PS946 Fundamental Statistics for Research

Coursework Test 2 Specimen Paper

Time Allowed: Three (3) Hours

UNIVERSITY REGISTRATION NUMBER:

THIS TEST CONSISTS OF FOURTEEN SHORT ANSWER QUESTIONS

Write your answers using the spaces provided in this booklet.

Full marks can be obtained by answering correctly ALL questions.

Each question indicates its total marks available.

- **You are permitted to bring to your place a university-approved calculator and SIX sides of A4 hard copy notes**
- **Critical values of *F* are given on page 16**
- **The total marks available for this paper is 80, this will be converted into a percentage**

THIS IS A RESTRICTED PAPER AND MUST NOT BE REMOVED FROM THE EXAMINATION ROOM

- Do not leave your seat unless you are given permission by an invigilator.
- Do not communicate in any way with any other candidate in the test room.
- Do not open this question booklet until told to do so.
- All answers must be written in this question booklet, sufficient space has been provided.
- All rough work should be written in the question booklet.
- At the end of the test, remain seated until your question booklet has been collected and you have been told you may leave.

Version 1.1, 19th December 2023

Question 1

For this question, there are a series of tables, each represents the scores from a small within-subjects design experiment, but with missing data.

1a) Fill in the missing numbers on the table below such that the *within-group variance* comes to zero, then tick the boxes on the right to answer the questions about other types of variance.

Subject	Level A_1	Level A_2	**For the data on the left:**	Zero	More than zero
S_1	3	1	The between-group variance is	☐	☐
S_2	☐	☐	The individual-differences variance is	☐	☐
S_3	☐	☐	The residual variance is	☐	☐
S_4	☐	☐			
S_5	☐	☐			

1b) Fill in the missing numbers on the table below such that the *between-group variance* comes to zero, USE POSITIVE NUMBERS ONLY, then tick the boxes on the right to answer the questions about other types of variance.

Subject	Level A_1	Level A_2	**For the data on the left:**	Zero	More than zero
S_1	3	1	The within-group variance is	☐	☐
S_2	4	☐	The individual-differences variance is	☐	☐
S_3	4	☐	The residual variance is	☐	☐
S_4	4	☐			
S_5	5	☐			

1c) Fill in the missing numbers on the table below such that the *residual variance* comes to zero, then tick the boxes on the right to answer the questions about other types of variance.

Subject	Level A_1	Level A_2	**For the data on the left:**	Zero	More than zero
S_1	3	1	The between-group variance is	☐	☐
S_2	4	☐	The within-group variance is	☐	☐
S_3	4	☐	The individual-differences variance is	☐	☐
S_4	4	☐			
S_5	5	☐			

[Question 1 continues on the next page]

1d) Fill in the missing numbers on the table below such that the *between-subject variance* comes to zero, then tick the boxes on the right to answer the questions about other types of variance.

Subject	Level A_1	Level A_2	**For the data on the left:**	Zero	More than zero
S_1	3	1	The between-group variance is	☐	☐
S_2	4	☐	The within-group variance is	☐	☐
S_3	4	☐	The residual variance is	☐	☐
S_4	4	☐			
S_5	5	☐			

[8 Marks]

[Question 2 commences on the next page]

Question 2

Two small separate within-subjects design experiments were performed and their means are given (along with standard deviations in brackets) in the tables below. Note that by coincidence these are identical between experiments. However, when statistical tests were performed (single factor within-subjects ANOVA) it was found that one pair of differences was significant but not the other.

	Experiment 1				Experiment 2		
	Level A_1	Level A_2	Significance (ANOVA)		Level A_1	Level A_2	Significance (ANOVA)
Mean (*SD*)	20 (*5*)	15 (*5*)	$p > .05$ NS	Mean (*SD*)	20 (*5*)	15 (*5*)	$p < .05$ sig

Briefly suggest and comment upon the possible reasons why, for a within-subjects design, identical pairs of means with identical standard deviations can have a significant difference in one instance, but a non-significant difference in another. Relate these suggestions to the two experiments shown here.
Assume that the statistical tests were performed correctly.

[4 Marks]

Question 3

Below are six datasets labelled A to F.
There are also six ANOVA tables.

Write the letter of each dataset in the space provided under the ANOVA table which corresponds to the dataset.
There is no need to perform ANY calculations.

Dataset A

	Level A_1		Level A_2
S_1	3	S_5	3
S_2	5	S_6	2
S_3	2	S_7	2
S_4	2	S_8	1
Means	4		2

Dataset B

	Level A_1	Level A_2
S_1	1	2
S_2	1	2
S_3	1	2
S_4	1	2
Means	1	2

Dataset C

	Level A_1		Level A_2
S_1	2	S_5	2
S_2	4	S_6	3
S_3	5	S_7	4
S_4	5	S_8	7
Means	4		4

Dataset D

	Level A_1	Level A_2
S_1	1	2
S_2	2	2
S_3	2	4
S_4	3	4
Means	2	3

Dataset E

	Level A_1		Level A_2
S_1	3	S_5	4
S_2	3	S_6	4
S_3	3	S_7	4
S_4	3	S_8	4
Means	3		4

Dataset F

	Level A_1	Level A_2	Level A_3
S_1	1	2	2
S_2	2	2	2
S_3	2	4	4
S_4	3	4	4
Means	2	4	4

Srce	SS	df	MS	*F*
A	2	1	2	1.5
S/A	8	6	1.3	

Dataset: A

Srce	SS	df	MS	*F*
A	0	1	0	0
S/A	20	6	3.3	

Dataset: C

Srce	SS	df	MS	*F*
A	2	1	2	∞
S/A	0	6	0	

Dataset: E

Srce	SS	df	MS	*F*
A	2	1	2	6
S	5	3	1.7	
AxS	1	3	0.3	

Dataset: D

Srce	SS	df	MS	*F*
A	2	1	2	∞
S	0	3	0	
AxS	0	3	0	

Dataset: B

Srce	SS	df	MS	*F*
A	2.7	2	1.3	6
S	8.7	3	2.9	
AxS	1.3	6	0.2	

Dataset: F

[6 Marks]

Question 4

The following follow-up tests can all be used in order to investigate significant effects found when performing a one-factor ANOVA: *Tukey test*; *Planned comparisons*.

For each test, give an example of ONE experiment/prediction where it is particularly appropriate to use the test, and say why.
You may use the same experimental design to illustrate each test.
You do not have to describe experimental details nor how to perform each test.

[4 Marks]

Question 5

Whenever a two-factor Analysis of Variance is performed, it is often the case that the main summary ANOVA table tells us very little about the pattern of results obtained, and further follow-up tests are required. Each of the three ANOVA summary tables below is from a two-factor fully between-subjects design. A and B on each table represent the main effects and AxB represents the interaction.

For each table, suggest suitable follow-up tests in order fully to understand the patterns of significances displayed. Explain your choice of test(s). If you believe that no follow-up tests are necessary, then also explain why.
Assume that no planned comparisons are intended.

Srce	SS	df	MS	*F*
A	60	3	20	20
B	1	1	1	1
AxB	1	1	1	1
S/AB	20	20	1	

Srce	SS	df	MS	*F*
A	20	1	20	20
B	20	1	20	20
AxB	1	1	1	1
S/AB	20	20	1	

[Question 5 continues on the next page]

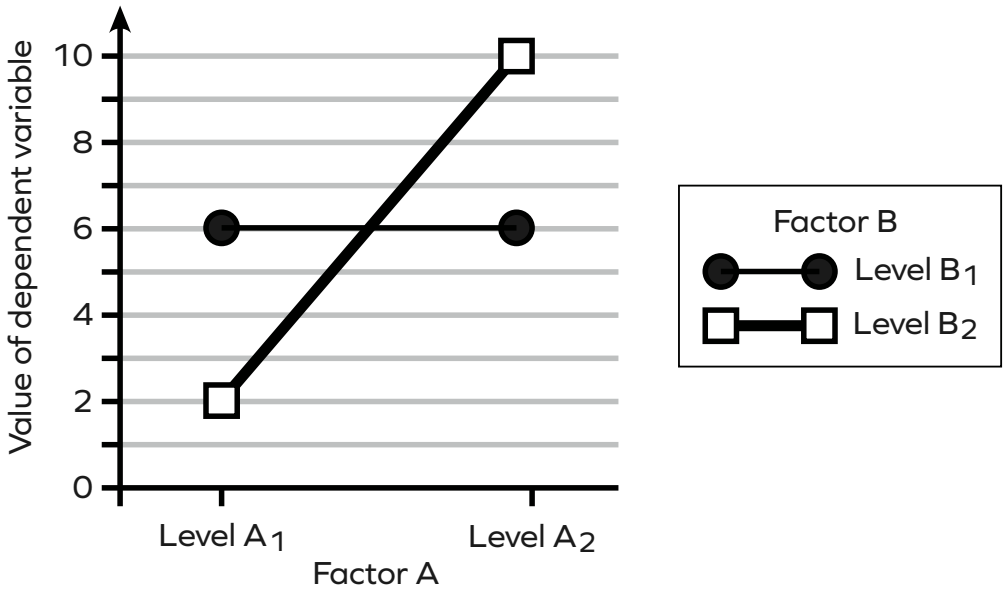
Srce	SS	df	MS	*F*
A	1	1	1	1
B	1	1	1	1
AxB	20	1	20	20
S/AB	20	20	1	

[6 Marks]

[Question 6 commences on the next page]

Question 6

Consider the following interaction plot and table of cell means and answer the following questions by writing the appropriate answers in the spaces provided.



A_1B_1 cell mean: 6	A_2B_1 cell mean: 6
A_1B_2 cell mean: 2	A_2B_2 cell mean: 10

6a) For each of the level means identify the pair of cell mean numerical values that comprise its components and calculate the value of the level mean

Level mean	Comprising cell means	Value of mean	Level mean	Comprising cell means	Value of mean
A_1			B_1		
A_2			B_2		

[Question 6 continues on the next page]

6b) For each the two Main Effects: Identify the appropriate pair of level means that are compared to test the effect; find the size of the Main Effect (the difference between the pair of means; determine whether the Main Effect is significant.

[Assume that a pair of means differs significantly if their difference is THREE or more]

Main Effect	Pair of means to be compared	Size of effect	Significant effect?
A (A_1 vs. A_2)			**Y | N**
B (B_1 vs. B_2)			**Y | N**

6c) For each of the four Simple Main Effects: Identify the appropriate pair of cell means that are compared to test the effect; find the size of the Simple Main Effect (the difference between the pair of means); determine whether the Simple Main Effect is significant.

[Assume that a pair of means differs significantly if their difference is THREE or more]

Simple Main Effect	Pair of means to be compared	Size of effect	Significant effect?
A at B_1 (A_1B_1 vs. A_2B_1)			**Y | N**
A at B_2 (A_1B_2 vs. A_2B_2)			**Y | N**
B at A_1 (A_1B_1 vs. A_1B_2)			**Y | N**
B at A_2 (A_2B_1 vs. A_2B_2)			**Y | N**

[Question 6 continues on the next page]

6d) For the effects of Factor A circle each appropriate answer

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Simple Main Effect of A at B_1 *and* Simple Main Effect of A at B_2?

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Main effect of A *and* Simple Main Effect of A at B_1

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Main effect of A *and* Simple Main Effect of A at B_2

Do the above three answers imply a **S***ignificant* **S | NS**
or a **N***on-***S***ignificant* interaction?

6e) For the effects of Factor B circle each appropriate answer

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Simple Main Effect of B at A_1 *and* Simple Main Effect of B at A_2?

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Main effect of B *and* Simple Main Effect of B at A_1

Do the following effects **A***gree* or **D***isagree* with each other? **A | D**
Main effect of B *and* Simple Main Effect of B at A_2

Do the above three answers imply a **S***ignificant* **S | NS**
or a **N***on-***S***ignificant* interaction?

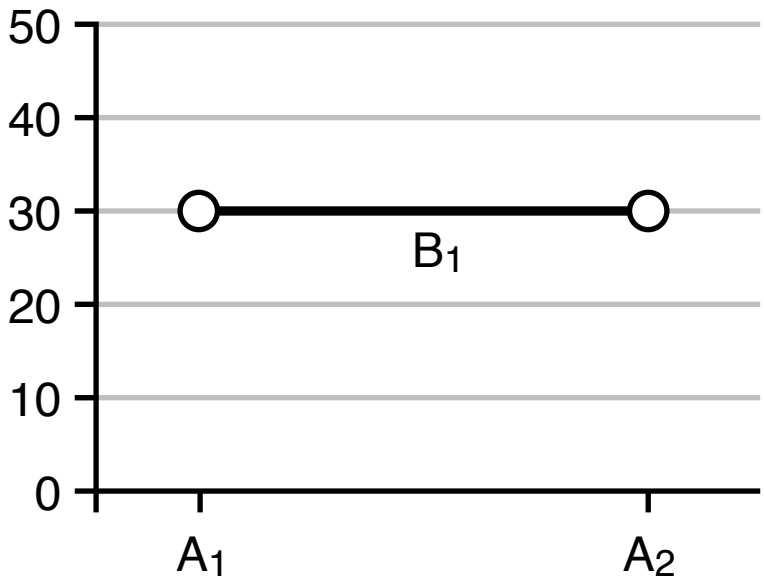
[13 Marks]

[Question 7 commences on the next page]

Question 7

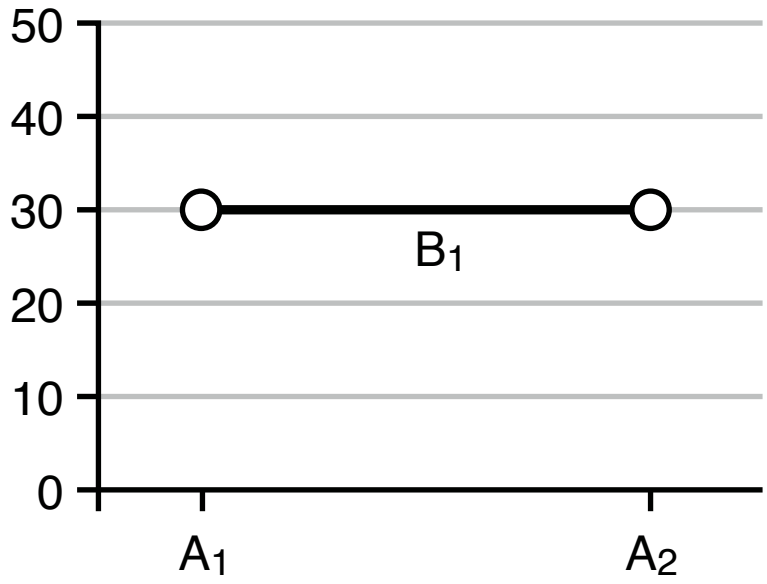
The three interaction plots below show just two of the means from a two-factor design, each factor with two levels. Below each graph, a table indicates which of the effects are significant. ASSUME THAT ANY DIFFERENCE BETWEEN MEANS THAT IS TEN OR GREATER IS STATISTICALLY SIGNIFICANT.

For each example, complete the interaction plot by adding two means, linked by a line. You need to make sure that the overall pattern matches the effects specified. NB. If you think that it is impossible to complete the plot to show the effects as specified, then write "impossible" on the plot.

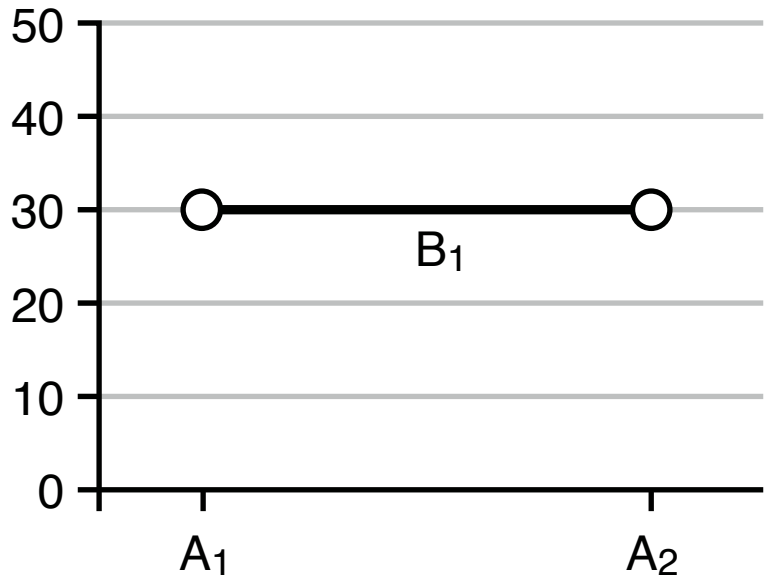


Main Effect of Factor A: Significant
Main Effect of Factor B: Significant
Interaction: Significant

[Question 7 continues on the next page]



Main Effect of Factor A: Significant
Main Effect of Factor B: Non-significant
Interaction: Non-significant



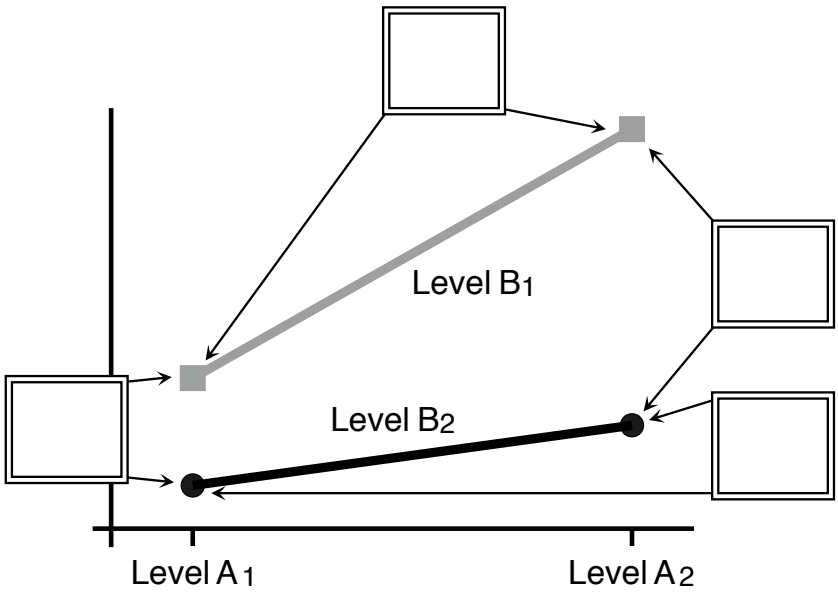
Main Effect of Factor A: Significant
Main Effect of Factor B: Non-significant
Interaction: Significant

[6 Marks]

Question 8

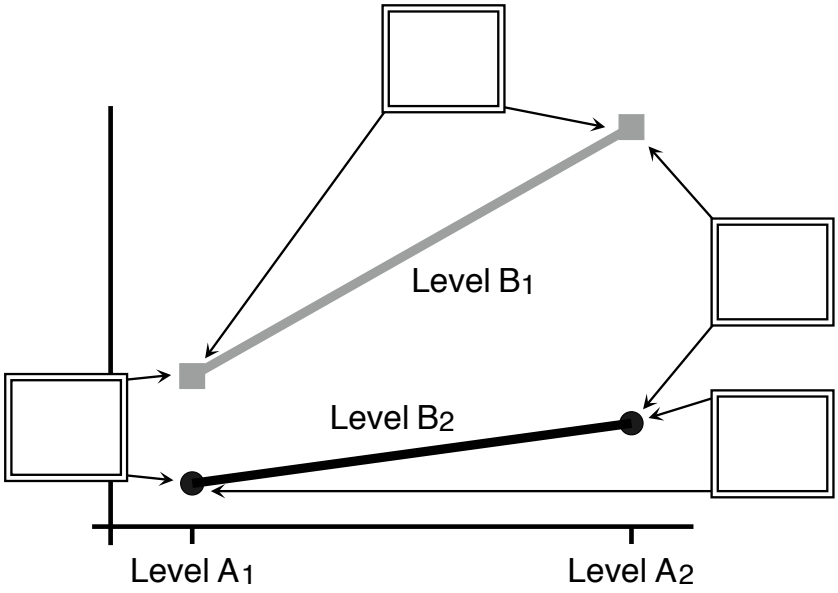
With a two-factor ANOVA, one, both or neither of the main effects may be significant. The interaction also may or may not be significant. These are related to which (if any) of the simple main effects are significant. On the next two pages are three interaction plots. Each has its four simple main effects marked with a pair of arrows. Below each plot is a table showing the significances of the main effects of Factor A, Factor B, and the interaction.

Using the information on the table for each plot, decide on which (if any) of the four simple main effects are likely to be significant. Show this clearly writing "S" in a box to denote a significant simple main effect and "NS" to denote a non-significant one.
There may be more than one acceptable combination of significances for any given plot, but you may not give the same answer combination twice for two different plots.

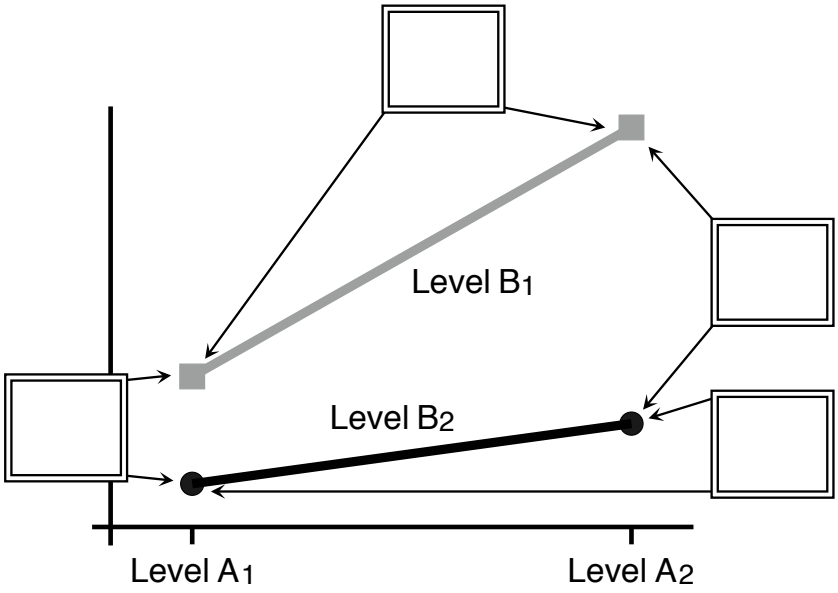


Factor A :	Non-significant effect
Factor B:	Non-significant effect
Interaction:	Non-significant effect

[Question 8 continues on the next page]



Factor A: Non-significant effect
Factor B: Significant effect
Interaction: Significant effect



Factor A: Significant effect
Factor B: Significant effect
Interaction: Non-significant effect

[6 Marks]

[Question 9 commences on page 17]

Critical Values of the F *Distribution*

.05 significance level in bold type

.01 significance level in plain type

degrees of freedom of numerator

degrees of freedom of denominator	1	2	3	4	5
1	**161**	**200**	**216**	**225**	**230**
	4052	4999	5403	5625	5764
2	**18.5**	**19.0**	**19.2**	**19.2**	**19.3**
	98.5	99.0	99.2	99.2	99.3
3	**10.1**	**9.55**	**9.28**	**9.12**	**9.01**
	34.1	30.8	29.5	28.7	28.2
4	**7.71**	**6.94**	**6.59**	**6.39**	**6.26**
	21.2	18.0	16.7	16.0	15.5
5	**6.61**	**5.79**	**5.41**	**5.19**	**5.05**
	16.3	13.3	12.1	11.4	11.0
6	**5.99**	**5.14**	**4.76**	**4.53**	**4.39**
	13.8	10.9	9.78	9.15	8.75
7	**5.59**	**4.74**	**4.35**	**4.12**	**3.97**
	12.2	9.55	8.45	7.85	7.46
8	**5.32**	**4.46**	**4.07**	**3.84**	**3.69**
	11.3	8.65	7.59	7.01	6.63
9	**5.12**	**4.26**	**3.86**	**3.63**	**3.48**
	10.6	8.02	6.99	6.42	6.06
10	**4.96**	**4.10**	**3.71**	**3.48**	**3.33**
	10.0	7.56	6.55	5.99	5.64
11	**4.84**	**3.98**	**3.59**	**3.36**	**3.20**
	9.65	7.21	6.22	5.67	5.32
12	**4.75**	**3.89**	**3.49**	**3.26**	**3.11**
	9.33	6.93	5.95	5.41	5.06
13	**4.67**	**3.81**	**3.41**	**3.18**	**3.03**
	9.07	6.70	5.74	5.21	4.86
14	**4.60**	**3.74**	**3.34**	**3.11**	**2.96**
	8.86	6.51	5.56	5.04	4.69
15	**4.54**	**3.68**	**3.29**	**3.06**	**2.90**
	8.68	6.36	5.42	4.89	4.56
16	**4.49**	**3.63**	**3.24**	**3.01**	**2.85**
	8.53	6.23	5.29	4.77	4.44
17	**4.45**	**3.59**	**3.20**	**2.96**	**2.81**
	8.40	6.11	5.18	4.67	4.34
18	**4.41**	**3.55**	**3.16**	**2.93**	**2.77**
	8.29	6.01	5.09	4.58	4.25
19	**4.38**	**3.52**	**3.13**	**2.90**	**2.74**
	8.18	5.93	5.01	4.50	4.17
20	**4.35**	**3.49**	**3.10**	**2.87**	**2.71**
	8.10	5.85	4.94	4.43	4.10
22	**4.30**	**3.44**	**3.05**	**2.82**	**2.66**
	7.95	5.72	4.82	4.31	3.99
24	**4.26**	**3.40**	**3.01**	**2.78**	**2.62**
	7.82	5.61	4.72	4.22	3.90
26	**4.23**	**3.37**	**2.98**	**2.74**	**2.59**
	7.72	5.53	4.64	4.14	3.82
28	**4.20**	**3.34**	**2.95**	**2.71**	**2.56**
	7.64	5.45	4.57	4.07	3.75
30	**4.17**	**3.32**	**2.92**	**2.69**	**2.53**
	7.56	5.39	4.51	4.02	3.70

Question 9

In the four boxes below, draw a scatterplot that is a good representation of the correlation coefficient given in the box, then by eye add a regression line that you estimate to be an accurate for the data that you have drawn. Use TEN observations for each scatterplot.

Do not perform any calculations, this question is assessing your estimation ability.

$r = .75$	$r = -.3$
$r = -.85$	$r = .1$

[8 Marks]

[Question 10 commences on the next page]

Question 10

Describe what are residuals in a simple linear regression. Why are residual values relevant?

[2 Marks]

Question 11

When comparing two groups (e.g., comparing women and men, natives and immigrants), why can it be more beneficial to report a similarity metric alongside differences, rather than only differences?

[2 Marks]

Question 12

A researcher recruited 5 women and 5 men to test whether men have on average higher levels of self-esteem. The mean level of self-esteem for women was 3.5, *SD* = 1.0, the mean level for men was 3.6, *SD* = 1.1. However, a single-factor between-subjects ANOVA was non-significant, *p* > .05. What is the most plausible explanation for the non-significant finding?

_____

_____

_____

_____

_____

[2 Marks]

[Question 13 commences on the next page]

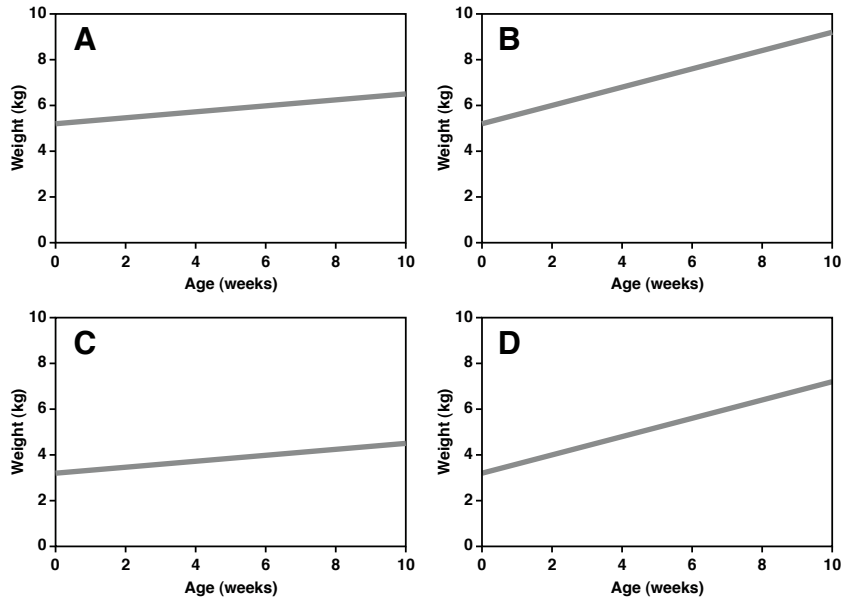
Question 13

A paediatric nurse is analysing baby weight (in kg) predicted as a function of age (in weeks). The nurse conducted a regression analysis, which revealed the following:

	Unstandardised Coefficient	*SE*	*t*	*p*
constant	3.20	0.891	3.59	< .01
age (weeks)	0.12	0.023	5.22	< .01

13a) Based on this information, which of the below panels shows the regression line? *Circle the appropriate answer*.

A B C D



Now assume that the gender of the babies also has an impact on their weight, with boys being heavier. The regression analysis conducted by the nurse revealed the following:

	Unstandardised Coefficient	*SE*	*t*	*p*
constant	3.10	0.632	4.91	< .01
gender – boy	0.25	0.047	5.32	< .01
age (weeks)	0.11	0.021	5.24	< .01

13b) What do the unstandardized coefficients represent in this context?

_____

_____

[3 Marks]

Question 14

A group of researchers collected data from 1000 employees from a large company. The researchers measured (1) *quality of relationship with the line manager*; (2) *job satisfaction*; and (3) *productivity*. Quality of relationship with the line manager is the predictor and productivity is the outcome variable.

14a) Formulate a hypothesis in plain English which postulates that job satisfaction *mediates* the association between relationship with the line manager and productivity. Also state which effect would need to be significant for the mediation hypothesis to be significant.

14b) Formulate a hypothesis which postulates that job satisfaction *moderates* the association between relationship with the line manager and productivity. Also state how you would test it.

[10 Marks]

END OF TEST