

```

#Data Load
CrimeData = read.csv(file.choose())
View(CrimeData)
library(ggplot2)
install.packages("ggmap")
library(ggmap)
map <- get_map(location = "chicago")
ggmap(map)

#Removing Empty Rows
sum(is.na(CrimeData$Latitude))
sum(is.na(CrimeData$Longitude))
CrimeData <- CrimeData[!is.na(CrimeData$Latitude),]
sum(is.na(CrimeData$Latitude))
sum(is.na(CrimeData$Longitude))
summary(CrimeData$Latitude)
summary(CrimeData$Longitude)

#Data Visualization
library(grid)
ggplot(data=CrimeData,aes(x=Longitude,y=Latitude))+geom_point()
ggplot(data=CrimeData,aes(x=Longitude,y=Latitude))+geom_point(size=.1)
ggplot(data=CrimeData,aes(x=Longitude,y=Latitude))+geom_point(size=.1,alpha=.5)

ggmap(map) + geom_point(aes(x=Longitude,y=Latitude),data=CrimeData,size=.1,alpha=.5)

ggmap(map) + geom_point(aes(x=Longitude,y=Latitude),data=CrimeData[CrimeData$District==1,],color="red")
ggmap(map) + geom_point(aes(x=Longitude,y=Latitude),data=CrimeData[CrimeData$District==25,],color="red")

ggmap(map,zoom=11) + geom_point(aes(x=Longitude,y=Latitude,color=CrimeData$District),data=CrimeData,size=.1,alpha=.5)
ggmap(map,zoom=11) + geom_point(aes(x=Longitude,y=Latitude,color=CrimeData$Arrest),data=CrimeData,size=.1,alpha=.5)

library(dplyr)
dist <- CrimeData %>% group_by(District) %>% summarise(count= n())
dist
ggplot(data=dist,aes(x=District,y=count))
ggplot(data=dist,aes(x=District,y=count)) + geom_bar(stat="identity")
CrimeData %>% group_by(Year,District)
CrimeData %>% group_by(Year,District) %>% summarise(count=n())
forecastdata <- CrimeData %>% group_by(Year,District) %>% summarise(count=n())
View(forecastdata)
forecastdata$Year=2001
forecastdata <- CrimeData %>% group_by(Year,District) %>% summarise(count=n())
str(forecastdata)
forecastdf <- as.data.frame(forecastdata)
forecastdf$Year = 2001
str(forecastdf)
View(forecastdf)
forecastdata <- CrimeData %>% group_by(Year,District) %>% summarise(count=n())
forecastdf <- as.data.frame(forecastdata)
str(forecastdf)
ggplot(forecastdf[forecastdf$District==1,],aes(Year,count))+geom_bar(stat="identity")
ggplot(forecastdf[forecastdf$District==25,],aes(Year,count))+geom_bar(stat="identity")
ggplot(forecastdf[forecastdf$District==1,],aes(Year,count))+geom_smooth()
ggplot(forecastdf[forecastdf$District==25,],aes(Year,count))+geom_smooth()

CrimeData5 = read.csv(file.choose())
sum(is.na(CrimeData5$Latitude))
CrimeData5 <- CrimeData5[!is.na(CrimeData5$Latitude),]
CrimeData5$DateMod <- as.POSIXlt(CrimeData5$Date,format= "%m-%d-%Y %H:%M")
head(CrimeData5$DateMod)
library(chron)
CrimeData5$TimeMod <- times(format(CrimeData5$DateMod, "%H:%M:%S"))
head(CrimeData5$TimeMod)
time.part <- chron(times= c("00:00:00", "06:00:00", "12:00:00", "18:00:00","23:59:00"))
time.part
CrimeData5$time.part <- cut(CrimeData5$TimeMod, breaks= time.part,
labels= c("00-06","06-12", "12-18", "18-00"), include.lowest=TRUE)

table(CrimeData5$time.part)
table(CrimeData5$Primary.Type)
length(unique(CrimeData5$Primary.Type))
CrimeData5$CrimeType <- as.character(CrimeData5$Primary.Type)
CrimeData5$CrimeType <- ifelse(CrimeData5$CrimeType %in% c("BURGLARY",
"Motor Vehicle Theft", "Robbery", "Theft", "Battery"), 'Theft', CrimeData5$CrimeType)
CrimeData5$CrimeType <- ifelse(CrimeData5$CrimeType %in% c("Other Narcotic Violation",
"Narcotics"), 'Drugs', CrimeData5$CrimeType)
CrimeData5$CrimeType <- ifelse(CrimeData5$CrimeType %in% c("Gambling",
"Interference with Public Officer", "Intimidation", "Liquor Law Violation",
"Obscenity", "Non-Criminal", "Public Peace Violation",
"Stalking"), 'Nonviolent', CrimeData5$CrimeType)
CrimeData5$CrimeType <- ifelse(CrimeData5$CrimeType %in% c("Weapons Violation",
"Assault", "Homicide"), 'Violent', CrimeData5$CrimeType)
CrimeData5$CrimeType <- ifelse(CrimeData5$CrimeType %in% c("Sex Offense",
"Crim Sexual Assault", "Prostitution"), 'SexualCrime', CrimeData5$CrimeType)

table(CrimeData5$CrimeType)
View(CrimeData5)

library(dplyr)
str(CrimeData5$time.part)
df <- CrimeData5[,-17] %>% group_by(time.part) %>% summarise(n())
ggplot(data=df,aes(df$time.part,df$n()))+geom_bar(stat="identity")
CrimeData5[,-17] %>% group_by(CrimeType) %>% summarise(n())
dfcrimetype <- CrimeData5[,-17] %>% group_by(CrimeType) %>% summarise(n())
ggplot(data=dfcrimetype,aes(dfcrimetype$CrimeType,dfcrimetype$n()))+geom_bar(stat="identity")

model <- lm(count ~., data=forecastdata)
predict(model,newdata = data.frame(Year=2017,District=1))
View(CrimeData)
summary(model)
forecastdata$LogCount=log(forecastdata$count,exp(1))
ExpReg = lm (LogCount~.-count, data=forecastdata)
summary(ExpReg)
predict(ExpReg,newdata = data.frame(Year=2017,District=2,count=30))

```