99.09731~    19.4
## # ... with 2,236 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>
```

##Crimes within a 250m radius

Summary of crimes within a 250 radius of any school

```
data(data_250, package = "DMwR2")
summary(data_250)
```

```
##       FID           Join_Count       TARGET_FID        nombre
##  Min.   :   0.0   Min.   :  0.00   Min.   :   0.0   Length:2246
##  1st Qu.: 560.2   1st Qu.: 31.00   1st Qu.: 560.2   Class :character
##  Median :1120.5   Median : 55.00   Median :1120.5   Mode  :character
##  Mean   :1120.5   Mean   : 67.69   Mean   :1120.5
##  3rd Qu.:1680.8   3rd Qu.: 82.00   3rd Qu.:1680.8
##  Max.   :2241.0   Max.   :691.00   Max.   :2241.0
##  NA's   :4                         NA's   :4
##   domicilio_         coordenada           latitud         longitud
##  Length:2246        Length:2246        Min.   :19.17   Min.   :-99.33
##  Class :character   Class :character   1st Qu.:19.34   1st Qu.:-99.16
##  Mode  :character   Mode  :character   Median :19.38   Median :-99.12
##                                        Mean   :19.39   Mean   :-99.12
##                                        3rd Qu.:19.45   3rd Qu.:-99.08
##                                        Max.   :19.58   Max.   :-98.95
```

```
##                                      NA's   :4       NA's   :4
##   domicilio           BUFF_DIST      ORIG_FID       ao_hechos
##  Length:2246        Min.   :250   Min.   :   0.0   Length:2246
##  Class :character   1st Qu.:250   1st Qu.: 595.2   Class :character
##  Mode  :character   Median :250   Median :1197.5   Mode  :character
##                     Mean   :250   Mean   :1193.7
##                     3rd Qu.:250   3rd Qu.:1788.2
##                     Max.   :250   Max.   :2382.0
##                     NA's   :4     NA's   :4
##   mes_hechos         fecha_hech           delito
##  Length:2246        Length:2246        Length:2246
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   categoria_          fiscalia           agencia
##  Length:2246        Length:2246        Length:2246
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   unidad_inv         alcaldia_h         colonia_he
##  Length:2246        Length:2246        Length:2246
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   ao_inicio          mes_inicio         fecha_inic
##  Length:2246        Length:2246        Length:2246
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##   calle_hech         calle_he_1
##  Length:2246        Length:2246
##  Class :character   Class :character
##  Mode  :character   Mode  :character
##
##
##
##
```

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

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.00   31.00   55.00   67.69   82.00  691.00
```

###We can review the vairance, standard deviation, and inter-quartile ranges:

```
#variance
var(data_250$Join_Count)
```

```
## [1] 4207.749
```

```
#standard deviation
sd(data_250$Join_Count)
```

```
## [1] 64.86717
```

```
#inter-quartile range
quantile(data_250$Join_Count)
```

```
##   0%  25%  50%  75% 100%
##    0   31   55   82  691
```

```
quantile(data_250$Join_Count, probs=c(0.2, 0.8))
```

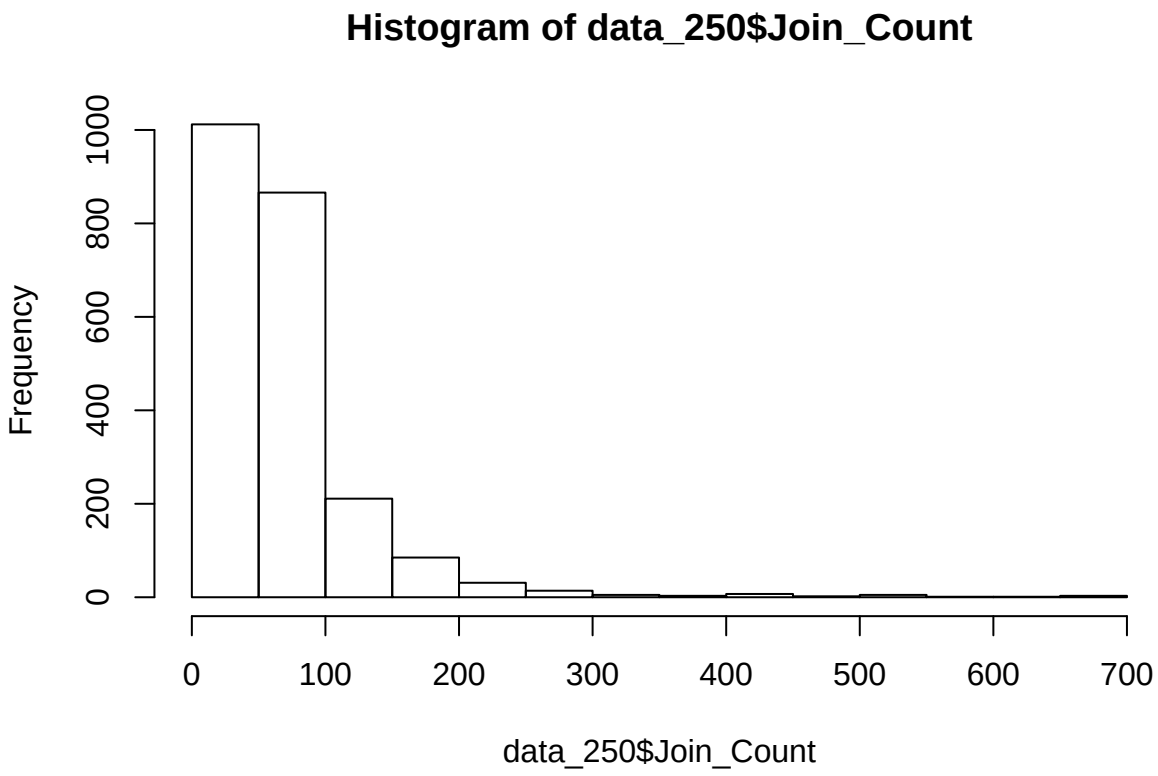
```
## 20% 80%
##  26  91
```

Joint Count Distribution

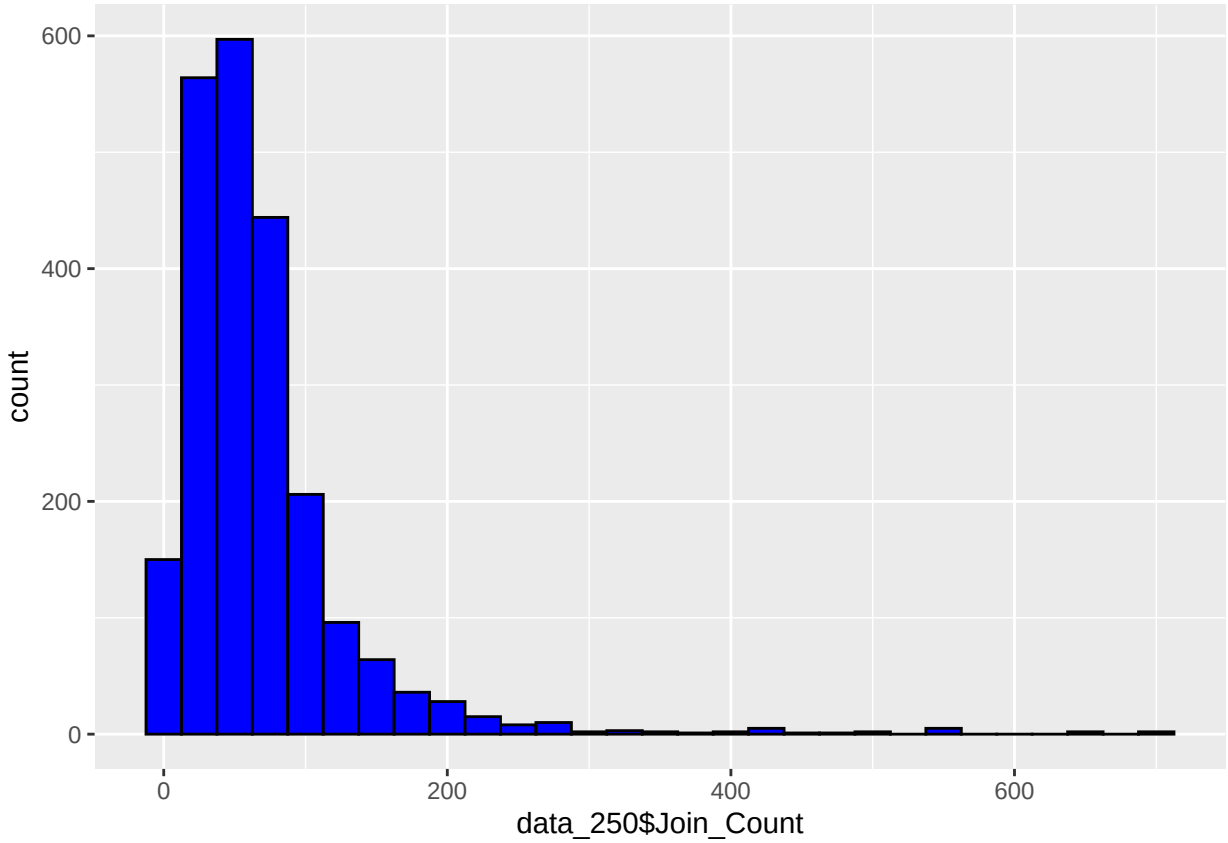
Our histogram is skewed to the lower end of the crime numbers, with most areas experiencing under 150 crimes.

When we look at the bandwith a little closer we can see that most areas with a 250 radius aruond a school have between 25 and 75 crimes on average, with 50 crimes

```
library(ggplot2)

hist(data_250$Join_Count)
```

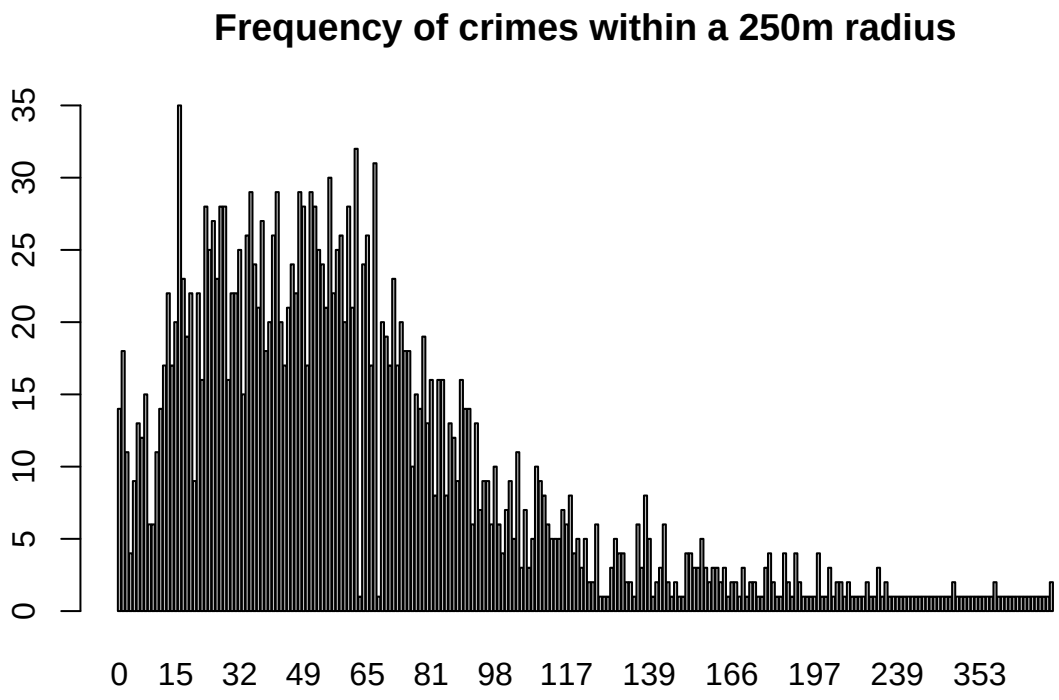



```
ggplot(data_250) +
    geom_histogram(aes(x = data_250$Join_Count),
                   binwidth = 25, fill = "blue", color = "black")
```

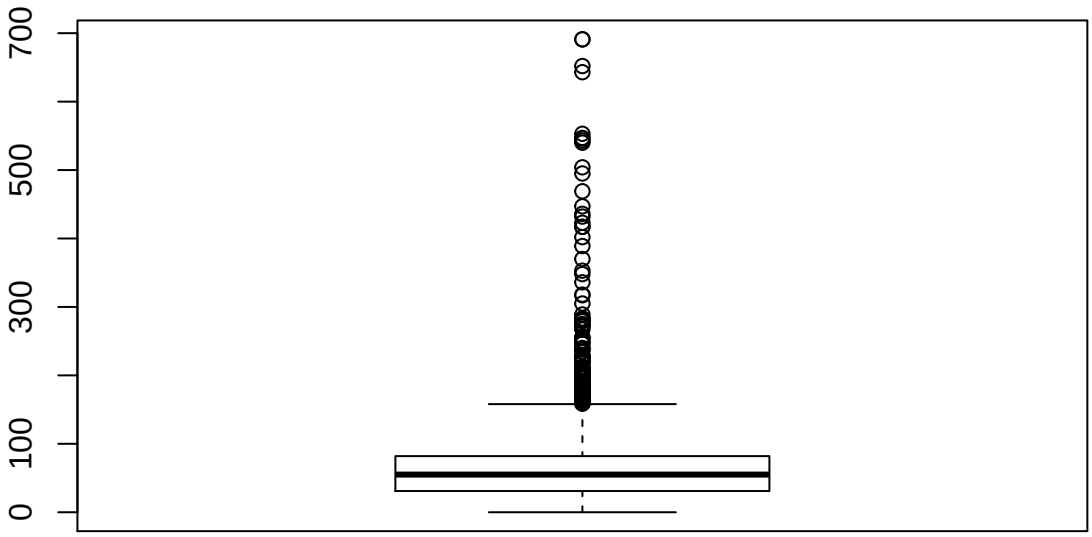


And look at the frequency of crimes

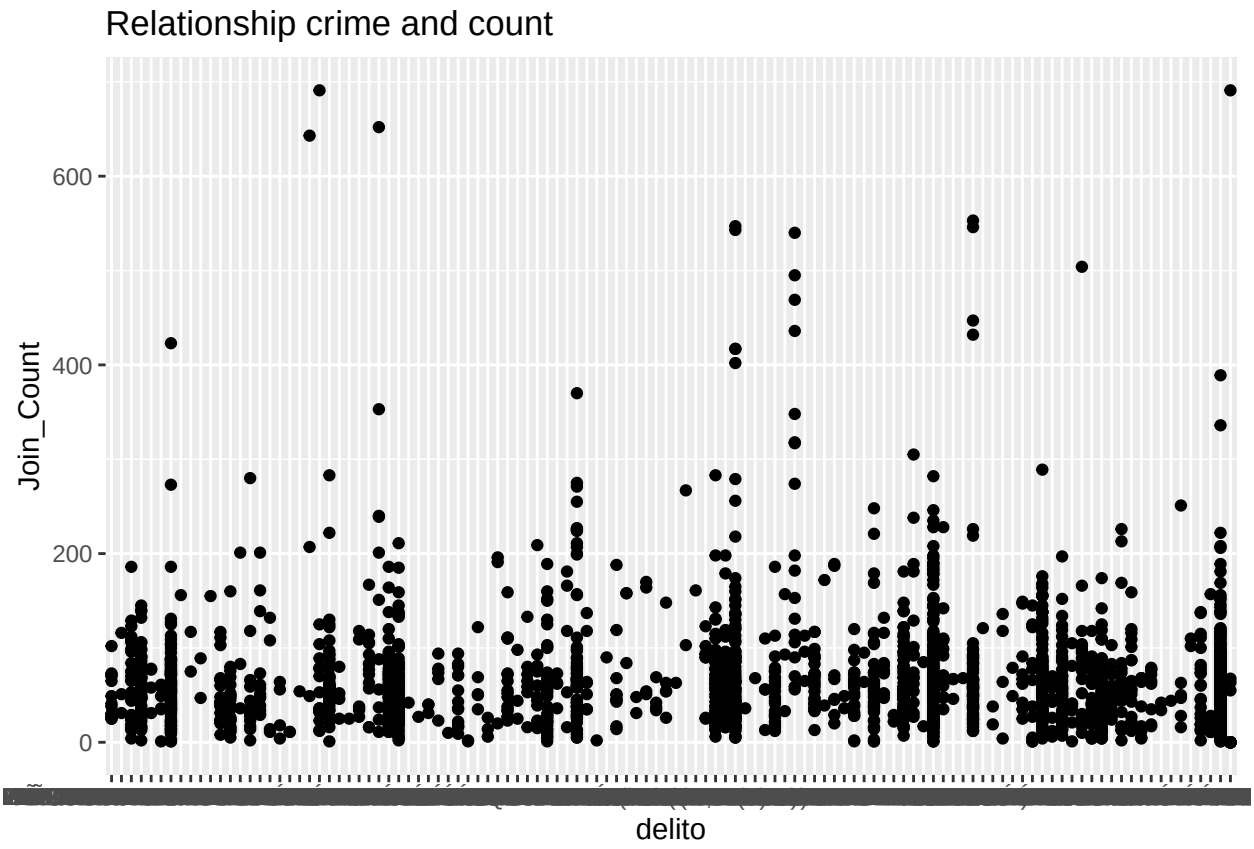
```
freq <- table(data_250$Join_Count)
barplot(freq, main="Frequency of crimes within a 250m radius")
```



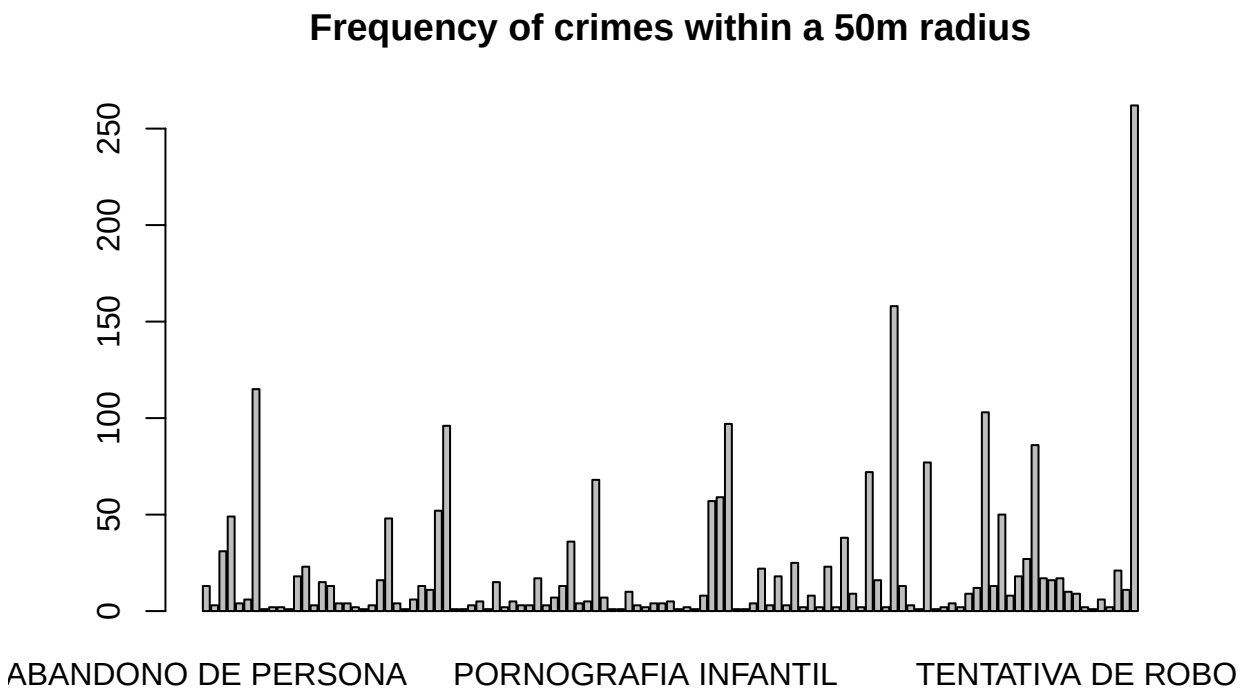

```r
boxplot(data_250$Join_Count)
```



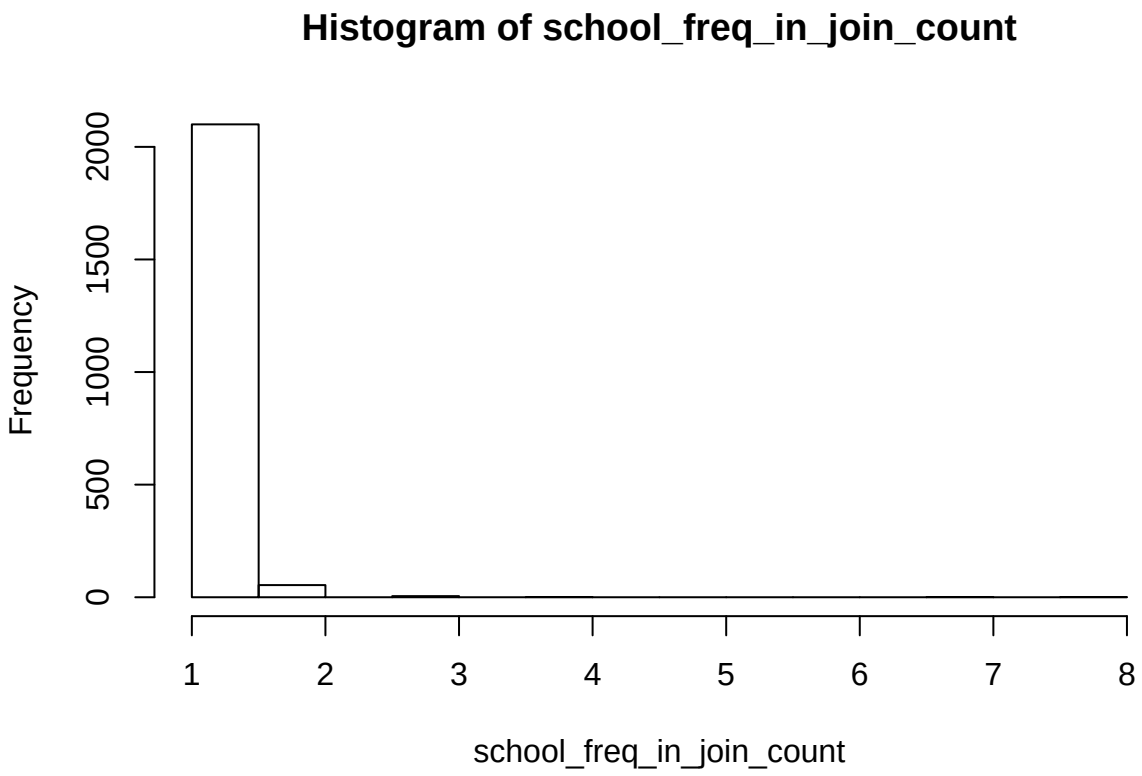

```
ggplot(data_250, aes(x=delito, y=Join_Count)) + geom_point()+ggtitle("Relationship crime and count")
```



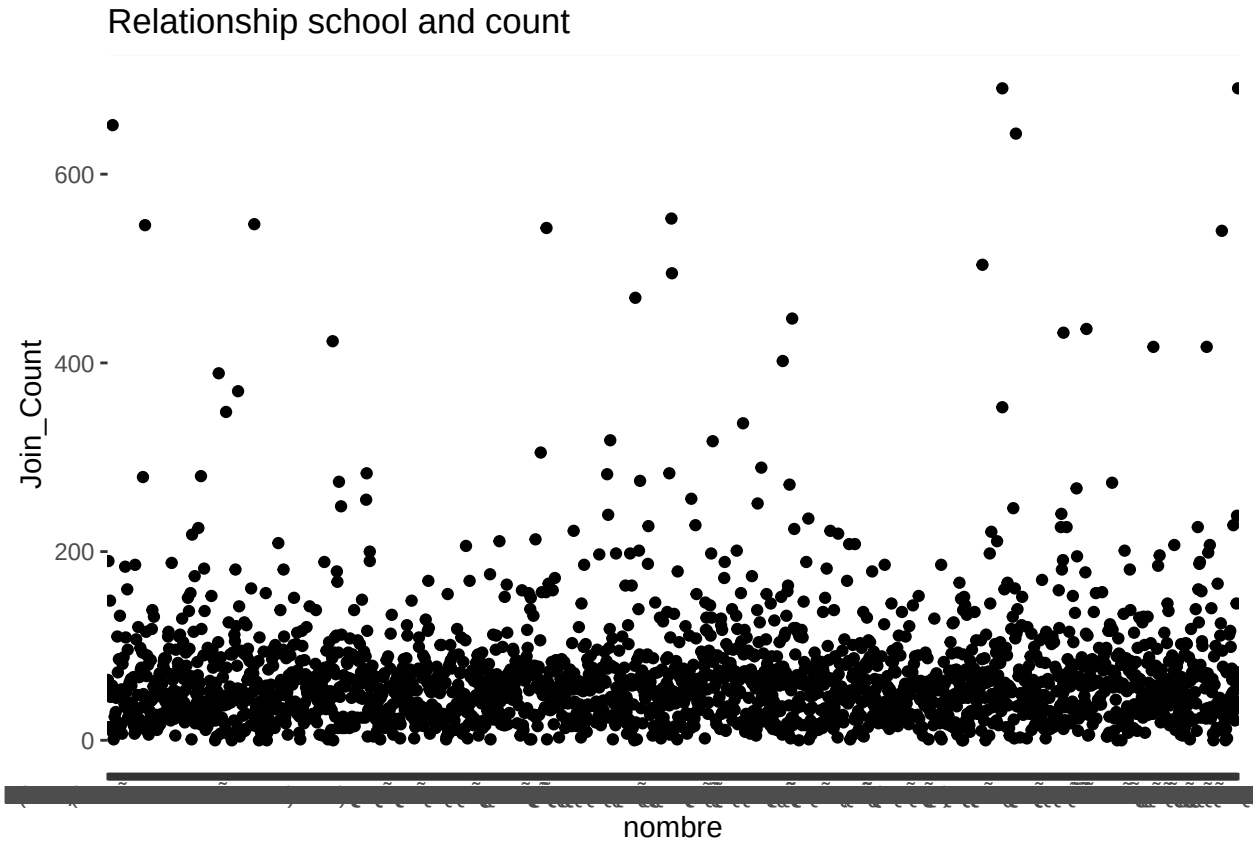

```
freq <- table(data_250$delito)
barplot(freq, main="Frequency of crimes within a 50m radius")
```



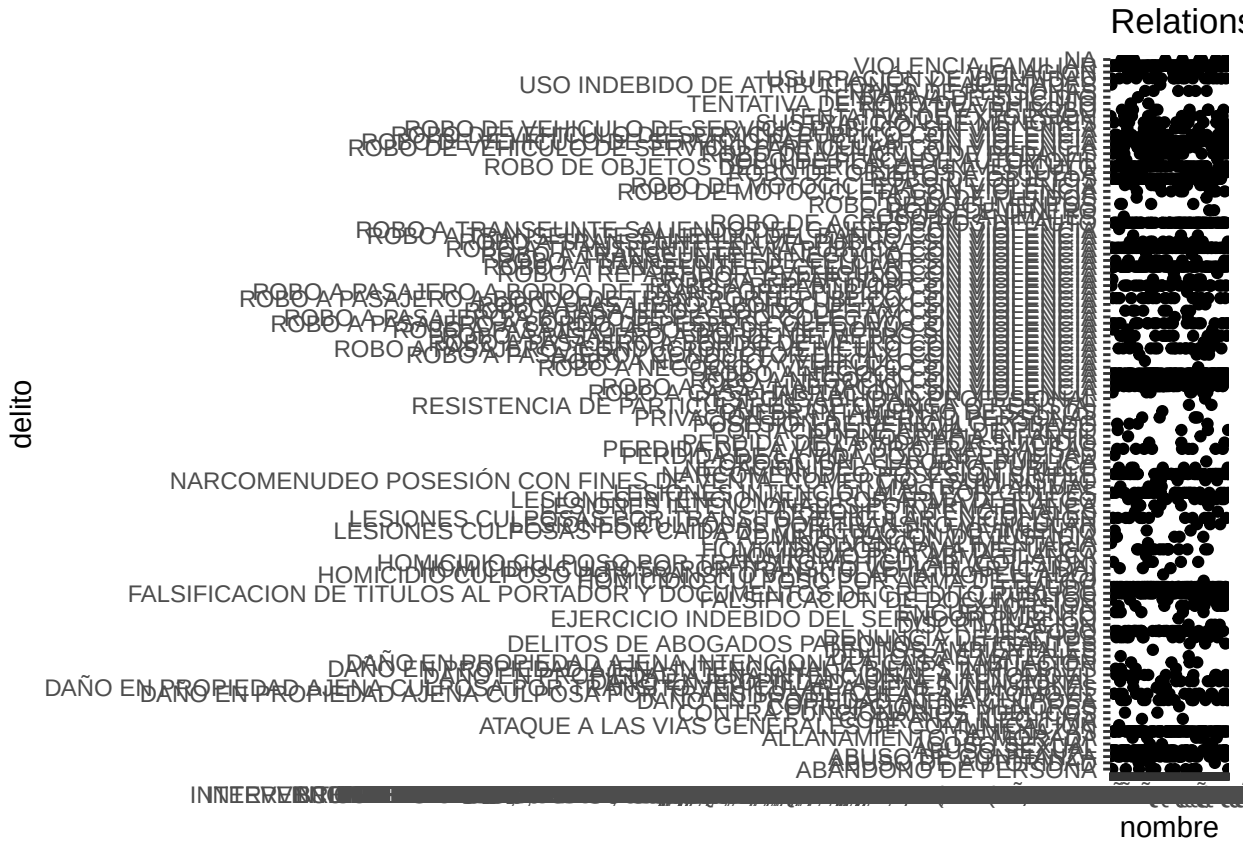

```
#table(data_250$Join_Count, data_250$nombre)
school_freq_in_join_count <- table(data_250$nombre)
hist(school_freq_in_join_count)
```




```
ggplot(data_250, aes(x=nombre, y=Join_Count)) + geom_point()+ggtitle("Relationship school and count")
```




```
ggplot(data_250, aes(x=nombre, y=delito)) + geom_point()+ggtitle("Relationship school and crime")
```




```
#missing.value.rows <- filter(data_250, !complete.cases(data_250))
#missing.value.rows


#apply(data_250, 1, function(x) sum(is.na(x)))

#sort(apply(data_250, 2, function(x) sum(is.na(x))))
```

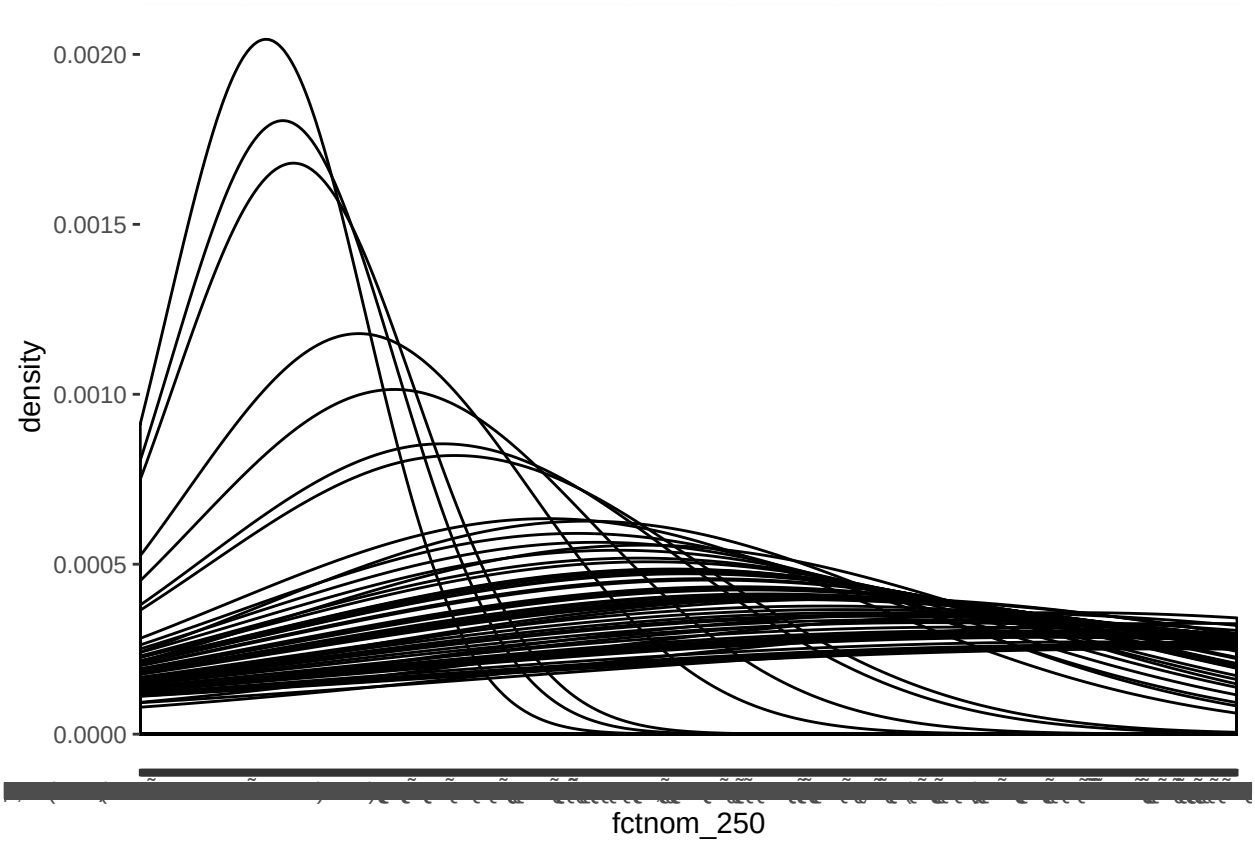
```
dfdata_250 <- data.frame(data_250)

fctdelt_250 <- as.factor(data_250$delito)

fctnom_250 <- as.factor(data_250$nombre)
```

```
#table2 <- table(data_250$delito, data_250$nombre, data_250$Join_Count)
#table2
```

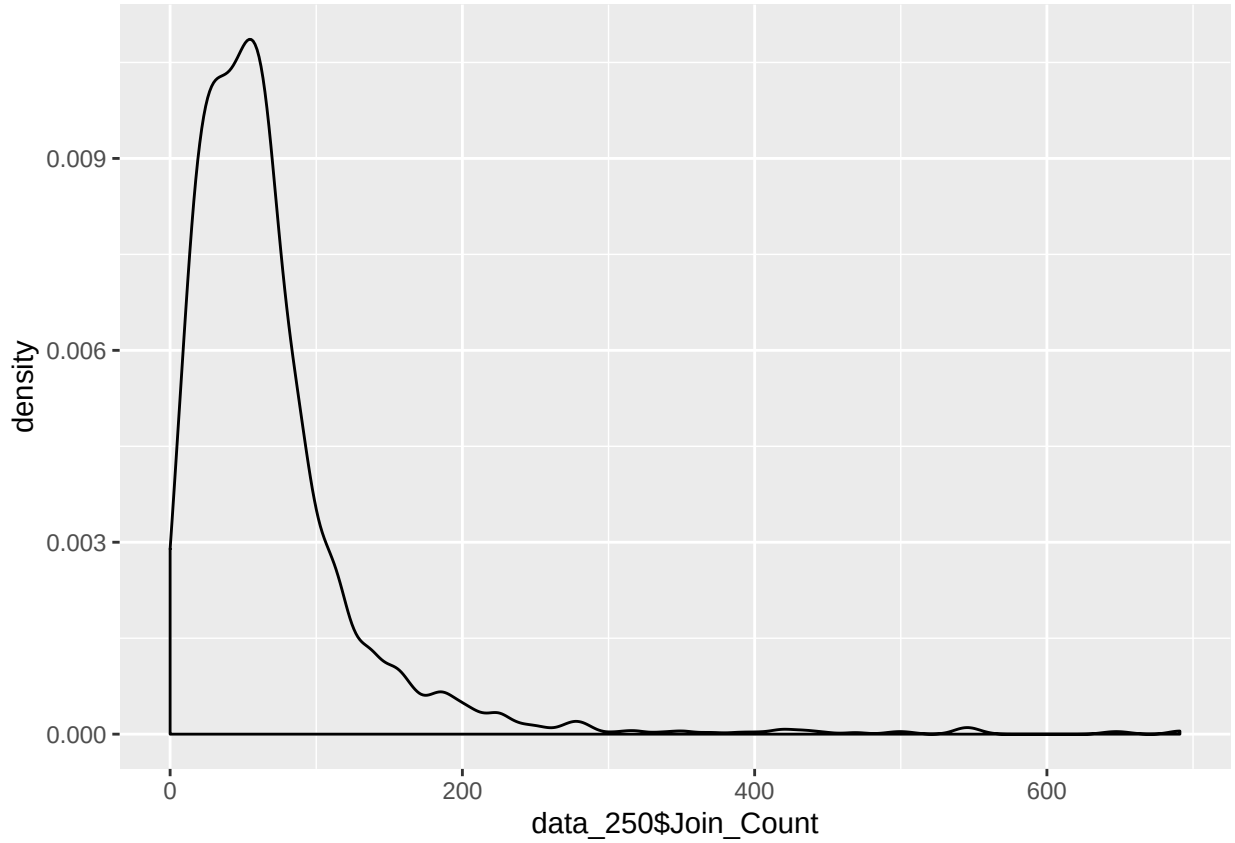
```
library(lattice)
densityplot(~ fctnom_250, group = fctdelt_250, data = data_250, auto.key = TRUE)
```

RATO ANIMAL
OMENUDEO POSESIÓN CON FINES DE VENTA, COMERCIO Y SUMINISTRO
OMENUDEO POSESION SIMPLE
CION DEL SERVICIO PUBLICO
ICION A LA OBRA PUBLICA
DA DE LA VIDA POR ENFERMEDAD
DA DE LA VIDA POR OTRAS CAUSAS
DA DE LA VIDA POR SUICIDIO
OGRAFIA INFANTIL
ACION DE ARMA DE FUEGO
SION DE VEHICULO ROBADO
CION DE LA LIBERTAD PERSONAL
RANTAMIENTO DE SELLOS
TENCIA DE PARTICULARES (SE OPONGA O RESISTA)
ONSABILIDAD PROFESIONAL
A CASA HABITACION CON VIOLENCIA
A CASA HABITACION SIN VIOLENCIA
A NEGOCIO CON VIOLENCIA
A NEGOCIO SIN VIOLENCIA
A NEGOCIO Y VEHICULO CON VIOLENCIA
A NEGOCIO Y VEHICULO SIN VIOLENCIA
A PASAJERO / CONDUCTOR DE TAXI CON VIOLENCIA
A PASAJERO / CONDUCTOR DE VEHICULO CON VIOLENCIA
A PASAJERO A BORDO DE METRO CON VIOLENCIA
A PASAJERO A BORDO DE METRO SIN VIOLENCIA
A PASAJERO A BORDO DE METROBUS SIN VIOLENCIA
A PASAJERO A BORDO DE PESERO COLECTIVO CON VIOLENCIA
A PASAJERO A BORDO DE PESERO COLECTIVO SIN VIOLENCIA

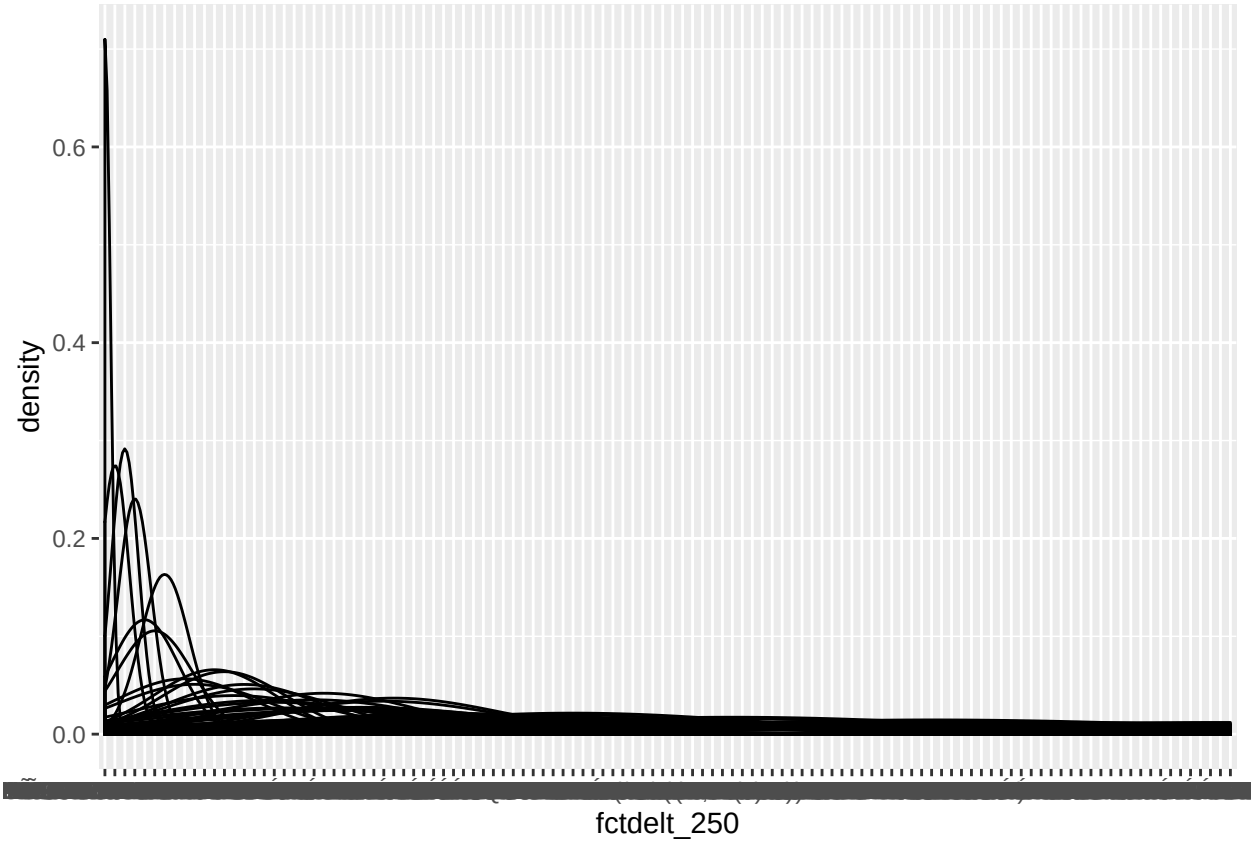
```
ggplot(data_250) + geom_density(aes(x = fctnom_250))
```




```
ggplot(data_250) + geom_density(aes(x = data_250$Join_Count))
```




```
ggplot(data_250) + geom_density(aes(x = fctdelt_250))
```




```
#ggplot(data_250) + geom_density(aes(x = fctdelt_250)) + facet_wrap(~fctnom_250)
```