- contains not too many questions; usually, there are 3–4 discussion points (main area of discussion) for an interview lasting between one and a half to two hours.

Table 1.6. The outline of a script

1. Introduction (5 min)
– moderator's self-introduction
– informing the participants on the research subject
– informing about the rules of the discussion
– informing the participants that the discussion is recorded
– participants' self-introductions
2. Introductory questions—„a warm-up"—easy questions which everybody should be able to answer. For example, in research on the quality of food it can be:
– *Please finish the following sentence: Good food means to me ....*
– *What does it mean to you that some food is of good quality?*
3. Questions essential for achieving the research objective and supporting techniques (50–60 min)
4. Closing questions (15 min)

Source: Own elaboration.

We must remember that the questions essential for achieving the research objective should be open, e.g.: *If you were the manager, what would you change first?*, and non-threatening as e.g. a question with reproach: *Why didn't you give up smoking?*

1.2.6. Ecological culture of Bulgarians: case study (Part 2)

In the previous part of this case study, a brief description of a qualitative research initiated by the Bulgarian Ministry of Environment and Water was presented. The goal of the research was to study in detail the attitudes of Bulgarian citizens to ecology, ecological issues, and the measures undertaken by the government to solve these issues, as well as to better understand everyday behaviour regarding ecological issues and personal involvement in environmental topics. In this part, the planning stage of the focus group interview—the creation of questions and script are focused on.[2]

Before the start of each session, participants received a document explaining the privacy policy, and a form to express their consent for researchers to record, store, analyse, and present their opinions. When the discussion started, the moderator followed the script shown in Table 1.7.

---

[2] Please note that the questions and the script are not the ones used by the research agency that conducted the research. The given script is created specifically to fit the learning objectives of this chapter.

Iwona Olejnik, Atanaska Reshetkova

Table 1.7. Ecological culture of Bulgarians—script of the focus group

Script item	Estimated time
Introduction The moderator introduces himself and describes the purpose of the interview. Then he asks for permission to record the session with an audio and / or video recorder. He kindly invites participants to introduce themselves	5 min
Warm-up questions – What is your definition of ecology? – What do you hear people around you say about today's ecological problems? – What kind of people care about the environment, and why? How do you know that they care?	10 min
Everyday life – How can households limit the waste they produce? – Which of the above-mentioned measures do you apply at your home? – What sources of information about ecological issues do you typically use? – Do you participate in any nature preserving initiatives? – Does your household engage in a separate collection of waste? – What do you think about energy-saving appliances? Do you have any?	20 min
Nationwide measures – Who do you think is responsible for the prevention of pollution of air, soil, and water? *Probe:* Do you think that people / government should also be held responsible? – Are you aware of any initiatives of the government that are aimed at solving a major environmental problem? *Probe:* Do you think that there are none or they are not well promoted? – Do you think that the government does enough to protect the environment? *Probe:* What more should be done? – Can you name a recent activity—a government or citizens' initiative—aimed to protect the environment that you remember taking place in your hometown? – Do Bulgarians participate in initiatives intended to preserve nature for future generations? *Probe:* What are the most popular initiatives you have heard of? – Are people in Bulgaria educated on the environmental issues in the country and the whole world? *Probe:* What could be done to make them more educated?	20 min
Ideological questions – How do you understand the relationship between people and the environment? – Is there a harmony between humans and nature? *Probe:* What is causing this lack of harmony? How can this harmony be restored? – How important is the protection of the environment for the personal well-being of most Bulgarians? – What is the reason why some people ignore environmental problems?	20 min
Closing questions – Is there anything else you would like to share? – Does anyone have something to add that was not commented so far?	15 min

Source: Own elaboration.

1.3. Conducting a focus group interview

1.3.1. The organizational aspects of FGI

Inviting respondents to the interview we must remember about:
- „back-up participants"—as there may be people who will not come (e.g., while conducting and FGI with people who buy clothing in second-hand stores, out of 10 invited women only 6 joined, even though they had confirmed their attendance by phone the previous evening);
- incentive tools to make people come and participate in a discussion—taking part in an interview is often rewarded with a gift or some amount of money—larger or smaller, e.g. the respondent can receive a gift voucher for the value of 20–25 euros.
- the time and location of the interview—it should finish before 8.00 p.m. If we invite housewives with low income, we should not arrange for it to be in a luxury hotel;
- preparation of recording equipment (along with backup equipment).

1.3.2. FGI moderation

A person who conducts an interview is a moderator. His / her preparation, experience and predisposition are essential for a properly conducted interview (McDaniel & Gates, 2018).

Features of a good moderator
- shows respect for respondents
- expresses himself / herself clearly and loudly
- asks simple questions
- actively listens to answers
- demonstrates creativity and flexibility when asking questions
- uses various techniques involving all respondents
- adopts an open body posture
- conducts the discussion in a logical sequence, from general to specific questions

When asking questions, a moderator should:
- make further inquiry about an ambiguous thing said by a respondent. *Could you explain it to me again, in other words?*
- extend the statements *Could you tell me more about it?*, or *So you mean the biggest advantage of this product is its simplicity?*

How to invite a person to give us some opinion and extend conversations? The moderator could ask *do you agree*?, but it is quite closed question. So, you should open up that question to get a more interesting answer.

There are some examples of a polite way of inviting respondents into the discussion:

- *What do you think…?*
- *What is your opinion…?*
- *How do you feel…?*
- *I am interested to hear your opinion about this.*
- *I would like to hear your thoughts on…*
- *Would you like to add anything?*

On the other hand, what should researcher say when they want to interrupt—when someone else is talking and talking. He or she could say:

- *Excuse me,*
- *Do you mind if someone else adds to that,*
- *Sorry to interrupt but…, sorry for interrupting but…,* or *if I may interrupt for a minute…*

And when somebody has gone off the topic, and started talking about something else—not necessarily connected with the main topic, the moderator has to lead the conversation back to the topic, he / she can say something like this:

- *So anyway, getting back to my question…*
- *So anyway, where were we?*
- *As I was saying / asking…*

In summary, there are a few techniques of conducting group discussion, including: paraphrasing, confirmation and further extending questions.

The beginnings of paraphrasing statements:
I understand that ....
What you mean is ....
What you are saying is .... *(and repeat back what you heard),*
I understand that you are asking about ....
Do you mean ....? *(and then paraphrase what you heard),*
In other words, what you are saying is .... *(and summarize what you heard),*
I would like to confirm what you said ....
Can I just confirm that ....?

Paraphrasing is shortened restating of a person's words by a moderator. It should not contain more details than you have heard. If you use a paraphrase, you show that you are paying attention and understand what has been said and it highlights your interest in what a respondent has to say. Moreover, it allows you to check if you have correctly understood a respondent's intentions (if you have misunderstood,

your respondent has a chance to correct your mistake and present the idea more clearly) and allows you to organize the content of a respondent's message.

On the other hand, we use clarification when we ask the respondent about some unclear answer. We then use sentences like: *Could you explain this to me again, in other words?* or *Can you explain once again what it means to you?*

The examples of further extending questions include: *Could you say something more about it? Could you give me more details?*

Finally, it is worth adding that when the topic is controversial, it does not mean, that moderator needs to avoid it. How to keep the conversation open and positive in this situation? The moderator could start like this: *I know not everyone agrees with it…* or *I know this is a contentious issue, but…*

When conducting FGIs we may come across various problems. These include the following: overeager debater, untalkative—'silent' participants, overactive or passive group and jokers. But of course, a well-prepared, qualified moderator can handle these issues.

Conducting face-to-face interviews is different from online projects. The basic differences in moderating interviews are presented in Table 1.8.

Table 1.8. Moderating online and face-to-face interviews

Online	Face-to-face
– the discussion may be conducted by more than 1 moderator; a second moderator is very useful – it allows you to generate more data in less time – length of the interview: maximum 90 minutes – much greater enthusiasm and involvement of the moderator is necessary – greater opportunities for communication between moderators, and moderators and the research client	– the interview is usually conducted by only 1 moderator – it allows you to generate less data in a longer time – length of the interview: about 1.5–2 hours – it is easier to control the involvement of all respondents – less communication opportunities between the moderator and his assistant, and between the moderator and the manager during the survey

Source: (Olejnik, Dębska & Zieliński, 2020).

1.4. Analysis of qualitative data using the CAQDAS programs

1.4.1. Fundamentals of qualitative data analysis

CAQDAS is an acronym for computer-assisted qualitative data analysis software. Qualitative data typically includes text, images, audio and video materials. The decision whether or not to use CAQDAS software in the process of analysing

Usually, we use numbers for coding the categories, and in the case of nominal level for measurement, the numbers have no meaning. Using codes like:

1. *Bulgaria,*
2. *Croatia,*
3. *Czech Republic,*
4. *Hungary,*
5. *Poland,*
6. *Slovakia,*
7. *Ukraine,*
8. *Other,*

or

1. *Ukraine,*
2. *Slovakia,*
3. *Poland,*
4. *Hungary,*
5. *Czech Republic,*
6. *Croatia,*
7. *Bulgaria,*
8. *Other*

gives the same possibilities in interpretation.

What should remain constant is the last position of "other". "Other" is used very often if we cannot name all possible answers.

Data from a nominal scale should not be used for arithmetical calculations, it would be meaningless. It does not make sense to replace Czech Republic (represented by value 3) by combination of added Bulgaria (1) and Croatia (2). And there is no point in presenting the arithmetic mean from the countries from the above example.

If the variable has only two possible categories we call it a dichotomous variable. The distinction is important, as some statistical methods can be used for dichotomous variables and cannot be used for other nominal variables. For example, the question about believing in human driven climate change, with two possible answers (yes, no) is a dichotomous variable.

The ordinal level of measurement offers us more than nominal data. Using the ordinal scale, we not only compare if the elements are the same or different, but we can also order them according to the intensity or time. However, the data tell us nothing about the differences between the values.

Imagine the three households A, B and C. The households compare produced waste. Household B produced the most, whereas household C produced the least. However, we know nothing about the differences. Was household A more similar in waste production to B or C? Maybe it produced almost the same number of waste bins as B with a huge difference to C? This could also be completely different.

Example of ordinal level of measurement:
– *Performance in waste production (first, second, third)*
– *Liking this e-book (dislike very much, rather dislike, rather like, like very much)*
– *Supporting state investments in coal mining (fully against, rather against, rather accepting, fully accepting)*
– *Attitude toward sorting waste: very negative, rather negative, rather positive, very positive*

The numbers are important and have to be assigned to the verbal levels in ascending or descending order.

Interval level of measurement is even more useful. We not only compare if the elements are the same or different, or not only can we order them, but we are able to measure the distances between them. Data are interval, if the differences can be measured in units, and the units have equal intervals between them. The interval variables can have negative values.

Example of interval level of measurement:
– *Bank account balance (e.g., –270 EUR, 50 EUR, 70 000 EUR)*
– *Temperature of the sewage in main sewer collector in Celsius (e.g., 17, 21)*
– *The year you have started sorting waste in your household: (e.g., 2001, 2002, 2003…)*

We can calculate that Samantha's household (1995) started sorting waste six years earlier than Bob's household (2001).

Ratio level of measurement goes a step further. Data collected on this level allow for comparing elements, ordering them and measuring the distances between them. And additionally, the ratio level allows for calculating meaningful ratios of values along the scale. This is possible only if the scale has true and meaningful zero point. The zero value indicates in this case, that there is none of that variable. Of course, the ratio level variables cannot have negative values.

Example of ratio level of measurement:
– *Number of bumblebees visiting a given flower in one hour (e.g., 0, 5, 35)*
– *Weight of the waste produced in one month by the household (15 kg, 45 kg, 70 kg)*
– *Tons of unsorted waste reported by municipality of your city in 2020: (e.g., 3, 250, 781)*

2.1.3. Measurement levels and data analysis methods

The level of measurement has serious impact on how we can analyze them. You will know more about it in the next part of the book. Table 2.1 shows the central tendency measures and basic calculations available for given levels.

Table 2.1. Levels of measurement and basic calculations

	Nominal	Ordinal	Interval	Ratio
Mode, frequency distribution	+	+	+	+
Median, presenting order		+	+	+
Calculating mean, adding, subtracting			+	+
Calculating ratio				+

Source: Own elaboration.

Levels of measurement imply the permissibility of statistical techniques. When forming hypotheses and designing the questionnaire for the primary research we always have to plan the method of future analysis and check its feasibility.

Table 2.2. Permissible descriptive and inferential statistics for certain levels of measurements

	Descriptive	Inferential
Nominal	percentage, mode	chi-square, binominal test
Ordinal	percentile (e.g., median)	rank-order correlation, Friedman ANOVA
Interval	range, mean, standard deviation	product-moment correlations, t-test, ANOVA, regression, factor analysis
Ratio	geometric mean, harmonic mean	coefficient of variation

Source: Own elaboration based on: (Malhotra & Birks, 2003).

2.2. Questionnaire design

2.2.1. Stages in creating a questionnaire

The questionnaire is the basic research tool used in the primary quantitative research, both in the interview methods (in which the questions are asked, and answers are written by the interviewer) and in the questionnaire (in which the respondent reads the questions and writes down the answers himself). It is an indispensable research tool in many types of research: door-to-door interviews, executive interviews, mall-intercept interviews, telephone interviews, self-administered questionnaires, and mail surveys

(McDaniels & Gates, 2018). The questionnaire is a set of written questions formulated to elicit the desired answers related to the examined problem (Neuman, 2014).

A properly constructed questionnaire should meet many rules concerning both the content of the questions, the way they are formulated, the scaling of the answers, the order of the questions, and the graphic design (Churchill & Iacobucci, 2018). The questionnaire should encourage the respondent to participate in the survey and make the survey interesting for them. At the same time, it is designed to facilitate the work of the researcher (especially the interviewer) through instructions on how to ask questions and write down the answers.

The following stages of building the questionnaire can be distinguished:
1) specification of the purpose and subject of the questionnaire,
2) formulation of an initial list of questions,
3) initial scaling of the response,
4) initial check of the questionnaire,
5) construction of the sample questionnaire,
6) pilotage and possible modifications of the questionnaire,
7) preparation of the final version of the questionnaire.

When starting to build the questionnaire, one should remember about the assumptions of the research project, in particular, about defining who the respondents will be (their level of education, age, etc.) and what their knowledge is in relation to the subject covered by the research (general or specialist?)—it determines form of asked questions and the use of specific phrases.

The construction of the questionnaire, including the number of questions and their form, depends on the research method. The form will look slightly different in the interview method than in the survey method. These methods differ from each other, inter alia, in the form of contact with the respondent (direct, indirect) and the number of questions asked and answers received. An important determinant of the selection of the type and construction of the questionnaire is also the data that is to be collected—their scope and details, as well as the costs to be incurred, time and planned methods of analysing the obtained data.

Regardless of the intended research method and the type of questionnaire, certain elements are present in each of them. The questionnaire therefore consists of three basic parts:
1) introductory part—the header,
2) the main part—containing questions aimed at solving a research problem,
3) respondent's particulars—personal information (data about the respondent).

The introductory part—the header of the questionnaire—should include such elements as: the name and contact details of the organizer of the study, the title of the questionnaire, a short description of the purpose, nature and scope of the study, and possible indication of benefits for the respondent resulting from the

participation in the study, ensuring that the survey is anonymous (optional), the date of return (in the case of forms returned by post), the name and surname of the interviewer and, possibly, information on how to complete the questionnaire.

2.2.2. Types of questions in the questionnaire

The main part of the questionnaire consists of the subject-related questions concerning the research problem. There should be enough questions to reflect the entire substantive scope of the planned research, and thus to enable the verification of the research hypotheses set out in the project, including the following categories of questions (Neuman, 2014, p. 317):
–   behaviour (*How frequently do you ...., When did you last ....*),
–   attitudes, beliefs, opinion (*What is the biggest problem facing ...., What do you think about ....*),
–   expectations (*Do you plan to buy .... in the next 3 months*),
–   self-classification (*Do you consider yourself to be ....*),
–   knowledge (*About what percentage of .... It is legal to own ....*).

The questions in the questionnaire can be open or closed. Open-ended questions do not have any given options, i.e., the respondent formulates the answers in a completely arbitrary way. They are used when it is difficult to predict the answers of the respondents, when the number of these answers may be too high, or when they concern difficult issues, e.g.:

> How should food waste be prevented? Please develop your answer / write down your suggestions ....
> ....
> Please, finish the statement: „In second hands the most irritating thing is ....,
> because ....”

Closed-ended questions (with a scale) provide for the selection of answers from a prepared set of options predicted and determined in advance by researchers, i.e., from the so-called cafeteria. These questions are also called scaled questions. It should be remembered that the consequence of the selection of a specific scale are—after the research is conducted—the methods of processing and analysing the data obtained using them (Aczel, 2009). As a result, the so-called one-dimensional and multi-dimensional scales are applied (Neuman, 2014).

An example of a one-dimensional closed question is, e.g.:

> Do you think that biodegradable waste should be sorted separately? *Please underline 1 answer*: a) definitely yes, b) rather yes, c) rather not, d) definitely not.

Statistics		
sustainability		
N	Valid	200
	Missing	0
Mean		6,33
Std. Error of Mean		,198
Median		6,00
Mode		6[a]
Std. Deviation		2,800
Variance		7,840
Skewness		,010
Std. Error of Skewness		,172
Kurtosis		–,162
Std. Error of Kurtosis		,342
Range		14
Minimum		0
Maximum		14
Sum		1266
Percentiles	25	4,25
	50	6,00
	75	8,00
a. Multiple modes exist. The smallest value is shown		

Figure 2.5. Descriptive statistics for ratio variable (Frequency command)

Source: Own elaboration.

The frequency tables can be visualized with the chart. Same frequency procedure allows for producing charts (clicking the button charts). We can choose between:
– a pie chart,
– a bar chart,
– a histogram.

Pie charts are designed for nominal data. They can show the percentage of categories, but they are difficult to compare. They are not clear with the bigger number of categories. And even with the smaller number of categories we can be mistaken assessing the structure. So even for the nominal data the bar chart is really worth considering as the default one. To explain the issue, compare the structure in two pie charts presented below.
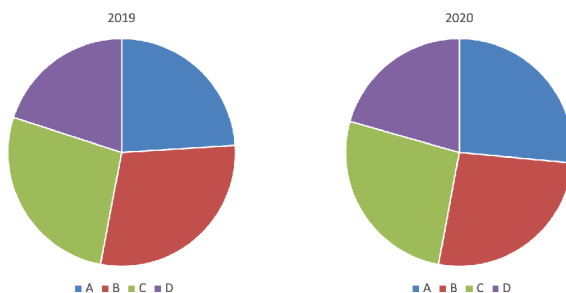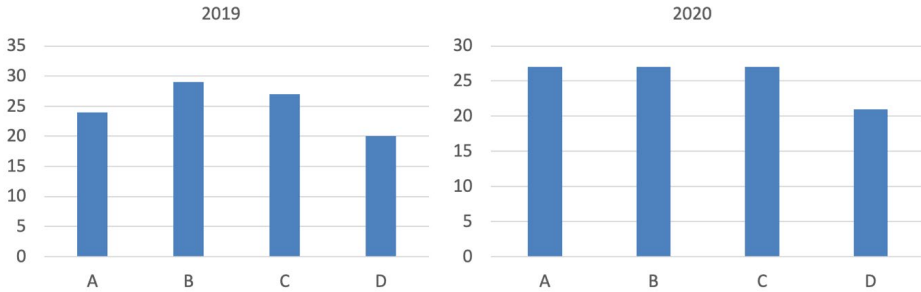

**Figure 2.6. Pie charts**

Source: Own elaboration.

Bar charts are designed for nominal and ordinal data. They show differences between frequencies of categories much better. The vertical axis can show counts or percentages. Compare the same data as presented above in the form of bar charts, the differences within the chart and between the charts are much clearer now.

Figure 2.7. Bar charts

Source: Own elaboration.

The histograms are designed for interval and ratio levels of measurement. They are similar to bar charts, but bars are touching (continuous data), and every bar shows the frequency of observation within the interval.
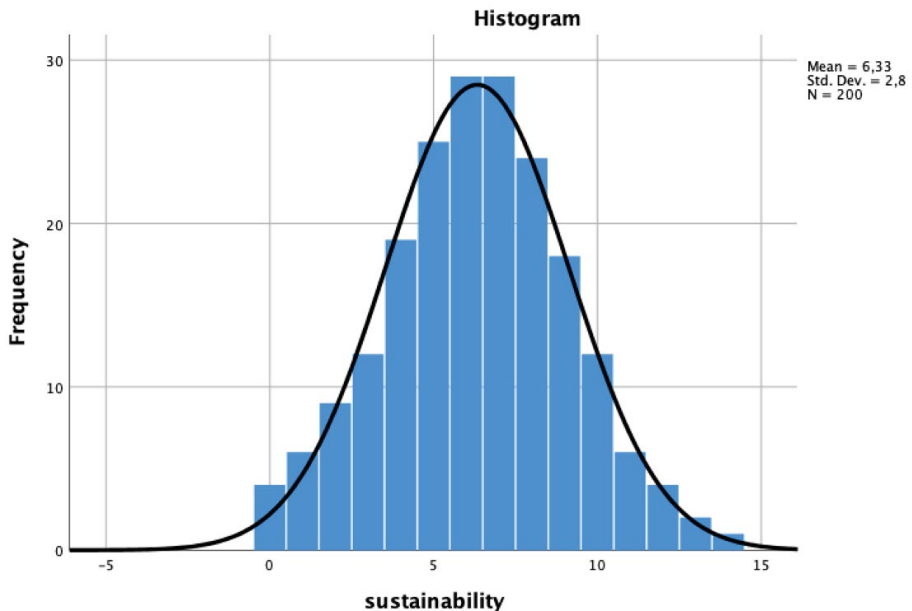
Figure 2.8. Histogram with the added normal curve

Source: Own elaboration.

2.4.2. Box and Whiskers chart

There is one more popular and very informative way to visualize variables. The chart is called box and whiskers chart, five numbers chart or box chart. What we need for constructing the chart is a minimum and a maximum, first, second and third quartile. The box is drawn having its bottom as the first quartile, the upper limit as the third quartile and the thick line within this range presenting the median.

The minimum and maximum are presented in the form of whiskers. It sounds very easy and can be hand-drawn, but SPSS definitely helps to produce it.

If the minimum (maximum) is considered an outlier, the whisker will present the smallest (highest) value not exceeding the 1.5 of interquartile range. And all the outliers will be shown as small circles (closer distance outliers) or asterisks (far distance outliers). You can produce the box and whiskers chart using the Chart menu in SPSS (both Charts / Legacy Dialog / Boxplot and Chart / Chart Builder / Boxplot lead to this), but a very handy way is available in Analyse / Descriptive Statistics / Explore. It produces a bunch of useful descriptive statistics, and two graphical summaries: Stem and Leaf (not described here), as well as Box and Whiskers chart. Just drop the investigated variable into the field of Dependent list. If you want to compare the results splitting your data (e.g., for females and males) you have to drop this variable into the factor list. There is one more field there, the Label Cases by. It tells SPSS how to name the outliers. If you leave this field empty, every outlier will get the number of rows from the data file, so you can spot it if needed. If you drop the variable there, the outliers will be named according to their value of the chosen variable.
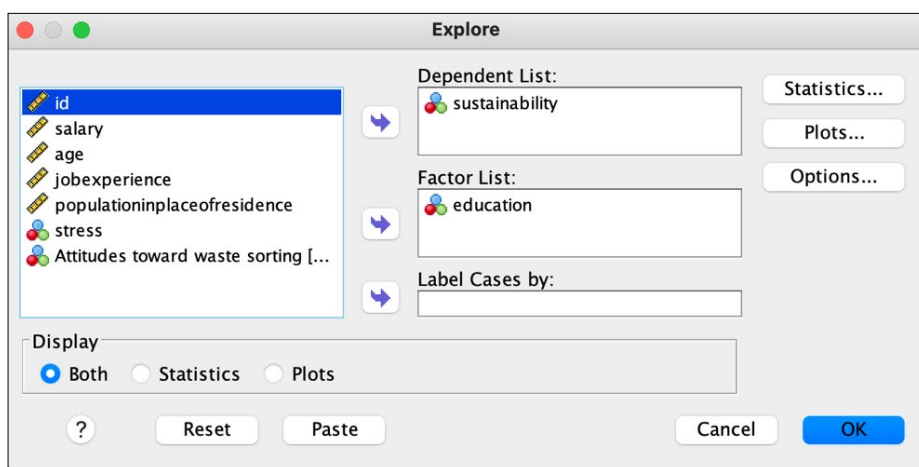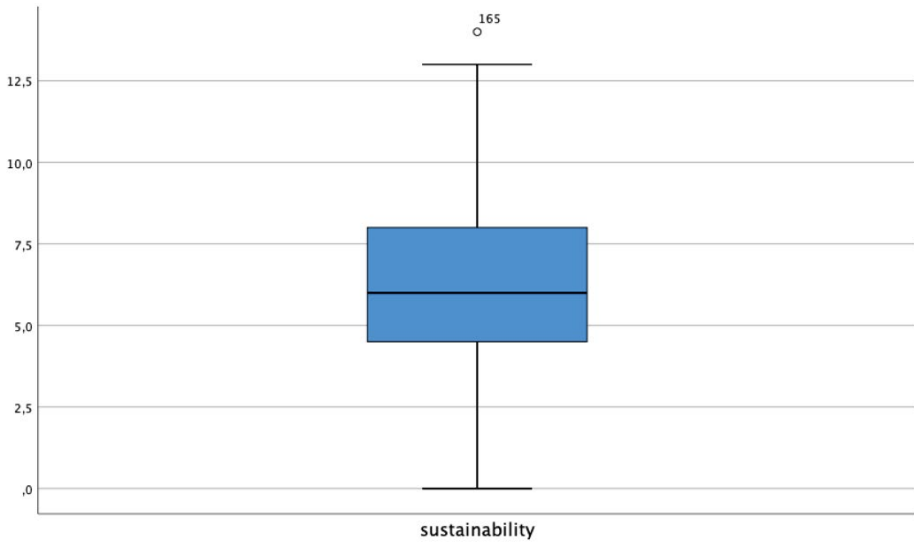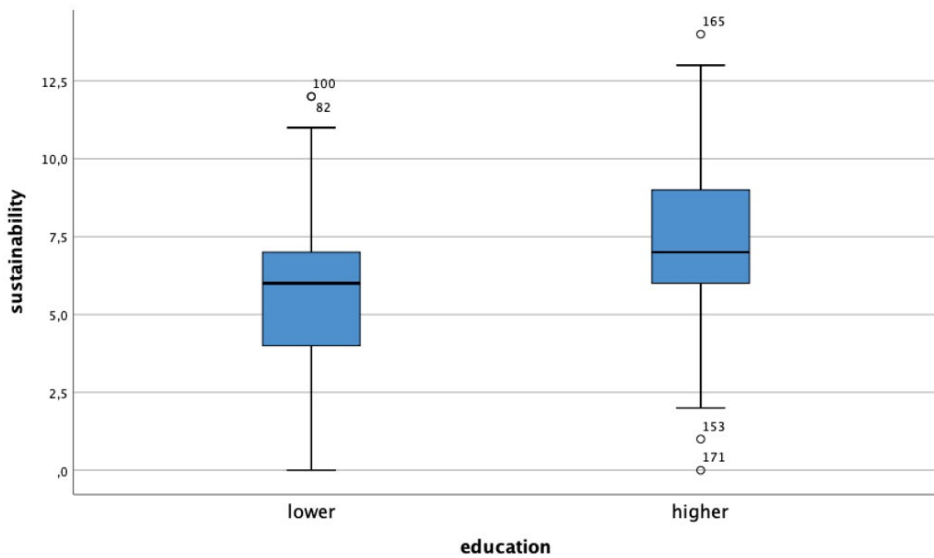

Figure 2.9. Dialog box for Explore (including Box and Whiskers chart)

Source: Own elaboration.

Figure 2.10. Box and Whiskers chart—single box

Source: Own elaboration.

Figure 2.11. Box and Whiskers chart—comparison (more boxes)

Source: Own elaboration.

The comparison box shows that even if the difference between medians is small, the typical 50% of the people (inside the box) with higher education have better sustainability knowledge compared to the people with lower education levels. We can observe two outliers in the group with lower education (having much higher knowledge) and three outliers in the group with higher education (two much lower, one much higher).

2.4.3. Crosstabs: percentage

Tables are the most popular way to present data. But this popularity does not mean, that readers always understand them correctly. In this chapter we will focus on the tables presenting relationships between two categorical variables.

A simple table usually presents the dependent variable and all its values as rows. The independent variable constitutes columns. This allows to check the counts or percentage in the cell where row and column cross.

The basic table may be obtained in SPSS in Analyse / Descriptive Statistics / Crosstabs.
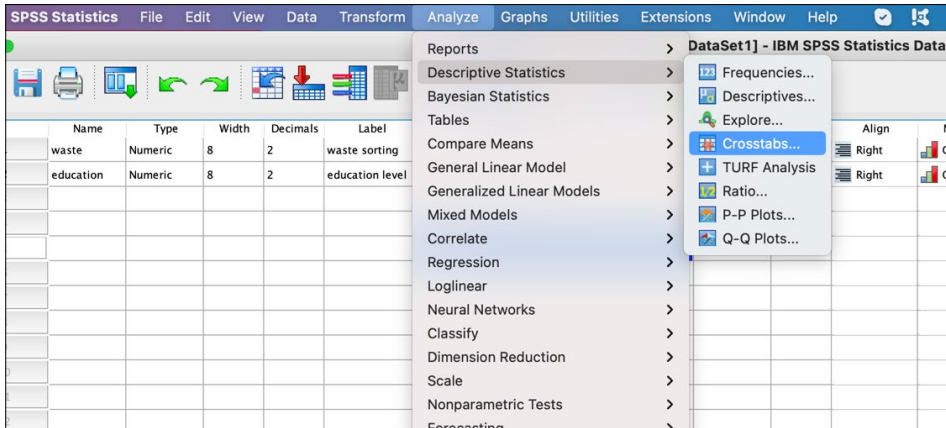

Figure 2.12. Obtaining the Crosstabs: path

Source: Own elaboration.

In the dialog box you have to drop the independent and dependent variable into fields called Row(s) and Column(s) as shown on the example in Figure 2.13.

Referring to the model in Figure 4.4, the measurement model could be defined as follows:

```
# Step ONE: Defining a measurement model
ECSIPLSmm <—matrix(c(
 "CUEX", "Cuex1",
 "CUEX", "Cuex2",
 "CUEX", "Cuex3",
 "PERQ", "Perq1",
 "PERQ", "Perq2",
 "PERQ", "Perq3",
 "PERQ", "Perq4",
 "PERQ", "Perq5",
 "PERV", "Perv1",
 "PERV", "Perv2",
 "IMAG", "Imag1",
 "IMAG", "Imag2",
 "IMAG", "Imag3",
 "IMAG", "Imag4",
 "CUSA", "Cusa1",
 "CUSA", "Cusa2",
 "CUSA", "Cusa3",
 "CUSCO","Cusco1",
 "CUSCO","Cusco2",
 "CUSL", "Cusl1",
 "CUSL", "Cusl2",
 "CUSL", "Cusl3" ), ncol=2, byrow=TRUE)
```

The definition of the structural model from Figure 4.4 is performed in a similar matrix format:

```
# Step TWO: Defining a structural model
ECSIPLSsm <—matrix(c(
 "CUEX", "PERQ",
 "CUEX", "PERV",
 "CUEX", "CUSA",
 "PERQ", "PERV",
 "PERQ", "CUSA",
 "PERV", "CUSA",
 "CUSA", "CUSL",
 "CUSA", "CUSCO",
 "CUSCO","CUSL",
 "IMAG", "CUEX",
 "IMAG", "CUSA",
 "IMAG", "CUSL" ), ncol=2, byrow=TRUE)
```

To fit the PLS model to the simulated 799-respondent data set, we use `plsm(data, strucmod, measuremod)` command line from semPLS package, based on matrices that previously were defined on step one and step two. Next, using `sempls(model, data)` command line we can estimate model parameters:

```
# Defining the whole model and fit it to data
library(semPLS)

ECSIPLS.mod <-plsm(data=ECSISimData , strucmod=ECSIPLSsm , measuremod
=ECSIPLSmm)
ECSIPLS.fit <-sempls(model=ECSIPLS.mod , data=ECSISimData)
```

The evaluation results are contained in the `ECSIPLS.fit` object. To estimate the factor structure (so-called factor loadings) of the measurement model, it is necessary to use the command plsLoadings (MODEL):

```
> plsLoadings(ECSIPLS.fit)
  IMAG CUEX PERQ PERV CUSA CUSCO CUSL
Imag1 0.71    .    .    .    .    .    .
Imag2 0.59    .    .    .    .    .    .
Imag3 0.70    .    .    .    .    .    .
Imag4 0.72    .    .    .    .    .    .
Cuex1    .  0.76    .    .    .    .    .
Cuex2    .  0.83    .    .    .    .    .
Cuex3    .  0.72    .    .    .    .    .
Perq1    .    .  0.75    .    .    .    .
Perq2    .    .  0.71    .    .    .    .
Perq3    .    .  0.72    .    .    .    .
Perq4    .    .  0.68    .    .    .    .
Perq5    .    .  0.70    .    .    .    .
Perv1    .    .    .  0.78    .    .    .
Perv2    .    .    .  0.86    .    .    .
Cusa1    .    .    .    .  0.86    .    .
Cusa2    .    .    .    .  0.83    .    .
Cusa3    .    .    .    .  0.81    .    .
Cusco1   .    .    .    .    .  0.85    .
Cusco2   .    .    .    .    .  0.87    .
Cusl1    .    .    .    .    .    .  0.83
Cusl2    .    .    .    .    .    .  0.81
Cusl3    .    .    .    .    .    .  0.85
```

In this example, all latent variables have strong loadings with its manifest indicator variables. If a latent variable has a factor loading below 0.3 for any indicator (which in this case is not present), or below 0.5 for all of its indicators—then the reliability of the measures will be debatable and further investigation and / or re-specification of the model structure is needed (Chapman & Feit, 2019, p. 288).

To obtain the values of the coefficients of the relationships between the latent variables (i.e., the structural model) it is necessary to use the command `pathCoeff(MODEL)`. The results are presented below:

```
> pathCoeff(ECSIPLS.fit)
       IMAG   CUEX   PERQ   PERV    CUSA  CUSCO    CUSL
IMAG    .   0.287     .      .     0.062     .    0.096
CUEX    .       .   0.370  0.334   0.114     .       .
PERQ    .       .      .   0.065   0.354     .       .
PERV    .       .      .      .    0.157     .       .
CUSA    .       .      .      .        .  -0.443  0.475
CUSCO   .       .      .      .        .      .   -0.161
CUSL    .       .      .      .        .      .       .
```

We see that 'Customer complaints' has a negative influence on 'Customer loyalty', and 'Customer satisfaction' has also negative influence on 'Customer complaints', which corresponds to the expectations (face validity).

In order to visualize the results of the evaluation through the familiar path diagram, it is necessary to take several additional technical steps, using the free available third-party software component Graphviz.[10] Using the program command `pathDiagram(MODEL, FILE, full-TRUE, …)` we can create a plot with the structural coefficients and loadings as an output as *.dot file. Once Graphiz software is installed, we can process this file as input and produces the corresponding image as a PDF or PNG file. For this purpose, first it is necessary to execute the following program line:

```
# Creating object for path diagram for Graphviz
pathDiagram(ECSIPLS.fit , file = "ECSIPLSfull", full = TRUE , digits = 2,
 edge.labels = "values", output.type = "graphics", graphics.fmt = "pdf")
```

Because PLS models do not assess global model fit, according to Chapman and Feit (2019, p. 289), „(…) there is not a general way to compare CB-SEM and PLS-SEM results apart from interpreting the models and their implications, so it is not advisable to compare the coefficients directly".

After fitting the model to the data, however, the question of its statistical evaluation remains open. Because by PLS-SEM there is no summary metric for global model assessment and comparison, there are three possible approaches to its validation. First, we can examine the model's coefficients for their face validity. *Face validity* is the extent to which a coefficient estimated corresponds to our subjective expectation (in size and sign). Second, we can calculate the overall coefficient of determination (R^2) for the model, which is a measure of overall variance explained

---

[10] Graphviz can be downloaded for free from: http://www.graphviz.org/download/

within each part of the model. Thirdly, one can think about applying a bootstrap method to examine coefficient stability (Hair, Sarsted et al., 2011, pp. 423–424; Chapman & Feit, 2019, p. 289).

Figure 4.10. Path diagram with PLS coefficient estimates for the ECSI model using `semPLS`
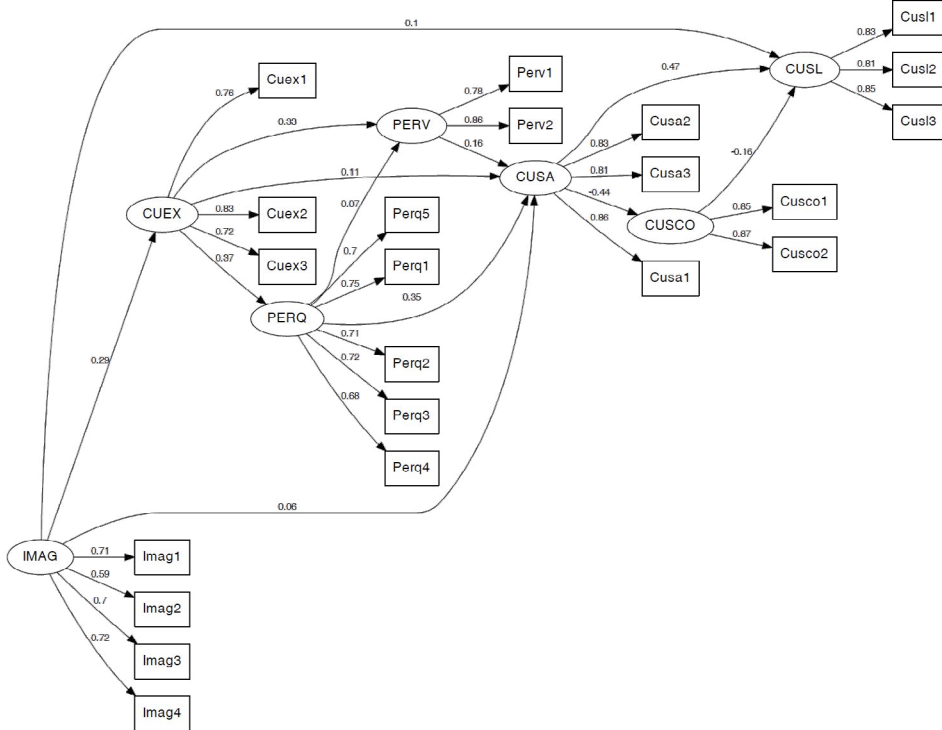
Source: Own work.

The calculation of the coefficient of determination R^2 for each of the latent variable is possible by the `rsquared()` function:

```
> rSquared(ECSIPLS.fit)
  R-squared
IMAG  .
CUEX 0.082
PERQ 0.137
PERV 0.132
CUSA 0.244
CUSCO 0.196
CUSL 0.345
```

There is no general "rule of thumb" for interpreting this measure, but since the R^2 values range is between 0 and 1, the closer the empirical score of R^2 to 1, the better the model.

A more general approach to access coefficient stability is to use a bootstrapping procedure. We can perform it with `bootsempls()` command. A key point in performing the procedure is setting a sufficiently large value for resample sets of observation. In this case, we set 1000:

```
# Bootstrapping
set.seed (5250)
ECSIPLS.boot <-bootsempls(ECSIPLS.fit , nboot =1000, start="ones")
```

After 1000 resample through command summary() we get the following results:[11]

```
> summary(ECSIPLS.boot)
Call: bootsempls(object = ECSIPLS.fit, nboot = 1000, start = "ones")

Lower and upper limits are for the 95 percent perc confidence interval
```

	Estimate	Bias	Std. Error	Lower	Upper
lam_1_1	0.7125	-4.38e-03	0.03923	0.620324	0.779
lam_1_2	0.5876	-2.54e-03	0.05264	0.464342	0.677
lam_1_3	0.7003	-2.02e-03	0.04066	0.613629	0.771
lam_1_4	0.7246	-2.90e-03	0.03804	0.643485	0.789
lam_2_1	0.7570	-5.85e-04	0.02085	0.712676	0.793
lam_2_2	0.8300	-4.54e-04	0.01547	0.797551	0.857
lam_2_3	0.7235	-2.77e-04	0.02533	0.670878	0.768
lam_3_1	0.7504	-8.70e-04	0.01963	0.708437	0.786
lam_3_2	0.7102	-1.42e-03	0.02290	0.659964	0.753
lam_3_3	0.7207	-7.41e-04	0.02166	0.674986	0.761
lam_3_4	0.6786	4.17e-05	0.02510	0.629882	0.723
lam_3_5	0.6991	-5.56e-04	0.02452	0.645462	0.745
lam_4_1	0.7815	-3.36e-03	0.03328	0.701759	0.832
lam_4_2	0.8624	9.87e-04	0.02183	0.817765	0.906
lam_5_1	0.8602	-2.73e-04	0.00911	0.840928	0.876
lam_5_2	0.8269	-1.02e-03	0.01329	0.798288	0.851
lam_5_3	0.8122	-7.55e-04	0.01537	0.780249	0.840
lam_6_1	0.8481	-9.06e-04	0.01594	0.813603	0.876
lam_6_2	0.8690	5.01e-05	0.01334	0.841386	0.894
lam_7_1	0.8345	-7.65e-04	0.01297	0.806461	0.857
lam_7_2	0.8090	-5.58e-04	0.01502	0.776243	0.836

---

[11] Note that the use of bootstrapping procedure with relatively small samples (e.g., < 100) is sometimes problematic and the model could be unstable because bootstrap iterations failed to converge.