Lab 7 – physics 101

R

This lab will begin with the examination of some scatterplots. The file named data.csv conatins fifty observations each of four variables named **a**, **b**, **c**, and **d**. Read the data into R. (see the R3 handout)

	A	B	C	D
1	a	b	c	d
2	110.2	270.4	220	498.4
3	118.7	288.9	148.9	510.2
4	136.1	375	153.3	499.8
5	124.1	376.6	108.3	522.4
6	182.3	371.1	115.2	529
7	163.2	403.7	77.3	517.6
8	138	323.4	221.2	526.6
9	158	365.1	169.8	505.5
10	125.9	330.5	197.2	508

To look at a scatterplot of **a** and **b** (with **a** on the horizontal axis and **b** on the vertical axis) use this code:

```
> plot(a,b)
```

Based on what the scatterplot looks like, do you think the correlation between **a** and **b** is positive, negative, or close to zero?

Now make a specific guess for the value of r:

What is the exact value of r? Use this code:

```
> cor(a,b)
```

Just for fun, switch **a** and **b**. In other words, put **b** on the horizontal axis and **a** on the vertical axis.

```
> plot(b,a)
> cor(b,a)
```

Did switching **a** and **b** change the appearance of the scatterplot?

Did switching **a** and **b** change the value of r?

Now play the guessing game again with the other pairs of variables. First look at the scatterplot, then guess the value of r, then have R compute the exact value.

variables	comments on scatterplot	your guess for r	exact value of r
a and **c**			
a and **d**			
b and **c**			
b and **d**			
c and **d**			

You can have R compute the correlations between all pairs of variables at once. Suppose you named your data frame "data1" when you read the .csv file.

```
> cor(data1)
```

Here is one more exercise for you to try: add a least-squares regression line to any scatterplot and see if it has about the same intercept and slope as what you would have guessed.

```
> plot(a,b)
> abline(lm(b~a))
> lm(b~a)
```

SAS

Professor Patel teaches Physics 101 to undergraduates at a technical college. She allows her students unlimited time when completing the midterm exam and the final exam. She wonders if X = time needed to complete the midterm (minutes) will be a good predictor of Y = time needed to complete the final (minutes). Last year she collected data from 100 randomly selected students enrolled in her Physics 101 courses. Professor Patel's data is in the file named ph101.txt

Read the data into SAS (see the SAS3 handout) then use **proc sgplot** and **proc reg** (see the SAS8 handout) to draw the scatterplot and find the equation of the least-squares regression line. You will also see a lot of information in the output that we may not have talked about in class yet.

What is the equation of the least-squares line? Round the intercept and slope off to the second decimal place.

What is the prediction for the mean number of minutes needed to complete the final among students who needed 60 minutes to complete the midterm?

Your output may not tell you the exact value of r, but you should be able to find the value of r^2. Take the square root of it to find r. (Remember: when you calculate a square root you get both a positive and negative answer. Look at the scatterplot to determine if you are expecting a positive or negative value for r.)

Round off to the second decimal place:

r^2 =

r =

After you have completed this handout, complete the Canvas quiz titled:
Lab 07 – physics 101