STAT3340 Ass't 1, Fall 2024, Due Thursday, September 12 11:59 PM

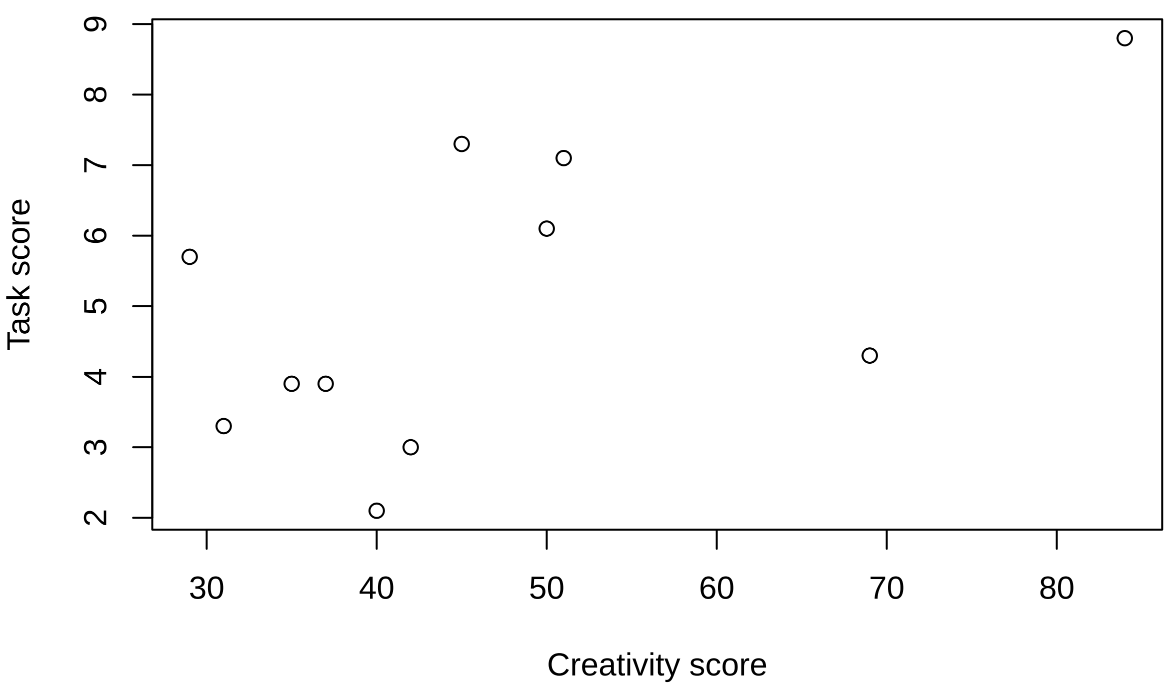
Your name

Your Banner # B00??????

This assignment is to be started using R Markdown. At minimum, process question 1 in R markdown. Add other answers in R Markdown or by hand, and submit a single pdf file to brightspace.

1. A random sample of 11 elementary school students is selected, and each student is measured on a creativity score (x) using a well-defined testing instrument and on a task score (y) using a new instrument. The task score is the mean time taken to perform several hand-eye coordination tasks. The data are as follows.

```
x=c(35,37,50,69,84,40,29,42,51,45,31)
y=c(3.9,3.9,6.1,4.3,8.8,2.1,5.7,3.0,7.1,7.3,3.3)
plot(x,y,xlab="Creativity score",ylab="Task score")
```



Use R to do the following questions. Make sure to include your commands in the R markdown document.

- 1a) Calculate the summaries S_{xx}, S_{xy}, S_{yy} and $\bar{X}$ and $\bar{Y}$.

```
N=length(x)
xbar=sum(x)/N ;  cat("xbar = ", xbar)
```

```
## xbar =  46.63636
```

```
Sxx=sum((x-xbar)^2) ; cat("Sxx = ",Sxx)
```

```
## Sxx =  2778.545
```

```r
ybar=0;cat("ybar = ",ybar) #use correct equation for ybar
```

```
## ybar =  0
```

```r
Syy=0; cat("Syy = ",Syy) #use correct equation for  Syy
```

```
## Syy =  0
```

```r
Sxy=0;cat("Sxy = ",Sxy) #use correct equation for Sxy
```

```
## Sxy =  0
```

- 1b) Use these summaries to calculate the least squares estimates of the intercept and slope.

```r
 beta1 = 0; cat("beta1 = ", beta1)  #use correct equation for the slope
```

```
## beta1 =  0
```

```r
 beta0 = ybar+0; cat("beta0 =", beta0) #use correct equation for the intercept
```

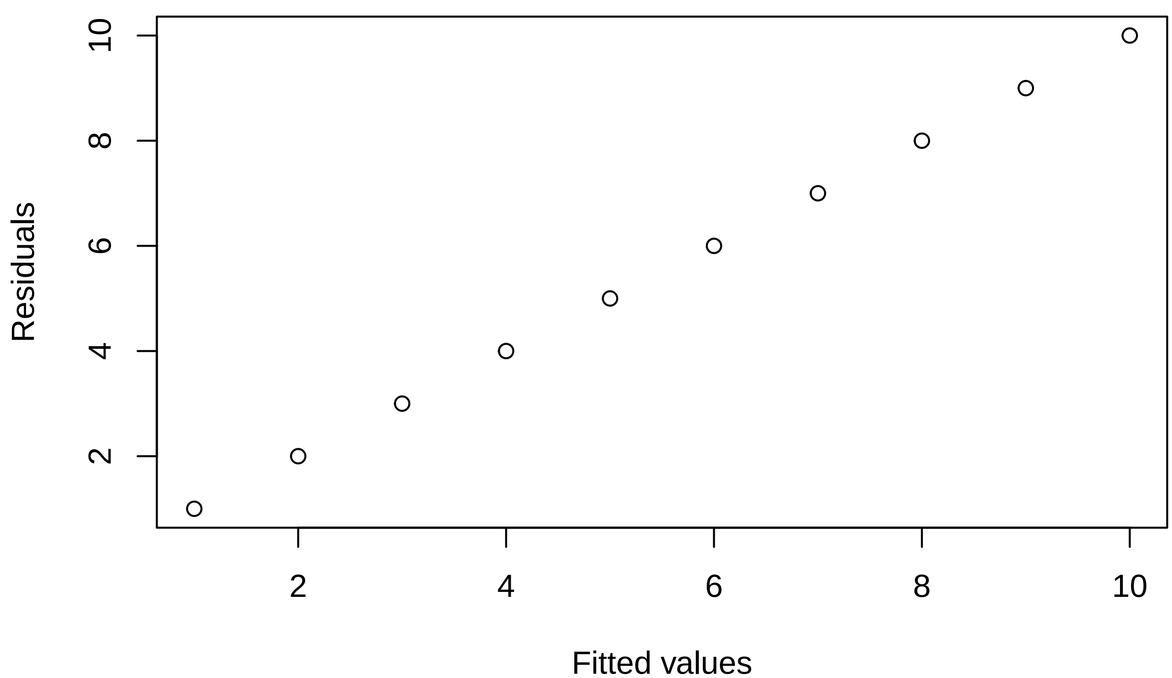
```
## beta0 = 0
```

- 1c) Calculate the predicted (fitted) values and the residuals.

```r
fit=beta0+y #use correct equation for the fitted values
resids=y ##use correct equation for the residuals
```

- 1d) Plot the residuals (y axis) vs fitted values (x axis).

```r
plot(1:10,1:10,xlab="Fitted values",ylab="Residuals") #correct the plot formula
```



- 1e) Calculate the mean of the residuals to verify it is zero (to machine precision), and the correlation of the residuals with X to verify it is also zero (to machine precision). The R command to calculate the correlation between u and v is cor(u,v).

```r
set.seed(43)  #drop this line
data=rnorm(3); data #drop this line
```

```
## [1] -0.03751376 -1.57460441 -0.48596752
```

```r
mean(data) #correct the call to mean.  this one averages 3 standard normals.
```

```
## [1] -0.6993619
```

```r
#you want to instead calculate the mean of the residuals.

#enter the call to cor here to get correlation of residuals with x.  see help(cor)

round(mean(data),5)  #round to five decimal digits. replace data with residuals.
```

```
## [1] -0.69936
```

```r
#enter the appropriate command here to round the correlation to 5 decimal digits.
```

```r
#the reason why we do the rounding is to get something that looks like 0. e.g.
z=1.23e-26
z; round(z,5)
```

```
## [1] 1.23e-26
```

```
## [1] 0
```

2. Some data gives the summaries: $n = 10$, $\sum_{i=1}^{10} x_i y_i = 100$, $\sum_{i=1}^{10} x_i = 20$ and $\sum_{i=1}^{10} y_i = 10$. Suppose that the response y is temperature in degrees Celsius.

   - $\bar{x} = 20/10 = 2$, $\bar{y} = 10/10 = 1$, so $S_{xy} = \sum_{i=1}^{10} x_i y_i - n\bar{x}\bar{y} = 100 - 10(2)(1) = 80$.
   - 2a) (behaviour of sums under linear transformation) If the response was converted to temperature in degrees Fahrenheit ($y' = 32 + 1.8y$), what is $\sum_{i=1}^{10} y_i'$?

$$\begin{aligned}\sum_{i=1}^{10} y_i' &= \sum_{i=1}^{10}(32 + 1.8y_i)\\ &=?\\ &=?\end{aligned}$$

(The final numerical answer is sufficient, but you may want to correct the above latex code.)

- 2b) (Behaviour of sums of squares under linear transformation of y.) If the response was converted to temperature in degrees Fahrenheit, what is $S_{xy'}$?

$$\begin{aligned}S_{xy'} &= \sum_{i=1}^{10}(x_i y_i') - (\sum_{i=1}^{10} x_i \sum_{i=1}^{10} y_i')/n\\ &= \sum_{i=1}^{10}(x_i(32 + 1.8y_i) - (\sum_{i=1}^{10} x_i \sum_{i=1}^{10} y_i')/n\\ &= 32\sum_{i=1}^{10} x_i + ?\\ &= 32(20) + ?\\ &=?\end{aligned}$$

(The final numerical answer is sufficient, but you may want to correct the above latex code.)

3. Find the equation of the line which passes through the points (1,1) and (4,5).

   Suppose that you found the slope to be (2/7) and the intercept to be -(1/3). You could enter your answer as

$$y = -\frac{1}{3} + \frac{2}{7}x$$

(Correct the above.)

4. Derive the partial derivative of $SSE = \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$ with respect to β_1.

5. Where $f(x) = log(x) + e^{x^2}$, derive $f'(x)$, the derivative of $f(x)$.
(Note: throughout this course $log(x) = log_e(x) = ln(x)$, the natural logarithm.)

6. Where $f(x) = arcsin(\sqrt{x})$, derive $f'(x)$.
(Note: *arcsin* is the inverse *sin* function. It will be clear later why this derivative is of interest.)

7.(Inverse of a diagonal matrix.)
Where

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

calculate M^{-1}, the inverse of M.

8. (Inverse of a 2×2 matrix.)
Where

$$M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

calculate M^{-1}, the inverse of M.