wiscLottery

Maggie Westerland

2/6/2020

## Descriptive Statistics

We are using the data “wiscLottery.csv” and are interested in the relationship between the population (POP) and the number of tickets sold (SALES)

wisc\_lottery <- read.csv("WiscLottery-1.csv", header = TRUE)  
  
attach(wisc\_lottery)  
names(wisc\_lottery)

## [1] "ZIP" "PERPERHH" "MEDSCHYR" "MEDHVL" "PRCRENT" "PRC55P"   
## [7] "HHMEDAGE" "MEDINC" "SALES" "POP"

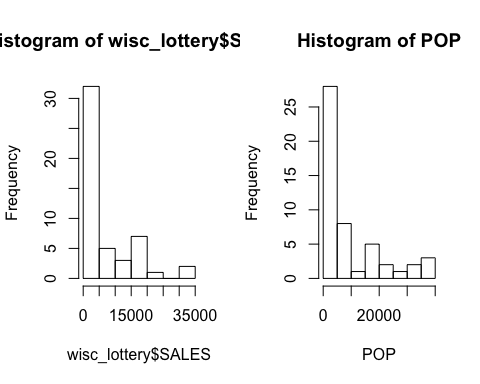
summary(wisc\_lottery$SALES)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 189.0 821.3 2426.4 6494.8 10016.5 33181.4

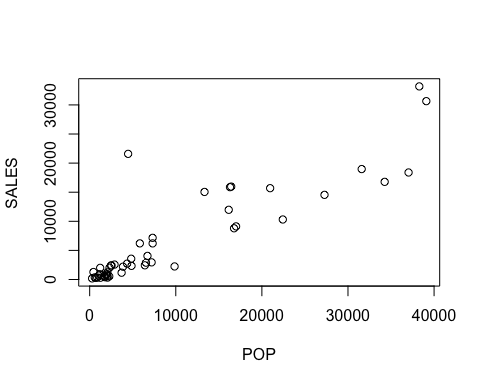
summary(SALES)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 189.0 821.3 2426.4 6494.8 10016.5 33181.4

par(mfrow = c(1, 2))  
hist(wisc\_lottery$SALES)  
hist(POP)



par(mfrow = c(1, 1))  
plot(POP, SALES)



## Comments

There seems to be a positive relationship between population and sales.

## Linear relation between population and sales

lm1 <- lm(SALES ~ POP, data = wisc\_lottery)  
lm2 <- lm(SALES ~ POP + ZIP, data = wisc\_lottery)  
lm1

##   
## Call:  
## lm(formula = SALES ~ POP, data = wisc\_lottery)  
##   
## Coefficients:  
## (Intercept) POP   
## 469.7036 0.6471

lm2

##   
## Call:  
## lm(formula = SALES ~ POP + ZIP, data = wisc\_lottery)  
##   
## Coefficients:  
## (Intercept) POP ZIP   
## 95511.9173 0.6378 -1.7577