



Espiral. Cuadernos del Profesorado
Vol. 16, Num. 32
Marzo 2023
ISSN 1988-7701

Cómo referenciar este artículo / How to reference this article:

Ibáñez-López, F. J., Ponce Gea, A. I., Pedreño Plana, M., & Sánchez-Martín, M. (2023). Basic survival manual for descriptive statistical analysis. *Espiral. Cuadernos del Profesorado*, 16(32), 118-126. <https://doi.org/10.25115/ecp.v16i32.9134>

Basic survival manual for descriptive statistical analysis

Manual de supervivencia básico para los análisis estadísticos descriptivos

Francisco Javier Ibáñez-López, Ana Isabel Ponce Gea, Marina Pedreño Plana,
and Micaela Sánchez-Martín

University of Murcia, Murcia, Spain

"Do it or don't do it, but don't try" (Yoda, Star Wars)

Abstract

Descriptive statistical analysis is the first step that every researcher must take when preparing to analyse the data collected during the fieldwork of their research. It is essential to know, in a descriptive way, the characteristics of the sample participating in the study and the initial results obtained, as a preliminary step to carrying out inferential analyses or adjusting statistical models that explain the data obtained in greater depth. This work presents an introduction to the different variables that can be found and the implementation of their descriptive analysis using the free statistical software *jamovi*. It shows, therefore, a tool at no cost, which offers powerful results without the need to know a statistical programming language.

Keywords: Statistical analysis; descriptive; variables; *jamovi*; Mathematics education

Resumen

El análisis estadístico descriptivo es el primer paso que cualquier persona que pretenda realizar una investigación cuantitativa debe dar cuando se dispone a realizar el análisis de los datos recogidos durante el trabajo de campo de su investigación. Es fundamental conocer de una forma descriptiva las características de la muestra participante en el estudio y los resultados iniciales obtenidos, como paso previo a la realización de análisis inferenciales o al ajuste de modelos estadísticos que expliquen de forma más profunda los datos obtenidos. Este trabajo presenta una introducción a las diferentes variables que nos podemos encontrar y a la implementación de su análisis descriptivo mediante el software estadístico libre *jamovi*. Se muestra, por tanto, una herramienta sin coste alguno que ofrece resultados potentes sin necesidad de conocer un lenguaje de programación estadístico.

Palabras clave: Análisis estadístico; descriptivos; variables; *jamovi*; enseñanza de las matemáticas

Received: February 04, 2023

Accepted: March 18, 2023

Correspondence: Ana Isabel Ponce Gea, University of Murcia, Murcia, Spain
Email: anaisabel.ponce@um.es

Key issues

What is known

- This is the first step that must be taken in order to understand the quantitative data collected during the field work.

What this work contributes

- Basic ideas for the easy descriptive analysis of the variables in our research using the free software *jamovi*.

Practical situation

A student of a Degree on Education is carrying out his first research work: his Final Degree Project. At the time, he decided to conduct a research project about the use of digital resources to promote learning in the attention to diversity and, after a review of the theoretical background, he decided to use a questionnaire published and validated by a reference author in the field (Porto Castro, 2022). He has already collected all the data in a document generated by the survey application that he has used to carry out his work. He has never analysed the data and he has not enough knowledge to do that. Then, how to carry out a descriptive analysis of the data collected in order to be able to obtain the first results and conclusions?

Descriptive analysis in context

We have to consider that the transversal and interdisciplinary nature of statistics implies that it is found in the studies of all fields of research. To define statistics in a simple way, we will say that it is the area of mathematics that deals with sets of data to represent the characteristics of a sample, in such a way that they can be analysed and allow inferences. These inferences, which are based on the calculation of probabilities (with what probability a characteristic or a situation is more likely to occur than another), can finally draw conclusions about the population (Suárez et al., 2017). That is why all students should have at least the basic knowledge to be able to carry out the analysis and interpretation of the data collected for their research, both phases necessary for the discussion of the results and the drafting of the research report.

Initially, a distinction must be made between descriptive statistics and inferential statistics. The former refers to the description, organisation, synthesis and graphical representation of the data that are the object of study, while the latter tries to generate models to reach certain conclusions based on premises (inferences) and predictions derived from the data under study (Field et al., 2012).

The challenge for teachers of research methodology is to provide prospective education graduates with the necessary knowledge to make appropriate use of statistical techniques. And the use of paid software does not necessarily guarantee the achievement of this goal (Şahin & Aybek, 2020). Following, we introduce the different types of variables and their descriptive statistical analysis through the use of the free software *jamovi* (The *jamovi* project, 2022; R Core Team, 2019), a free access and cross-platform software supported with R programming language (Santabárbara and Lasheras, 2020). Among the different free software available, *jamovi* provides a very user-friendly environment with simple and intuitive menus, without the need of programming knowledge. In addition, *jamovi* can read and load data created with Excel, JASP, R, SAS, SPSS or STATA, thus showing its versatility.

We will start by defining what a variable is and their types, and then deal with their analysis in *jamovi*.

Variables

A variable is a set of values obtained by measuring a particular characteristic of each of the individuals who belong to a sample or population.

In most statistical programmes we represent, in the data matrix, the variables in the columns and the cases (individuals measured) in the rows (Figure 1). Thus, columns A, B, C, etc. will be the measured characteristics (gender, age, group, opinion, ...) and rows 1, 2, 3, ..., the individuals.

Figure 1

Data matrix in jamovi

Qualitative or categorical variables

These are variables whose measured values are not numbers but names or words. These values are called categories. In addition, a process of categorisation is followed, consisting of assigning numbers to each of these values. For instance, the variable *gender* has three categories: *male*, *female* and *other*. If we categorise this variable, we assign 1 to the *male gender*, 2 to the *female gender* and 3 to *other genders*. This establishes relationships of the type:

$$\text{Male (1)} \neq \text{Female (2)} \neq \text{Other (3)}$$

When the variable has only two categories, it is said to be a dichotomous variable. If it has more than two categories, such as eye colour, it is said to be a polytomous variable.

Ordinal variables

In ordinal variables, symbols or numbers are assigned to designate the order of each of the values of the variable. Therefore, relationships of the type more than and less than are established. For instance, to provide an opinion on the statement "I consider that the quality of studies offered in higher education is very good", we will have the following options:

- Strongly disagree (SD; 1)
- Disagree (D; 2)
- Agree (A; 3)
- Strongly agree (SA; 4)

The following relationships can be established to establish an order:

$$\text{Strongly disagree (1)} < \text{Disagree (2)} < \text{Agree (3)} < \text{Strongly agree (4)}$$

$$SD (1) < D (2) < A (3) < SA (4)$$

Logically, strongly disagreeing is not the same as strongly agreeing with the proposed statement. These variables are also known as quasi-quantitative variables or ordinal categorical variables.

Quantitative variables

The values measured with these variables are numbers. The height or weight of an individual are examples of quantitative variables. If one individual is 1.80 metres tall and another is .90 metres tall, we can say that the former is twice as tall as the latter.

These variables, also known as scalar or numerical variables, can be classified into two types:

- Continuous quantitative variables: those that have infinite numbers or options between two assigned values. For example, between a height of 1.70 and 1.75 metres there are infinite values, as many as decimals you want to consider.
- Discrete quantitative variables: those that do not have any options between two consecutive values assigned to them. They are the result of counting, so there is no sense in considering intermediate values. An example is the number of children.

Variables in jamovi

In *jamovi*, by following the path *Data > Setup*, we find the different types of variables depending on the type of measurement (Measure type; Figure 2).

Figure 2

Types of variables in jamovi

Thus, *jamovi* differentiates between the following types of variables:

- Nominal variable: qualitative variables.
- Ordinal variable: ordinal variables.
- Continuous variable: continuous and discrete quantitative variables.
- ID: variable used exclusively to give an identifier to each of the individuals involved in the research, thus omitting confidential data such as their names or identity card number.

Furthermore, in the configuration shown in Figure 2, you can name and describe the variable, specify the type of data (whole numbers, decimal numbers or text) or define the value for missing data (missing data that could not be measured during data collection, e.g. data not provided in a questionnaire by the person filling in the questionnaire, either intentionally or by mistake).

The following is an example of the different types of variable classification (Table 1) and the *jamovi* display of these data, obtained through the filling in of the questionnaire by 30 students from the

Faculty of Education of a state university (Figure 3). By default, depending on the origin of the database, *jamovi* does not configure the type of measurement and data for each of the variables. This must be configured from the menu *Data > Setup*.

Table 1

Examples of classification of variables

Variable	Values that can be taken	Classification
Gender	Male, Female, Other	Polytomous nominal qualitative
Age	18, 19, 20, ...	Continuous quantitative
Number of siblings (including you)	1, 2, 3, ...	Discrete quantitative
Do you study and work at the same time?	Yes/No	Dichotomous nominal qualitative
I1. I have enough technological resources at my disposal to attend to students with difficulties in the areas of communication and language/specific educational needs.	Strongly Disagree / Disagree / Indifferent / Agree / Strongly Agree	Ordinal
I2. I use ICT in my intervention with learners with difficulties in the areas of communication and language/with specific educational needs.	Strongly Disagree / Disagree / Indifferent / Agree / Strongly Agree	Ordinal
I3. In the response to diversity, work with technological media facilitates the teaching and learning process of students.	Strongly Disagree / Disagree / Indifferent / Agree / Strongly Agree	Ordinal

Note. Adapted from Porto Castro (2022).

Figure 3

View of data in jamovi

	Género	Edad	Número d...	¿Estudias ...	I1. He teni...	I2. Dispon...	I3. Compa...
1	2	23	2	2	5	5	1
2	1	20	3	2	4	3	1
3	2	23	2	2	5	5	1
4	1	20	3	2	4	3	1
5	1	20	3	2	4	3	1
6	1	20	3	2	4	3	1
7	2	21	1	1	4	4	2
8	2	23	2	2	5	5	1
9	2	18	3	2	4	5	1
10	2	19	3	1	5	5	4
11	2	20	4	2	2	5	3
12	1	17	1	1	5	5	1
13	1	18	2	1	5	5	3
14	2	23	2	2	5	5	1
15	2	41	2	1	5	5	1
16	1	23	2	2	4	3	5
17	1	25	2	2	5	5	2
18	2	23	2	1	4	5	2
19	1	18	2	1	5	5	3
20	2	22	2	1	5	5	2
21	1	22	3	2	4	4	1
22	2	21	2	2	4	5	1
23	2	26	1	1	5	5	1
24	2	28	3	1	4	4	4
25	2	23	4	1	4	4	2
26	2	22	2	2	5	5	2
27	1	19	4	2	4	4	5
28	1	23	3	2	5	5	1
29	1	19	4	2	4	4	5
30	2	20	1	1	5	5	1

Next, we should categorise the variables in the configuration menu (Figure 2), which consists of transforming a continuous or discrete quantitative variable into a dichotomous (dichotomise) or polytomous (polytomise) qualitative variable. Again, in the menu *Data > Setup*, we will categorise the variables *Gender*, *Do you study and work at the same time?*, and the three perception items, I1, I2 and I3, assigning each of the numerical levels to its category. The result is shown in Figure 4.

Figure 4

View of categorised data in jamovi

	Género	Edad	Número d...	¿Estudias ...	I1. He tenido acces...	I2. Dispongo de di...	I3. Comparto esto...
1	Femenino	23	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
2	Masculino	20	3	No	Acuerdo	Indiferente	Totalmente desacuerdo
3	Femenino	23	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
4	Masculino	20	3	No	Acuerdo	Indiferente	Totalmente desacuerdo
5	Masculino	20	3	No	Acuerdo	Indiferente	Totalmente desacuerdo
6	Masculino	20	3	No	Acuerdo	Desacuerdo	Totalmente desacuerdo
7	Femenino	21	1	Sí	Acuerdo	Acuerdo	Desacuerdo
8	Femenino	23	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
9	Femenino	18	3	No	Acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
10	Femenino	19	3	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Acuerdo
11	Femenino	20	4	No	Desacuerdo	Totalmente de acuerdo	Indiferente
12	Masculino	17	1	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
13	Masculino	18	2	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Indiferente
14	Femenino	23	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
15	Femenino	41	2	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
16	Masculino	23	2	No	Acuerdo	Indiferente	Totalmente de acuerdo
17	Masculino	25	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Desacuerdo
18	Femenino	23	2	Sí	Acuerdo	Totalmente de acuerdo	Desacuerdo
19	Masculino	18	2	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Indiferente
20	Femenino	22	2	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Desacuerdo
21	Masculino	22	3	No	Indiferente	Indiferente	Indiferente
22	Femenino	21	2	No	Acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
23	Femenino	26	1	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
24	Femenino	28	3	Sí	Acuerdo	Acuerdo	Acuerdo
25	Femenino	23	4	Sí	Acuerdo	Acuerdo	Desacuerdo
26	Femenino	22	2	No	Totalmente de acuerdo	Totalmente de acuerdo	Desacuerdo
27	Masculino	19	4	No	Totalmente desacuerdo	Totalmente desacuerdo	Totalmente desacuerdo
28	Masculino	23	3	No	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo
29	Masculino	19	4	No	Acuerdo	Acuerdo	Totalmente de acuerdo
30	Femenino	20	1	Sí	Totalmente de acuerdo	Totalmente de acuerdo	Totalmente desacuerdo

Conclusions

This paper provides a basic explanation of the different types of variables that can be used in quantitative research. The aim is to make it easier to understand for people interested in getting started in educational research, by adopting a didactic perspective based on practical assumptions and concrete examples. This type of initial analysis is essential to ensure that the following steps of the study can be carried out correctly, as errors in the measurement of variables can hinder the efforts made in other phases of the research.

In addition, a tool has been presented which facilitates the execution and interpretation of descriptive analyses and which also has the support of a large community of users who, altruistically, improve the analysis offer of the programme and its correct execution on a daily basis. In fact, jamovi has different modules with more advanced analyses that can be installed according to the demand of each user, which makes it an analysis programme with enough power to cover the statistical needs of any research, not only at the undergraduate level, but also at the postgraduate level.

Therefore, students, teaching staff and educational researchers have the opportunity to know, use and spread this programme, without the need to master a specific programming language and at no cost.

Contribution of each Author: The contribution of each author to this publication was as follows: conceptualisation, FJIL and MSM; manuscript writing, FJIL, MPP and AIPG; writing, revising and editing, MPP and AIPG; supervision, MSM.

Funding: This research did not receive any external funding.

Conflict of Interest: The authors declare that they have no conflict of interest.

References

- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE.
- Porto Castro, A. M. (2022). El uso de los recursos tecnológicos como facilitadores del aprendizaje en la atención a la diversidad. *Espiral. Cuadernos del Profesorado*, 15(31), 71-81, <https://doi.org/10.25115/ecp.v15i31.7822>
- R Core Team (2019). *R: A Language and environment for statistical computing*. (Version 3.6) [Computer Software]. <https://cran.r-project.org>
- Şahin, M., & Aybek, E. (2020). Jamovi: An Easy to Use Statistical Software for the Social Scientists. *International Journal of Assessment Tools in Education*, 6(4), 670-692. <https://doi.org/10.21449/ijate.661803>
- Santabárbara Serrano, J., & Lasheras, I. (2020). Docencia de Bioestadística en Medicina con software gratuito jamovi: una ventana de oportunidad. *Revista Española de Educación Médica*, 1(1), 9–10. <https://doi.org/10.6018/edumed.421421>
- Suárez J. C., Recio P., San Luís M. C., & Pozo, M. P (2017). *Introducción al análisis de datos. Aplicaciones en Psicología y Ciencias de la Salud*. Sanz y Torres.
- The jamovi project (2022). *Jamovi*. (Version 2.3) [Computer Software]. <https://www.jamovi.org>



BASIC SURVIVAL MANUAL FOR DESCRIPTIVE STATISTICAL ANALYSIS

What is statistics?

An **area of mathematics** that deals with data sets to **represent** the characteristics of a population in such a way that they can be **analysed** and allow inferences to be made based on the calculation of probabilities, until a conclusion is reached.



And... descriptive statistics?

In contrast to inferential statistics, which aims to generate models, descriptive statistics refers to the **description, organisation, synthesis and graphical representation** of data.



Statistics works with variables

A variable is a **set of values** obtained by measuring a **certain characteristic** of each of the individuals belonging to a sample or population.

TYPE OF VARIABLE	DEFINITION	EXAMPLE
Qualitativa or categorial	It is one whose measured values are not numbers, but words.	For the item "gender", answers such as male, female, ..., which do not correspond to a number, are expected.
Ordinal	It is one for which symbols or numbers are assigned to order the values of the variable.	To the item "I consider the quality of studies offered in higher education to be very good", the individual answers between (1) strongly disagree and (4) strongly agree.
Quantitative	It is one whose measured values are numbers. It can be discrete or continuous.	If the individual is asked, by means of an item, his or her age or height, the answer will be a number.