

How Your Work Was Done: The Decisions and How to Repeat Them

How to read this document

This document tells the story of your project in the order it happened: what we received from you, what we decided during the consultation and why, which analysis answers each of your hypotheses, how your data were prepared, what changed along the way, and what we did not do.

Its last part shows, step by step, how to repeat every analysis yourself in SPSS. You do not need any background in statistics to follow it: each technical word is explained the first time it appears.

The numbers themselves (the results) are in your findings chapter; this document never repeats them. Where it tells you which number to look at in SPSS, it also tells you where the same number appears in your findings chapter.

1. What we received from you

Each file you uploaded, and what we understood it to be.

File	What it is	Version notes	How sure we are
data.csv	Data	—	Certain
study-brief.md	Instructions	—	Likely
questionnaire.docx	Questionnaire	—	Certain
codebook.md	Instructions	—	Certain

We opened your data file and checked that it can be read in full.

Item	Detail
The file could be read	Yes
Rows (usually one per respondent)	180
Columns (usually one per question or detail)	11

2. What we decided during the consultation, and why

Before any analysis ran, we read your materials and your data and built a plan. A statistical consultant decides by themselves whatever the materials and the data already make clear, and explains each decision; where there was real doubt we asked you. These are the decisions, as they appeared in the consultation report you approved.

Decisions we made for you

How we understood your study

For each thing your study measures: the part it plays in your study and what kind of answers it holds. Together these decide which test fits each hypothesis.

- **Your study design** — a single group of respondents, measured once.
- **d1** — a factor that may explain or predict the outcome; its answers are ordered categories (for example, low / medium / high).

- **q1_1** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q1_2** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q1_3** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q1_4** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q1_5** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q2_1** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q2_2** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q2_3** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q2_4** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **q2_5** — one question among a group of questions analysed together; its answers are ordered categories (for example, low / medium / high).
- **academic_self_efficacy** — a factor that may explain or predict the outcome; its answers are numbers on a scale (for example, the average of 1-to-5 answers).
- **learning_persistence** — the outcome: what your study tries to explain; its answers are numbers on a scale (for example, the average of 1-to-5 answers).

How the questions were handled, round by round

We did not need to ask you any question.

3. The analysis plan: which test answers each hypothesis

For each of your hypotheses we chose one statistical test (a standard calculation that checks whether a pattern in your data is larger than chance alone would produce). Below: what each test checks, which of your measures it uses, why it was chosen, and the other tests we considered.

Analysis 1: hypothesis H1

- The test: Pearson correlation.
- What it checks: Measures how strongly two numeric measures move together, and in which direction.
- Your measures in it: the first measure: «Academic self efficacy» (`academic_self_efficacy`); the second measure: «Learning persistence» (`learning_persistence`).
- Why this test: it is the standard test for this kind of question, given how your measures were understood in the consultation (part 2).

Other tests we considered:

- Spearman correlation was kept as a backup. It is used only if your data do not meet the conditions Pearson correlation needs, which we check before the test is run.

4. How your data were prepared

Before the tests ran, your data file was prepared. Each step below changed the data in a way the analysis needed.

- Your columns were given short, standard names (a letter and the question's number). The list of old and new names is in your codebook.

- «academic_self_efficacy»: one score per respondent, the average of its questions (q1_1, q1_2, q1_3, q1_4, q1_5).
- «learning_persistence»: one score per respondent, the average of its questions (q2_1, q2_2, q2_3, q2_4, q2_5).

In SPSS, the first part of syntax.sps (under DATA PREPARATION) repeats every one of these steps on your original file, and data_analysis.sav is the file it produces.

5. What changed along the way, and why

Nothing in the plan changed after you approved it.

6. Repeat the analysis yourself in SPSS

You did not choose a program when you ordered, so the steps below are for SPSS, the program most universities use.

What you need

IBM SPSS Statistics (version 25 or later), and these files from the «Software Outputs (Appendix)» part of your package: data_analysis.sav (your data, already prepared), data_original.sav (your data as you sent it), syntax.sps (the commands) and output.spv (our output, to compare with yours).

The quickest way: everything at once

Open syntax.sps in SPSS (File > Open > Syntax...), choose Run > All, and SPSS repeats the whole analysis, from your original file to the last table. Keep the files of the package in one folder. If SPSS says it cannot find a file, add the line CD 'C:\my-folder'. at the very top of syntax.sps, with your folder's location instead of the example, and run it again.

Step by step, one analysis at a time

Open data_analysis.sav: File > Open > Data....

Analysis 1: hypothesis H1 — Pearson correlation

1. From the menus at the top, choose Analyze > Correlate > Bivariate....
2. Move «Academic self efficacy» (academic_self_efficacy) into the box Variables.
3. Move «Learning persistence» (learning_persistence) into the box Variables.
4. Click OK. The results open in the Output window.

The same analysis as a command: this is its block in syntax.sps, which starts with the comment line shown first. To run only it: open syntax.sps, select these lines, and choose Run > Selection.

```
* ==== Analysis 1 of 1 (hypothesis H1) ====.  
CORRELATIONS /VARIABLES=academic_self_efficacy learning_persistence  
/PRINT=TWOTAIL SIG.  
GRAPH /SCATTERPLOT(BIVAR)=academic_self_efficacy WITH learning_persistence.
```

In the table Correlations, look at the column Sig. (2-tailed). This number is the p-value (a number every test produces; a small p-value means the result is unlikely to be a coincidence). It is the same number your findings chapter reports for hypothesis H1. Small differences in the last digit come from rounding.

7. What we did not do

Everything in your approved plan was done.

8. Good to know

Institutional AI-use policies were not verified. Statikali does not check whether your academic institution, journal, or supervisor permits the use of AI-assisted analysis tools. Verifying those policies and complying with them remains your responsibility.

An AI language model assisted only by writing the explanatory prose in this package. Every statistical computation, every number, and every test-selection decision was produced by a deterministic analysis engine — never by the AI.