#Project

setwd("/Users/Seema/Documents/Project")

install.packages("ggplot2")

install.packages("reshape")

library(reshape)

#City\_HomesSoldAsForeclosures-Ratio\_AllHomes

forecl <- read.csv("Foreclosure.csv", header = T)

#forecl1 <- subset(forecl, select=c(Metro, X2013.09))

#transpose the data

foreclt <- melt(forecl, id=c("RegionName", "State", "Metro", "CountyName" ))

#remove columns I don’t want

#foreclt$Metro <-NULL

foreclt$State <-NULL

foreclt$CountyName <-NULL

#selecting NY city

#f.set <- subset(foreclt, RegionName == "New York")

#This pulls up cities

#f.ny <- subset(forecl, RegionName == ("New York"))

#plot

#ggplot(f.set, aes(x = variable, y = value, colour = variable)) + geom\_point()

#rename the value variable

names(foreclt)[4] <- "valueforecl"

head(foreclt, n=10)

#### Second file ###

#NumberOfHomesSold

NHomesSold <- read.csv("NumberOfHomesSold.csv", header = T)

nsoldt <- melt(NHomesSold, id=c("RegionName", "State", "Metro", "CountyName" ))

#nsoldt$Metro <-NULL

nsoldt$State <-NULL

nsoldt$CountyName <-NULL

#print first 10 rows

head(nsoldt, n=10)

#merge the two files

merge1 <- merge(foreclt, nsoldt)