

BUSINESS DATA ANALYSIS, A15, TEST 1

Name: _____

Student number _____

Instructions:

- Please show all of your work on the test paper. Partial marks will be awarded in the event of an incorrect answer. Conversely, if the incorrect method has been implemented, or no work is shown, then no marks will be awarded even though the answer is correct.
- Round all decimals to two digits after the decimal point.
- All unauthorized materials must be put away such as cell phones, listening devices, etc. Programmable calculators and formula sheets are not allowed.
- **Cheating is a serious academic offense. Anyone caught cheating or aiding in the act of cheating will be given a mark of zero for this assessment, and a note will be placed in his or her file.**

Formulas:

$$\bar{x} = \frac{\sum x}{n}$$

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

$$s = \sqrt{s^2}$$

$$z = \frac{x-\bar{x}}{s}$$

$$IQR = Q_3 - Q_1$$

$$r = \frac{\sum(x-\bar{x})(y-\bar{y})}{(n-1)s_x s_y}$$

$$b_1 = r \frac{s_y}{s_x}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

$$\hat{y} = b_0 + b_1 x$$

- (1) (3 marks) Sandi teaches a section of Business Data Analysis for Business Administration this semester. In the second week of classes, she gave a quiz, and the marks are shown below:

50 60 65 72 75 83 85 94

- (a) Determine the mean.
- (b) Determine the variance and standard deviation.
- (c) What is the z -score for the student with a mark of 60?

- (2) (6 marks) The table below contains data on the results on a Term Test, x , for five students versus the results for the same five students on a Final exam, y :

x	y
25	42
50	73
60	64
77	75
88	96

For all the values below compute with precision of two digits after the decimal point.

- Draw a scatterplot.
- Compute the means and the standard deviations both for x and for y .
- Compute the correlation coefficient r .
- Compute the intercept and the slope of the regression line.
- Draw the regression line on your scatterplot.
- What percentage of the variance in the data is captured by the regression model.
- Based on R^2 is this an acceptable model?
- Make a prediction for the Final Exam mark for a student with a mark of 70 on the Term Test.

- (3) (4 marks) A survey was done at Center Hospital to determine how long (in months) clerical staff had been in their current positions. The responses (in months) of 20 clerical staff were

3 6 7 14 15 17 18 20 22 22
24 25 25 26 29 31 31 32 42 92

Determine the following

- (a) The five number summary.
- (b) The *IQR*.
- (c) Compute the values for the upper and lower fences on the box-and-whiskers plot.
- (d) Draw a box-and-whiskers plot for the data above. Clearly label values for the 5 point summary. Do not forget to draw the outliers, if any.

- (4) (5 marks) Alice and Bob are comparing data for two airlines. Here are the numbers of flights on time and delayed for two airlines at three airports during a one-month period.

	United Airlines		Delta Airlines	
	On time	Delayed	On time	Delayed
New York	497	62	694	117
Chicago	221	12	4840	345
San Diego	212	20	383	65
Total	930	94	5917	527

Alice observes that overall, Delta Airlines has a higher percentage of on time flights over United Airlines. Bob, on the other hand, notices that Delta has more delays than United when comparing the data by each airport.

- Show that Alice's argument is correct, that Delta performs better overall than United.
- Show that Bob's argument is correct, that Delta performs worse at every one of the three airports.
- Explain what is happening, and give your own interpretation of which airline is better in terms of on time flights.

- (5) (3 marks) The following table shows attendance data collected during the period of 2002 to 2006. Figures are in millions of movie admission.

Year	Patron Age						Total
	12 to 24	25 to 29	30 to 39	40 to 49	50 to 59	60 and Over	
2002	553	125	245	247	104	104	1378
2003	565	111	271	185	164	110	1406
2004	573	131	240	229	126	115	1414
2005	496	133	183	200	120	115	1247
2006	500	139	262	200	124	108	1333
Total	2687	639	1201	1061	638	552	6778

- (a) What percent of admissions in 2003 were bought by people between the ages of 40 and 49?
- (b) What percent of admissions bought by people over 60 years old were bought in 2005?
- (c) What is the conditional distribution for the year 2004 by age group?