

Assignment 2 Answers

This assignment focuses on graphs and distributions; it is based on the information covered in module 3.

First, do the steps you need to do in Stata (parts A-D). Using the General Social Survey 2012 data (gss2012.dta), conduct the following analyses. Make sure to open a log file before starting (use .log extension). Upon finishing, open your Stata log file in Word, insert graphs into that same file (make sure to export the graphs as .png right after generating them; do not use screenshots), and save it as Word file. You will submit that file on Canvas together with your answers to the question (see below).

Next, answer the questions below; record your answers in this word file. When you are done, please save the resulting file and submit that word file together with the file containing your log and graphs to Canvas.

Part A. Get the codebook information and frequency distributions for variables *satfin*, *fefam*, *fechld*, and *educ*. Identify the level of measurement for each variable:

1. *satfin* _____ **ordinal** _____
2. *fefam* _____ **ordinal** _____
3. *fechld* _____ **ordinal** _____
4. *educ* _____ **ratio** _____

Part B. Construct a bar chart for the variable *satfin*. Fill the blanks based on these data.

5. About **29** % of people are not at all satisfied with their financial situation.
6. The largest percentage of people report that they are **more or less** satisfied with their financial situation.

Part C. Construct pie charts for the variables *fefam* and *fechld*; include percentage labels. Answer the following questions.

7. In these two survey questions, are people more likely to agree that men should work and women should tend home, or that mother working doesn't hurt children?
People are more likely to agree with the statement that mother working doesn't hurt children than with the statement that men should work and women should tend home.

8. Based on these two questions, how would you summarize what people think about women's place in the family and in the job market (make sure to use specific percentages from the data in your summary)?

Most respondents think that women should be able to combine their roles in the family with paid employment. The vast majority either disagree (49.49%) or strongly disagree (19.38%) that it is better for men to work and women to tend home. Less than one third of respondents (31.13%) agree with that statement. Similarly, most people agree that mother's employment doesn't hurt children – only a little bit over a quarter (22.52+5.11=27.63%) disagree with that statement.

Part D. Obtain measures of skewness and kurtosis to describe the distribution of variable *educ*. Also construct a histogram and a boxplot for this variable. Fill the blanks and answer the questions below based on these data.

9. Based on measure of skewness, is this distribution skewed, and if so, is it a positive or negative skew?

Skewness statistic (-.398) is pretty small in terms of its absolute value, definitely below 2. So we can conclude that the distribution is approximately normal with regard to its skewness.

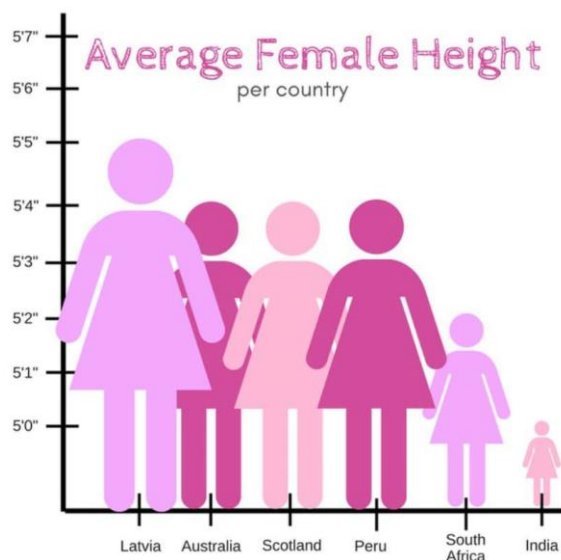
10. Based on the kurtosis statistic, is this distribution leptokurtic, platykurtic, or normal?

Kurtosis value is $4.163187 - 3 = 1.16$, which is below 2. We can say that the distribution is approximately normal with regard to both kurtosis and skewness.

11. Are there outliers in this variable's distribution, and if so, which observations are those? Which best helps us determine that?

Yes, there are outliers – those are observations with fewer than 6 years of education. We can clearly see them in the boxplot (better than in a histogram).

Part E. Please carefully examine this graph:



12. Is this a univariate or a bivariate graph?

Bivariate – variables are country and height.

13. This is a stylized version of what kind of graph (bar, pie, line, boxplot, histogram, stem-and-leaf, frequency polygon, ogive)?

Bar graph.

14. What are the two key ways in which this graph is misrepresenting the data?

Y axis doesn't start from 0 so the differences look larger than they are; proportional ink principle is violated – bars are different width which further misrepresents the size of these differences.

Log file:

```
-----
log:  A2.log
log type:  text
opened on:  24 Sep 2025, 18:48:37

. codebook satfin fefam fechl d educ
-----
satfin                                SATISFACTION WITH FINANCIAL SITUATION
-----
                                Type: Numeric (byte)
                                Label: SATFIN

                                Range: [1,3]                                Units: 1
                                Unique values: 3                                Missing .: 7/1,974

                                Tabulation: Freq.    Numeric    Label
                                                537            1    satisfied
                                                857            2    MORE OR LESS
                                                573            3    NOT AT ALL SAT
                                                7              .

-----
fefam                                BETTER FOR MAN TO WORK, WOMAN TEND HOME
-----
                                Type: Numeric (byte)
                                Label: LABAG

                                Range: [1,4]                                Units: 1
                                Unique values: 4                                Missing .: 689/1,974

                                Tabulation: Freq.    Numeric    Label
                                                87             1    STRONGLY AGREE
                                                313            2    agree
                                                636            3    disagree
                                                249            4    STRONGLY DISAGREE
                                                689            .

-----
fechl d                                MOTHER WORKING DOESNT HURT CHILDREN
-----
                                Type: Numeric (byte)
                                Label: LABAG

                                Range: [1,4]                                Units: 1
```

Unique values: 4

Missing .: 682/1,974

Tabulation:	Freq.	Numeric	Label
	342	1	STRONGLY AGREE
	593	2	agree
	291	3	disagree
	66	4	STRONGLY DISAGREE
	682	.	

educ

HIGHEST YEAR OF SCHOOL COMPLETED

Type: Numeric (byte)

Label: LABAB, but 21 nonmissing values are not labeled

Range: [0,20]

Units: 1

Unique values: 21

Missing .: 2/1,974

Examples: 12

12

14

16

. tabl satfin fefam fechld educ

-> tabulation of satfin

SATISFACTION WITH FINANCIAL SITUATION	Freq.	Percent	Cum.
satisfied	537	27.30	27.30
MORE OR LESS	857	43.57	70.87
NOT AT ALL SAT	573	29.13	100.00
Total	1,967	100.00	

-> tabulation of fefam

BETTER FOR MAN TO WORK, WOMAN TEND HOME	Freq.	Percent	Cum.
STRONGLY AGREE	87	6.77	6.77
agree	313	24.36	31.13
disagree	636	49.49	80.62
STRONGLY DISAGREE	249	19.38	100.00
Total	1,285	100.00	

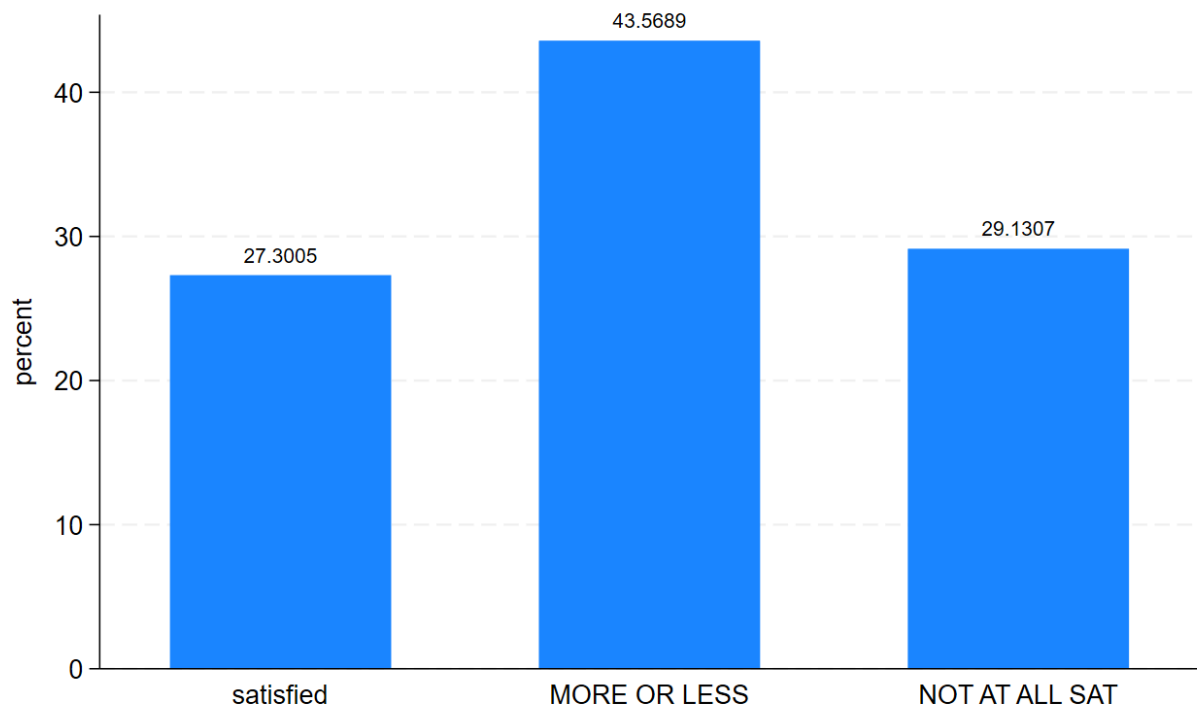
-> tabulation of fechld

MOTHER WORKING DOESNT HURT CHILDREN	Freq.	Percent	Cum.
STRONGLY AGREE	342	26.47	26.47
agree	593	45.90	72.37
disagree	291	22.52	94.89
STRONGLY DISAGREE	66	5.11	100.00
Total	1,292	100.00	

```
-> tabulation of educ
```

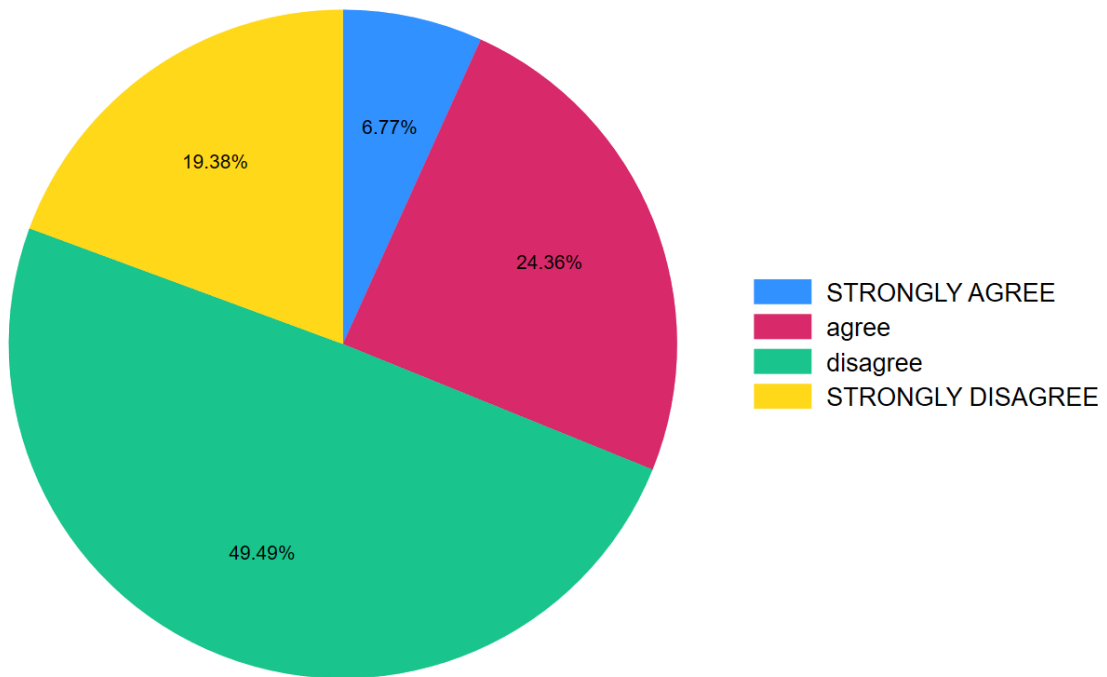
HIGHEST YEAR OF SCHOOL COMPLETED	Freq.	Percent	Cum.
0	3	0.15	0.15
1	2	0.10	0.25
2	3	0.15	0.41
3	6	0.30	0.71
4	10	0.51	1.22
5	4	0.20	1.42
6	27	1.37	2.79
7	8	0.41	3.19
8	48	2.43	5.63
9	47	2.38	8.01
10	59	2.99	11.00
11	101	5.12	16.13
12	540	27.38	43.51
13	163	8.27	51.77
14	261	13.24	65.01
15	99	5.02	70.03
16	307	15.57	85.60
17	80	4.06	89.66
18	92	4.67	94.32
19	41	2.08	96.40
20	71	3.60	100.00
Total	1,972	100.00	

```
. graph bar, over(satfin) blabel(bar)
```



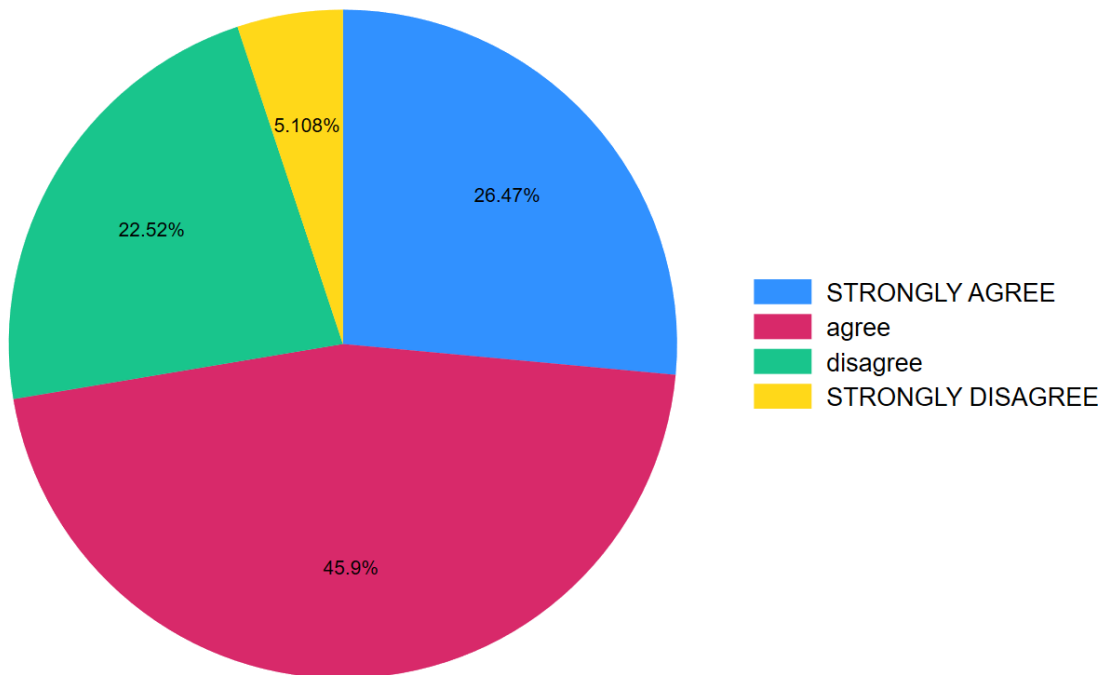
```
. graph export satfin.png, replace
(file satfin.png not found)
file satfin.png saved as PNG format
```

```
. graph pie, over(fefam) plabel(_all percent)
```



```
. graph export fefam.png, replace  
(file fefam.png not found)  
file fefam.png saved as PNG format
```

```
. graph pie, over(fechld) plabel(_all percent)
```

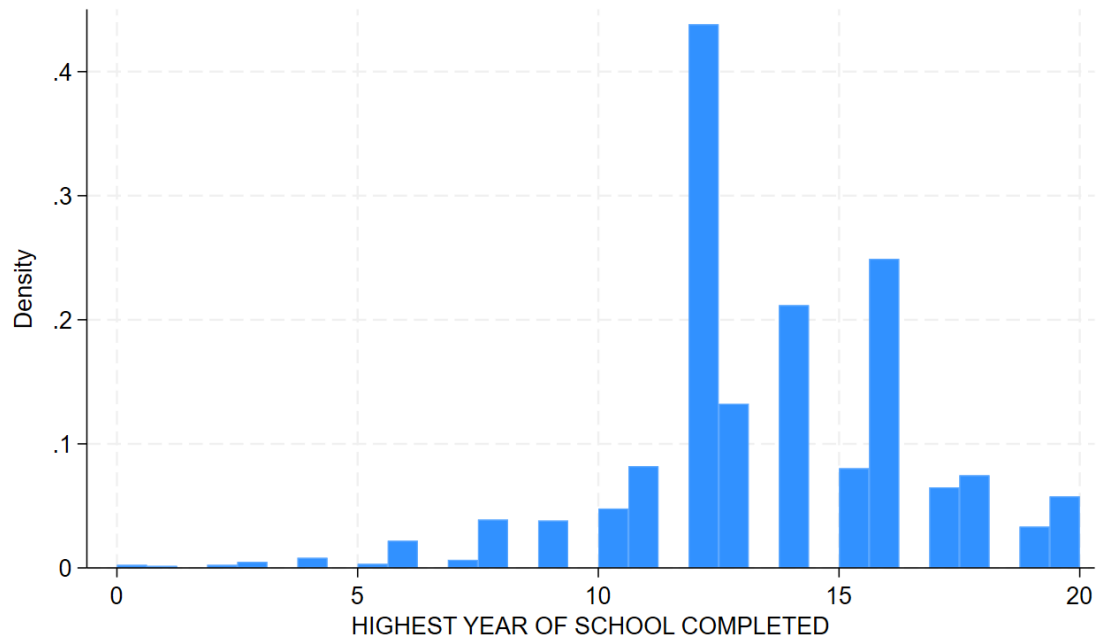


```
. graph export fechld.png, replace  
(file fechld.png not found)  
file fechld.png saved as PNG format
```

```
. tabstat educ, stats(skewness kurtosis)
```

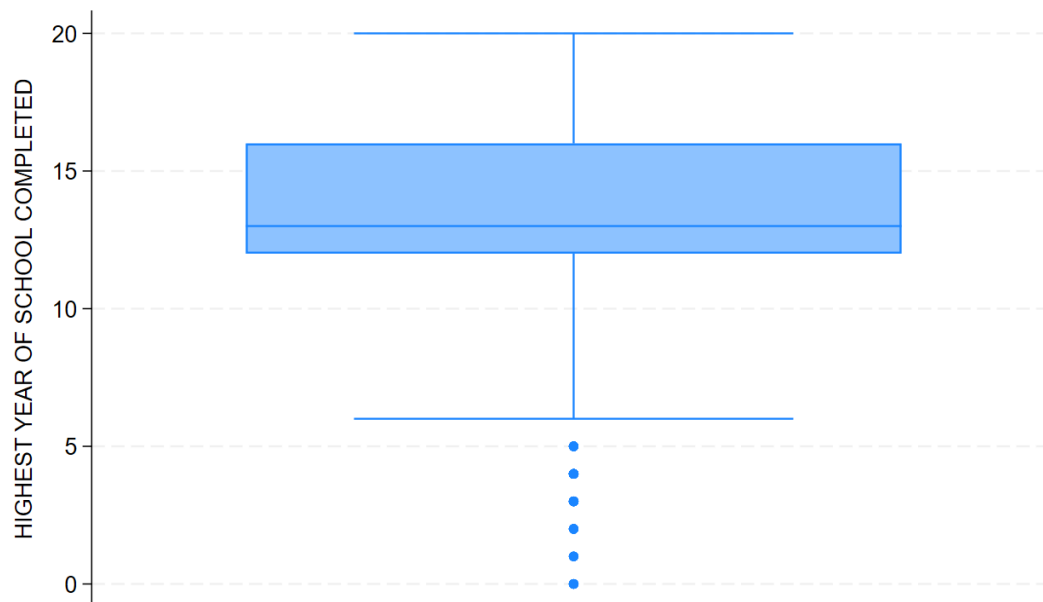
Variable	Skewness	Kurtosis
educ	-.3976432	4.163187

```
. histogram educ
(bin=32, start=0, width=.625)
```



```
. graph export educ_histogram.png, replace
(file educ_histogram.png not found)
file educ_histogram.png saved as PNG format
```

```
. graph box educ
```



```
. graph export educ_box.png, replace  
(file educ_box.png not found)  
file educ_box.png saved as PNG format
```

```
. log close  
    log:  A2.log  
    log type:  text  
closed on:  24 Sep 2025, 18:59:10
```
