BI-In class 5-1-19

> setwd("C:/Users/nmandya/Desktop")

> getwd()

[1] "C:/Users/nmandya/Desktop"

>

> data<-read.csv("data1.csv")

> View(data)

> summary(data)

household age\_husband age\_wife income\_husband income\_wife bedrooms electricity gas number\_children internet

Min. : 48 Min. :17.00 Min. :19.00 Min. : -6300 Min. : -6300 Min. : 0.000 Min. : 4.0 Min. : 3.00 Min. : 0.0000 No : 642

1st Qu.: 389236 1st Qu.:42.00 1st Qu.:40.00 1st Qu.: 24000 1st Qu.: 6225 1st Qu.: 3.000 1st Qu.: 80.0 1st Qu.: 3.00 1st Qu.: 0.0000 Yes:7169

Median : 764131 Median :55.00 Median :53.00 Median : 43000 Median : 18110 Median : 3.000 Median :110.0 Median : 20.00 Median : 0.0000

Mean : 759565 Mean :54.32 Mean :52.08 Mean : 59831 Mean : 28985 Mean : 3.117 Mean :131.9 Mean : 45.09 Mean : 0.6655

3rd Qu.:1137444 3rd Qu.:66.00 3rd Qu.:63.00 3rd Qu.: 70500 3rd Qu.: 39600 3rd Qu.: 4.000 3rd Qu.:160.0 3rd Qu.: 70.00 3rd Qu.: 1.0000

Max. :1492278 Max. :95.00 Max. :95.00 Max. :756000 Max. :421000 Max. :10.000 Max. :500.0 Max. :360.00 Max. :12.0000

mode own language decade\_built

followup:1586 Occupied without payment of rent: 76 English only:6751 Min. :1930

internet:4004 Owned free and clear :1896 Other : 572 1st Qu.:1950

mail :2221 Owned with mortgage or loan :4505 Spanish : 488 Median :1970

Rented :1334 Mean :1971

3rd Qu.:1990

Max. :2010

Histogram Visualization:

> mean(data$age\_husband)

[1] 54.31776

>

> hist(data$number\_children)

Barplot Visualization:

> counts<-table(data$bedrooms)

> barplot(counts,main="Bedroom Distribution",xlab="Number of Bedrooms")

> mpg<-c(50,60,80)

> cyl<-c(5,6,5)

> hp<-c(110,140,150)

> cars<-data.frame(mpg,cyl,hp)

> cars

mpg cyl hp

1 50 5 110

2 60 6 140

3 80 5 150

Inserting ROwnames

> rownames(cars)<-c("BMW","Saburu","Tesla")

> cars

mpg cyl hp

BMW 50 5 110

Saburu 60 6 140

Tesla 80 5 150

> install.packages(("ggplot2"))

> library("ggplot2")

> ggplot(cars,aes(x=cyl,y=mpg))+geom\_point()

> p<-ggplot(cars,aes(x=cyl,y=mpg))

> p+geom\_point()