Quiz 2

Name: _____

Red ID: _____

Use the following to answer questions 1 and 2

We are interested in predicting the cortisol levels of individuals, using either the average number of hours slept per night, or the average number of hours worked per week. We collect data on several subjects and obtain the following linear models and associated coefficients of determination:

cortisol = 32 - 1.8(sleep) $R^2 = 83\%$

cortisol = 43 + 0.3(work) $R^2 = 72\%$

1. Which variable would you use to predict subjects cortisol levels, sleep or work, and why?

a) Work, as the slope is positive, and so more accurately models the relationship
b) Sleep, the value of the coefficient of determination is higher, and so has higher predictive power
c) Work, the value of the coefficient of determination is higher, and so has higher predictive power
d) Sleep, as the slope is more significant and stronger than the slope for work
e) Impossible to determine from the information provided

2. What is the correct interpretation of the value of the correlation coefficient between cortisol and hours sleep?

a) There is a weak, positive linear relationship between cortisol and hours sleep
b) There is a strong, negative linear relationship between cortisol and hours sleep
c) There is a strong, positive linear relationship between cortisol and hours sleep
d) There is a weak, negative linear relationship between cortisol and hours sleep

$r = -\sqrt{0.83}$
$= -0.91$

Use the following for questions 3 through 5

To study the effect of aspirin on blood clotting times, we measure the clotting time of 4 subjects given varying amounts of aspirin. Data obtained is provided below.

Aspirin (mg)	Clotting time (seconds)
0	62
15	133
30	257
45	278

$\hat{y} = 5.1467(\text{aspirin}) + 66.7$

3. What is the predicted clotting time of a subject given 20 mg of aspirin?

a) 157.29
b) 181.85
c) 176.04
d) 169.63
e) 161.77

$\hat{y} = 5.1467(20) + 66.7$
$= 169.63$

4. One subject given 25 mg had a clotting time of 207 seconds. Find the residual for this observation, and state whether we made an over or an under prediction.

a) -11.63, under-prediction
b) -11.63, over-prediction
c) 11.63, under-prediction
d) 11.63, over-prediction

$\hat{y} = 5.1467(25) + 66.7$
$= 195.37$

$e = y - \hat{y}$
$= 207 - 195.37$
$= 11.63$

5. One subject given 40 mg of aspirin had a residual of -2.8. What is the actual clotting time for this subject?

a) 280.21
b) 260.07
c) 275.37
d) 269.77
e) 272.57

$\hat{y} = 5.1467(40) + 66.7$
$= 272.57$

$e = y - \hat{y}$
$-2.8 = y - 272.57$
$y = 269.77$

6. How do we determine which linear model best fits the data?

a) maximize $\sum (y_i - \hat{y}_i)^2$
b) minimize $\sum (y_i - \hat{y}_i)^2$
c) maximize $\sum (y_i - \bar{y})^2$
d) minimize $\sum (y_i - \bar{y})^2$

7. What is the sum of squared errors for the following linear model applied to the data given below?

blood pressure = 152 - 3(dosage)

dosage	blood pressure	$\hat{y}_i$	$y_i - \hat{y}_i$
7	136	131	5
9	124	125	-1
14	118	110	8

$\sum e_i^2 = (5)^2 + (-1)^2 + (8)^2$

$= 90$

a) 87
b) 90
c) 106
d) 82
e) 95

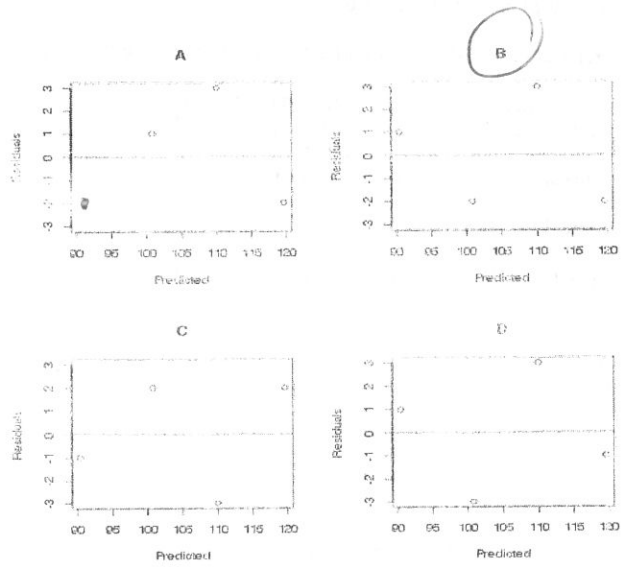
8. Soybean plants were raised in soil containing varying amounts of nitrogen and phosphorous. When fully grown their height was recorded. Data obtained was used in creating the following linear model. A portion of the data is also provided below. Using this data and linear model, which of the following is the correct residual plot for this data?

height = 122 - 1.2(phosphorus) - 0.8(nitrogen)

phosphorous	nitrogen	height	$\hat{y}_i$
0	3	117.6	119.4
4	9	113.0	110
9	13	98.8	100.8
15	17	91.4	90.4

$y_i - \hat{y}_i$

-2
3
-2
1



9. Consider the following possible residual plots. In which of the residual plots does it appear that the model used to create that residual plot is appropriate? That is, the model used fits the data well.

small residuals, no pattern or fanning