lotto <- read.csv(file.choose(),header=TRUE)

attach(lotto)

lotto

summary(POP)

summary(lotto[,c("POP","SALES")])

# {rows you want,columns you want]

#the blank before the comma means I want all the rows since I didn't specify

plot(POP,SALES)

# scatterplot(x,y)

text(5000, 24000, "Kenosha")

cor(POP,SALES)

[1] 0.8862827

# cor for correlation

#.886 is a strong positive cor.,which corresponds to

# the positive slope in the plot and the fact that most points are

# close to the lines,thus the high or strong decimal.

cor(lotto)

ZIP PERPERHH MEDSCHYR MEDHVL PRCRENT PRC55P

ZIP 1.00000000 -0.25043757 -0.1840536 -0.55502196 -0.1881527 0.3568658

PERPERHH -0.25043757 1.00000000 -0.1889199 0.06791617 -0.3683181 -0.5178883

MEDSCHYR -0.18405361 -0.18891986 1.0000000 0.70780357 0.6422949 -0.4792305

MEDHVL -0.55502196 0.06791617 0.7078036 1.00000000 0.3726068 -0.6213228

PRCRENT -0.18815274 -0.36831807 0.6422949 0.37260677 1.0000000 -0.3096679

PRC55P 0.35686577 -0.51788827 -0.4792305 -0.62132281 -0.3096679 1.0000000

HHMEDAGE 0.30556073 -0.47822064 -0.4810156 -0.54337571 -0.3695101 0.9636132

MEDINC -0.50253198 0.33397632 0.5222065 0.84133554 0.2001336 -0.7762837

SALES -0.21661170 -0.13625569 0.4674995 0.57150508 0.4964811 -0.3857442

POP -0.09552086 -0.18038028 0.5895641 0.52577941 0.6433933 -0.4035664

HHMEDAGE MEDINC SALES POP

ZIP 0.3055607 -0.5025320 -0.2166117 -0.09552086

PERPERHH -0.4782206 0.3339763 -0.1362557 -0.18038028

MEDSCHYR -0.4810156 0.5222065 0.4674995 0.58956409

MEDHVL -0.5433757 0.8413355 0.5715051 0.52577941

PRCRENT -0.3695101 0.2001336 0.4964811 0.64339330

PRC55P 0.9636132 -0.7762837 -0.3857442 -0.40356640

HHMEDAGE 1.0000000 -0.6809147 -0.3648614 -0.41572416

MEDINC -0.6809147 1.0000000 0.4687388 0.43299180

SALES -0.3648614 0.4687388 1.0000000 0.88628266

POP -0.4157242 0.4329918 0.8862827 1.00000000

# fins the correltaion between all positive pairs of variables

cor(cbind(POP, SALES,MEDINC))

POP SALES MEDINC

POP 1.0000000 0.8862827 0.4329918

SALES 0.8862827 1.0000000 0.4687388

MEDINC 0.4329918 0.4687388 1.0000000

# IF YOU have any missing values in the dataset,cor will not skip over it

# It will put NA's in ur output,to avoid that ,see below

cor(cbind(POP, SALES,MEDINC),use= "pairwise.complete.obs")

POP SALES MEDINC

POP 1.0000000 0.8862827 0.4329918

SALES 0.8862827 1.0000000 0.4687388

MEDINC 0.4329918 0.4687388 1.0000000

# "xxxx.complete...could wipe out all incomplete data,

#yet pairwise did eliminate this pattern and only rule out the selected pairs