```
################################################################################
# list the packages we need and loads them, installs them automatically if we don't have them
# add any package that you need to the list
need <- c('glue', 'dplyr','readxl',  'ggplot2','tidyr','AER','scales','mvtnorm',
          'stargazer','httr', 'repmis','gridExtra')

have <- need %in% rownames(installed.packages())
if(any(!have)) install.packages(need[!have])
invisible(lapply(need, library, character.only=T))

# Save the R script to the assignment 1 folder before this
# To set up the working directory
getwd()
setwd(getwd()) #change getwd() here is you need to set a different working directory


#this clears the workspace
rm(list = ls())
#this sets the random number generator seed to my birthday for replication


options(scipen=999)


data(anscombe)

stargazer(anscombe, type="text", digits=2)

cor(anscombe$x1,anscombe$y1)
cor(anscombe$x2,anscombe$y2)
cor(anscombe$x3,anscombe$y3)
cor(anscombe$x4,anscombe$y4)
```

Using the Anscombe's Quartet as instructed in the header code, for four pairs of (x1,y1),(x2,y2),(x3,y3 ),(x4,y4)

a. Estimate and interpret the sample covariance and correlation.

b. Use four scatterplots to visually observe the data.

c.  Using the first series of Anscombe's quartet, choose the intercept and the slope of the line that v:

d.  Using the data set calculate the predicted values for the outcome, residuals, and the sum of squarec

e.  Calculate the OLS estimators for the intercept and the slope using the formula.

f.  Compare the sum of the squared residuals from the OLS estimators to the one that you used in part (l

g.  Interpret the estimated OLS coefficients.

h.  Calculate and interpret the R-squared of the regression.