

# SO5041 Assignment 1, Autumn 2023/4

Autumn 2023

## Instructions

Answer each question. Use a wordprocessor to prepare your assignment and, where relevant, cut and paste material from Stata into the wordprocessor document. Read the questions carefully to make sure you are answering everything.

*Deadline* 4:00pm Friday October 13. Please submit assignments via Sulis, as a **single** PDF, Open Document Format or Word file, and please don't forget the cover-sheet declaration at <http://www.ulsites.ul.ie/sociology/student-resources>

## 1 “Levels of Measurement”

A certain data set contains the following variables:

- (a) Sex (1 = male; 2 = female)
- (b) Age group (1 = 16–24, 2 = 25–39, 3 = 40–59, 4 =60+)
- (c) Income (in 1000 Euros per annum)
- (d) Percentage of income saved each month
- (e) Party political affiliation (1 = Sinn Féin; 2 = Fine Gael; 3 = Fianna Fáil; 4 = Other; -9 = No preference)
- (f) Frequency of newspaper reading (1 = more than one per day; 2 = one per day; 3 = less than 1/day but more than once per week; 4 = once per week; 5 = less than once per week; 6 = do not read newspapers)

For each variable,

1. indicate at which of the following levels of measurement it is:
  - Nominal
  - Ordinal
  - Interval
  - Ratio
2. briefly explain why and
3. suggest a suitable way of summarising the variable

## 2 Data entry

The data below is an excerpt from the 1999 School Leavers Survey. This survey is taken of people leaving second level education, about a year and half after they have left.

Sex is coded 1 for male and 2 for female, and the “Last exam taken” variable codes the last certificate examination sat as 1 – Junior Cert, 2 – Leaving Cert, 3 – LC Applied/Vocational, 4 – none, -9 – don’t know/missing.

	Sex	Last exam taken
1.	2	3
2.	2	2
3.	1	2
4.	2	2
5.	2	1
6.	1	2
7.	1	4
8.	2	2
9.	1	1
10.	2	4
11.	1	2
12.	1	2
13.	1	1
14.	1	-9
15.	1	3
16.	2	2
17.	2	2
18.	1	1
19.	1	3
20.	1	4
21.	2	2
22.	2	1
23.	1	1
24.	1	2
25.	2	1
26.	2	2
27.	1	2
28.	1	2
29.	2	2
30.	2	4
31.	2	4
32.	1	1
33.	1	2
34.	2	1
35.	1	-9
36.	2	3
37.	1	2
38.	1	2
39.	2	2
40.	1	1

Using Stata, enter this data, using variable labels and value labels based on the information at the

start of this question. Take account of the missing values (-9), and create a cross-tabulation. Include the output from the cross-tabulation in your answer, and make an assessment whether or not there is a relationship between gender and last exam taken.

Would you expect to see such a relationship? If so, what form should it take?

### 3 Univariate Summaries

Download the file `assign1.dta` from <http://teaching.sociology.ul.ie/so5041/assign1.dta> (or access it from within Stata with the command `use http://teaching.sociology.ul.ie/so5041/assign1`). This is a larger extract from the Dept of Education/ESRI School Leavers' survey (for 1993). Variable and value labels have been added.

Variable List	
SEX	School Leavers Sex
EMPSTAT	Usual Employment Situation
STAGELFT	At What Stage of Education did you Leave
LASTEXAM	Last Certificate Examination
GRSEARN	Gross Earnings Last Week
NETEARN	Net Earnings Last Week
FEELEMP	Feeling re Present Employment Situation

For the variables SEX, EMPSTAT, GRSEARN and FEELEMP do the following:

- present a numerical univariate summary
- present an appropriate graphical summary, and
- **briefly describe what the numerical and graphical summaries tell about the variable.**

### 4 Bivariate Summaries

With the same file as in question 3, present numerical and graphical summaries, on the following pairs of variables:

- Sex and usual employment situation
- Sex and net earnings
- Usual employment situation and feelings about present situation
- Gross earnings and net earnings (note: do just the graphical summary, as we haven't covered appropriate numerical summaries yet)

In each case,

- **Briefly indicate** what sort of relationship you might have expected to see, and why
- **Describe what you do see**, and compare it with your expectations.

## **5 Sampling**

Write a brief note (c.300–500 words) on the role of sampling in quantitative research.

## **6 Earnings and age**

With the data in <http://teaching.sociology.ul.ie/so5041/ass1q6.dta>, conduct a simple analysis of how income varies over the life-cycle (do use <http://teaching.sociology.ul.ie/so5041/ass1q6> in Stata). Do this by breaking age into groups and calculating the mean income in each group. Take account of missing values. Interpret the pattern you observe.

Include in your answer the Stata commands used (this should not be more than a few lines).