

UNIVERSITY OF ESSEX

PS115 Statistics Assignment 2 (Computer-Based Assignment) 2022

PS115 STATISTICS FOR PSYCHOLOGISTS

This is the PS115 Computer-Based Assignment. The assignment is designed to test students' skills and knowledge about performing statistics, using SPSS and reporting their results.

You should complete this assignment offline and submit your answers via the link that you have been given. For the longer responses (e.g., reporting results), you may wish to type your responses into a Word document, so that you can copy and paste them into the survey.

You must submit your answers before the deadline: Friday, 29th April at 5pm UK time. It will take some time to enter all your responses into the survey, so be sure to leave yourself enough time.

You have been provided with two SPSS files that you need to complete this assignment: *Box.sav* and *Styles.sav*. These files were emailed to you and can also be downloaded from the survey link.

IMPORTANT: These files are associated with YOUR student registration number, **so you must only use these files and not those that you have obtained from other students or elsewhere.** Different students will have different questions and data, and if you give answers for the incorrect data you may receive a mark of zero.

The answers that you enter should be your own work. Work which is plagiarised or identical to other students may receive a mark of zero. **Collusion with other students, including chatting about your questions and answers during the week, is not permitted and may be an academic offence. Sharing this paper, or the associated data files, is strictly prohibited.**

This assignment has two sections.

In **Section A**, there are two questions. You will be asked to write a short results section based on some SPSS analysis, as well as answering questions about the tests involved. **Full details of this study are on page 2.**

In **Section B**, there are some additional questions based on a second set of data. **Full details of this study are on pages 6-7.**

There are 75 marks available. Section A = 34 marks, Section B = 41 marks. The number of marks available for each question is shown in brackets.

For determining *statistical significance* you should use a significance level of .05. Unless otherwise stated, you should report numerical values *to three significant figures*.

SECTION A (INFORMATION)

The question in Section A is about the Closed Box Study (Data file: Box.sav)

Information on the Closed Box Study

A psychologist has conducted a study into children's problem solving. The task in this study required children to release a catch on a box to allow a doll to 'escape' from the closed box. Children could not touch the box, but had to use an object to release the catch. Four objects were on the same table as the box (a key, a tennis ball, a long pencil, and a cup), but only one of these (the long pencil) could actually be used to successfully release the catch. This long pencil is referred to as the *target object* (as it is the object that children needed to use to solve the task). Ninety-six children participated, all of who were in the third year of primary school (aged 7 to 8 years). Children were assigned alternately to one of two experimental conditions, and were tested individually. In one condition, the task was explained to the participant and then the experimenter drew the child's attention to the four objects and asked him/her to say what each of the four objects is usually used for. This condition is called the *object question condition*. In the other condition, the task was explained to the participant and then the experimenter drew the child's attention to the four objects as in the other condition. However, the child was not asked to say what each of the four objects is usually used for. This condition is called the *no question condition*. After the instructions, the child was told to start (and try to open the box). The experimenter recorded a number of measures of each child's task performance. First, what the first object picked up by the child was. Second, the number of seconds from the start of the task before the child picked up the target object. Third, the number of seconds from the start of the task before the child opened the box. All children did successfully open the box. In addition, the age in *months* and the sex of each child was recorded.

The data were entered in an SPSS data file. The variables are as follows:

Variable Name, Description [with coding]

cond	Experimental condition [1 = object question condition, 2 = no question condition]
sex	Sex of child [1 = male, 2 = female]
age	Age in months
pick.1st	First object picked up [1 = key, 2 = tennis ball, 3 = long pencil, 4 = cup]
pick.tar	Was first object picked the target object? [1 = no, 2 = yes]
time.tar	Time (seconds) taken to pick up the target object
time.box	Time (seconds) taken to open the box

The data are available for all 96 participants. There are no missing values.

The experimenter has made a number of predictions (**P**) concerning the two experimental conditions:

- P1** The proportion of boys and girls will *not* differ significantly between the two experimental conditions.
- P2** The mean age will *not* differ significantly between the two experimental conditions.
- P3** Children in the no question condition will be more likely to pick the target object as their first object than children in the object question condition.
- P4** The average time taken to pick up the target object will be lower for children in the no question condition than for children in the object question condition.
- P5** The average time taken to open the box will be lower for children in the no question condition than for children in the object question condition.
- P6** Older children will take less time to open the box than younger children.

Note: You may assume that, unless a question explicitly states otherwise, parametric statistical tests can be used for analyses that involve measurement data (i.e., for analyses that involve at least one NON-categorical variable).

SECTION A (RESULTS SECTION)

The questions in Section A are about the Closed Box Study. Use the following *SPSS* data file for Section A: **Box.sav**

Question A1**[29 marks]**

For this question, you need *only* consider the following two predictions concerning the Closed Box Study:

- P3** Children in the no question condition will be more likely to pick the target object as their first object than children in the object question condition.
- P4** The average time taken to pick up the target object will be lower for children in the no question condition than for children in the object question condition.

Use the data in the file that you have been provided with to write a summary appropriate for the results section in a formal report in psychology. Analyse and report only the information that is appropriate for each prediction. You should assume that the only variables of interest for your results section are the ones needed to investigate these two predictions.

You can answer this question with text only (i.e., you cannot include tables or graphs). **You should write in full sentences. Marks will be awarded for clarity and correct interpretation of the statistical data.** In order to save space and time, you may use the abbreviations P3 and P4 to refer to the predictions. Any other abbreviations that you use should be defined first. You may assume that parametric statistical tests can be used for analyses that involve measurement data (i.e., for analyses that involve at least one NON-categorical variable). Depending on your web browser, you may find that you have problems entering some symbols (e.g., “<”) – this is a problem with how the text is displayed only, and will not affect your mark. If you wish to use Greek letters you can use the Roman equivalents or write them in words (“alpha”, “beta”, “chi” etc.).

Report your results for P3 here:

Report your results for P4 here:

Question A2

[5 marks]

For this question, you need to think about the procedures appropriate for testing the following predictions/assumptions. You do *not* need to perform these tests, though it may help to look at the relevant variables when considering which test is most appropriate.

(a) Which statistical test would be most appropriate for testing prediction P1?

(CIRCLE ONE ANSWER)

Independent
samples t -test

Paired samples t -
test

Pearson's
Correlation

Chi-squared
goodness of fit test

Chi-squared test for
contingency tables

(b) Which statistical test would be most appropriate for testing prediction P2?

(CIRCLE ONE ANSWER)

Independent
samples t -test

Paired samples t -
test

Pearson's
Correlation

Chi-squared
goodness of fit test

Chi-squared test for
contingency tables

(c) Which statistical test would be most appropriate for testing prediction P6?

(CIRCLE ONE ANSWER)

Independent
samples t -test

Paired samples t -
test

Pearson's
Correlation

Chi-squared
goodness of fit test

Chi-squared test for
contingency tables

(d) After examining prediction **P4** in the original group of children, the researcher decides to repeat this analysis with a NEW sample of 200 7-year olds (100 in each group).

(i) Without knowing the outcome, which of these analyses is likely to have more statistical power? (CIRCLE ONE ANSWER)

The analysis from the original study

The analysis with the NEW sample

(ii) Explain briefly *why* the analysis that you chose is likely to have greater statistical power:

END OF SECTION A

SECTION B (INFORMATION)

The questions in Section B are about the Cognitive Styles Study (Data file: Styles.sav)

Information on the Cognitive Styles Study

A questionnaire study was conducted to investigate two measures of ‘cognitive style’ and a number of behavioural measures. One measure of cognitive style obtained was the *need for cognition* (NFC). People with high NFC have a desire to engage in challenging intellectual activity. A second measure of cognitive style obtained was the *consideration of future consequences* (CFC). People with high CFC scores tend to consider the distant rather than the immediate consequences of their actions. Both these measures of cognitive style were assessed using 5-item questionnaire scales as follows.

For need for cognition, study participants provided ratings for the following questions on a 6-point scale [1 = strongly disagree, 2 = moderately disagree, 3 = slightly disagree, 4 = slightly agree, 5 = moderately agree, 6 = strongly agree]:

- A1 I would prefer complex to simple problems.
- A2 I find satisfaction in deliberating hard and for long hours.
- A3 I really enjoy a task that involves coming up with new solutions to problems.
- A4 I prefer my life to be filled with puzzles that I must solve.
- A5 The notion of thinking abstractly appeals to me.

For consideration of future consequences, study participants provided ratings for the following questions on a 5-point scale [1 = extremely uncharacteristic (of me), 2 = somewhat uncharacteristic, 3 = uncertain, 4 = somewhat characteristic, 5 = extremely characteristic (of me)]:

- B1 I consider how things might be in the future, and try to influence those things with my day-to-day behaviour.
- B2 Often I engage in a particular behaviour in order to achieve outcomes that may not result for many years.
- B3 I am willing to sacrifice my immediate happiness or well-being in order to achieve future outcomes.
- B4 I think it is important to take warnings about negative outcomes seriously even if the negative outcomes will not occur for many years.
- B5 I think it is more important to perform a behaviour with important distant consequences than a behaviour with less important immediate consequences.

For each participant, an overall NFC score was obtained by finding the mean of his/her five responses to questions A1-A5 above. Similarly, for each participant, an overall CFC score was obtained by finding the mean of his/her five responses to questions B1-B5 above.

Participants also answered the following questions about their past and present behaviour:

- | | | |
|----|---|-------------|
| C1 | Do you contribute to a pension fund for your retirement? | No/Yes |
| C2 | Do you have life insurance? | No/Yes |
| C3 | Do you insure the contents of your home against theft? | No/Yes |
| C4 | Do you invest in shares? | No/Yes |
| C5 | How many years have you spent in post-compulsory full-time education? | _____ years |
| C6 | How many cigarettes do you smoke per week? | _____ |

Participants were also asked to indicate their sex and their age in years.

The data were entered in an *SPSS* data file. The details of the variables in this file are as follows:

Variable Name, Description [with coding]

a1_nfc1	Response to A1 above (first NFC item)
a2_nfc2	Response to A2 above (second NFC item)
a3_nfc3	Response to A3 above (third NFC item)
a4_nfc4	Response to A4 above (fourth NFC item)
a5_nfc5	Response to A5 above (fifth NFC item)
mean_nfc	Overall NFC score (mean of responses to items A1 to A5)
b1_cfc1	Response to B1 above (first CFC item)
b2_cfc2	Response to B2 above (second CFC item)
b3_cfc3	Response to B3 above (third CFC item)
b4_cfc4	Response to B4 above (fourth CFC item)
b5_cfc5	Response to B5 above (fifth CFC item)
mean_cfc	Overall CFC score (mean of responses to items B1 to B5)
c1_pen	Response to: "Do you contribute to a pension fund for your retirement?" [1 = No, 2 = Yes]
c2_life	Response to: "Do you have life insurance?" [1 = No, 2 = Yes]
c3_home	Response to: "Do you insure the contents of your home against theft?" [1 = No, 2 = Yes]
c4_share	Response to: "Do you invest in shares?" [1 = No, 2 = Yes]
c5_educ	Number of years spent in post-compulsory full-time education
c6_cig	Number of cigarettes smoked per week
sex	Sex of participant [1 = male, 2 = female]
age	Age of participant in years

The above data are available for the 120 people who participated in the study. There are no missing values.

The researcher who conducted the study has made a number of methodological assumptions (**M**) and has specified some study hypotheses (**H**), as follows:

- M1** Together, the five NFC items (A1-A5) form an internally consistent scale.
- M2** Together, the five CFC items (B1-B5) form an internally consistent scale.
- H3** Participants who contribute to a pension fund will have higher average CFC scores than those that do not contribute.
- H4** Participants who have life insurance will have higher average CFC scores than those that do not have life insurance.
- H5** Participants who insure the contents of their home will have higher average CFC scores than those that do not insure the contents of their home.
- H6a** Participants who invest in shares will have higher average NFC scores than those that do not invest in shares.
- H6b** Participants who invest in shares will have higher average CFC scores than those that do not invest in shares.
- H7a** There will be a positive linear relationship between NFC score and the number of years spent in post-compulsory full-time education.
- H7b** There will be a positive linear relationship between CFC score and the number of years spent in post-compulsory full-time education.
- H8** There will be a negative linear relationship between CFC score and the number of cigarettes smoked per week.

Note: You may assume that, unless a question explicitly states otherwise, parametric statistical tests can be used for analyses that involve measurement data (i.e., for analyses that involve at least one NON-categorical variable).

SECTION B (SHORT-ANSWER QUESTIONS)

The questions in Section B are about the Cognitive Styles Study

Use the following *SPSS* data file for Section B: **Styles.sav**.

Question B1

For this question you will need to obtain information that could be reported in the *Participants* section of a formal report of this study.

GIVE ANSWERS TO THIS QUESTION TO **ONE** DECIMAL PLACE.

(a) For the age of the participants (in years), state:

(i) The mean: _____ **years**

(ii) The standard deviation: _____ **years**

(iii) The median: _____ **years**

(b) What percentage of the participants spent **three or more years** in post-compulsory education?

_____ **%**

[4 marks]

Question B2

This question is about study hypothesis **H4**:

H4 *Participants who have life insurance will have higher average CFC scores than those that do not have life insurance.*

- (a) Find the mean, median and standard deviation overall CFC score for people who have life insurance and for people who do not have life insurance. Complete the values below.

ANSWERS SHOULD BE ROUNDED TO **TWO** DECIMAL PLACES.

Overall CFC score	Mean	Median	Standard Deviation
Participants without life insurance			
Participants with life insurance			

- (b) A *t*-test can be used to examine whether there is a statistically significant difference between the pair of mean CFC scores that you reported in part (a). Undertake the appropriate *t*-test, and report the outcome of the test.

- (i) Which kind of *t*-test did you use? (CIRCLE ONE ANSWER)

One-sample *t*-test

Related-samples *t*-test

Independent samples *t*-test

- (ii) Report the test statistic (TO **THREE** DECIMAL PLACES):

***t* =** _____

- (iii) Report the degrees of freedom

***df* =** _____

- (iv) Report the exact *p*-value (TO **THREE** DECIMAL PLACES):

***p* =** _____

- (v) Are the results statistically significant? (CIRCLE ONE ANSWER):

Yes / No

- (vi) In a sentence, state what you conclude from the test:

[13 marks]

Question B3

This question is about study hypothesis **H6b**:

H6b *Participants who invest in shares will have higher average CFC scores than those that do not invest in shares.*

This hypothesis could be examined using a *t*-test, to determine whether the mean overall CFC score differed significantly between participants who do invest in shares and participants who do not invest shares. However, in this question, you are to examine this hypothesis using a **non-parametric test** (instead of using a *t*-test).

Carry out the appropriate **non-parametric test** to examine the difference that is described in hypothesis H6b, and give the results of this analysis below.

(i) Name the test you used: _____

(ii) Report the test statistic:
(GIVING THE VALUE OF THE STATISTIC TO **ONE** DECIMAL PLACE) _____

(iii) Report the total number of participants analysed in this test: **N** = _____

(iv) Report the exact *p*-value (TO **THREE** DECIMAL PLACES): **p** = _____

(v) Are the results statistically significant? (CIRCLE ONE ANSWER) **Yes / No**

(vi) In a sentence, state what you conclude from the test:

[8 marks]

Question B4

This question is about study hypothesis **H7a**:

H7a *There will be a positive linear relationship between NFC score and the number of years spent in post-compulsory full-time education.*

You will examine this hypothesis using simple linear regression.

Undertake the appropriate analysis that will allow you to obtain the equation for the regression line that can be used to predict overall NFC score from the number of years spent in post-compulsory full-time education.

- (a)(i) Write down the equation for the regression line based on the results of the linear regression analysis, using **N** for the overall NFC score and **P** for the number of years spent in post-compulsory full-time education.

VALUES IN THE EQUATION SHOULD BE GIVEN TO **THREE** DECIMAL PLACES.

Equation: _____

- (ii) Use this equation to answer the following question. Imagine that there are two people, and that one of these has spent one more year in post-compulsory full-time education than the other. What is the *expected* difference in overall NFC score between these two people?

THIS ANSWER SHOULD BE GIVEN TO **THREE** DECIMAL PLACES

Expected difference in NFC score: _____

- (iii) Use the equation for the regression line to calculate the expected overall NFC score for a participant who has spent five years in post-compulsory education.

YOUR FINAL ANSWER SHOULD BE WRITTEN CORRECT TO **ONE** DECIMAL PLACE

Answer (expected NFC score): _____

- (b) What is the percentage of variance in overall NFC score that is accounted for by the number of years spent in post-compulsory full-time education?

GIVE THE PERCENTAGE TO **ONE** DECIMAL PLACE.

_____ %

Question B4 (continued)

- (c)(i) For the linear relationship examined by this regression analysis, write down the unstandardized regression coefficient (b), the standardized regression coefficient (β) and the exact p -value for the regression.

GIVE THESE VALUES TO **THREE** DECIMAL PLACES

$b =$ _____

$\beta =$ _____

$p =$ _____

- (ii) Is the regression statistically significant? (CIRCLE ONE ANSWER)

Yes / No

- (iii) One of these three values listed in part (c)(i) above is also the Pearson correlation (r) for the relationship between the overall NFC score and the number of years spent in post-compulsory education. Which one of the three is this? (CIRCLE ONE ANSWER)

b

β

p

[12 marks]

Question B5

- (a) Obtain Cronbach's alpha for the five CFC items (B1-B5), and write the value below.
GIVE THE VALUE OF ALPHA TO **THREE** DECIMAL PLACES

$\alpha =$ _____

- (b) Which one of the ten methodological assumptions or study hypotheses does this value (α) provide information for? (CIRCLE ONE ANSWER)

M1 M2 H3 H4 H5 H6a H6b H7a H7b H8

- (c) With respect to this methodological assumption or study hypothesis the value that you have recorded above would usually be regarded as (CIRCLE ONE ANSWER):

Poor Almost adequate Good Very good Excellent

[4 marks]

END OF SECTION B

END OF ASSIGNMENT

You must submit your answers ONLINE before the assignment deadline. It will take some time to enter all your responses into the survey, so be sure to leave yourself enough time.

At the end of your submission, we will ask you for your email address so we can send you a confirmation that your assignment has been received. If you don't receive an email, it means that we have not received your response - please try again. Please save this email- it is your proof that you have submitted your response, in case something goes wrong.