**Exercise September 28. Statistics-I**

We want to find out if increasing the number of friends of a person in a social network make them more “socially attractive”. The researchers presented exactly the same Facebook profile to five different groups, the only difference being the number of friends of the profile (102, 302, 502, 702, or 902 friends). The assignment of the participants to each of the groups was at random.

***What is the independent variable?***

The independent variable is the "number of friends", which has 5 levels (102, 302, 502, 702, 902)

***What is the dependent variable?***

The “social attractiveness” of the person.

***Is this an experiment or a quasi-experiment?***

It is an experiment: the assignment of the participants to the levels of the independent variable was random.

***How many rows are necessary to analyze the data?***

As many as participants

***And how many columns?***

Two columns. One to indicate the group to which the participant belongs, and another to indicate the value of the dependent variable.

***Is this a "between-subjects" or a "within-subjects" design? Why?***

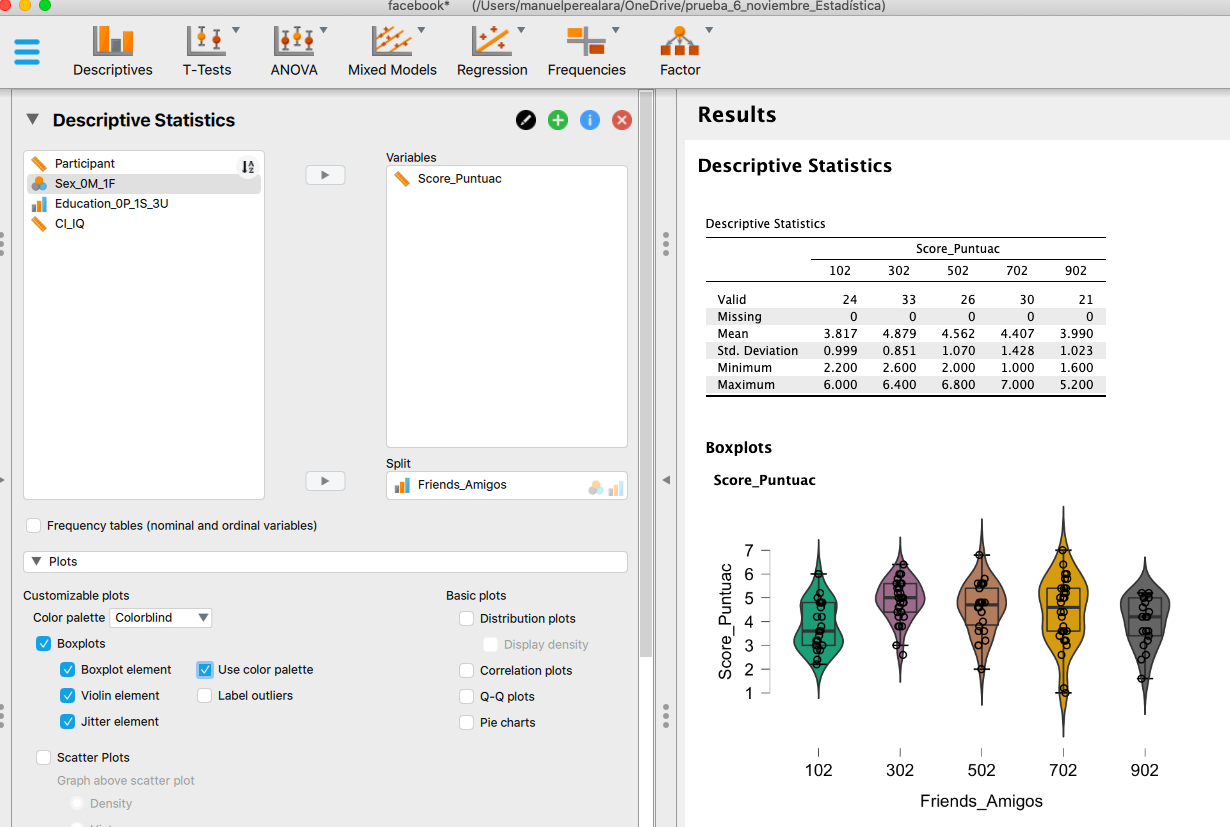
Between-subjects design: Each participant was assigned to ONE and only ONE of the groups.

***What type of plots can you use to describe to analyze the data? What would be your preference?***

The box plot is the most compact option, and surely the preferable one. Another option is to obtain one histogram for each of the levels of the independent variable, but it is less compact (it would be 5 histograms).

***What conclusions could you draw from this research with the following data (***[***https://www.uv.es/mperea/facebook.csv***](https://www.uv.es/mperea/facebook.csv)***)? (Remember that the file must be saved as .csv and then opened from JASP or SPSS; please, add the labels of the values ​​where appropriate)***

We have chosen the box plot:



The number of friends on Facebook modulates the perceived “social attractiveness”. Higher scores can be seen from 102 to 302 friends (M = 3.8 vs. 4.9, respectively), after which an increase in friends produces lower scores after the peak of 302 (502: M = 4.6: 702, M = 4.4; 902, M = 4.0).

In short, the increase in the number of friends does not lead to a monotonic improvement in social attractiveness: there are higher scores from 102 to 302 friends, but, after that, a greater number of friends (502, 702, 902) leads to a slow but progressive lower social attractiveness compared to 302 friends.

***Can you think of a way to improve this study? Do you think the employed design was the best choice? Could you improve the procedure?***

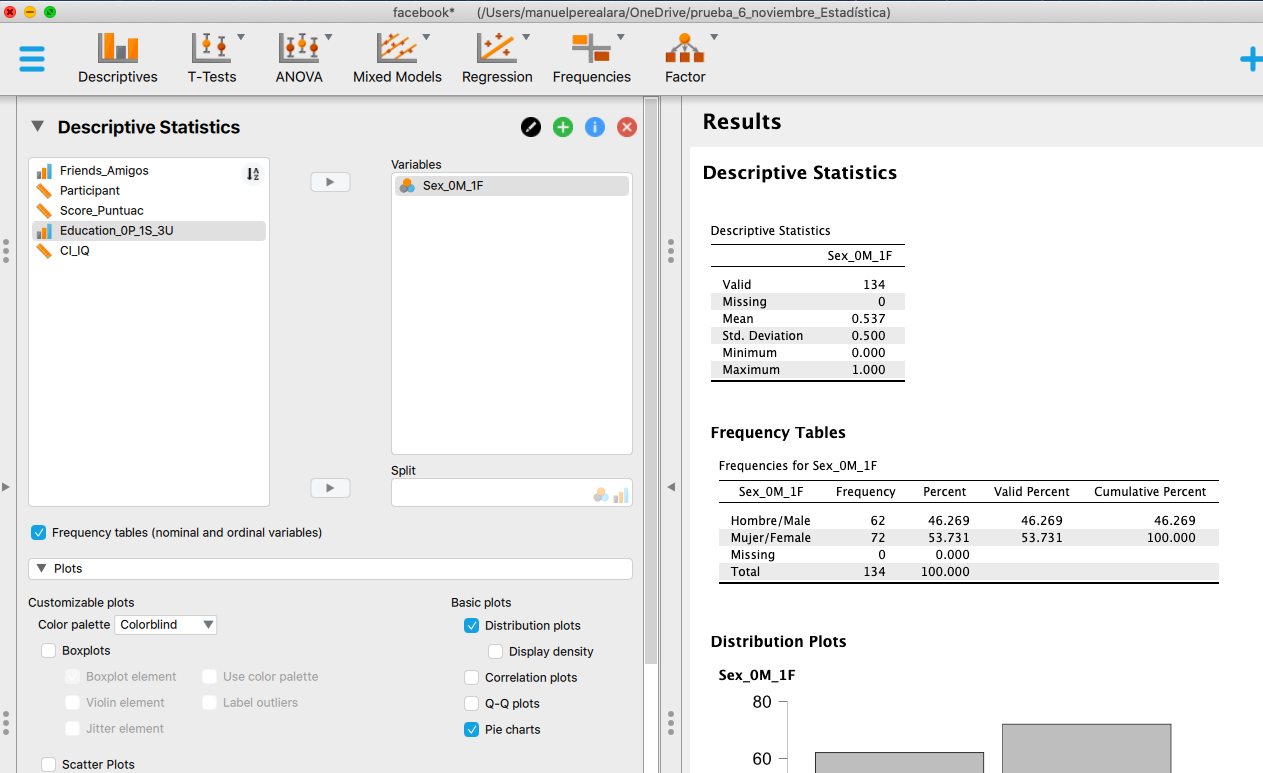
One option would be a within-subject design (that is, FIVE profiles are presented to each participant, one of 102 friends, another of 302 friends,…) because more data is obtained per participant (five instead of one) and because “each subject is her/his own control ”(if someone tends to be very positive or very negative, the participant will be in all conditions).

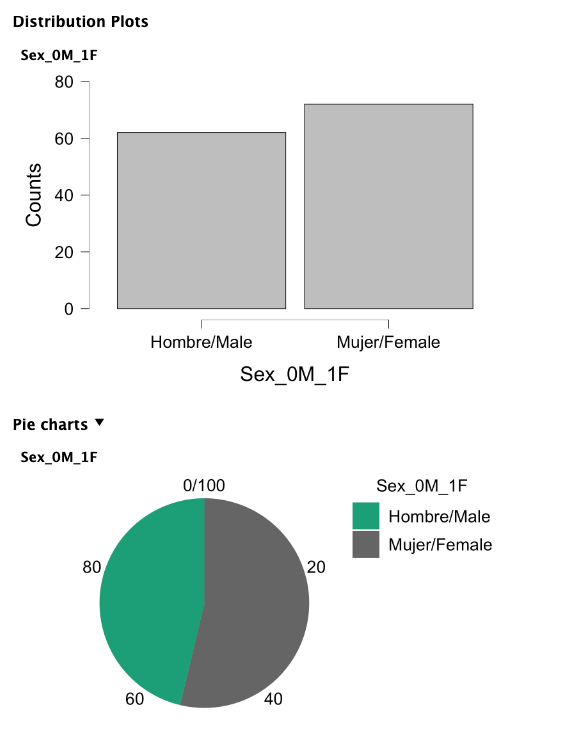
You may want to check whether the same pattern holds for the number of “followers” vs “following” in other social networks.

*As you will see, in the Excel file there were some other variables….*

***What is the appropriate plot for the variable Sex? Is it balanced in the sample?***

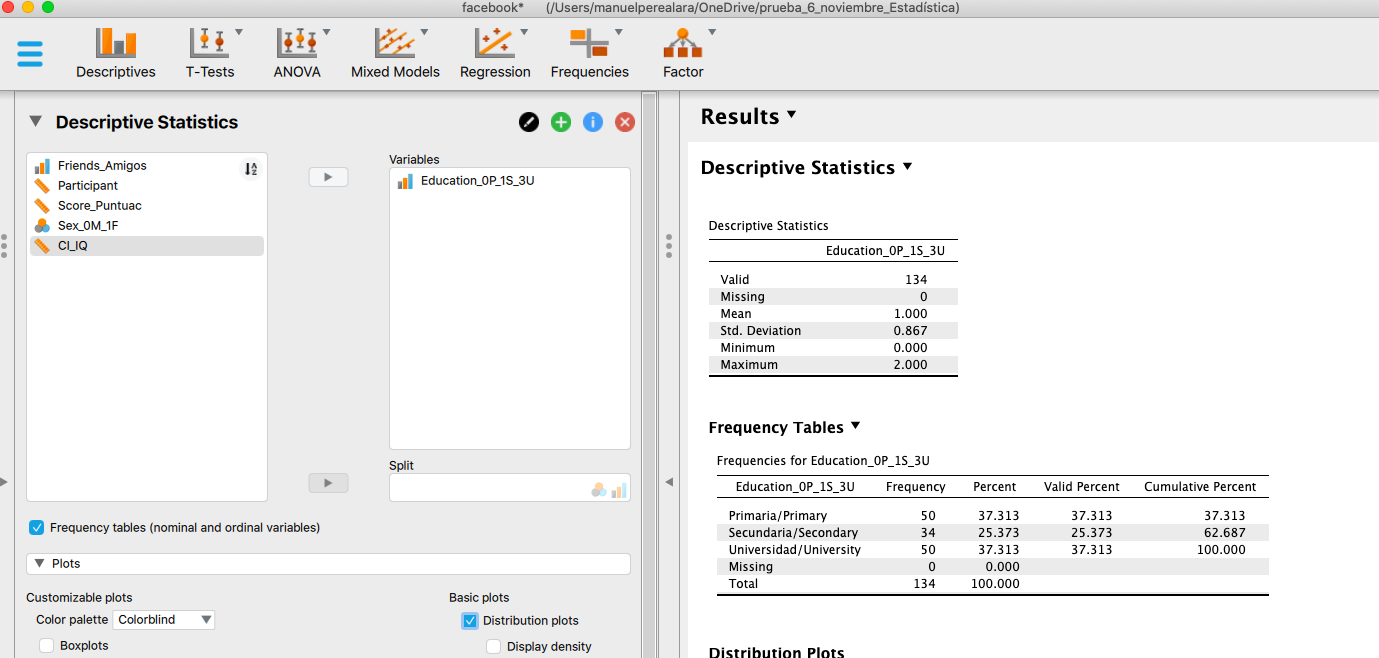
Bar chart or pie chart. You can also get the frequency table.

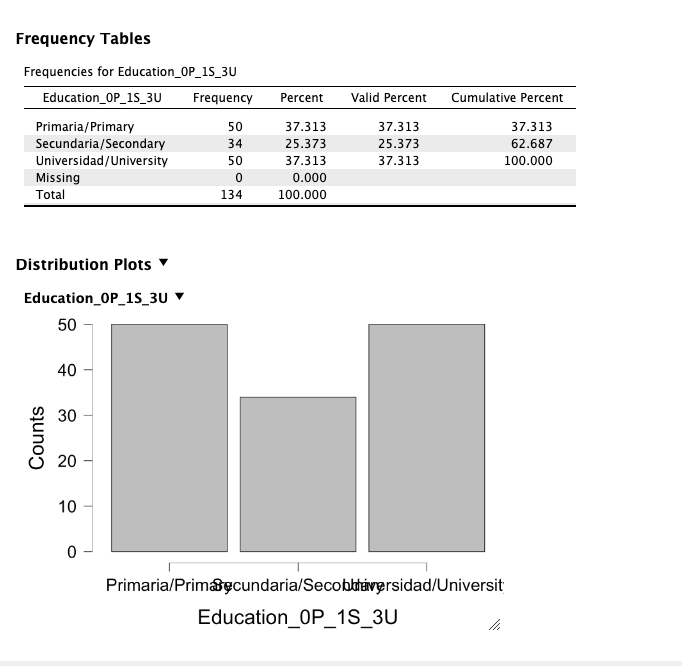




The variable “sex” is not completely well balanced, but it does not stray too far from 50/50, as can be seen in the frequency table or in the graphs.

*Idem for Educational level. Is it balanced in the sample?*



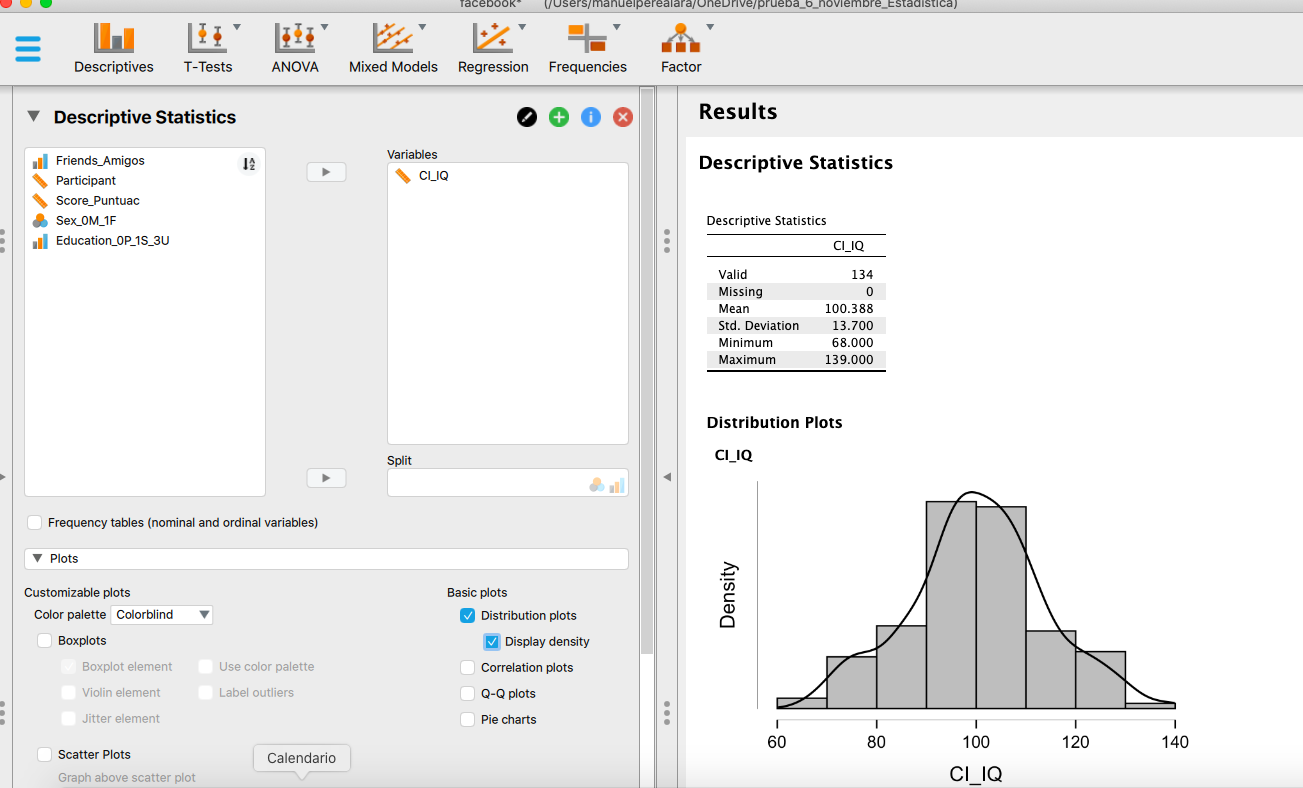


We chose the bar chart. All the information is actually in the frequency table. (The abscissa axis looks a bit weird when putting too much information on the value “label”.)

As can be seen in the chart (or in the frequency table), secondary school participants are under-represented in the sample.

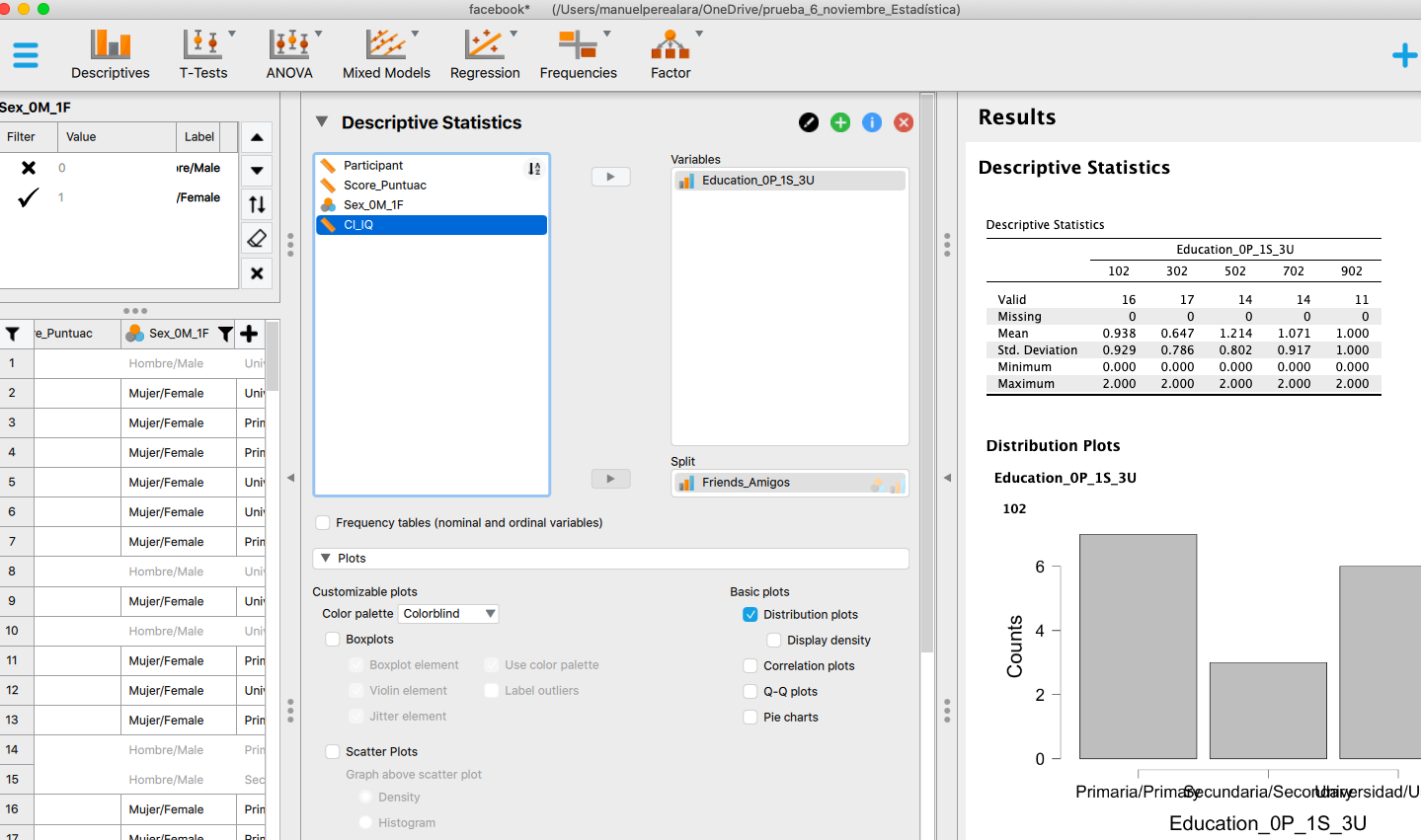
***Idem for IQ. Also, does it show how an approximately normal distribution would be expected?***

Note that "Distribution Plot", being IQ a quantitative variable, offers the histogram (instead of a bar graph).



The distribution appears to be approximately normal, with a mean of 100.

***Select only the women and plot the graph for educational level.***



You have to select/filer only the women (see panel on the left). The results are very similar to those of the global sample: high school people are under-represented.

References (taken from JASP, but they are not in APA format ...):

D. S. Moore, G. P. McCabe, and B A. Craig. Introduction to the practice of statistics (7th. ed). New York: Freeman.

Stephanie Tom Tong, Brandon Van Der Heide, Lindsey Langwell, Joseph B. Walther; Too Much of a Good Thing? The Relationship between Number of Friends and Interpersonal Impressions on Facebook, Journal of Computer-Mediated Communication, 13(3), 2008.