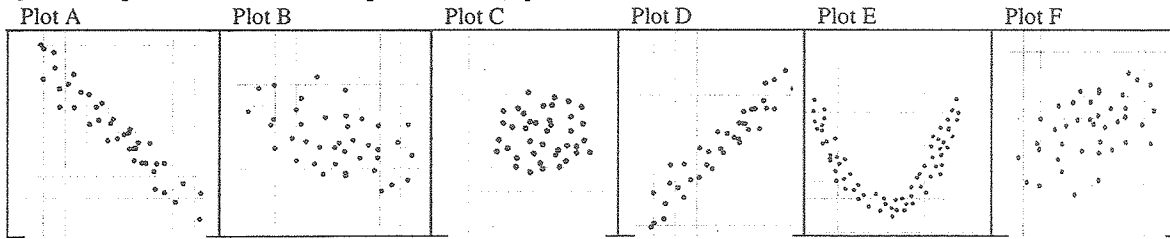


Question 1 pertains to the 6 scatter plots below: (6 pts.)



Determine whether the correlation coefficient is appropriate for analyzing the plots. If so, choose the r which best represents the plot.

Check only one box per row!

READ THIS- Each column will be used exactly once!

	Not appropriate to use r	$r = +0.9$	$r = +0.5$	$r = 0$	$r = -0.5$	$r = -0.9$
Plot A	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plot B	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plot C	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plot D	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plot E	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Plot F	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Question 2 (6 pts.)

For each of the following pairs of variables, check the box that best describes its correlation coefficient r .

Check only one box per row!

READ THIS- Each column will be used exactly once!

	Exactly +1	Between 0 and +1	About 0	Between 0 and -1	Exactly -1	Not Enough Information
As X increases, Y tends to increase.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
As X increases, Y tends to decrease.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
X and Y have no relation, they are just randomly paired.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Y is always less than X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Y is always exactly 1 less than X .	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
X and Y always add up to 100.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

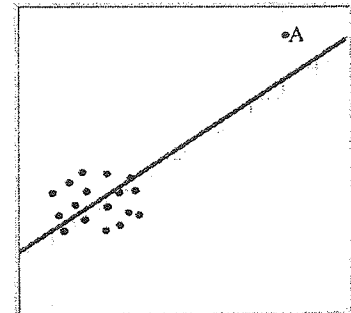
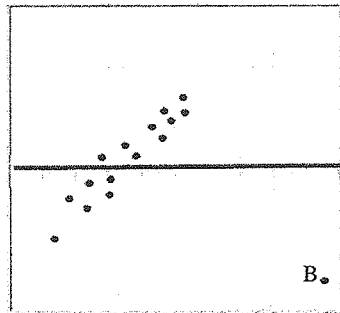
Question 3 pertains to the 2 scatter plots below. The regression line is shown for each plot. (4 pts.)

a) If we removed point B the correlation coefficient (r) would.... Choose one:

- i) Increase ii) Decrease
iii) Stay the Same iv) Not enough info

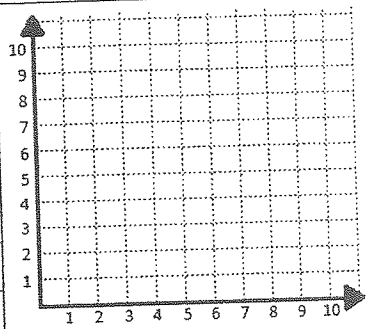
b) If we removed point A the correlation coefficient (r) would.... Choose one:

- i) Increase ii) Decrease
iii) Stay the Same iv) Not enough info



Question 4 (18 pts.)**Part A:** Calculate the correlation coefficient (r) by filling in the table. The averages of X and Y are both 7.The SD of X and Y are both 2. (13 pts.)**Plot the points on the graph below.**
(2.5 pts.)

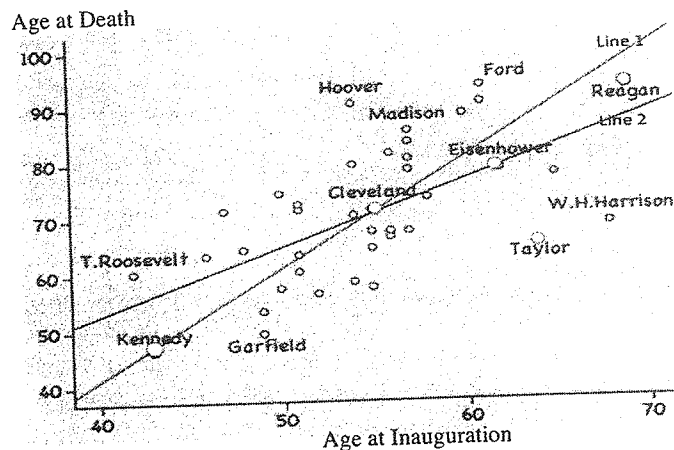
X	Y	Z-score for X	Z-score for Y	Products
4	7			
6	6			
7	10			
8	4			
10	8			
Totals		Total should = _____	Total should = _____	Total = _____


Part B: From above, find the correlation coefficient r . $r = \underline{\hspace{2cm}}$ (1pt)**Part C:** What happens to the correlation coefficient r if ... (Note: X and Y below refer to the original X and Y values given in part A) (Use the r you calculated in Part B to answer Part C. If you didn't answer Part B, then use $r = 0.3$ to answer Part C) (4 pts.)

- i) We subtract 2 from each Y value. $r = \underline{\hspace{2cm}}$ ii) We multiply each Y value by -3 . $r = \underline{\hspace{2cm}}$
- iii) We divide each Y value by 2. $r = \underline{\hspace{2cm}}$ iv) We swap the X and Y columns. $r = \underline{\hspace{2cm}}$

Question 5: The graph below plots the age at inauguration vs the age at death of 39 US Presidents. (14 pts.)

- a) The SD Line and the Regression Line are shown. Which is the SD Line?
Choose One: i) Line 1 ii) Line 2
- b) The correlation is closest to ... Choose One:
i) 0.2 ii) 0.6 iii) 0.9 iv) -0.2 v) -0.6
- c) About what is the average age at Death?
Choose One: i) 50 ii) 55 iii) 60 iv) 65 v) 70
- d) About what is the average age at Inauguration?
Choose One: i) 50 ii) 55 iii) 60 iv) 65 v) 70
- e) Taylor is an outlier that _____ the correlation.
Choose One:
i) strengthens ii) weakens iii) does not affect



- f) Which president(s) died exactly when predicted by the regression estimate?
Circle all that died when predicted: i) Kennedy ii) Cleveland iii) Eisenhower iv) Reagan
- g) Which president(s) have the same z-scores for age at Inauguration and age at Death?
Circle all that have the same z-scores: i) Kennedy ii) Cleveland iii) Eisenhower iv) Reagan
- h) The regression equation for predicting Age at Death from Age at Inauguration is: $\text{Age}_{\text{death}} = 1.2 (\text{Age}_{\text{Inauguration}}) + 5.4$
Fill in the 2 blanks below with numbers.
- i) Ford was inaugurated at age 61, what is the regression estimate for his age at death? _____ years (Don't round.)
- ii) Ford actually died at age 93. What is his residual (also called his prediction error)? _____ years (Don't round.)

Question 6 (8 pts.)

Suppose the correlation between IQ scores and ACT scores have the following summary statistics among the population of students who take both exams and that the scatter plot is football-shaped:

	Average	SD
<i>IQ</i>	100	15
<i>ACT</i>	20	5

Correlation: $r = 0.8$

- a) What is the slope of the regression equation when predicting ACT scores from IQ scores? Choose one:
 i) 3 ii) 2.4 iii) 0.8 iv) 0.333 v) 0.267
- b) What is the SD of the prediction errors (the RMSE) when predicting ACT scores from IQ scores? Choose one:
 i) 5 ii) 15 iii) $\sqrt{1 - 0.8^2} \times 15$ iv) $\sqrt{1 - 0.8^2} \times 5$ v) 4
- c) The regression equation predicts an ACT score of 24 for those who score 115 on the IQ. Of course not everyone with an IQ of 115 will score 24 on the ACT test. Instead there's a range of ACT scores, with about 68% of them scoring between ...
 Choose one:
 i) 15 and 25 ii) 21 and 27 iii) 23 and 25 iv) 24 and 33
- d) The regression equation for predicting IQ scores from ACT scores is: $IQ = 2.4 \times (ACT) + \underline{\hspace{1cm}}$. What is the y-intercept? (the y-intercept is the blank in the above equation.)
 Choose one: i) 52 ii) -52 iii) -21.6 iv) 21.6 v) 0

Question 7 (12 pts.)

To the right are the 5 summary statistics for the ages of fathers and mothers of a large group of students (roughly based on our Survey data).

	Average	SD
<i>Father's Age</i>	52	6
<i>Mother's Age</i>	48	5

Correlation: $r = 0.8$

- a) In the table below you are given the age of the fathers for 2 students. For each, compute the regression estimate for the age of the student's mother. Show work, where indicated.

Father's Age	Father's Z-score	r	Mother's Z-score	Mother's Age
Father is 64 years. Show work for full credit.	$Z = \underline{\hspace{1cm}}$ (2pt)	$r = 0.8$	$Z = \underline{\hspace{1cm}}$ (1pt)	Mother's Age = $\underline{\hspace{1cm}}$ (1pt) Show work for full credit.
Father is 46 years. Show work for full credit.	$Z = \underline{\hspace{1cm}}$ (2pt)	$r = 0.8$	$Z = \underline{\hspace{1cm}}$ (1pt)	Mother's Age = $\underline{\hspace{1cm}}$ (1pt) Show work for full credit.

- b) One student has a father who is 1.5 SD's above average in age. The regression estimate for his mother's age is $\underline{\hspace{1cm}}$ SD's above average. Fill in the blank in the sentence with the correct number. (2 pt.)
- c) Students with 58 year-old fathers have mothers who are 52 years old on the average. Students with 52 year-old mothers have fathers who are about $\underline{\hspace{1cm}}$ years old on the average. (2 pt.)
 Choose the closest answer: i) 52 ii) 56 iii) 58 iv) 60 v) not enough info given.

Question 8 (2 pts)

The average yearly wine consumption and average life span in 50 countries was computed. The correlation for the 50 pairs of averages was 0.8. Can you conclude that the correlation between wine consumption and life span for the millions of individuals in these 50 countries is also 0.8? Choose one:

- i) Yes, since it's based on the same info.
 ii) No, it's likely to be higher than 0.8 since it's based off millions of data points instead of just 50.
 iii) No, it's likely to be lower than 0.8 since the millions of data points will produce scatter within each country.

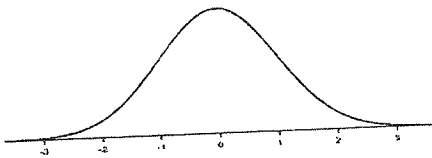
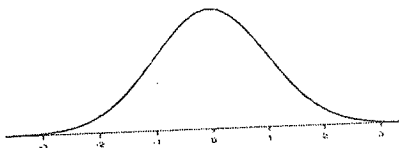
Question 9 (9 pts.)

Suppose reading skills and math skills of grade school children follow the normal curve but have different correlations among different populations.

- a) Consider 5 populations where the correlation coefficients between reading and math skills are as given in the table below. If a child is in the 75th percentile in reading skills, estimate her percentile in math skills in each population.

Reading Skills Percentile	r	Math Skills Percentile
75 th	0	Choose One: 20 th 25 th 30 th 50 th 70 th 75 th 80 th
75 th	1	Choose One: 20 th 25 th 30 th 50 th 70 th 75 th 80 th
75 th	-1	Choose One: 20 th 25 th 30 th 50 th 70 th 75 th 80 th
75 th	0.8	Choose One: 20 th 25 th 30 th 50 th 70 th 75 th 80 th
75 th	-0.8	Choose One: 20 th 25 th 30 th 50 th 70 th 75 th 80 th

- b) If a student is in the 42nd percentile in reading skills where $r = 0.5$, estimate her math skills percentile *by filling in the table below*. You may round areas and z-scores to fit the nearest line on the table. Include *negative* sign for *negative* Z scores!

Reading Percentile	Reading Z	r = 0.5	Math Z	Math Percentile
Student is in the 42 nd percentile for reading skills. (That means 42% scored lower.)	Z = _____	r = 0.5	Z = _____	Math Skills Percentile = _____ (1 pt) (Percentile is --What % scored <i>lower</i>)
What middle area on the table should you look up to find the Z score? _____ % (1 pt)			(1 pt)	Mark the Z score on the graph below. Use the normal table to find the middle area. Write percentile in blank above.
				
Correctly mark graph and write the correct Z in the next column. (1pt.)				Round the middle area given in the NormalTable to the nearest WHOLE number and then calculate the percentile.

Question 10 pertains to a drug screening test for bus drivers. Suppose only 5% of bus drivers who take the test are really using drugs. Suppose 80% of those using drugs will correctly test positive, but 20% of those *NOT* using drugs will also test positive. Fill in the table for a typical sample of 100 people who get tested. (6 pts)

	Tests Positive	Tests Negative	Total
Using Drugs	(1pt)		Hint: Fill in this blank first using the underlined info above. (1pt)
Not Using Drugs	(1pt)		
Total			100

Check that rows and columns sum correctly. (1pt)

- a) A person tests positive, what's the chance (s)he is really using drugs? _____ (leave answer as a fraction) (1pt)
- b) A person tests negative, what's the chance (s)he is really using drugs? _____ (leave answer as a fraction) (1pt)

Question 11 pertains to the table below that shows last semester's survey responses to 2 questions: "Have you ever tried marijuana?" and "Should the regulations on marijuana be more strict than those for alcohol, the same as those for alcohol, or less strict as those for alcohol?"

	More Strict	Same	Less Strict	Totals
Tried Marijuana	59	188	162	409
Never Tried Marijuana	178	173	39	390
Totals	237	361	201	799

Suppose you randomly draw a student from those who answered this survey: (4 pts)

- What is the chance that you'll get someone who answered "Less Strict"?
i) 162/201 ii) 162/409 iii) 409/799 iv) 201/799 v) 162/799
- What is the chance you'll get someone who answered "Less Strict" if you draw only from those who tried marijuana?
i) 162/201 ii) 162/409 iii) 409/799 iv) 201/799 v) 162/799
- What is the chance that you'll get someone who tried marijuana if you draw only from those who answered "Less Strict"?
i) 162/201 ii) 162/409 iii) 409/799 iv) 201/799 v) 162/799
- What is the chance that a randomly drawn student will either have tried marijuana *or* have answered "Less Strict"?
i) 162/201+162/409 ii) 162/201+162/409-162/799 iii) (409 + 201-162)/799 iv) (409 + 201)/799 v) 162/799

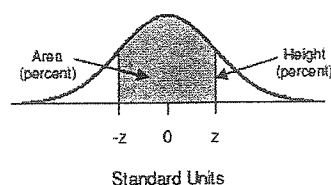
Question 12 pertains to a well-shuffled deck of 52 cards. A deck of cards has 4 suits: clubs, diamonds, hearts and spades. There are 13 cards in each suit: 2 through 10, jack, queen, king, ace. (So there are 4 Kings, 4 Queens and 13 Hearts.) (4 pts.)

- Draw 2 cards with replacement. What is the chance that the first card is a King and the second is a Heart?
i) $4/52 + 13/52$ ii) $4/52 + 13/52 - 1/52$ iii) $4/52 * 3/51$ iv) $4/52 * 13/52$ v) $4/52 * 13/51$
- Draw 2 cards without replacement. What is the chance that the both cards are Queens?
i) $4/52 + 4/52$ ii) $4/52 + 4/52 - 1/52$ iii) $4/52 * 4/52$ iv) $4/52 * 3/51$ v) $4/52 * 4/51$
- Draw *one* card. What is the chance that it's *either* a King *or* a Queen?
i) $4/52 + 4/52 - 1/52$ ii) $4/52 + 4/52$ iii) $4/52 * 4/52$ iv) $4/52 * 3/51$ v) $4/52 * 4/51$
- Draw *one* card. What is the chance that it's *either* a King *or* a Heart?
i) $4/52 + 13/52 - 1/52$ ii) $4/52 + 13/52$ iii) $4/52 * 3/51$ iv) $4/52 * 13/51$ v) $4/52 * 13/52$

Question 13 pertains to rolling fair dice. (7 pts.)

- Two dice are rolled. What is the chance that the sum of the spots is 5?
i) 2/36 ii) 3/36 iii) 4/36 iv) 5/36 v) $1/6 * 1/6$ vi) $1/6 + 1/6$
- One die is rolled 4 times. What is the chance of getting *no* 4's?
i) $(5/6)^4$ ii) $(1/6)^4$ iii) $1 - (5/6)^4$ iv) $1 - (1/6)^4$ v) 4/6
- One die is rolled 4 times. What is the chance of *at least one* 4?
i) $(5/6)^4$ ii) $(1/6)^4$ iii) $1 - (5/6)^4$ iv) $1 - (1/6)^4$ v) 4/6
- One die is rolled 3 times. What is the chance of getting *all* 2's?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1 - (5/6)^3$ iv) $1 - (1/6)^3$ v) 3/6
- One die is rolled 3 times. What is the chance of getting *not all* 2's?
i) $(5/6)^3$ ii) $(1/6)^3$ iii) $1 - (5/6)^3$ iv) $1 - (1/6)^3$ v) 3/6
- One die is rolled twice. What is the chance that the first roll is a 2 *or* the second roll is a 3?
i) 1/36 ii) 2/36 iii) 6/36 iv) 11/36 v) 12/36
- One die is rolled twice. What is the chance that the first roll is a 2 *and* the second roll is a 3?
i) 1/36 ii) 2/36 iii) 6/36 iv) 11/36 v) 12/36

STANDARD NORMAL TABLE



z	Height	Area	z	Height	Area	z	Height	Area
0.00	39.89	0.00	1.50	12.95	86.64	3.00	0.443	99.730
0.05	39.84	3.99	1.55	12.00	87.89	3.05	0.381	99.771
0.10	39.70	7.97	1.60	11.09	89.04	3.10	0.327	99.806
0.15	39.45	11.92	1.65	10.23	90.11	3.15	0.279	99.837
0.20	39.10	15.85	1.70	9.40	91.09	3.20	0.238	99.863
0.25	38.67	19.74	1.75	8.63	91.99	3.25	0.203	99.885
0.30	38.14	23.58	1.80	7.90	92.81	3.30	0.172	99.903
0.35	37.52	27.37	1.85	7.21	93.57	3.35	0.146	99.919
0.40	36.83	31.08	1.90	6.56	94.26	3.40	0.123	99.933
0.45	36.05	34.73	1.95	5.96	94.88	3.45	0.104	99.944
0.50	35.21	38.29	2.00	5.40	95.45	3.50	0.087	99.953
0.55	34.29	41.77	2.05	4.88	95.96	3.55	0.073	99.961
0.60	33.32	45.15	2.10	4.40	96.43	3.60	0.061	99.968
0.65	32.30	48.43	2.15	3.96	96.84	3.65	0.051	99.974
0.70	31.23	51.61	2.20	3.55	97.22	3.70	0.042	99.978
0.75	30.11	54.67	2.25	3.17	97.56	3.75	0.035	99.982
0.80	28.97	57.63	2.30	2.83	97.86	3.80	0.029	99.986
0.85	27.80	60.47	2.35	2.52	98.12	3.85	0.024	99.988
0.90	26.61	63.19	2.40	2.24	98.36	3.90	0.020	99.990
0.95	25.41	65.79	2.45	1.98	98.57	3.95	0.016	99.992
1.00	24.20	68.27	2.50	1.75	98.76	4.00	0.013	99.9937
1.05	22.99	70.63	2.55	1.54	98.92	4.05	0.011	99.9949
1.10	21.79	72.87	2.60	1.36	99.07	4.10	0.009	99.9959
1.15	20.59	74.99	2.65	1.19	99.20	4.15	0.007	99.9967
1.20	19.42	76.99	2.70	1.04	99.31	4.20	0.006	99.9973
1.25	18.26	78.87	2.75	0.91	99.40	4.25	0.005	99.9979
1.30	17.14	80.64	2.80	0.79	99.49	4.30	0.004	99.9983
1.35	16.04	82.30	2.85	0.69	99.56	4.35	0.003	99.9986
1.40	14.97	83.85	2.90	0.60	99.63	4.40	0.002	99.9989
1.45	13.94	85.29	2.95	0.51	99.68	4.45	0.002	99.9991