

Name: \_\_\_\_\_

Student Number: \_\_\_\_\_

### QUESTION 1 (worth 5 marks)

A work colleague approaches you and asks for advice about a survey they want to run on fatigue amongst mothers with infant children (less than 1 year old). They want to understand why some mothers are fatigued and some are not. But they are not sure how many people they need in their survey.

You tell them to go and find a study that has already addressed this question (or as close to as they can find) and bring it to you.

The following day they showed you a paper which measured fatigue of mothers with young children (aged 0 to 4 years). The authors reported the following multiple regression model (dependent variable is mother's fatigue):

Predictor variable	B	SE B	$\beta$	t	p
<i>Child sleep factors</i>					
How many nights a week the child wakes	-.20	.31	-.04	-0.64	.523
Extent to which the child's sleep is a problem	.46	.30	.13	1.58	.115
Duration of sleep problem	.27	.37	.05	0.73	.463
<i>Maternal factors</i>					
Expectations about sleep	-.22	.10	-.22	-2.21	.027
Sleep quality	.76	.47	.14	1.61	.107
Health and self-care behaviour	.31	.12	-.17	-2.58	.010
Social support – needs	.22	.06	-.22	3.40	.001
Social support – satisfaction	-.08	.05	-.11	-1.62	.106

The model accounted for 50% of the variance in maternal fatigue,  $R^2 = .50$ , Adj  $R^2 = .47$ ,  $F(8,163) = 8.25$ ,  $p < .001$ .

Your colleague says they would like to do the same type of analysis. Based upon this model, how many mothers does your colleague need for them to feel confident that they have enough for their analysis? Run a statistical power analysis with the above data and write up that output. Include all the necessary information for someone else to replicate it. What is your interpretation of that power analysis and what advice would you give to your colleague?

## QUESTION 2 (worth 15 marks)

Four months later you bump into your work colleague and they excitedly tell you that they have run their survey. They ask for your help in analysing their data and email you a .csv data file with a description of what it contains.

----first part of email starts here---  
Hi!  
Thanks again for your help with this.

The dependent variable is the **Fatigue assessment scale** (called *FAS* in the dataset I sent you). The total score ranges from 10 to 50, with a higher score indicating more fatigue.

I also collected data on the following measures:

- **Age of child** (called *age* in dataset). Measured in months from 0 to 12.
- **Gender of focus child** (called *gender* in dataset).
- **Age of mother** (called *mother.age* in dataset). Measured in years.
- **Total number of children aged below 18 years** (called *children.number in dataset*).
- **Times awake during the night** (called *night\_away\_week* in dataset). Mothers were asked to indicate the mean number of nights per week the child wakes (0 to 7 nights).
- **Perception of child's sleep as a problem** (called *sleep\_problem\_perception* in dataset). Extent to which the child's sleep is a problem, rated on a Likert scale from 1=not a problem to 21=severe problem.
- **Duration of sleep problem** (called *sleep\_duration\_problem* in dataset). Measured in months from 0 to 12.
- **Sleep quality** (called *PSQI\_total* in dataset). Subjective sleep quality in the past month was taken from the Pittsburgh Sleep Quality Index (PSQI). Score ranges from 0 to 21. Higher scores indicate worse sleep quality.
- **The Multidimensional Scale of Perceived Social Support** (called *MSPSS\_n* in dataset). This is a 12-item scale measuring perceptions of adequacy of an individual's current levels of social support, rated on a Likert scale from 1=very strongly disagree to 7=very strongly agree. You can calculate three sub scores and a total score:
  - Significant Other Subscale: Sum across items 1, 2, 5, & 10, then divide by 4.
  - Family Subscale: Sum across items 3, 4, 8, & 11, then divide by 4.
  - Friends Subscale: Sum across items 6, 7, 9, & 12, then divide by 4.
  - Total Scale: Sum across all 12 items, then divide by 12.
- **Health and self-care behaviour** (3 items; *diet\_quality*, *physical\_activity* and *self\_care*). Sum of three items asking parents to rate the quality of their diet, overall level of physical activity, and the extent to which they engage in self-care behaviours (i.e., relaxation, taking time out) on a 5-point Likert scale ranging from 0-poor to 4-excellent.

I haven't had the chance to tidy this dataset up – could you do it for me? Oh and could you create the summary scores for last two measures (details above) – the subscores for the MSPSS might be interesting – thanks! You're a star 😊  
----first part of email ends here---

At this point you need to perform three tasks:

- (1) Perform a basic data clean (see wk 6 powerpoint, slide 4).
- (2) Check for univariate and multivariate outliers (wk 6 powerpoint) and deal with accordingly
- (3) Create the summary scores for the MSPSS and health and self-care variables and check/report reliability accordingly (Cronbach's alpha).

Make a note of each change you do, why you did it, and make sure all edits are present in the jamovi file you submit. **If you do decide to remove any participants – DO NOT DELETE – use the filter function.** If you delete cases from your jamovi file you will lose marks (see rubric).

### QUESTION 3 (worth 60 marks)

-----second part of email starts here-----

Once you have cleaned up my dataset could you have a look at the relationship between all these variables I collected and the fatigue (FAS) score at an univariate level? Could you make me a table of descriptives that show this relationship please? Also, I need a visualization of each of these relationships. Not sure what would be the best way of doing that – what do you think? If you can do it – can you tell me what each graph shows exactly – just so I'm clear.

After that I need to know what's related to mother's fatigue? What's the best combination of variables to include? Are these predictors reasonable to use here? Is there anything we shouldn't include?

-----second part of email ends here-----

Write up the analysis and your interpretation that you see as appropriate to answer this question (explain why you have chosen this approach to your friend). Please use APA format when presenting data. Marks will be deducted for typos, poor expression and confusing presentation.

Please indicate appropriate use of data analysis technique you use (e.g, assumptions). All analysis run in jamovi should be present in the jamovi file you submit. Any missing analysis in your jamovi file, even if it is in the word document, means you will lose marks (see rubric).

### QUESTION 4 (worth 20 marks)

-----third part of email starts here-----

Oh, sorry, one last favour – I need one extra graph. It would show the number of people who score above the cutoff for the FAS for being really fatigued (details below) – I haven't had a chance to make that variable – could you do it?

- FAS scores 10 – 21: no fatigue (normal)
- FAS scores 22 – 50: substantial fatigue

If you do that – could you also graphically show the relationship between this FAS cutoff variable with times awake during the night (aka night\_away\_week) in some way? Again – I have no idea how to do this. What does it show?

Oh...(tiny extra favour) could you also give me a table showing the percentage breakdown for the FAS cutoff by times awake during the night as well? Are they related? You are a star 😊

Thank you! I owe you a post-lockdown coffee.

-----third part of email ends here-----

All plots and statistics should be performed in jamovi should be present in the jamovi file you submit. Any missing output in your jamovi file, even if it is in the word document, means you will lose marks (see rubric). Plots can be copied from jamovi into your word doc report. You will not lose marks due to axis label formatting (these cannot be edited directly in jamovi).