Housing Permits and New House Value

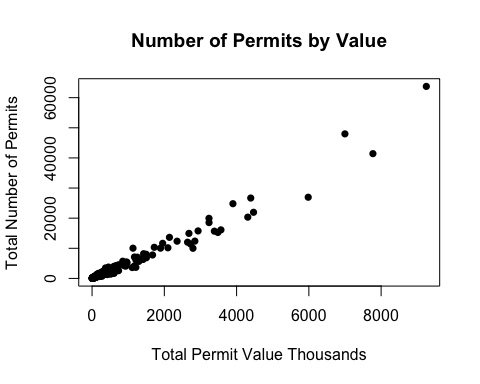
Ryan Thomas

December 30, 2015

What is the relationship of new housing permits and the price for new housing? Are places with the highest number of permits also having the highest new housing value? This study will use 2014 census data to examine the relationship between new housing units permits and the value of new housing units.

## 'data.frame': 381 obs. of 14 variables:  
## $ CSA : int 999 184 999 440 104 106 999 408 999 108 ...  
## $ CBSA : int 10180 10420 10500 10540 10580 10740 10780 10900 11020 11100 ...  
## $ Name : Factor w/ 381 levels "AbileneTX","AkronOH",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ TotalV : int 58424 181440 30971 77351 406696 436320 57600 261907 17545 173848 ...  
## $ OneV : int 58424 173540 28674 65175 296663 407739 56565 221371 16215 128990 ...  
## $ TwoV : int 0 0 1651 409 8161 1700 1035 350 640 575 ...  
## $ ThreeFourV : int 0 1900 458 0 16113 5618 0 2164 0 2279 ...  
## $ FiveMoreV : int 0 6000 188 11767 85759 21263 0 38022 690 42004 ...  
## $ TotalN : int 284 763 271 393 2231 2543 323 1801 100 918 ...  
## $ OneN : int 284 684 235 270 1203 2128 311 1051 88 496 ...  
## $ TwoN : int 0 0 22 2 62 12 12 2 4 6 ...  
## $ ThreeFourN : int 0 16 9 0 127 67 0 14 0 44 ...  
## $ FiveMoreN : int 0 63 5 121 839 336 0 734 8 372 ...  
## $ FiveStructN: int 0 1 1 11 70 15 0 38 1 13 ...

plot(permits$TotalV/1000,permits$TotalN,   
 pch = 16,   
 xlab = "Total Permit Value Thousands",   
 ylab = "Total Number of Permits",   
 main = "Number of Permits by Value")



Now, to find the perunit value of a unit in each type of housing, we will divide the value of housing permits for each type of building by the number of units.

one <- permits$OneV/permits$OneN  
two <- permits$TwoV/(permits$TwoN\*2)  
three <- permits$ThreeFourV/(permits$ThreeFourN\*3.5)  
five <- permits$FiveMoreV/(permits$FiveMoreN\*6)  
  
perunit <- data.frame(cbind(one, two, three, five))  
colnames(perunit) <- c("one", "two", "three", "five")  
str(perunit)

## 'data.frame': 381 obs. of 4 variables:  
## $ one : num 206 254 122 241 247 ...  
## $ two : num NaN NaN 37.5 102.2 65.8 ...  
## $ three: num NaN 33.9 14.5 NaN 36.2 ...  
## $ five : num NaN 15.87 6.27 16.21 17.04 ...

permits$single <- perunit$one  
permits$multi <- apply(perunit[,2:4], 1, mean, na.rm = TRUE)  
permits$perunit <- apply(perunit, 1, mean, na.rm = TRUE)  
summary(permits$single)

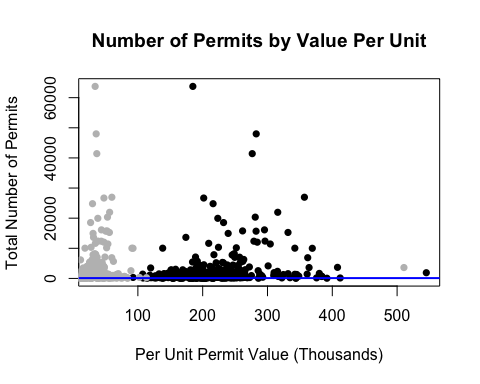
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 29.35 184.30 212.60 219.10 247.80 545.10

summary(permits$multi)

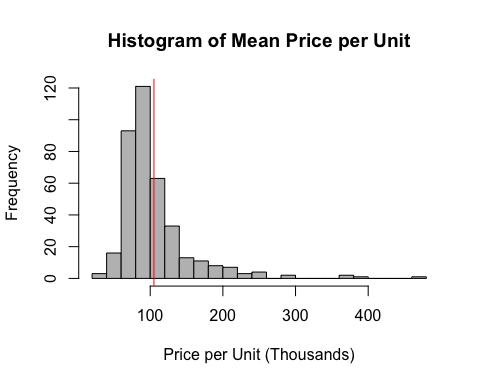
## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 5.556 25.460 34.480 37.110 43.890 510.600 33

What is the relationship between the total number of permits and the perunit value?

plot(permits$single,permits$TotalN,   
 pch = 16,   
 xlab = "Per Unit Permit Value (Thousands)",   
 ylab = "Total Number of Permits",   
 main = "Number of Permits by Value Per Unit")  
points(permits$multi,permits$TotalN,   
 pch = 16,  
 col = "grey")  
fit <- glm(permits$perunit~permits$TotalN)  
co <- coef(fit)  
abline(fit, col="blue", lwd=2)



hist(permits$perunit, 20,   
 col = "grey",   
 main = "Histogram of Mean Price per Unit",   
 xlab = "Price per Unit (Thousands)",   
 axes = TRUE,   
 plot = TRUE)  
 abline(v = mean(permits$perunit),col = "red")

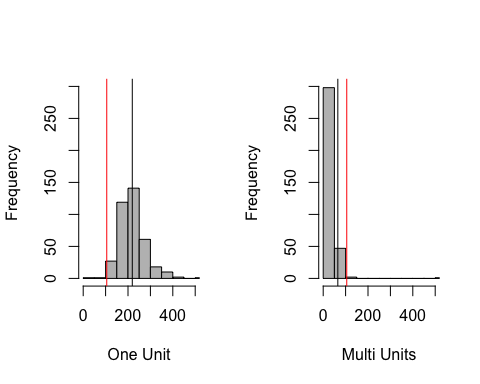


Compare this to the median income to find out whether the average price per unit is affordable. #Need to incorporate data on median income.

permits <- cbind(permits, one, two, three, five)  
  
par(mfrow = c(1,2),  
 main = "Number of Permits by Value Per Unit")

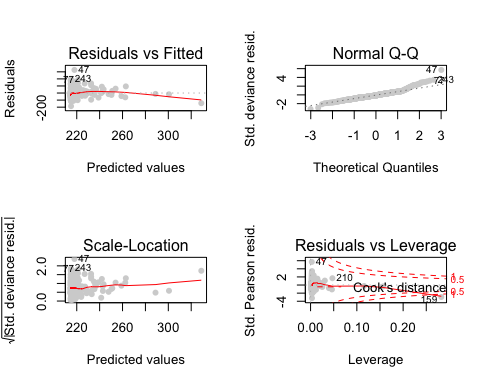
## Warning in par(mfrow = c(1, 2), main = "Number of Permits by Value Per  
## Unit"): "main" is not a graphical parameter

hist(permits$single, breaks = seq(from = 0, to = 1500, by = 50),   
 xlim = c(0,500),  
 ylim = c(0,300),  
 xlab = "One Unit",  
 main = "",  
 col = "grey")  
 abline(v = mean(permits$perunit),col = "red")  
 abline(v = mean(permits$one, na.rm = TRUE),col = "black")  
  
hist(permits$multi, breaks = seq(from = 0, to = 1500, by = 50),   
 xlim = c(0,500),  
 ylim = c(0,300),  
 xlab = "Multi Units",   
 main = "",  
 col = "grey")  
 abline(v = mean(permits$perunit),col = "red")  
 abline(v = mean(permits$two, na.rm = TRUE),col = "black")

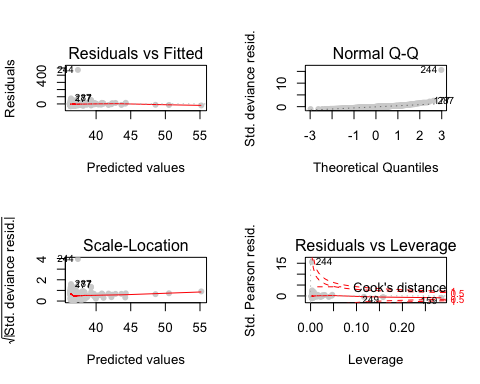


# hist(permits$three,breaks = seq(from = 0, to = 1500, by = 50),   
# xlim = c(0,500),  
# ylim = c(0,300),  
# xlab = "Three and Four Units",   
# main = "")  
# abline(v = mean(permits$perunit),col = "red")  
# abline(v = mean(permits$three, na.rm = TRUE),col = "black")  
#   
# hist(permits$five, breaks = seq(from = 0, to = 1500, by = 50),   
# xlim = c(0,500),  
# ylim = c(0,300),  
# xlab = "Five and Over",  
# main = "")  
# abline(v = mean(permits$perunit),col = "red")  
# abline(v = mean(permits$five, na.rm = TRUE),col = "black")

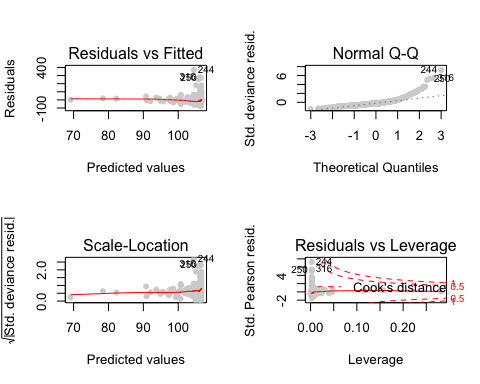
par(mfrow = c(2,2))  
plot(glm(permits$single ~ permits$TotalN), pch = 16, col = "light grey")



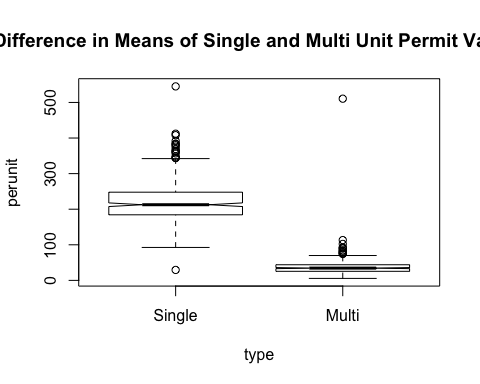
par(mfrow = c(2,2))  
plot(glm(permits$multi ~ permits$TotalN), pch = 16, col = "light grey")



par(mfrow = c(2,2))  
plot(glm(permits$perunit ~ permits$TotalN), pch = 16, col = "light grey")



perunit.single <- data.frame(permits$single)  
perunit.single$type <- as.factor("Single")  
colnames(perunit.single) <- c("perunit", "type")  
  
perunit.multi <- data.frame(permits$multi)  
perunit.multi$type <- as.factor("Multi")  
colnames(perunit.multi) <- c("perunit", "type")  
  
perunit.anova <- rbind(perunit.single, perunit.multi)  
  
plot(perunit ~ type, data = perunit.anova, na.rm = TRUE,   
 notch = TRUE,  
 main = "Difference in Means of Single and Multi Unit Permit Values")



lm.perunit.anova <- lm(perunit.anova)  
anova(lm.perunit.anova)

## Analysis of Variance Table  
##   
## Response: perunit  
## Df Sum Sq Mean Sq F value Pr(>F)   
## type 1 6025249 6025249 2714.3 < 2.2e-16 \*\*\*  
## Residuals 727 1613826 2220   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1