EUROSTUDENT V

MANUAL FOR DATA PROCESSING AND DELIVERY

Final draft (21.11.2013)

1. Definition of the Target Groups

Following a survey among administrators, researchers and users of the data and the discussions at the workshops in Berlin (24.-25.5.12), an intensive seminar in Hainburg (25.-27.6.12), various discussions within the consortium and with attached experts and a final seminar in Vienna (15.-17.10.12), we have defined a standard target group to be surveyed by all participating countries and optional groups that might be surveyed. The core report of EUROSTUDENT V will only include data on the standard target group. However, if a minimum number of countries can also provide data on any optional groups, we will consider including special chapters in the Synopsis of indicators on these groups or producing special (online) reports on these groups (e.g. PhD students).

In defining the standard target group we have particularly taken note of previous rounds of EUROSTUDENT and of standard international practice (e.g. by Eurostat).

Standard target group to be covered by all participating countries ("minimum"):

- All students in your country, i.e. national and foreign students who are pursuing their studies for a degree in the country of the survey[1], except students on leave and excluding students on incoming and outgoing credit mobility.
- Full-time and part-time students by status.
- Students in all ISCED 5, 6 and 7 programmes, regardless of their character as general or professional, as long as the programmes are considered higher education in your national context.[2]
- All higher education institutions offering programmes considered "normal". In many cases this means only public, non-specialist institutions of higher education.
- All national degrees corresponding to ISCED 5, 6 and 7 (e.g. BA, MA, traditional diploma, Lizentiat, national degrees in medicine. Short courses only if they are based on ISCED 5.)
- Distance students who study at a "normal" higher education institution, i.e. excluding institutions solely for long distance students like open universities, Fernuniversität Hagen and similar.

Optional groups:

- ISCED 8, Doctoral/PhD-Students
- Higher education institutions not considered for the standard target group (e.g. specialist institutions).

This document provides instructions on data cleaning and processing needed for the focus groups and preparing the dataset for weighting.

---

[1] Widening the focus of EUROSTUDENT towards students who obtained their prior education in another country but the country of the survey is a major change in the definition of the standard target group. We have thereby taken into account the increasing mobility throughout Europe on all educational levels.

[2] EUROSTUDENT V now refers to the new ISCED 2011 classification. For further information see http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx

2. Prepare your data

Dataset

This manual refers to the E:V core questionnaire and can be used as a coding instruction. We advise you to organise your data according to the order of the questionnaire, as follows:

1. **Save a copy of your original (raw) dataset.**

2. **Exclude or delete all the questions/ variables you won't need for data delivery (i.e. additional questions that are not part of the core questionnaire).**

SPSS Syntax:

EXE.
DELETE VARIABLES #CountryspecifiekQuestion_01 #CountryspecifiekQuestion_02 #Etc.
/*Delete all variables you don't need for data delivery.
EXE.

---

3. **Then bring your variables in the same order as they are in the E:V questionnaire and rename them according to the number of the question. In the Annex of this document we have added the structure of the basic EUROSTUDENT V database.**

SPSS Syntax:

RENAME VARIABLES
   #oldvarname1 = v1.1
   #oldvarname2 = v1.2
   #oldvarname3 = v1.3
   Etc.

MATCH FILES / FILE = * /KEEP ID v1.1 v1.2 v1.3 etc. /*Restructure your data.

EXE.

---

4. **Label your variables according to the question text and label the values of each variable according to the corresponding category. Again, for this you can make use of the document with the structure of the basic EUROSTUDENT V database in the Annex of this document.**

In general there are a few preconditions your data need to meet before you can apply the data cleaning rules that are discussed in this paper. Apart from the order and consistency of the data according to the core questionnaire it is also necessary to complete some basic adjustments and corrections, like for example

Define missing values: A variable may contain values that are not valid for analysis. E.g., if your survey was conducted online, the survey client may have defined different

	missing values for filtered cases, missing responses etc. There might also be 0 or "strange" values coded (from typing errors?). In all of these cases, these values must be defined as missings in the respective variables.
Define the correct variable type:	String variables (i.e. variables containing text from open questions) in contrast to numeric variables (i.e. variables containing values that can be labelled) must be defined as such. To transform a string variable into a numeric one, we recommend to use the AUTO RECODE command in SPSS.
Incorporate open answers:	Sometimes respondents use open fields (e.g. *"Other, please specify"*) in multiple answer scales to specify something they already reported in a pre-set category, or simply over read the respective category and report their answer in the open question. Then the case should be coded in the respective category and simultaneously deleted in the category "other".

These are only examples of what is necessary to set up and organise the data properly. If you have any questions on how to prepare your data for processing and analyses, please contact eurostudent@his.de.

Missing Values

Missing answers may either be coded as system missing values or with a value <0. However, it is reasonable to define values also for missing data, to be able to identify them. Non-response missings shall therefore be coded as **-99** and labelled **"No answer"**. Missings due to filtering should be allocated the value **-77** and labelled **"Filtered *[specify filter]*"**, because these respondents did not refuse to answer but were not asked the filtered question(s) at all. Missings due to data cleaning shall be coded **-55** and labelled **"Implausible answer"**. To definitely exclude missing values from the analysis (but keeping them in the dataset at the same time), values to be treated as **missings must be defined for each variable separately!** Next to recoding the strange values to -99, -77 or -55 they have to be set as missing value. If not, they will be incorporated in the analyses. [mis val VARIABLENAME (-99, -77, -55)]. The following instructions will always contain notes on missing treatment.

We strongly advise you to save a copy of your original dataset and to never change an original variable. Always produce a copy of your variables and correct values only in the copy. Thus it is always possible to check for the original responses within the same dataset to monitor the cleaning process and avoid errors.

3. Standard focus groups and valid Cases: "4+2 Criterion"

To define valid cases for the Eurostudent V data analysis there are 4 variables, where no missing values are allowed. That means, all cases must be valid in all of these 4 variables. Therefore we can assure that the variables needed for data weighting contain only valid cases.

Additionally, there must be only valid values and no missing values in at least 2 of the other focus group variables, which are:

A case from the sample is therefore only valid if there is logically consistent information on 4 compulsory focus groups (age, sex, qualification, type of HEI) and at least 2 out of the remaining 6 focus groups.

This definition of valid cases is a basic precondition for data weighting, which is only possible with a cleansed dataset. As the 4 compulsory focus groups are also the preferred variables for the weighting procedure, there must not be any missing values in any of these 4 variables.

The final definition of valid cases is the last step after all the data cleaning is done!

The rules on data cleaning are only the minimum of necessary data cleaning. We advise you to conduct further data cleaning procedures in order to ensure best quality of your data.

4. Data cleaning and definition of focus groups

4 compulsory focus groups: no missings allowed

Note: All variables created for the focus groups of E:V will have the praefix "e_" included in the variable name. Thus all focus group variables will (if alphabetically sorted) be assorted all together in your data set.

Age (Question 5.1)

Age is defined by students' age at the time of the survey and is computed from the respondent's year and month of birth. These two variables must be defined as numeric variables.

Used variables in this cleaning step			
Name	Label		
v5.1year	Year of birth		
v5.1month	Month of birth		
New output variables			
Name	Label	Categories	Source Variables
e_age	Age	1 "up to 21 years" 2 "22 to <25 years" 3 "25 to <30 years" 4 "30 years or over"	v5.1year v5.1month

Data processing

Year of birth as label or as value: If the question on year of birth has been an open question (text box), the values of the variable may represent the years of birth right away. If, in an online survey, the question had a drop-down menu, the years of birth might be represented by values (e.g. 1 for 2001, 2 for 2000 etc.). If so, compute a variable that contains years of birth as values (see the syntax below).

Implausible responses: Check whether the latest year of birth indicates an age that corresponds to the earliest possible entrance into higher education in your country. If there are students younger than the youngest possible age, allocate the value -55 for "implausible answer" in the variable for year of birth. Do also execute this procedure if there are years of birth that lie implausibly far in the past. If year and month are obviously swapped, correct the data.

Missing responses: If there has been no response on the month of birth, replace it with the value representing June (if January is coded 1 and December is coded 12, then June would be 6). If there has been no response on the year of birth, these cases should be defined as missings (e.g. -99 for "no answer"or sysmis in both variables).

1. **Computation of respondent's age:** The respondent's age is computed as the sum of year and month of birth. Therefore execute the following steps:

    1a. Compute the month of birth as a decimal number:[3] $\frac{\text{Month}}{12} - \frac{1}{24}$

    1b. Sum up the year of birth and the months as decimals: $\text{Year} + \frac{\text{Month}}{12} - \frac{1}{24}$

    1c. Then subtract this figure from the year and month your survey was conducted.[4]

    1d. Check again if there are implausibly high or low ages. If there are, code them -55 for "implausible answer".

    1e. Define a range of missing values.

Example: The survey was conducted in May 2013. A respondent was born in June 1989: He/ she would then be $\left(2013 + \frac{5}{12} - \frac{1}{24}\right) - \left(1989 + \frac{6}{12} - \frac{1}{24}\right) = 23,917$ years.

Creating the focus group

The focus group age consists of four categories: **up to 21,99 years, 22,0-24,99 years, 25,0-29,99 years, 30,0 years and over**. Those age limits might look a bit unconventional at the first glance. Nevertheless, as long as a person has not had his or her 22nd birthday, he or she would report the age of 21 years if directly asked. Therefore the category "up to 21 years" must include students up to the age of 21,99 years.

SPSS Syntax:

MISSING VALUES v5.1year v5.1month (lo thru 0). /*Define all values <=0 as missing values.

FREQUENCY v5.1year v5.1month. /*Have a look at the raw data and check for implausible answers. If necessary, correct the data (preferably by using syntax; Copy the variables and only correct the data in the copies - never in the originals!.

RECODE v5.1year (else=copy) into v5.1year_r.
EXE. /*Copy the variable year of birth for further processing.

RECODE v5.1month (else=copy) into v5.1month_r.
EXE. /*Copy the variable month of birth for further processing.

******Alternatively: If year of birth is coded with representative values (e.g."1" for "2001") compute the variable as follows:

---

3 Dividing the month by 12 transforms the month into a decimal number referring to one year. E.g. One month (or January) would then be $\frac{1}{12} = 0,0833$. However, using this formula December would result in $\frac{12}{12} = 1$ and will therefore add 1 year instead of a decimal. Therefore subtract $\frac{1}{24}$, representing half a month, assuming the respondents' birthdays are in the middle of the month.

4 If your fieldphase was longer than 1 month, use the month in which the majority has answered.

```
******COMPUTE v5.1year_r=2002-v5.1year. /*By subtracting the value of the variable, the
new variable will contain the year of birth as a value (2002-1=2001)5.

IF v5.1year>2001 OR v5.1year<1910 v5.1year_r=-55. /*Example for data cleaning. If a
respondent reported a year of birth after 2001 or before 1910 it is considered implausible.
Adapt to national context!.
IF MIS(v5.1month) v5.1month_r=6. /*Replace missing answers on the month of birth with
June (6).
IF MIS(v5.1year) v5.1year_r=-99.
IF MIS(v5.1year) v5.1month_r=-99. /*If information on year is missing both variables are
missing.
EXE.

MISSING VALUES v5.1year_r v5.1month_r (lo thru 0). /* Define all values <=0 as missing
values.

COMPUTE v5.1month_dec=(v5.1month_r/12)-1/24.
EXE. /*Transform the month of birth into a decimal.

COMPUTE v5.1birthyear=v5.1year_r+v5.1month_dec.
EXE. /*Add the decimal month to the year of birth to get the date of birth as decimal number.

COMPUTE v5.1ageyears=#decimal date of survey- v5.1birthyear.
EXE. /*If all missing values are defined correctly, sums will only be calculated if both variables
are valid. All other cases will become system missings.6

MISSING VALUES v5.1ageyears (lo thru 0). /*Define all values <=0 as missing values.

RECODE v5.1ageyears (MISSING=COPY) (lo thru 21.99=1) (22.0 thru 24.99=2) (25 thru 29.99=3)
(30 thru hi=4) into e_age.
EXE. /*Categorise respondent's age into 4 categories to create the focus group age.

MISSING VALUES e_age (lo thru 0). /*Define all values <=0 as missing values.

VARIABLE LABELS e_age 'Age'.
VALUE LABELS e_age 1 'up to 21 years' 2 '22 to <25 years' 3 '25 to <30 years' 4 '30 years or
over'. /*Define Labels for the variable and its values.

EXE.
```

---

[5] The minuend (here: 2002) depends on how the values are allocated to your variable. In a drop-down menu the first category (code 1) may be "please select" and the first year is then coded "2". According to our example the equation would then be 2003-v5.1year=2001.

[6] Assuming the survey was conducted in May 2013: Decimal date of survey = $2013 + \frac{5}{12} - \frac{1}{24}$. If your survey was active for more than one month (e.g. May and June), choose the month in which the majority of respondents answered.

Sex (Question 5.2)

This question asks for the sex (= biological make-up of an individual's reproductive anatomy), not the gender (= an individual's lifestyle, often culturally learned).

Used variables in this cleaning step			
Name	Label		
v5.2	What is your sex?		
New output variables			
Name	Label	Categories	Source Variables
e_sex	Sex	1 "Female" 2 "Male"	v5.2

SPSS Syntax:

RECODE v5.2 (ELSE=COPY) into e_sex.
EXE. /*Copy the variable for sex to create the focus group.

MISSING VALUES e_sex (lo thru 0). /*Define all values <=0 as missing values.

VARIABLE LABELS e_sex 'Sex'.
VALUE LABELS e_sex 1'Female' 2 'Male'. /*Labelling according to E:V questionnaire. Check for original values and labels in your data!

Qualification being studied for (Question 1.1)

All students studying at ISCED (2011) 5 (short cycle), 6 (Bachelor) and 7 (Master) shall be coded as valid cases. PhD students (ISCED 8) are considered an optional group and hence must be excluded from the E:V data set. For ISCED 2011 classification see:
http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf

Used variables in this cleaning step			
Name	Label		
v1.1	Which study programme are you currently enrolled in?		
New output variables			
Name	Label	Categories	Source Variables
e_qualification	Qualification studied for	1 "Bachelor" 2 "Master" 3 "other" -77 "PHD"	v1.1

> **SPSS Syntax:**
>
> RECODE v1.1 (2=1) (3=2) (1,4,5,6=3) (7=-77) (ELSE=COPY) into e_qualification.
> EXE. /*Coding according to E:V questionnaire. Check for original values and labels in your data!
>
> MISSING VALUES e_qualification (lo thru 0). /*Define all values <=0 as missing values.
>
> VARIABLE LABELS e_qualification 'Qualification studied for'.
> VALUE LABELS e_qualification 1 'Bachelor' 2 'Master' 3 'other' -77 'PhD'. /*Define Labels for the variable and its values.

Type of HEI (Question 1.0)

This variable separates students studying at universities from students studying at other HEIs. Students studying at any of these HEI in the respective country shall be included in the sample.

Used variables in this cleaning step			
Name	Label		
v1.0	Do you actively pursue your studies in the current semester?		
New output variables			
Name	Label	Categories	Source Variables
e_ hei	Type of HEI	1 "University" 2 "Other HEI" -77 "not in E:V target group"	v1.0

Data processing and focus group

Copy category 1 (university) and collate categories 2 and 3 (university of applied sciences and other HEI). Categories 4-7 define persons that must not be included in the standard target group of E:V. Therefore recode these categories into -77 for "Filtered: not in the target group of E:V".

Alternative if only 1 type of HEI is given:
Compute a dichotomous variable and allocate the value 1 to all cases in your dataset.

> **SPSS Syntax:**
>
> RECODE v1.0 (1=1) (2,3=2) (ELSE=-77) into e_hei.
> EXE. /*Coding according to E:V questionnaire. Check for original values and labels in your data!
>
> MISSING VALUES e_hei (lo thru 0). /*Define all values <=0 as missing values.

VARIABLE LABELS e_hei 'Type of HEI'.
VALUE LABELS e_hei 1 'University' 2 'Other HEI' -77 'not in E:V target group'. /*Define Labels for the variable and its values.

******Alternative if only 1 type of HEI exists.

*COMPUTE e_hei=1.
* VARIABLE LABELS e_hei 'Type of HEI'.
*VALUE LABELS e_hei 1 'University'. /*Define Labels for the variable and its values.

2 out of 6 focus groups

Study intensity (Questions 3.9, 3.13, 3.14)

The purpose of this variable is to categorise students according to the time they spend on study-related activities. In order to check the plausibility of the data on time budget, information on time spent on paid jobs must as well be considered.

Used variables in this cleaning step			
Name	Label		
v3.9	Do you have a paid job during the current semester?		
v3.13	How many hours do you spend on paid jobs in a typical week during this semester?		
v3.14	How many hours do you spend in a typical week in taught courses and on personal study time?		
New output variables			
Name	Label	Categories	Source Variables
e_intens	Study intensity	1 "Low intensity" 2 "Medium intensity" 3 "High intensity"	v3.14

Data processing

2. Non-response:

2a. If all fields on study related activities (i.e. all days of the week with hours spent on taught studies or personal study time) are empty or filled with 0, then allocate the value -99 for "no answer" in all respective variables (i.e. all days of the week with hours spent on taught studies or personal study time).

2b. If at least one of the fields is filled with a value above 0, then allocate 0 to all empty fields.

3. Implausible answers:

3a. If a student has responded that he/she does not work (question 3.9=3 or missing), and the value for "paid jobs" in question 3.13 is not 0, set it to 0.[7]

3b. If a student has responded that he/she works "during the whole semester" (question 3.9=1) and the field for "paid jobs" in question 3.13 is empty or 0, then allocate the value -99 for "no answer" in 3.13.[8]

3c. Compute the total workloads per day for study related activities. If total hours per day exceed 24 hours, allocate the value -55 for "implausible answer" in all variables in 3.14.[9]

3d. Compute the total weekly workload for study related activities and paid employment.[10] If the weekly workload of study related activities and paid employment exceeds 120 hours, allocate the value -55 for "implausible answer" in all variables in 3.13 and 3.14.[11]

Include non-working students: For all students that reported to be not working during this semester (3.9=3 or 3.9=missing) allocate the value 0 in 3.13.7

Please quantify the number of all excluded cases and the number of all cases where empty fields were replaced with 0 in the metadata and the respective subtopic's comment box.

4. Computing weekly sums for each activity:

4a. Sum up daily hours spent on taught studies to compute the **weekly workload for taught studies**.

---

[7] All non-working students shall be represented by the value "0" in the variable on time spent on paid job, in order to calculate the average time spent on paid employment for all students (either employed or unemployed). If you want to calculate the average time spent on paid job of working students only, define 0 occasionally as missing value in the respective analysis. For the focus groups however, 0 should not be defined as missing.

[8] As the given information is not consistent, the "easier" question should be trusted. Answering a yes/no question is easier (i.e. there is less reflecting necessary to find the answer) for the respondent than reporting the amount of hours worked per week. For students who work occasionally during the semester this rule does not apply, because for these students it is possible to work 0 hours in a <u>typical week</u> of the semester.

[9] The total of hours spent for study-related activities per day is allowed to reach 24 hours, because there are subjects where compulsory courses or placements include e.g. shift work (e.g. for students of medicine). Therefore the total of hours spent on study related activities can reach 24 hours per day.

[10] Sum up all the daily sums calculated for (2c), i.e. sum of 7 daily workloads for study related activities and paid employment (already reported as weekly amount of hours).

[11] Though a total workload of 24h per day is allowed, the total workload per week must not exceed 120h, as we assume that in a <u>typical week</u> there must be some time for sleeping, eating and other activities to satisfy basic needs.

4b. Sum up daily hours spent on personal study time to compute the **weekly workload for personal study time**.

4c. Sum up the weekly sums for taught studies and personal study time to compute the **weekly workload for study related activities**.

4d. Sum up the weekly workload for study related activities and for paid employment to compute the **total weekly workload**.

Creating the focus group

Collate the variable on total **weekly workload for study related activities** into **students with low** (i.e. 0-20,0 hours), **medium** (20,1-40,0 hours) and **high study intensity** (40,1 or more hours).

SPSS Syntax:

MISSING VALUES v3.14.1mo v3.14.1tu v3.14.1we v3.14.1th v3.14.1fr v3.14.1sa v3.14.1su v3.14.2mo v3.14.2tu v3.14.2we v3.14.2th v3.14.2fr v3.14.2sa v3.14.2su (lo thru -1). /*Define all values <=-1 missing in the time budget. 0 shall remain valid. [12]

COUNT count_time=v3.14.1mo v3.14.1tu v3.14.1we v3.14.1th v3.14.1fr v3.14.1sa v3.14.1su v3.14.2mo v3.14.2tu v3.14.2we v3.14.2th v3.14.2fr v3.14.2sa v3.14.2su (0, MISSING).
/*Counts how many missings or 0 there are in all 14 variables.

/*Define a variable list for the originals and the copies to perform a loop[13].
/*Copy the original variables for further processing.
/*If all variables are missing or 0, all variables on study related activities are allocated the value -99.
/*If not all variables are missing but one single variable is missing, it shall be 0.
/* Define all values <=-1 as missing values. 0 shall remain valid.
/*Stop loop.
DO REPEAT var= v3.14.1mo v3.14.1tu v3.14.1we v3.14.1th v3.14.1fr v3.14.1sa v3.14.1su v3.14.2mo v3.14.2tu v3.14.2we v3.14.2th v3.14.2fr v3.14.2sa v3.14.2su
/newvar= v3.14.1mo_r v3.14.1tu_r v3.14.1we_r v3.14.1th_r v3.14.1fr_r v3.14.1sa_r v3.14.1su_r v3.14.2mo_r v3.14.2tu_r v3.14.2we_r v3.14.2th_r v3.14.2fr_r v3.14.2sa_r v3.14.2su_r.
COMPUTE newvar=var.
IF count_time=14 newvar=-99.
IF count_time<14 AND MIS(var) newvar=0.
MISSING VALUES var newvar (lo thru -1).

---

[12] Taught studies have an added „1" to the variable name (e.g. v3.14.**1**mo), personal study time has an added "2" to the variable name (e.g. v3.14.**2**mo).

[13] Using a "DO REPEAT" loop in SPSS can often reduce the amount of commands that must be written, and may help reduce the likelihood of error. It is possible to create or modify many variables with fewer lines of syntax.

```
END REPEAT.
EXE.

RECODE v3.13 (ELSE=COPY) into v3.13_r. /*Copy the variable for further processing.
IF v3.9=3 OR MIS(v3.9) v3.13_r=0. /*If a student does not work or at least did not report that
he/she works, the workload for paid employment shall be 0.
IF v3.9=1 AND (v3.13=0 OR MIS(v3.13)) v3.13_r =-99. /*If a student works during the whole
semester but did not report any working hours the case is allocated the value-99 in the field
"time spent on paid employment".
EXE.
MISSING VALUES v3.13_r (lo thru -1). /*Define missing cases for time spent on paid
employment.

COMPUTE v3.14mo= v3.14.1mo_r+ v3.14.2mo_r.
COMPUTE v3.14tu= v3.14.1tu_r+ v3.14.2tu_r.
COMPUTE v3.14we= v3.14.1we_r+ v3.14.2we_r.
COMPUTE v3.14th= v3.14.1th_r+ v3.14.2th_r.
COMPUTE v3.14fr= v3.14.1fr_r+ v3.14.2fr_r.
COMPUTE v3.14sa= v3.14.1sa_r+ v3.14.2sa_r.
COMPUTE v3.14su= v3.14.1su_r+ v3.14.2su_r.
EXE. /*Compute daily sums for study related activities.

/*If the daily workload exceeds 24 hours per day allocate the value -55 to all variables related
to study time.
/*Define missing cases for all variables in the list.
DO REPEAT varlist= v3.14.1mo_r v3.14.1tu_r v3.14.1we_r v3.14.1th_r v3.14.1fr_r v3.14.1sa_r
v3.14.1su_r v3.14.2mo_r v3.14.2tu_r v3.14.2we_r v3.14.2th_r v3.14.2fr_r v3.14.2sa_r
v3.14.2su_r.
IF v3.14mo>24 OR v3.14tu>24 OR v3.14we>24 v3.14th>24 OR v3.14fr>24 OR v3.14sa>24 OR
v3.14su>24 varlist=-55.
MISSING VALUES varlist (lo thru -1).
END REPEAT.
EXE.

COMPUTE
v3.14week_test=v3.13_r+v3.14mo+v3.14tu+v3.14we+v3.14th+3.14fr+v3.14sa+v3.14su./*Com
pute the weekly sum of daily workloads on study related activity plus paid employment. If the
missing values are defined properly, the command will only be executed for cases where all
variables are valid. A missing in at least one variable will produce a system missing in
v3.14week_test.
EXE.

/*If the weekly sum of hours spent on study related activities and paid employment exceeds
120h, allocate the value -55 for "implausible answer" to all time related variables (i.e. study
related activities and paid employment).
```

```
/*Define missing cases for all variables in the list.
DO REPEAT varlist= v3.13_r v3.14.1mo_r v3.14.1tu_r v3.14.1we_r v3.14.1th_r v3.14.1fr_r
v3.14.1sa_r v3.14.1su_r v3.14.2mo_r v3.14.2tu_r v3.14.2we_r v3.14.2th_r v3.14.2fr_r
v3.14.2sa_r v3.14.2su_r.
IF v3.14week_test>120 varlist=-55.
MISSING VALUES varlist (lo thru -1).
END REPEAT.
EXE.


***************************************.
*Computing weekly sums for each activity.

COMPUTE v3.14.1week=v3.14.1mo_r+v3.14.1tu_r+v3.14.1we_r+v3.14.1th_r+v3.14.1fr_r+
v3.14.1sa_r+v3.14.1su_r. /*Add up the daily workload for taught studies to a weekly sum. If
the missing values are defined properly, the command will only be executed for cases where
all variables are valid. A missing in at least one variable will produce a system missing
inv3.14.1week.
COMPUTE v3.14.2week= v3.14.2mo_r+v3.14.2tu_r+v3.14.2we_r+v3.14.2th_r+v3.14.2fr_r+
v3.14.2sa_r+v3.14.2su_r. /*Add up the daily workload for personal study time to a weekly
sum. If the missing values are defined properly, the command will only be executed for cases
where all variables are valid. A missing in at least one variable will produce a system missing
inv3.14.2week.
EXE.
COMPUTE v3.14week=v3.14.1week+v3.14.2week. /*Add up the weekly workload for taught
studies and personal study time to calculate the total study related workload. If the missing
values are defined properly, the command will only be executed for cases where all variables
are valid. A missing in at least one variable will produce a system missing in v3.14week.
COMPUTE v3.13.14week= v3.13_r+v3.14.1week+v3.14.2week. /*Add up the weekly workload
for taught studies, personal study time and paid employment to calculate the total workload.
If the missing values are defined properly, the command will only be executed for cases where
all variables are valid. A missing in at least one variable will produce a system missing
inv3.13.14week.
EXE.

RECODE v3.14week (lo thru 20=1) (20.1 thru 40=2) (40.1 thru hi=3) into e_intens.
EXE. /*Categorise the weekly workload for study related activities to create the focus group
variable.
VARIABLE LABELS e_intens 'Study intensity'.
VALUE LABELS e_intens 1 'Low intensity' 2 'Medium intensity' 3 'High intensity'.

DELETE VARIABLES count_time v3.14 week_test. /*Delete variables only used temporarily for
data processing, unnecessary for data delivery.
```

Field of study (Question 1.5)

This focus group aims at comparing the fields of study, using international standard classifications (ISCED 2011) to ensure comparability across all countries. For ISCED 2011 classification see: http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

Used variables in this cleaning step			
Name	Label		
v1.5	What is the subject you study?		
New output variables			
Name	Label	Categories	Source Variables
e_field	Field of study	0 "General Programmes" 1 "Education" 2 "Humanities and arts" 3 "Social sciences, business and law" 4 "Science" 5 "Engineering, manufacturing and construction" 6 "Agriculture" 7 "Health and welfare" 8 "Services" 9 "Not known/ unspecified"	v1.5

Data processing

Depending on the type of question you used for collecting information on the field of study (open question, drop-down menu or other), different coding procedures will be needed. Therefore it is not possible to provide a general syntax for this. In the end, every subject must be allocated to the proposed ISCED 2011 classification. The focus group variable should be named *e_field* and consist of the 10 categories listed above according to ISCED 2011.

Educational attainment of parents (Question 6.1)

This question aims at providing data on the social make-up of the student body. This indicator is used to measure the inclusiveness of higher education systems and their capacity to deal with underprivileged social groups.

Used variables in this cleaning step	
Name	Label
v6.1father	What is the highest level of education your father has obtained?
v6.1mother	What is the highest level of education your mother has obtained?

New output variables			
Name	Label	Categories	Source Variables
e_edupar	Highest educational attainment of parents	1 "no higher education (ISCED 0-4)" 2 "higher education (ISCED 5-8)" 3 "Don't know"	v6.1father v6.1mother

Data processing

The educational attainment is asked separately for mother and father. For computing this indicator use the highest educational level attained of either the mother or the father. If the information on one parent is missing, use the information given for the other parent. If mother or father have the same educational level take any of them. Each of the national educational qualifications must be allocated to the corresponding ISCED 2011 category.

Creating the focus group

The focus group compares three groups of students: those with parents who attained **no higher education** (secondary, non-tertiary education, ISCED 0-4) or **higher education** (tertiary education, ISCED 5-8). For ISCED 2011 classification see:
http://www.uis.unesco.org/Education/Documents/isced-2011-en.pdf.

SPSS Syntax:

MISSING VALUES v6.1father, v6.1mother (lo thru 0). /*Define all values <=0 as missing values.

COMPUTE v6.1parents =MAX(v6.1father, v6.1mother). /*Allocates the maximum value of the two variables to the new variable. If information on one parent is missing, the information on the other parent is automatically adopted. If information on both parents is missing, system missing will be allocated. If the values of mother and father are equal, the given value is adopted[14].

IF (v6.1father=8 OR MIS(v6.1father)) AND v6.1mother>0 v6.1parents =v6.1mother. /*If the educational attainment of the father is unknown and there is valid information about the mother, the value of the mother shall be adopted.

IF (v6.1mother=8 OR MIS(v6.1mother) AND v6.1father>0 v6.1parents =v6.1father. /*If the educational attainment of the mother is unknown and there is valid information about the father, the value of the father shall be adopted.

EXE.

---

[14] CAUTION! This procedure includes also the value 8 for "don't know". If the highest educational attainment for one parent is unknown, but the other was reported, v6.1parent will contain "don't know". Therefore the following steps are crucial!

******Now label this variable using the same labels as for each parent separately. This variable will be used for subtopic "D Social Background"

RECODE v6.1parents (1,2,3=1) (4 thru 7=2) (8=3) into e_edupar.
EXE. /*Categorise the highest educational attainment of parents to create the focus group.

MISSING VALUES e_edupar (3).
VARIABLE LABELS e_ edupar 'Highest educational attainment of parents'.
VALUE LABELS e_edupar 1 'no higher education (ISCED 0-4)' 2 'higher education (ISCED 5-8)' 3 'Don't know'.

International students (Question 2.1)

For educational statistics, the educational origin of students is of more interest than their nationality. Thus, international students are defined as students who obtained their prior education outside of the country of survey (2.1=6).

Used variables in this cleaning step			
Name	Label		
v2.1	What was the highest level of education you obtained on graduating from the school system for the first time?		
v2.2	Where did you graduate from the educational level mentioned above?		
New output variables			
Name	Label	Categories	Source Variables
e_international	International Students	0 "Domestic student" 1 "International student"	v2.1

SPSS Syntax:

COMPUTE e_international=0.
IF v2.1=6 e_international=1.

IF v2.2country <> #country e_international=1. /*Countercheck with v2.2. If v2.2country is not the country of survey, the student shall be recorded as international student.

IF MIS(v2.1) e_international=-99. /*If the value in 2.1. is missing, allocate -99 for "no answer".
MISSING VALUES e_international (lo thru -1).
VARIABLE LABELS e_international 'International Students'.
VALUE LABELS e_international 1 'international student' 0 'domestic student'.

Direct and delayed transition students (Questions 2.1, 2.3 - 2.7)

This indicator combines information on former education and higher education entrance qualification. Direct transition students are defined as those students who entered tertiary education for the first time within two years after graduating from school for the first time

(usually after upper secondary education). Delayed transition students are defined as students, who entered higher education for the first time more than two years after leaving the school system for the first time[15].

This focus group is only calculated for students with national school leaving certificate/ higher education qualification! Students who graduated from a foreign school system must be excluded for this part of the analyses!

Used variables in this cleaning step			
Name	**Label**		
v2.1	What was the highest level of education you obtained on graduating from the school system for the first time?		
v2.3	When did you obtain the qualification mentioned in 2.1?		
v2.4	What qualifications, examinations or measures qualified you for entry into higher education?		
v2.5	When did you obtain/complete the most recent qualification, examination or measure, which qualified you for entry into higher education?		
v2.6	When did you enter higher education for the first time?		
v2.7	When did you start your current study programme?		
New output variables			
Name	**Label**	**Categories**	**Source Variables**
e_transit	Direct and delayed transition	1 "Direct" 2 "Delayed" -77 "Foreign qualification"	v2.3 v2.6

Data processing

Reported year as label or as value: If the questions on the dates of school leaving, HE entrance qualification, HE entrance and start of current study has been an open question (text box), the values of the variable may represent the years right away. If, in an online survey, the questions had a drop-down menu, the years might be represented by values (e.g. 1 for 2013, 2 for 2012 etc.). If so, compute a variable that contains years as values (see instructions on Age p. 7 et seqq.).

Missing responses: If there has been no response on the month, replace it with the value of the month, the majority reported. If there has been no response on the year, these cases should be defined as missings (allocate -99 for "no answer" or system missing) in the respective variable.

---

[15] In E:IV the definition of direct and delayed transition students also referred to the type of HE qualification. As in E:V we have better information on the chronology of the education career, it is sufficient to refer to the date of the first school leave (i.e. graduation) and the first entry into HE, without losing information according to the old definition.

Computation of dates: Transform questions 2.3, 2.5, 2.6 and 2.7 into decimal numbers (see also instructions on Age on p. 7 et seqq.). If the vast majority (>= 95%) has reported the same month(s) within one question (e.g. 95% of the students entered HE for the first time in September/ October) it is possible to only refer to the year and ignore the month completely.

1. **Implausible responses:**

   1a. Check whether the dates are in a consistent chronological order: graduating from school for the first time (2.3) <= attaining higher education entrance qualification (2.5) <= entering higher education for the first time (2.6) <= start of current study programme. If there are cases with an inconsistent order of dates, allocate the value -55 for "implausible answer" to all variables (2.3, 2.5, 2.6, 2.7). Do also execute this procedure if there are years that lie implausibly far in the past and for future dates.

   1b. Check also whether the year of birth, resp. the age of a person is consistent with the reported dates in 2.3, 2.5, 2.6, 2.7. If not, allocate the value -55 for "implausible answer" to all dates referring to the education career (2.3, 2.5, 2.6, 2.7).

   1c. If year and month are obviously swapped, correct the data.

CAUTION! There may be possibilities to obtain the school leaving certificate/ entrance qualification <u>after</u> entering higher education for the first time (e.g. programmes in arts/ military service). If so, please define a reasonable range within which an "inverse" order of dates is possible (e.g. obtaining entrance qualification within 4 months <u>after</u> entering HE for the first time). Cases that are covered by this range shall not be defined as missing cases.

Creating the focus group

Delayed transition students: Students that *for the first* time entered higher education more than two years after graduating from school *for the first time* (2.6 *minus* 2.3 *is more than* 2 years).

Direct transition students: Students that *for the first* time entered higher education within two years after graduating from school *for the first time* (2.6 *minus* 2.3 *is less or equal to* 2 years).

SPSS Syntax:

MISSING VALUES v2.3year v2.3month v2.5year v2.5month v2.6year v2.6month v2.7year v2.7month (lo thru 0). /*Define all values <=0 as missing values.

FREQUENCY v2.3year v2.3month v2.5year v2.5month v2.6year v2.6month v2.7year v2.7month. /*Have a look at the raw data and check for implausible responses.

RECODE v2.3year v2.3month v2.5year v2.5month v2.6year v2.6month v2.7year v2.7month (else=copy) into v2.3year_r v2.3month_r v2.5year_r v2.5month_r v2.6year_r v2.6month_r v2.7year_r v2.7month_r.
EXE. /*Copy the variables month for further processing.

******Alternatively: If year is coded with representative values (e.g."1" for "2001") compute the variable as follows:

******COMPUTE v2.6year _r=2014-v2.6year. /*By subtracting the value of the variable, the new variable will contain the year as a value (2002-1=2001)[16]. Apply for 2.3, 2.5 and 2.7.

/*Define variable lists for months and years.
/*Replace missing answers with the month, the majority reported. Assumtion: June (6). Check with national data!
/*If information on year is missing both variables shall be missing (-99 for "no answer").
/*Define all values <=0 as missing values.

```
DO REPEAT sourcevarM= v2.3month v2.5month v2.6month v2.7month
/targetvarM= v2.3month_r v2.5month_r v2.6month_r v2.7month_r
sourcevarY= v2.3year v2.5year v2.6year v2.7year
targetvarY= v2.3year_r v2.5year_r v2.6year_r v2.7year_r.
IF MIS(sourcevarM) targetvarM=6.
IF MIS(sourcevarY) targetvarY=-99.
IF MIS(sourcevarY) targetvarM=-99.
MISSING VALUES targetvarM targetvarY (lo thru 0).
END REPEAT.
EXE.
```

```
COMPUTE v2.3month_dec=( v2.3month_r /12)-1/24.
COMPUTE v2.5month_dec=( v2.5month_r /12)-1/24.
COMPUTE v2.6month_dec=(v2.6month_r /12)-1/24.
COMPUTE v2.7month_dec=(v2.7month_r /12)-1/24.
```
EXE. /*Transform the month into a decimal. See also coding instructions on age, p. 7 et seqq.

```
COMPUTE v2.3date=v2.3year_r + v2.3month_dec.
COMPUTE v2.5date=v2.5year_r + v2.5month_dec.
COMPUTE v2.6date=v2.6year_r + v2.6month_dec.
COMPUTE v2.7date=v2.7year_r + v2.7month_dec.
```
EXE. /*If all missing values are defined correctly, sums will only be calculated if both variables are valid. All other cases will become system missings.

```
RECODE v2.3date v2.5date v2.6date v2.7date (ELSE=COPY) into v2.3date_r v2.5date_r
v2.6date_r v2.7date_r.
```
/*Copy the date variables for further cleaning and processing.
```
EXE.
```

```
COMPUTE date_valid=1.
```
/*Compute a test variable to mark cases with inconsistent answers.
```
IF v2.3date <= v2.5date <= v2.6date <= v2.7date date_valid=0.
```
/*If the dates are in a non-consistant order, mark the cases as not valid by allocating the value 0.

---

[16] The minuend (here: 2014) depends on how the values are allocated to your variable. In a drop-down menu the first category (code 1) may be "please select" and the first year is then coded "2". According to our example the equation would then be 2015-2.6year =2013.

```
IF v2.6date - v2.5date>= #-0,29 date_valid=1. /*If the HE entrance qualification was not
obtained more than 4 months after entering HE for the first time, the case shall be valid again.
Define national rules and if applicable, apply the command to other dates! 17.

IF (v2.6date- v5.1birthyear)<17 date_valid=0. /*If a respondent's age at the first entry to
higher education is lower than 17, mark the cases as not valid by allocating the value 0. Define
national rules and if applicable, apply the command to other dates!.

/*For all cases defined non-valid in the test-variable, allocate -55 for "implausible answer" to
all dates related to the education career.
/*Define all values <=0 as missing values.
DO REPEAT varlist= v2.3date_r v2.5date_r v2.6date_r v2.7date_r.
IF date_valid=0 varlist=-55. MISSING VALUES varlist (lo thru 0).
END REPEAT.
EXE.

COMPUTE gap=v2.6date_r- v2.3date_r. /*Calculate the gap between first school leave and
first HE entry.
EXE.

COMPUTE e_transit=0.
IF v2.4.1=1 OR gap <= 2 e_transit=1. /*If entrance qualification is upper secondary
qualification (ISCED 34/35) and the gap between first school leave and first HE entry is lower
or equal to 2 years, the case is defined as direct transition student.
IF gap>2 e_transit=2. /*If the gap between first school leave and first HE entry is greater than
2 years, the case is defined as delayed transition student.
IF v2.1=6 e_transit=-77. /*Exclude foreign students by allocating -77 for "Filtered" and define
them as missings.
EXE.

MISSING VALUES e_transit (lo thru 0). /*Define all values <=0 as missing values.
VARIABLE LABELS e_transit 'Direct and delayed transition'.
VALUE LABELS e_transit 1 'direct' 2 'delayed' -77 'foreign qualification'.

EXE.
```

Dependency on income source (Questions 3.6, 3.7, 3.9)

Students are dependent on an income source, if the share of the respective source of the total
income amounts to more than 50% of their income.

---

[17] Assuming the national context allows to enter higher education before gaining the HE entrance
qualification (national researchers must define the possible overlap!); e.g. an overlap of one
Semester (=4 months) is possible: $0,29 = \frac{4}{12} - \frac{1}{24}$; i.e. 4 months as decimal; see also instructions on
age p.8 et seqq.

Definitions:

Total cash income: Sum of all income sources (3.6).

Transfers in kind: Expenses directly paid by parents/ partner/ others (i.e. no money transfer to the student!).

Total income: Sum of all income sources in cash (3.6) plus transfers in kind (i.e. expenses paid by parents/ partner/ others; 3.7).[18]

In general, it is hardly possible to give a sufficient instruction on data cleaning of financial variables. The following instructions aim at guiding you through the data cleaning procedure suggesting cleaning procedures. It is nonetheless inevitable to extend these instructions and conduct further data cleaning steps according to the national circumstances and the answers responded.

Used variables in this cleaning step			
Name	Label		
v3.6	What is the average monthly amount at your disposal from the following sources during the current semester?		
v3.7	What are your average expenses for the following items during the current semester?		
v3.9	Do you have a paid job during the current semester?		
New output variables			
Name	Label	Categories	Source Variables
e_depend	Dependency on source of income	1 'Dependent on family' 2 'Dependent on self-earned income' 3 'Dependent on public student support' 4 'Other'	v3.6 v3.7

Data processing

1. **Non-response:**

   1a. If all fields in question 3.6 (i.e. every single income source listed) are empty or filled with 0, then allocate the value -99 for "no answer" in all income variables in 3.6. If at least one of the fields is filled with a value above 0, then allocate 0 to all empty fields.

---

[18] Transfers in kind must be included in total income (resp. to cash income received from family/partner). Not adding them to the student's income would lead to an underestimation of the total amount of money utilized for sustenance. Nevertheless, transfers in kind must also be included into students' total expenses, because excluding them would lead to an underestimation of the total monthly cost of sustenance.

1b. If any fields in question 3.7.2.1a-i or 3.7.2.2a-d (transfers in kind) are empty, fill them up with 0.[19]

2. Implausible answers:

2a. Some respondents may have reported obviously wrong figures, such as 666, 88/ 888 or similar in more than one field. Track down these cases and allocate -55 in all income and expense variables in 3.6, 3.7. You may also think about excluding those cases totally from the dataset.[20]

2b. If a student has reported that he/she works during the whole semester (3.9=1), but no income is reported in the field "self-earned income from current job", allocate the value -55 for "implausible answer" to the variable "self-earned income from current job".[21]

2c. For some income sources you may be able to define ceiling amounts. The maximum amount of public support for instance, is legally defined in most countries. In those fields you may correct the numbers given. e.g. if the students' information obviously refers to a semester whilst you need monthly amounts, divide the amount reported by the respective number of months.

Include non-working students: For all students that reported to be not working during this semester (3.9=3 or 3.9=missing) allocate the value 0 in 3.6.[22]

Study related costs per semester: Study related costs were asked per semester, whereas living costs were asked per month. Therefore convert the study related costs paid by parents/ partner/ others (i.e. transfers in kind) in 3.7.2 into monthly amounts, by dividing them by 6.[23]

---

[19] It is plausible, that a student does not receive any transfers in kind at all, whereas 0€ expenses or 0€ income is not plausible.

[20] In this case, the answers are obviously faked. Therefore you should not trust the respondent's answers on any variable related to financing. You may also conclude not to trust this case at all and define it as non-valid case (see section on "final definition on valid cases").

[21] As the given information is not consistent, the "easier" question should be trusted. Answering a yes/no question is easier (i.e. there is less reflecting necessary to find the answer) for the respondent than reporting the amount of hours worked by week. For students who work occasionally during the semester this rule does not apply, because for these students it is possible to be working 0 hours in a <u>typical week</u> of the semester.

[22] All non-working students shall be represented by the value "0" in the variable on income from paid job, in order to calculate the average income for all students (either employed or unemployed). If you want to calculate the average income from paid job of working students only, define 0 occasionally as missing value in the respective analysis. For the focus groups however, 0 should not be defined as missing.

[23] The division by 6 refers to 6 months per Semester. If, in the national context, the answers refer to other periods, adapt the divisor.

3. **Adapting the distribution:** The income distribution of your data might contain outliers that determine the mean income.

    3a. Look at the income distribution of all sources of income and whether the lowest and highest value seems plausible. If you detect outliers (implausibly high or low values), correct the data if possible. E.g.: A common mistyping is an additional 0 (e.g. 10.000€ instead of 1.000€) or income is obviously given in months instead of semesters or vice versa.[24]

    3b. Sum up cash income from family/partner (3.6) plus all transfers in kind (3.7.2) and repeat the procedure described in (3a) for the new sum variable.

    3c. Sum up all cash income sources (3.6, excluding transfers in kind!) and repeat the procedure described in (3a) for the new sum variable.

    3d. Sum up all cash income sources (3.6) plus all transfers in kind (3.7.2) and cut off the distribution at a reasonable point (e.g. where the number of respondents cumulate) between 0.25% and 2% from each end of the distribution.

Please quantify the number of all excluded cases and the number of all cases where empty fields were replaced with 0 in the metadata and the respective subtopic comment box.

> **For data cleaning of expenses, apply the procedure instructed for income.**

Creating the focus group

Calculate the share of income from family support (incl. transfers in kind), public sources and self-earned income (incl. savings from previous income) of the total income (incl. transfers in kind). Define the students according to the share that accounts for more than 50% of the total income (incl. transfers in kind) as either **dependent on family support**, **dependent on self-earned income** or **dependent on public student support**. If the 50%-criterion is not met define the case as **other**.

> **SPSS Syntax:**
>
> MISSING VALUES v3.6.1 v3.6.2a v3.6.2b v3.6.3a v3.6.3b v3.6.4a v3.6.4b v3.6.5a v3.6.5b (lo thru -1). /*Define all values <=-1 missing.
>
> FREQUENCIES v3.6.1 v3.6.2a v3.6.2b v3.6.3a v3.6.3b v3.6.4a v3.6.4b v3.6.5a v3.6.5b. /*Look at the income distribution of each variable and search for outliers, mistyping etc.
>
> COUNT count_income= v3.6.1 v3.6.2a v3.6.2b v3.6.3a v3.6.3b v3.6.4a v3.6.4b v3.6.5a v3.6.5b (0, MISSING). /*Counts how many missings or 0 there are in all income variables.

---

[24] If you are not able to correct the case, allocate -55 for "implausible answer" for all those cases.

```
/*Define a variable list for the originals and the copies to perform a loop.
/*Copy the original variables for further processing.
/*If all variables are missing or 0, all variables on income are allocated the value -99.
/*If not all variables are missing but one single variable is missing, it shall be 0.
/*Define all values <=-1 as missing values. 0 shall remain valid.
/*Stop loop.
DO REPEAT var= v3.6.1 v3.6.2a v3.6.2b v3.6.3a v3.6.3b v3.6.4a v3.6.4b v3.6.5a v3.6.5b
/newvar= v3.6.1_r v3.6.2a_r v3.6.2b_r v3.6.3a_r v3.6.3b_r v3.6.4a_r v3.6.4b_r v3.6.5a_r
v3.6.5b_r.
COMPUTE newvar=var.
IF count_income=9 newvar=-99.
IF count_income<9 AND MIS(var) newvar=0.
MISSING VALUES var varlist (lo thru -1).
END REPEAT.
EXE.

MISSING VALUES v3.7.2.1a v3.7.2.1b v3.7.2.1c v3.7.2.1d v3.7.2.1e v3.7.2.1f v3.7.2.1g v3.7.2.1h
v3.7.2.1i v3.7.2.2a v3.7.2.2b v3.7.2.2c v3.7.2.2d (lo thru -1) /*Define all values <=-1 missing[25].

FREQUENCIES v3.7.2.1a v3.7.2.1b v3.7.2.1c v3.7.2.1d v3.7.2.1e v3.7.2.1f v3.7.2.1g v3.7.2.1h
v3.7.2.1i v3.7.2.2a v3.7.2.2b v3.7.2.2c v3.7.2.2d. /*Look at the distribution of each variable on
transfers in kind.

/*Defines a variable list for the originals and the copies to perform a loop.
/*Copies the original variables for further processing.
/*If there is an empty field in one variable, replace it with 0.
/*Define all values <=-1 missing.
DO REPEAT var=v3.7.2.1a v3.7.2.1b v3.7.2.1c v3.7.2.1d v3.7.2.1e v3.7.2.1f v3.7.2.1g v3.7.2.1h
v3.7.2.1i v3.7.2.2a v3.7.2.2b v3.7.2.2c v3.7.2.2d
/newvar=v3.7.2.1a_r v3.7.2.1b_r v3.7.2.1c_r v3.7.2.1d_r v3.7.2.1e_r v3.7.2.1f_r v3.7.2.1g_r
v3.7.2.1h_r v3.7.2.1i_r v3.7.2.2a_r v3.7.2.2b_r v3.7.2.2c_r v3.7.2.2d_r.
COMPUTE newvar=var.
IF MIS(var) newvar=0.
MISSING VALUES newvar (lo thru -1).
END REPEAT.
EXE.

LIST #casenumber v3.6.1 v3.6.2a v3.6.2b v3.6.3a v3.6.3b v3.6.4a v3.6.4b v3.6.5a v3.6.5b. /*To
check for obviously wrong answers (e.g. "88" or "666" in more than one variable), produce a
list of all variables. [26]
```

---

[25] Variable numbering as follows:
Living costs paid out of own pocket: v3.7.**1**.1a (accommodation) - v3.7.**1**.1i (other)
Living costs paid by parents/partner/other: v3.7.**2**.1a (accommodation) - v3.7.**2**.1i (other)
Study-related costs paid out of own pocket: V3.7.**1**.2a (fees) - v3.7.**1**.2d (other)
Study-related costs paid by parents/partner/other: V3.7.**2**.2a (fees) - v3.7.**2**.2d (other)

```
/*Define lists of variables, containing all income and transfers in kind, resp. cash expense
variables.
/*Copy cash expense variables for further processing
/*If the case number is "1234" allocate the value -55 to all variables on income, transfers in
kind (included in "varlist") and cash expenses included in "newvar").
/*Define all values <=-1 missing.
DO REPEAT varlist= v3.6.1_r v3.6.2a_r v3.6.2b_r v3.6.3a_r v3.6.3b_r v3.6.4a_r v3.6.4b_r
v3.6.5a_r v3.6.5b_r v3.7.2.1a_r v3.7.2.1b_r v3.7.2.1c_r v3.7.2.1d_r v3.7.2.1e_r v3.7.2.1f_r
v3.7.2.1g_r v3.7.2.1h_r v3.7.2.1i_r v3.7.2.2a_r v3.7.2.2b_r v3.7.2.2c_r v3.7.2.2d_r.
/var=v3.7.1.1a v3.7.1.1b v3.7.1.1c v3.7.1.1d v3.7.1.1e v3.7.1.1f v3.7.1.1g v3.7.1.1h v3.7.1.1i
v3.7.1.2a v3.7.1.2b v3.7.1.2c v3.7.1.2d
/newvar= v3.7.1.1a_r v3.7.1.1b_r v3.7.1.1c_r v3.7.1.1d_r v3.7.1.1e_r v3.7.1.1f_r v3.7.1.1g_r
v3.7.1.1h_r v3.7.1.1i_r v3.7.1.2a_r v3.7.1.2b_r v3.7.1.2c_r v3.7.1.2d_r
COMPUTE newvar=var.
IF #casenumber=1234 varlist=-55.
IF #casenumber=1234 newvar=-55.
MISSING VALUES varlist newvar (lo thru -1).
END REPEAT.
EXE.


IF v3.9=1 AND MIS(v3.6.4a) OR v3.6.4a=0 v3.6.4a_r=-55.
EXE. /*If a student works during the whole semester but did not report any income from
current paid job, allocate -55 for "implausible answer".

*Assume the highest possible amount for non-repayable, public student grant is 300€ per
month.
IF v3.6.2a=1800 v3.6.2a_r=v3.6.2a/6. /*If the amount reported obviously refers to another
period than one month (e.g. one semester), convert it to a monthly amount.

IF v3.6.2a>300 v3.6.2a_r=300. /*If the reported amount exceeds the highest possible amount,
replace it with the highest possible amount.

EXE.

IF v3.9=3 OR MIS(v3.9) v3.6.4a_r=0. /*If a student does not work or at least did not report that
he/she works, the workload for paid employment shall be 0.
EXE.

/*Defines a list of variables on transfers in kind for study-related costs.
/*Convert the already processed variables (_r) on transfers in kind into monthly amounts
```

---

[26] "#casenumber" stands for the sequential number of the case. Add it to the variable list to identify cases with obviously wrong answers. Carefully check your data for obviously wrong answers and identify the cases by their sequential number. NB. only use this command when total N is lower than 5.000.

(_m).

/*Define all values -1 as missing values.

```
DO REPEAT var= v3.7.2.2a_r v3.7.2.2b_r v3.7.2.2c_r v3.7.2.2d_r
/newvar= v3.7.2.2a_mv3.7.2.2b_m v3.7.2.2c_m v3.7.2.2d_m.
COMPUTE newvar=var/6.
MISSING VALUES newvar (lo thru -1).
END REPEAT.
EXE.
```

*Applying (3a) of the instructions on data cleaning of income sources.

```
FREQUENCIES v3.6.1_r v3.6.2a_r v3.6.2b_r v3.6.3a_r v3.6.3b_r v3.6.4a_r v3.6.4b_r v3.6.5a_r
v3.6.5b_r. /*Look for outliers in your income distribution.
IF v3.6.1_r=10000 v3.6.1_r=1000. /*If reasonable, correct data that resembles a mistyping.
You may also use #casenumber to identify single cases.
IF v3.6.1_r=0,5 v3.6.1_r=-55. /*Allocate -55 to implausible outliers[27].
MISSING VALUES v3.6.1_r (lo thru -1).
EXE.
```

*Repeat this procedure (3a) for every income source!.

```
COMPUTE v3.6.1_cashkind=v3.6.1_r + v3.7.2.1a_r + v3.7.2.1b_r + v3.7.2.1c_r + v3.7.2.1d_r +
v3.7.2.1e_r + v3.7.2.1f_r + v3.7.2.1g_r + v3.7.2.1h_r + v3.7.2.1i_r + v3.7.2.2a_m + v3.7.2.2b_m
+ v3.7.2.2c_m + v3.7.2.2d_m.
EXE. /*Sum up cash income from family/ partner and all transfers in kind (monthly amounts!).
Then repeat the correction/ cleaning procedure (3a)!.
```

```
COMPUTE v3.6cashsum= v3.6.1_r + v3.6.2a_r + v3.6.2b_r + v3.6.3a_r + v3.6.3b_r + v3.6.4a_r +
v3.6.4b_r + v3.6.5a_r + v3.6.5b_r.
EXE. /*Sum up all cash income sources. Then repeat the correction/ cleaning procedure (3a)!.
```

```
COMPUTE v3.6cashkindsum=v3.6cashsum+ v3.7.2.1a_r + v3.7.2.1b_r + v3.7.2.1c_r +
v3.7.2.1d_r + v3.7.2.1e_r + v3.7.2.1f_r + v3.7.2.1g_r + v3.7.2.1h_r + v3.7.2.1i_r + v3.7.2.2a_m
+ v3.7.2.2b_m + v3.7.2.2c_m + v3.7.2.2d_m.
EXE. /*Sum up the sum of cash income sources plus all transfers in kind. Then repeat the
correction/ cleaning procedure (3a)!.
```

```
COMPUTE v3.6_familyshare=v3.6.1_cashkind/ v3.6cashkindsum. /*Compute the share of
income from family/ partner (incl. transfers in kind) of the total income (incl. transfers in kind).
COMPUTE v3.6_incomeshare=(v3.6.4a_r+v3.6.4b_r)/ v3.6cashkindsum. /*Compute the share
of self-earned income (from current job and from previous job) of the total income (incl.
transfers in kind).
```

---

[27] Assume the lowest 0,2%-2% of the income distribution receive less than 10€ and the top 0,2%-2% of
the income distribution receive more than 1.000€. Therefore cut off at <10€ and >1.000€.

```
COMPUTE v3.6_publicshare=(v3.6.2a_r+v3.6.2b_r)/ v3.6cashkindsum. /*Compute the share of
income from public student support (non-repayable and repayable) of the total income (incl.
transfers in kind).
EXE.

COMPUTE e_depend=4. /*Compute the focus group variable and allocate 4 for "other".
IF v3.6_familyshare>0.5 e_depend=1.
IF v3.6_incomeshare>0.5 e_depend=2.
IF v3.6_publicshare >0.5 e_depend=3. /*If the share of one income source is >50%, this source
is the source the respondent depends on.
IF MIS(v3.6_familyshare) AND MIS(v3.6_incomeshare) AND MIS(v3.6_publicshare)
e_depend=-99. /*If all variables are missing, exclude the case from the focus group variable.
EXE.
MISSING VALUES e_depend (lo thru -1).
VARIABLE LABELS e_depend 'Dependency on income source'.
VALUE LABELS e_depend 1 'Dependent on family' 2 'Dependent on self-earned income' 3
'Dependent on public student support' 4 'Other'.

EXE.
DELETE VARIABLES  count_income /*Delete variables only used temporarily for data
processing, unnecessary for data delivery.
```

5. Final definition of valid cases

In order to exclude all non-valid cases according to the "4+2 criterion", it is best to compute a filter variable, called e_valid, where all valid cases are coded 1 and all non-valid cases are coded 0.

SPSS Syntax:

*Check, if missing values for all focus group variables are defined correctly!

COMPUTE e_valid=1. /*This Code produces a new variable where all cases in the dataset are allocated the value 1.

IF MIS(e_age) OR MIS(e_sex) OR MIS(e_qualification) OR MIS(e_hei) e_valid=0. /*This Code allocates the value 0 to all cases that are not valid (i.e. where information on either of the 4 variables is missing).

COUNT mis_groups=e_intens e_field e_ edupar e_transit e_international e_depend (MISSING). /*Counts how many missings occur per case in all six variables.

IF mis_groups>4 e_valid=0. /*If there are more than 4 missings (i.e. at least 2 variables valid) in either of the additional 6 focus group variables, the case is not valid and therefore coded 0 in the filter variable.

SELECT IF e_valid=1. /*Deletes all non-valid cases (where e_valid ≠ 0) from the dataset. This command is IRRETRIEVABLE!.

SAVE OUTFILE='*D:\Eurostudent\E:V\eurostudent_data.sav*'
/COMPRESSED. /*Save the file with a new name. Insert the path to the respective folder you want to save the E:V data file to.

6. Instructions on data weighting

A sample - whether randomly chosen or not - will in most of the cases not mirror the given population exactly. That means that certain groups, either systematically or randomly, might be overrepresented, while other groups might be underrepresented in the sample. In this case, the sample is not, or only to a certain extent, representative for the given population. Hence, data weighting is inevitable for producing representative results.

The weighting procedure

Data weighting aims at adjusting the distribution of the sample to the actual distribution of the population. Originally, each case has a weight of 1. Cases in underrepresented groups have to be allocated a weight higher than 1, because the size of this particular group in your sample is smaller than the actual size of the group in the given population. Cases in overrepresented groups must analogously be allocated a weight lower than 1, because the size of that group in the sample is greater than the actual size of the group in the given population.

Data weighting is only possible with a valid dataset! Therefore complete the data cleaning before you start the weighting procedure.

Calculating the weights

In E:V the recommended weighting variables are **sex, age, study programme, type of HE institution and field of study**. The more detailed the weighting tables are, the more appropriate the weighting will be. Therefore we encourage you to expand the list of weighting variables. For the instructions below we used sex, age, HEI and study programme.

Most important in the weighting procedure is, that you have the exact same table you produced from your sample data, also available from administrative data. Usually there is administrative data on the student body collated by the ministry or national statistical office. If you have access to the micro-data you can produce the tables from the administrative data yourself by following the instructions below (applicable for sample data <u>and</u> admin data). If not, you have to order the table from the ministry or national statistical office. In this case, follow the instructions below, produce a weighting table from your sample data and decide on the final make-up of this table. Only then, order a table of exactly the same make-up from the ministry.

Why is it that complicated?

For data weighting, you have to use **nested** tables. That means, that all variables used for data weighting are interlaced together in <u>one</u> table (see Table 1 - Table 4). If you don't have access to the micro administrative data, you need to order a nested table. Therefore, you have to calculate the weighting tables from your sample data first, in order to know exactly what to order from the ministry (keep in mind that both tables must look exactly the same!).

If it is not possible to order nested weighting tables from administrative data, please contact <u>eurostudent@his.de</u>.

1. **Weighting Tables (assuming access to micro administrative data)**

1a. Select the relevant target group (according to the E:V target group) in the administrative data. E.g. Students in PhD programmes must be excluded (see p. 3 in this document).

1b. Compute the weighting variables equally in both the administrative and the survey data. The variables **sex, age, study programme and HE institution** should **at least** be considered for data weighting. The categorisation of these variables depends on the size of each category. Make sure to have at least 30 cases in each category.

1c. Produce two tables of identical make-up: one table containing administrative data on the whole student body (Table 1), one table containing the cleansed survey data (Table 2). In these tables combine variables to produce nested tables. This is of high importance since variables like sex and field of study are interdependent.

1d. There might be empty fields or fields with too few (<30) cases in the sample table. If so, recode the respective variable and merge two groups to create one bigger subgroup. If there are a lot of cases in one field, split this category into two or more subcategories if possible.

1e. Copy the tables to Excel for further processing.

1f. Create a third table in Excel and calculate the expected frequency in each cell of the table by dividing the actual number of students (admin data Table 1) by the sum of all students (admin data Table 1). This equation will give you the share of students in the respective cell. Multiply this figure with the total number of students in the sample. This second equation will give you the expected frequency of students, that would have been surveyed, if the actual distribution was mirrored by the sample.

1g. Calculate the weights for every subgroup (Table 4) by dividing the expected frequency in each cell (Table 3) by the number of sampled cases in each cell (survey data Table 2). After weighting the data, the total n of the survey must be equal to the unweighted n.[28]

Trimmed weights:

    i. The weights should not be lower than 0,2 or greater than 5. If they are, replace values lower, resp. higher than that with the respective minimum or maximum (i.e. 0.2, resp. 5). Correct these weight factors by allocating the minimum or maximum weights. For example, if a weight factor is '0.0004' you replace this weight with '0.2'. If necessary, adapt these limits to your national data, but never produce weights <0,2.

---

[28] **CAUTION!** For empty cells (must be empty in both tables from admin data <u>and</u> survey data) no weight will be calculated!

ii.   The range of the lowest and the highest weight should not exceed the factor 10. Together with the rules described in (i), you have to find the perfect range for your weights.

iii.   Adjusting the weight factors (replacing the outliers with your set minimum or maximum weight factor) can cause an imbalance in your mean weight, since the mean weight factor for the entire response group has to be '1'. Compute the arithmetic mean of your weights and use it as a divisor for each weight. Hence, the average of all weights will be '1' again.

Table 1:   Weighting table - administrative data

e_hei_new	e_sex	e_age_new	e_qualification		
			1 BA	2 MA	3 Other
1 University	1 Female	1 under 21	6571	4391	0
		2 21-25	11056	6876	5827
		2.5 23-25J.	10140	9162	0
		3 26-30	6804	7981	3677
		4 over 30	6862	7064	1986
	2 Male	1 under 21	3032	1543	0
		2 21-25	9907	4224	3933
		2.5 23-25J.	12401	5977	0
		3 26-30	9242	6650	4665
		4 over 30	7531	7070	2502
2 University of applied science	1 Female	2 21-25	4338	1508	779
		3 26-30	1184	1499	753
	2 Male	2 21-25	3650	1072	385
		3 26-30	1492	1325	865
3 Other HEI	1 Female	1 under 21	956	387	0
		2 21-25	4575	1871	1208
		3 26-30	2080	1357	1228
		4 over 30	1255	995	599
	2 Male	2 21-25	4950	1134	770
		3 26-30	3026	1682	1554
			Total N		**240.654**

As an example: Additional age group for more precision (value 2,5); only for BA, MA - not for Dipl and only for university students!

Table contains values and value labels of the variables.
Administrative data, fictitious.

Table 2:   Weighting table - survey data

e_hei_new	e_sex	e_age_new	e_qualification		
			1 BA	2 MA	3 Other
1 University	1 Female	1 under 21	150	80	0
		2 21-25	246	141	162
		2.5 23-25J.	168	203	0
		3 26-30	70	137	84
		4 over 30	63	55	30
	2 Male	1 under 21	51	17	0
		2 21-25	156	56	98
		2.5 23-25J.	168	105	0
		3 26-30	72	91	74
		4 over 30	47	69	23
2 University of applied science	1 Female	2 21-25	83	43	23
		3 26-30	8	32	21
	2 Male	2 21-25	40	21	12
		3 26-30	8	16	15
3 Other HEI	1 Female	1 under 21	8	11	0
		2 21-25	54	40	21
		3 26-30	17	16	15
		4 over 30	8	11	4
	2 Male	2 21-25	31	14	13
		3 26-30	10	13	10
			Total n	**3.507**	

> As an example: Additional age group for more precision (value 2.5); only for BA, MA - not for Dipl and only for university students!

Table contains values and value labels of the variables.
Survey data, fictitious.

Table 3: **Weighting table - expected frequencies, assuming equal distributions in N and n[1)]**

e_hei_new	e_sex	e_age_new	e_qualification		
			1 BA	2 MA	3 Other
1 University	1 Female	1 under 21	96	64	0
		2 21-25	161	100	85
		2.5 23-25J.	148	134	0
		3 26-30	99	116	54
		4 over 30	100	103	29
	2 Male	1 under 21	44	22	0
		2 21-25	144	62	57
		2.5 23-25J.	181	87	0
		3 26-30	135	97	68
		4 over 30	110	103	36
2 University of applied science	1 Female	2 21-25	63	22	11
		3 26-30	17	22	11
	2 Male	2 21-25	53	16	6
		3 26-30	22	19	13
3 Other HEI	1 Female	1 under 21	14	6	0
		2 21-25	67	27	18
		3 26-30	30	20	18
		4 over 30	18	14	9
	2 Male	2 21-25	72	17	11
		3 26-30	44	25	23
			Total n:		**3.507**

Overlay box:

$$\frac{N_{ij}}{\text{Total N}} \times \text{sample n} = \frac{6571}{240654} \times 3507$$

[1)] N: number of persons in the given population; n: number of respondents in the sample.
Table contains values and value labels of the variables.

Table 4: **Weighting table - case weights for each subgroup**

e_hei_new	e_sex	e_age_new	e_qualification 1 BA	2 MA	3 Other
1 University	1 Female	1 under 21	0,6384	0,7999	
		2 21-25	0,6549	0,7107	0,5242
		2.5 23-25J.	0,8796	0,6577	
		3 26-30	1,4165	0,8489	0,6379
		4 over 30	1,5873	1,8717	0,9647
	2 Male	1 under 21	0,8664	1,3227	
		2 21-25	0,9255	1,0992	0,5848
		2.5 23-25J.	1,0757	0,8295	
		3 26-30	1,8706	1,0649	0,9187
		4 over 30	2,3351	1,4932	1,5853
2 University of applied science	1 Female	2 21-25	0,7616	0,5111	0,4936
		3 26-30	2,1568	0,6826	0,5225
	2 Male	2 21-25	1,3298	0,7439	0,4675
		3 26-30	2,7178	1,2068	0,8404
3 Other HEI	1 Female	1 under 21	1,7414	0,5127	
		2 21-25	1,2346	0,6816	0,8383
		3 26-30	1,7830	1,2360	1,1930
		4 over 30	2,2861	1,3182	2,1823
	2 Male	2 21-25	2,3269	1,1804	0,8632
		3 26-30	4,4097	1,8855	2,2646

(callout box, pointing to the cells): $$\frac{\text{Expected n}}{\text{Observed n}} = \frac{96}{150}$$

Table contains values and value labels of the variables.
4 decimals displayed. For data weighting, use all decimals!

SPSS Syntax:

*Recode your variables as applicable to you data. For this fictive example, we recoded e_hei and e_age into e_hei_new (university, university of applied sciences, other) and e_age_new (additional category: 21-23y; value 2.5; see Table 1 - Table 4).

CTABLES
/VLABELS VARIABLES=e_hei_new e_sex e_age_new e_qualification DISPLAY=DEFAULT
/TABLE e_hei_new > e_sex > e_age_new [C][COUNT F40.0] BY e_qualification [C]
/CATEGORIES VARIABLES= e_hei_new e_sex e_age_new e_qualification ORDER=A KEY=VALUE
EMPTY=EXCLUDE. /*Custom Tables for nested tables. If necessary, recode your variables to produce more or less groups (e.g. age groups); adapt to national data!.

*If Custom Tables is not featured by your version of SPSS, use the following command:.
*CROSSTABS e_age_new BY e_qualification BY e_sex BY e_hei_new.

```
***********************************.
*******Copy the weighing table from sample data and the one from administrative data into
an excel sheet and calculate expected frequencies and the cell weights in new tables.
***********************************.

COMPUTE weight=$sysmis. /*Compute the weighting variable. In the first place, all cases are
system missing in this variable.

IF e_age_new=1 AND e_qualification=1 AND e_hei_new=1 AND e_sex=1
weight=0,638385316678717. /*Copy the respective cell weight from Excel into this code line. Repeat
for every cell in the weighting table, adapting the conditions in the code line.

EXE.

DESCRIPTIVES weight/STATISTICS = MINIMUM MAXIMUM. /*Check the minimum and maximum
value of the weight factor.

RECODE weight (0 thru 0.2 = 0.2)(5 thru hi=5)(else=copy). /*Correct the outliers.

EXE.

DESCRIPTIVES weight/STATISTICS = MEAN. /*Check for minimum, maximum and mean value of the
weight factor.

RENAME weight=weight_old

COMPUTE weight=weight_old*(1/#weight_mean). /*Correct the mean to equal one, replace
"#weight_mean" with the arithmic mean you found in the descriptives above.

EXE.

WEIGHT BY weight. /*Weights the data by the variable weight. "WEIGHT OFF" turns off the weight.

EXE.
```

7. Tricky issues in data analysis

Gini coefficient

In the subtopic on *funding and state assistance*, countries are asked to deliver data on the number of students in each income decile (once for total income, once only for income from paid employment), in order to calculate the gini-coefficient. However, the SPSS-command *FREQUENCIES /ntiles* 10 will give you the deciles, but will not allocate cases with the same income (i.e. duplicates) randomly to the respective income decile. That means, this procedure would give you deciles of different sizes. For the calculation of the gini-coefficient, every decile must be of the exact same size (i.e. number of persons in the decile). Therefore, we need to sort the cases by income and find the correct cut-off point. The following, quite complex syntax command will produce income deciles of equal sizes. The comments will guide you through the procedure.

SPSS Syntax:

COMPUTE total_dummy=1.
EXE. /*Computes a new variable, where every case is allocated the value 1.

SORT CASES BY total_dummy (A) v3.6cashkindsum (A). /*Sorts the cases by random total and by the total income (inc. transfers in kind) in Ascending order.

COMPUTE MatchSequence=$casenum.

RANK MatchSequence /ntiles (10). /*Will automatically produce a new variable "NMatchSe". Its values will represent the decile membership of the cases.
/*CAUTION! If you repeat this command, it will produce another variable called NTI001, then NTI002 etc!. Therefore, if you need to run the command twice, replace the variable in all following analyses!

CTABLES
/VLABELS VARIABLES=NMatchSe v3.6cashkindsum DISPLAY=DEFAULT
/TABLE *NMatchSe* [C] BY v3.6cashkindsum [MEAN, MAXIMUM, COUNT F40.0]
/CATEGORIES VARIABLES=NMatchSe ORDER=A KEY=VALUE EMPTY=EXCLUDE. /*Custom Tables will show you the Mean of each income decile, the Maximum (i.e. the decile limit) and the weighted number of respondents in each decile group. Make sure, you are using the correct decile-variable!

Data cleaning on "form of housing"

In the E:V questionnaire students are asked „whom they live with" in a question with multiple answers possible (Question 3.1). The next question asks, whether the students live in a student accommodation or not (Question 3.2). You have to check these two questions for consistency, thinking of which scenarios might be possible and which are not (e.g. living with parents and living in a student accommodation Monday to Friday during the term is not quite plausible). Having identified those implausible cases, exclude them (i.e. define them as missing

values) from 3.1, 3.2 and also 3.3 (assessment of accommodation), as implausible answers on the form of housing are not supposed to be validly assessed.

SPSS Syntax:

CTABLES
/VLABELS VARIABLES=v3.1a v3.1b v3.1c v3.1d v3.1e v3.2 DISPLAY=DEFAULT
/TABLE v3.1a [C] + v3.1b [C] + v3.1c [C] + v3.1d [C] + v3.1e [C] + v3.2 [C] BY
v3.1a [C][COUNT] + v3.1b [C][COUNT] + v3.1c [C][COUNT] + v3.1d [C][COUNT] + v3.1e [C][COUNT]
/CATEGORIES VARIABLES= v3.1a v3.1b v3.1c v3.1d v3.1e v3.2 ORDER=A KEY=VALUE
EMPTY=INCLUDE MISSING=INCLUDE. /*Check out, what combinations the students reported for their form of housing. "COUNT" produces weighted counts, if you have already weighted your data.

/*Defines a list of variables regarding the form of housing.
/*Copies the original variables for further processing.
/*If a respondent reported that he/she lives with his/her parents and that he/she lives in a student accommodation allocate -55 for "implausible answer" to all variables on form of housing.
/*If a respondent reported to live alone but quoted any of the other options, allocate -55 for "implausible answer" to all variables on form of housing
/*Define all values <=-1 missing (0 stands for "not quoted").
DO REPEAT oldvar= v3.1a v3.1b v3.1c v3.1d v3.1e v3.2 v3.3
/newvar= v3.1a_r v3.1b_r v3.1c_r v3.1d_r v3.1e_r v3.2_r v3.3_r.
COMPUTE newvar=oldvar.
IF v3.1a_r=1 AND v3.2_r=1 newvar=-55.
IF v3.1e_r=1 AND (v3.1a_r=1 OR v3.1b_r=1 OR v3.1c_r=1 OR v3.1d_r=1) newvar=-55..
MISSING VALUES newvar (lo thru -1).
END REPEAT.
EXE.

In general, please check your data carefully regarding the plausibility and consistency of the given answers in order to ensure top quality for data delivery.

Annex

E:V core data set: list of variables

VariableName	VariableLabel	Values	Subtopic
v1.0	Do you actively pursue your studies in the current semester?	1 = Yes, at a university in country 2 = Yes, at a university of applied sciences in country 3 = Yes, at other HEI in country 4 = Yes, (temporarily) at a University abroad, outside country (e.g. on Erasmus) 5 = No, I'm currently interrupting my studies (either officially or not) 6 = No, I stopped studying 7 = No, I already graduated	Current Study Situation
v1.1	Which study programme are you currently enrolled in?	1 = Short Cycle [ISCED 5] 2 = Bachelor [ISCED 6] 3 = Master [ISCED 7] 4 = Short national degree [up to 3 years, ISCED 6] 5 = Long national degree [more than 3 years, ISCED 7] 6 = Other postgraduate programmes [ISCED 7] 7 = PhD/Doctoral Programme [ISCED 8]	Current Study Situation
v1.2	What is your current formal status as a student?	1 = Full-time student 2 = Part-time student 3 = Other [optional, formulate regarding national context]	Current Study Situation
v1.3	Is your study programme formally defined as a distance learning programme?	1 = Yes 2 = No	Current Study Situation
v1.4.1	Please provide the **location** of the higher education institution you attend.	[Name of the city/town/place]	Current Study Situation
v1.4.2	Please provide the **name** of the higher education institution you attend.	[Name of the higher education institution]	Current Study Situation
v1.5	What is the subject you study?	[Name of subject]	Current Study Situation
v1.6	Are you planning to continue studying after finishing your current study programme(s)?	1 = Yes, within a year after graduating from my current study programme 2 = Yes, but not within a year after graduating from my current study programme 3 = No, I do not plan to continue studying at all 4 = I don't know yet	Current Study Situation
v1.7	On what level do you plan to continue studying?	1 = Bachelor 2 = Master 3 = PhD 4 = Another programme not mentioned here 5 = I don't know yet	Current Study Situation
v1.8.1	Where do you plan to continue studying?	1 = In [country] 2 = Abroad 3 = I don't know yet	Current Study Situation
v1.8.2	Where do you plan to continue studying? [Name country]	[Name of the country]	Current Study Situation

v1.9	What do you plan to do within a year after finishing your current study programme?	1 = I will continue my current paid employment/job/occupation 2 = I will have or look for a new paid employment 3 = I will set-up my own business (self-employed) 4 = I will do something else 5 = I don't know yet	Current Study Situation
v1.10	What is the main teaching language of your study programme?	1 = [Official language(s)] 2 = [English] 3 = [Common language in your country] 4 = [Common language in your country] 5 = Other	Current Study Situation
v1.11.1	How satisfied are you with your studies concerning the **quality of teaching**	1 = Very well 2 = - 3 = - 4 = - 5 = Not at all	Current Study Situation
v1.11.2	How satisfied are you with your studies concerning the **possibility to select from a broad variety of courses**	1 = Very well2 = -3 = -4 = -5 = Not at all	Current Study Situation
v1.11.3	How satisfied are you with your studies concerning the **university administration's attitude towards students**	1 = Very well 2 = - 3 = - 4 = - 5 = Not at all	Current Study Situation
v1.11.4	How satisfied are you with your studies concerning the **teaching staff's attitude towards students**	1 = Very well 2 = - 3 = - 4 = - 5 = Not at all	Current Study Situation
v1.11.5	How satisfied are you with your studies concerning the **study facilities**	1 = Very well 2 = - 3 = - 4 = - 5 = Not at all	Current Study Situation

v1.12.1	How do you rate your chances on the labour market after graduating from your current study programme? - on **national** level	1 = Very good 2 = - 3 = - 4 = - 5 = Very poor 6 = Unable to rate	Current Study Situation
v1.12.2	How do you rate your chances on the labour market after graduating from your current study programme? - on **international** level	2 = Very good 2 = - 3 = - 4 = - 5 = Very poor 6 = Unable to rate	Current Study Situation
v2.1	What was the highest level of education you obtained on graduating from the school system for the first time?	1 = [Name of national qualification] (up to lower secondary (ISCED 24, 25)) 2 = [Name of national qualification] (upper secondary academic track (ISCED 34)) 3 = [Name of national qualification] (upper secondary dual track (ISCED 34/35)) 4 = [Name of national qualification] (upper secondary vocational track (ISCED 35)) 5 = [Name of other national qualification] 6 = Foreign Qualification	Study Background
v2.2prov	Where did you graduate from the educational level mentioned above? - province	[Province] [drop-down or open question]	Study Background
v2.2country	Where did you graduate from the educational level mentioned above? - country	[Country] [drop-down or open question]	Study Background
v2.3month	When did you obtain the qualification mentioned in 2.1? - month	[Month]	Study Background
v2.3year	When did you obtain the qualification mentioned in 2.1? - year	[Year]	
v2.4.1	What qualifications, examinations or measures qualified you for entry into higher education? [Name of national qualification according to **upper secondary qualification (ISCED 34/35)**]	1 = Yes 0 = No	Study Background
v2.4.2	What qualifications, examinations or measures qualified you for entry into higher education? [Name of **central higher education entrance** examination for all students]	1 = Yes 0 = No	Study Background
v2.4.3	What qualifications, examinations or measures qualified you for entry into higher education? [Name of national qualification according to **upper secondary school academic track** through adult learning (ISCED 34/35/44/45)]	1 = Yes0 = No	Study Background
v2.4.4	What qualifications, examinations or measures qualified you for entry into higher education? [Name of national qualification according to **special examinations for certain student groups** (e.g. for mature students)]	1 = Yes 0 = No	Study Background

v2.4.5	What qualifications, examinations or measures qualified you for entry into higher education? [Name of **special access courses**]	1 = Yes 0 = No	Study Background
v2.4.6	What qualifications, examinations or measures qualified you for entry into higher education? [Name of national qualification based on **accreditation of prior learning and/or vocational experience** (APR)]	1 = Yes 0 = No	Study Background
v2.4.7	What qualifications, examinations or measures qualified you for entry into higher education? [Name of national qualification according to **special entrance qualification for certain fields of study** (e.g. music, arts, sports)]	1 = Yes 0 = No	Study Background
v2.4.8	What qualifications, examinations or measures qualified you for entry into higher education? **Foreign qualification**	1 = Yes 0 = No	Study Background
v2.5month	When did you obtain/complete the most recent qualification, examination or measure, which qualified you for entry into higher education? - **Month**	[Month]	Study Background
v2.5year	When did you obtain/complete the most recent qualification, examination or measure, which qualified you for entry into higher education? - **Year**	[Year]	Study Background
v2.6month	When did you enter higher education for the first time? - **Month**	[Month]	Study Background
v2.6year	When did you enter higher education for the first time? - **Year**	[Year]	Study Background
v2.7month	When did you start your current study programme? - **Month**	[Month]	Study Background
v2.7year	When did you start your current study programme? - **Year**	[Year]	Study Background
v2.8.1	Did you ever interrupt your education career after entering higher education for at least one year? - **between entering higher education and graduating from higher education for the first time**	1 = Yes 0 = No	Study Background
v2.8.2	Did you ever interrupt your education career after entering higher education for at least one year? - **between graduating from higher education and re-entering higher education**	1 = Yes 0 = No	Study Background
v2.8.3	Did you ever interrupt your education career after entering higher education for at least one year? - **No interruption after entering higher education**	1 = Yes [no interruption] 0 = No	Study Background
v2.9	Did you have a paid job before entering higher education for the first time?	1 = Yes, working/vocational training for at least one year and at least 20h per week 2 = Yes, less than 1 year or less than 20h a week 3 = No	Study Background
v3.1.1	Who do you live with during the study term/semester (Monday until Friday)? - Parents/other relatives	1 = Yes 0 = No	Living Conditions

v3.1.2	Who do you live with during the study term/semester (Monday until Friday)? - Partner/spouse	1 = Yes 0 = No	Living Conditions
v3.1.3	Who do you live with during the study term/semester (Monday until Friday)? - My child(ren)/my partner's child(ren)	1 = Yes 0 = No	Living Conditions
v3.1.4	Who do you live with during the study term/semester (Monday until Friday)? - With (an)other person(s) not mentioned above	1 = Yes 0 = No	Living Conditions
v3.1.5	Who do you live with during the study term/semester (Monday until Friday)? - I live alone	1 = Yes0 = No	Living Conditions
v3.2	Do you live in a student accommodation, i.e dormitory or halls of residence?	1 = Yes 0 = No	Living Conditions
v3.3	How satisfied are you with your accommodation?	1 = Very satisfied 2 = - 3 = - 4 = - 5 = Not satisfied at all	Living Conditions
v3.4.1a	What are your frequently used modes of transportation to get from your home to your HE institution during the current semester? Please indicate all frequently used modes. - Foot	1 = Frequently used 0 = Not frequently used	Living Conditions
v3.4.1b	What are your frequently used modes of transportation to get from your home to your HE institution during the current semester? Please indicate all frequently used modes. - Bicycle	1 = Frequently used 0 = Not frequently used	Living Conditions
v3.4.1c	What are your frequently used modes of transportation to get from your home to your HE institution during the current semester? Please indicate all frequently used modes. - Car/motorbike or similar	1 = Frequently used 0 = Not frequently used	Living Conditions
v3.4.1d	What are your frequently used modes of transportation to get from your home to your HE institution during the current semester? Please indicate all frequently used modes. - Public transport	1 = Frequently used 0 = Not frequently used	Living Conditions
v3.4.1e	What are your frequently used modes of transportation to get from your home to your HE institution during the current semester? Please indicate all frequently used modes. - Other	1 = Frequently used 0 = Not frequently used	Living Conditions
v3.4.2	What is your **most frequently** used mode of transportation to get from your home to your HE institution during the current semester?	1 = Foot 2 = Bicycle 3 = Car/motorbike or similar 4 = Public transport 5 = Other	Living Conditions

v3.5	On a typical day, how much time does it take you to get from your home to your higher education institution during the current semester?	[One way minutes on average]	Living Conditions
v3.6.1	What is the average monthly amount at your disposal from the following sources during the current semester? - **Provision from family/partner**	[Average monthly amount]	Living Conditions
v3.6.2a	What is the average monthly amount at your disposal from the following sources during the current semester? - **From public sources: non-repayable student grant**	[Average monthly amount]	Living Conditions
v3.6.2b	What is the average monthly amount at your disposal from the following sources during the current semester? - **From public sources: repayable student loan**	[Average monthly amount]	Living Conditions
v3.6.3a	What is the average monthly amount at your disposal from the following sources during the current semester? - **From private sources (foundation etc.): non-repayable student grant**	[Average monthly amount]	Living Conditions
v3.6.3b	What is the average monthly amount at your disposal from the following sources during the current semester? - **From private sources: repayable student loan**	[Average monthly amount]	Living Conditions
v3.6.4a	What is the average monthly amount at your disposal from the following sources during the current semester? - **Self-earned income: from current paid job**	[Average monthly amount]	Living Conditions
v3.6.4b	What is the average monthly amount at your disposal from the following sources during the current semester? - **Self-earned income: from previous job**	[Average monthly amount]	Living Conditions
v3.6.5a	What is the average monthly amount at your disposal from the following sources during the current semester? - **Other: public sources (incl. child benefit, pension)**	[Average monthly amount]	Living Conditions
v3.6.5b	What is the average monthly amount at your disposal from the following sources during the current semester? - **Other: private sources**	[Average monthly amount]	Living Conditions
v3.6.6	What is the average monthly amount at your disposal from the following sources during the current semester? - **Total**	[Total monthly amount]	Living Conditions
v3.7.1.1a	What are your average **monthly** expenses for the following items during the current semester? - **Total accommodation cost** - I pay out of my own pocket	[Living costs per month]	Living Conditions
v3.7.1.1b	What are your average **monthly** expenses for the following items during the current semester? **Food** - I pay out of my own pocket	[Living costs per month]	Living Conditions

v3.7.1.1c	What are your average **monthly** expenses for the following items during the current semester? **Transportation** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1d	What are your average **monthly** expenses for the following items during the current semester? **Communication** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1e	What are your average **monthly** expenses for the following items during the current semester? **Health costs (e.g. medical insurance)** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1f	What are your average **monthly** expenses for the following items during the current semester? **Childcare** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1g	What are your average **monthly** expenses for the following items during the current semester? **Debt payment (except mortgage)** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1h	What are your average **monthly** expenses for the following items during the current semester? **Social and leisure activities** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1i	What are your average **monthly** expenses for the following items during the current semester? **Other regular living costs** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.1.1j	What are your average **monthly** expenses for the following items during the current semester? **Total living costs** - I pay out of my <u>own pocket</u>	[Living costs per month]	Living Conditions
v3.7.2.1a	What are your average **monthly** expenses for the following items during the current semester? **Total accommodation cost (rent/mortgage including utilities, water, electricity,...)** - Paid by <u>parents/partner/ others for me</u>	[Living costs per month]	Living Conditions
v3.7.2.1b	What are your average **monthly** expenses for the following items during the current semester? **Food** - Paid by <u>parents/partner/ others for me</u>	[Living costs per month]	Living Conditions
v3.7.2.1c	What are your average **monthly** expenses for the following items during the current semester? **Transportation** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1d	What are your average **monthly** expenses for the following items during the current semester? **Communication (telephone, internet etc.)** - Paid by <u>parents/partner/ others for me</u>	[Living costs per month]	Living Conditions

v3.7.2.1e	What are your average **monthly** expenses for the following items during the current semester? **Health costs (e.g. medical insurance)** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1f	What are your average **monthly** expenses for the following items during the current semester? **Childcare** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1g	What are your average **monthly** expenses for the following items during the current semester? **Debt payment (except mortgage)** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1h	What are your average **monthly** expenses for the following items during the current semester? **Social and leisure activities** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1i	What are your average **monthly** expenses for the following items during the current semester? **Other regular living costs (clothing, toiletries tobacco, pets, insurance [except medical insurance])** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.2.1j	What are your average **monthly** expenses for the following items during the current semester? **Total living costs** - Paid by parents/partner/ others for me	[Living costs per month]	Living Conditions
v3.7.1.2a	What are your average expenses **per semester** for the following items during the current semester? **Tuition fees, registration fees, examination fees, administrative fees** - I pay out of my own pocket	[Study-related costs per semester]	Living Conditions
v3.7.1.2b	What are your average expenses **per semester** for the following items during the current semester? **Social welfare contributions to the university/ college and student associations** - I pay out of my own pocket	[Study-related costs per semester]	Living Conditions
v3.7.1.2c	What are your average expenses **per semester** for the following items during the current semester? **Learning materials (e.g. books, photocopying, DVDs, field trips)** - I pay out of my own pocket	[Study-related costs per semester]	Living Conditions
v3.7.1.2d	What are your average expenses **per semester** for the following items during the current semester? **Other regular study-related costs (e.g. private tutoring, additional courses)** - I pay out of my own pocket	[Study-related costs per semester]	Living Conditions
v3.7.1.2e	What are your average expenses **per semester** for the following items during the current semester? **Total study-related costs** - I pay out of my own pocket	[Study-related costs per semester]	Living Conditions

v3.7.2.2a	What are your average expenses **per semester** for the following items during the current semester? **Tuition fees, registration fees, examination fees, administrative fees** - Paid by parents/partner/ others for me	[Study-related costs per semester]	Living Conditions
v3.7.2.2b	What are your average expenses **per semester** for the following items during the current semester? **Social welfare contributions to the university/ college and student associations** - Paid by parents/partner/ others for me	[Study-related costs per semester]	Living Conditions
v3.7.2.2c	What are your average expenses **per semester** for the following items during the current semester? **Learning materials (e.g. books, photocopying, DVDs, field trips)** - Paid by parents/partner/ others for me	[Study-related costs per semester]	Living Conditions
v3.7.2.2d	What are your average expenses **per semester** for the following items during the current semester? **Other regular study-related costs (e.g. private tutoring, additional courses)** - Paid by parents/partner/ others for me	[Study-related costs per semester]	Living Conditions
v3.7.2.2e	What are your average expenses **per semester** for the following items during the current semester? **Total study-related costs** - Paid by parents/partner/ others for me	[Study-related costs per semester]	Living Conditions
v3.8	To what extent are you currently experiencing financial difficulties?	1 = Very seriously 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions
v3.9	Do you have a paid job during the current semester?	1 = Yes, I work during the whole semester 2 = Yes, I work from time to time during the semester 3 = No, I don't work during the semester	Living Conditions
v3.10.1	To what extent do the following statements apply to your situation? **I work to fund my living**	1 = Totally 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions
v3.10.2	To what extent do the following statements apply to your situation? **I work to improve my living standard**	1 = Totally 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions

v3.10.3	To what extent do the following statements apply to your situation? **I work to gain experience on the labour market**	1 = Totally 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions
v3.10.4	To what extent do the following statements apply to your situation? **I work because I have free time to spend**	1 = Totally 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions
v3.11	How closely related is your job to the content of your study programme?	1 = Very closely 2 = - 3 = - 4 = - 5 = Not at all	Living Conditions
3.12	Which of the following describes your current situation best?	1 = Primarily, I am a student besides other activities (e.g. paid job, care responsibilities). 2 = Primarily, I work and I study alongside working. 3 = Primarily, I am occupied with other duties/activities (e.g. care responsibilities) besides my studies.	Living Conditions
v3.13	How many hours do you spend on paid jobs in a typical week during this semester?	[Paid jobs in hours per week]	Living Conditions
v3.14.1mo	How many hours do you spend in a typical week in **taught courses**? **Monday**	[Taught courses in hours per day]	Living Conditions
v3.14.1tu	How many hours do you spend in a typical week in **taught courses**? **Tuesday**	[Taught courses in hours per day]	Living Conditions
v3.14.1we	How many hours do you spend in a typical week in **taught courses**? **Wednesday**	[Taught courses in hours per day]	Living Conditions
v3.14.1th	How many hours do you spend in a typical week in **taught courses**? **Thursday**	[Taught courses in hours per day]	Living Conditions
v3.14.1fr	How many hours do you spend in a typical week in **taught courses**? **Friday**	[Taught courses in hours per day]	Living Conditions
v3.14.1sa	How many hours do you spend in a typical week in **taught courses**? **Saturday**	[Taught courses in hours per day]	Living Conditions
v3.14.1su	How many hours do you spend in a typical week in **taught courses**? **Sunday**	[Taught courses in hours per day]	Living Conditions
v3.14.2mo	How many hours do you spend in a typical week on **personal study time**? **Monday**	[Personal studytime in hours per day]	Living Conditions

v3.14.2tu	How many hours do you spend in a typical week on **personal study time**? **Tuesday**	[Personal studytime in hours per day]	Living Conditions
v3.14.2we	How many hours do you spend in a typical week on **personal study time**? **Wednesday**	[Personal studytime in hours per day]	Living Conditions
v3.14.2th	How many hours do you spend in a typical week on **personal study time**? **Thursday**	[Personal studytime in hours per day]	Living Conditions
v3.14.2fr	How many hours do you spend in a typical week on **personal study time**? **Friday**	[Personal studytime in hours per day]	Living Conditions
v3.14.2sa	How many hours do you spend in a typical week on **personal study time**? **Saturday**	[Personal studytime in hours per day]	Living Conditions
v3.14.2su	How many hours do you spend in a typical week on **personal study time**? **Sunday**	[Personal studytime in hours per day]	Living Conditions
v3.15.1	Looking at your workload based on the time you spend on **study-related activities**, please rate your satisfaction with your workload.	1 = I want less 2 = - 3 = - 4 = - 5 = I want more	Living Conditions
v3.15.2	Looking at your workload based on the time you spend on **paid jobs**, please rate your satisfaction with your workload.	1 = I want less 2 = - 3 = - 4 = - 5 = I want more	Living Conditions
v3.15.3	Looking at your **total workload**, please rate your satisfaction with your workload.	1 = I want less 2 = - 3 = - 4 = - 5 = I want more	Living Conditions
v4.1	Have you ever been enrolled abroad as a student in higher education?	1 = Yes 2 = No, but I plan to go 3 = No, and I do not plan to go	International Mobility
v4.2a	What degree were you studying for, when you went abroad? **Bachelor [ISCED 6]**	1 = Yes 0 = No	International Mobility
v4.2b	What degree were you studying for, when you went abroad? **Master [ISCED 7]**	1 = Yes 0 = No	International Mobility
v4.2c	What degree were you studying for, when you went abroad? **PhD/ Doctoral Programme [ISCED 8]**	1 = Yes 0 = No	International Mobility
v4.2d	What degree were you studying for, when you went abroad? **Short national degree [up to 3 years, ISCED 6]**	1 = Yes 0 = No	International Mobility

v4.2e	What degree were you studying for, when you went abroad? **Long national degree [more than 3 years, ISCED 7]**	1 = Yes0 = No	International Mobility
v4.2f	What degree were you studying for, when you went abroad? **Other postgraduate programmes [ISCED 7]**	1 = Yes 0 = No	International Mobility
v4.2g	What degree were you studying for, when you went abroad? **Other**	1 = Yes 0 = No	International Mobility
v4.3country	What country did you go to?	[open answer or drop down list with countries]	International Mobility
v4.3months	For how long did you study abroad?	[Duration in months]	International Mobility
v4.4	Were the credits (ECTS, certificates) you gained for your studies abroad recognized by your home institution?	1 = Yes, all of the credits were recognised 2 = Yes, the credits were partly recognised 3 = No, none of the credits were recognised 4 = I don't know (yet) 5 = I did not gain any credits for my studies abroad	International Mobility
v4.5	Within which of the following organisational frameworks was your study abroad organised?	1 = EU-programme 2 = Other programme 3 = No programme	International Mobility
v4.6.1a	Which of the following sources did you use to fund your study period abroad? - **Contribution from parents/family/partner**	1 = Yes 0 = No	International Mobility
v4.6.1b	Which of the following sources did you use to fund your study period abroad? - **Own income from previous job or own savings**	1 = Yes 0 = No	International Mobility
v4.6.1c	Which of the following sources did you use to fund your study period abroad? - **Income from paid job during my studies abroad**	1 = Yes 0 = No	International Mobility
v4.6.1d	Which of the following sources did you use to fund your study period abroad? - **Study grants/loans from host country**	1 = Yes 0 = No	International Mobility
v4.6.1e	Which of the following sources did you use to fund your study period abroad? - **Regular study grants/loans from home country**	1 = Yes 0 = No	International Mobility
v4.6.1f	Which of the following sources did you use to fund your study period abroad? - **Special study grant/loan from home country for going abroad**	1 = Yes 0 = No	International Mobility
v4.6.1g	Which of the following sources did you use to fund your study period abroad? - **EU study grants (e.g. Erasmus, other)**	1 = Yes 0 = No	International Mobility
v4.6.1h	Which of the following sources did you use to fund your study period abroad? - **Funding from private businesses or non-governmental organisations (NGOs)**	1 = Yes 0 = No	International Mobility

v4.6.1i	Which of the following sources did you use to fund your study period abroad? - **Other**	1 = Yes 0 = No	International Mobility
v4.6.2	Which of the following sources was your primary source of funding?	1 = Contribution from parents/family/partner 2 = Own income from previous job or own savings 3 = Income from paid job during my studies abroad 4 = Study grants/loans from host country 5 = Regular study grants/loans from home country 6 = Special study grant/loan from home country for going abroad 7 = EU study grants (e.g. Erasmus, other) 8 = Funding from private businesses or non-governmental organisations (NGOs) 9 = Other	International Mobility
v4.7.1	What country do you plan to go to for your temporary studies abroad?	[open answer or drop down list with countries]	International Mobility
v4.7.2	What country do you plan to go to for your temporary studies abroad? - **Don't know yet**	1 = Yes 0 = No	International Mobility
v4.8	Within which of the following organisational frameworks do you plan to organize your temporary study abroad?	1 = EU-programme (e.g. Erasmus, other …..)2 = Other programme (e.g. national programmes)3 = No programme4 = I don't know yet	International Mobility
v4.9a	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Insufficient skills in foreign language**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9b	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Lack of information provided by my HE institution**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9c	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Separation from partner, child(ren), friends**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9d	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Additional financial burden**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility

v4.9e	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Loss of paid job**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9f	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Lack of motivation**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9g	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Low benefit for my studies at home**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9h	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Difficult integration of studying abroad into the structure of my home study programme**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9i	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Problems with recognition of results achieved abroad**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.9j	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Problems with access regulations to the preferred country (visa, residence permit)**	1 = Big obstacle2 = -3 = -4 = -5 = No obstacle	International Mobility
v4.9k	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Insufficient marks for studying abroad**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility

v4.9l	To what extent are or were the following aspects an obstacle for studying abroad to you? - **Limited admittance to mobility programmes (of home/host institution)**	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	International Mobility
v4.10.1a	Have you ever been abroad for other study-related activities as a student in higher education? - **Research / Fieldtrip**	1 = Yes 0 = No	International Mobility
v4.10.1b	Have you ever been abroad for other study-related activities as a student in higher education? - **Internship / work placement**	1 = Yes 0 = No	International Mobility
v4.10.1c	Have you ever been abroad for other study-related activities as a student in higher education? - **Summer/winter school**	1 = Yes 0 = No	International Mobility
v4.10.1d	Have you ever been abroad for other study-related activities as a student in higher education? - **Language course**	1 = Yes 0 = No	International Mobility
v4.10.1e	Have you ever been abroad for other study-related activities as a student in higher education? - **Other**	1 = Yes 0 = No	International Mobility
v4.10.2a	Duration in months - **Research / Fieldtrip**	[Duration in months]	International Mobility
v4.10.2b	Duration in months - **Internship / work placement**	[Duration in months]	International Mobility
v4.10.2c	Duration in months - **Summer/winter school**	[Duration in months]	International Mobility
v4.10.2d	Duration in months - **Language course**	[Duration in months]	International Mobility
v4.10.2e	Duration in months - **Other**	[Duration in months]	International Mobility
v4.10.3a	Country - **Research / Fieldtrip**	[open answer or drop down list with countries]	International Mobility
v4.10.3b	Country - **Internship / work placement**	[open answer or drop down list with countries]	International Mobility
v4.10.3c	Country - **Summer/winter school**	[open answer or drop down list with countries]	International Mobility
v4.10.3d	Country - **Language course**	[open answer or drop down list with countries]	International Mobility
v4.10.3e	Country - **Other**	[open answer or drop down list with countries]	International Mobility
v4.10.4a	Was it part of your study programme? - **Research / Fieldtrip**	1 = Yes 0 = No	International Mobility
v4.10.4b	Was it part of your study programme? - **Internship / work placement**	1 = Yes 0 = No	International Mobility
v4.10.4c	Was it part of your study programme? - **Summer/winter school**	1 = Yes 0 = No	International Mobility
v4.10.4d	Was it part of your study programme? - **Language course**	1 = Yes 0 = No	International Mobility
v4.10.4e	Was it part of your study programme? - **Other**	1 = Yes 0 = No	International Mobility
v4.10.5a	Did you gain ECTS with it? - **Research / Fieldtrip**	1 = Yes 0 = No	International Mobility

v4.10.5b	Did you gain ECTS with it? - **Internship / work placement**	1 = Yes 0 = No	International Mobility
v4.10.5c	Did you gain ECTS with it? - **Summer/winter school**	1 = Yes 0 = No	International Mobility
v4.10.5d	Did you gain ECTS with it? - **Language course**	1 = Yes0 = No	International Mobility
v4.10.5e	Did you gain ECTS with it? - **Other**	1 = Yes 0 = No	International Mobility
v5.1year	When were you born?	[year of birth]	Personal Details
v5.1month	When were you born?	[month of birth]	Personal Details
v5.2	What is your sex?	1 = Female 2 = Male	Personal Details
v5.3self	In which country were you born?	[open answer or drop down list with countries]	Personal Details
v5.3mother	In which country was your mother born?	[open answer or drop down list with countries]	Personal Details
v5.3father	In which country was your father you born?	[open answer or drop down list with countries]	Personal Details
v5.4a	What are your written and spoken language skills? - **[Official language(s) in your country]**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details
v5.4b	What are your written and spoken language skills? - **English**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details
v5.4c	What are your written and spoken language skills? - **[Other common language in your country]**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details
v5.4d	What are your written and spoken language skills? - **[Other common language in your country]**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details

v5.4e	What are your written and spoken language skills? - **[Other living language 1]**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details
v5.4f	What are your written and spoken language skills? - **[Other living language 2]**	1 = Mother tongue 2 = Very good 3 = - 4 = - 5 = - 6 = No knowledge	Personal Details
v5.4g	What are your written and spoken language skills? - **[Other living language 3]**	1 = Mother tongue2 = Very good3 = -4 = -5 = -6 = No knowledge	Personal Details
v5.5.1	Do you have any children, if yes how many?	1 = Yes 0 = No	Personal Details
v5.5.2	How many children do you have?	[number of children]	Personal Details
v5.6	How old is your youngest child?	[age in years]	Personal Details
v5.7a	Please indicate if you have a disability, long standing health problems or functional limitations - **chronic diseases**	1 = Yes 0 = No	Personal Details
v5.7b	Please indicate if you have a disability, long standing health problems or functional limitations - **mental health problems**	1 = Yes 0 = No	Personal Details
v5.7c	Please indicate if you have a disability, long standing health problems or functional limitations - **mobility impairment**	1 = Yes 0 = No	Personal Details
v5.7d	Please indicate if you have a disability, long standing health problems or functional limitations - **sensory impairment (vision or hearing)**	1 = Yes 0 = No	Personal Details
v5.7e	Please indicate if you have a disability, long standing health problems or functional limitations - **learning disability (ADHD, Dyslexia)**	1 = Yes 0 = No	Personal Details
v5.7f	Please indicate if you have a disability, long standing health problems or functional limitations - **other long standing health problems**	1 = Yes 0 = No	Personal Details
v5.7g	Please indicate if you have a disability, long standing health problems or functional limitations - **No**	1 = Yes [no impairment] 0 = No	Personal Details

v5.8	Overall, to what extent are your impairments an obstacle to your studies?	1 = Big obstacle 2 = - 3 = - 4 = - 5 = No obstacle	Personal Details
v5.9	Overall, how well are you publicly or institutionally supported with regard to your impairments in your studies?	1 = very well 2 = - 3 = - 4 = - 5 = not at all 6 = I do not need / want any support	Personal Details
v6.1father	What is the highest level of education your father has obtained?	1 = Up to lower secondary (ISCED 0, 1, 2) 2 = Upper secondary (ISCED 3) 3 = Post-secondary non-tertiary (ISCED 4) 4 = Short-cycle tertiary education (ISCED 5) 5 = Bachelor or equivalent (ISCED 6) 6 = Master or equivalent (ISCED 7) 7 = Doctoral or equivalent (ISCED 8) 8 = Do not know	Family Background
v6.1mother	What is the highest level of education your mother has obtained?	1 = Up to lower secondary (ISCED 0, 1, 2)2 = Upper secondary (ISCED 3)3 = Post-secondary non-tertiary (ISCED 4)4 = Short-cycle tertiary education (ISCED 5)5 = Bachelor or equivalent (ISCED 6)6 = Master or equivalent (ISCED 7)7 = Doctoral or equivalent (ISCED 8)8 = Do not know	Family Background
v6.2father	What is the present or most recent occupation of your father?	1 = Managers 2 = Professionals 3 = Technicians and associate professionals 4 = Clerical support workers 5 = Service and sales workers 6 = Skilled agricultural, forestry and fishery workers 7 = Craft and related trades workers 8 = Plant and machine operators, and assemblers 9 = Elementary occupations 10 = Armed forces occupations 11 = Do not know 12 = Never been active in the labour market	Family Background

v6.2mother	What is the present or most recent occupation of your mother?	1 = Managers 2 = Professionals 3 = Technicians and associate professionals 4 = Clerical support workers 5 = Service and sales workers 6 = Skilled agricultural, forestry and fishery workers 7 = Craft and related trades workers 8 = Plant and machine operators, and assemblers 9 = Elementary occupations 10 = Armed forces occupations 11 = Do not know 12 = Never been active in the labour market	Family Background
v6.3	Thinking about your family background, where would you place your parents on this scale if the top indicated high social standing and the bottom indicated low social standing?	1 = High social standing 2 = - 3 = - 4 = - 5 = - 6 = - 7 = - 8 = - 9 = - 10 = Low social standing	Family Background