

Basic population estimates with British Social Attitudes Survey data using SPSS

UK Data Service

June 2024

This exercise is part of the [‘Introduction to the British Social Attitudes Survey \(BSA\)’](#) online module. In the exercise, we examine data from the 2020 British Social Attitudes survey to find out:

- what proportion of respondents said they voted remain in the EU Referendum?
- whether people think the government should raise taxes and spend more or reduce tax and cut social expenditures?
- how much people think they’ll get from the State pension?

Answers to the questions asked throughout the exercise can be found at the end of the page.

Getting started

Data can be downloaded from the [UK Data Service website](#) following [registration](#). Download the compressed folder, unzip and save it somewhere accessible on your computer.

The examples below assume that the dataset has been saved in a new folder named *UKDS* on your Desktop (Windows computers). The path would typically be `C:\Users\YOUR_USER_NAME\Desktop\UKDS`. Feel free to change it to the location that best suits your needs.

You need to set the folder as your working directory in SPSS. To do this, you need to add the correct file path to the folder on your computer to the code below.

- * Setting up the working directory
- * Change the command below to match yours:

```
cd "C:\Users\YOUR_USER_NAME\Desktop\UKDS".  
show DIRECTORY.
```

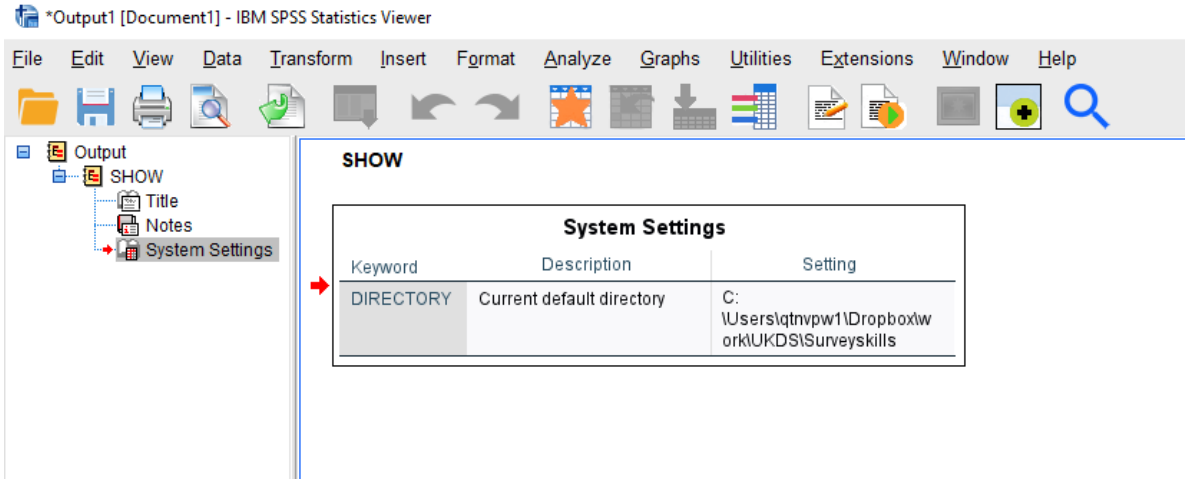


Figure 1: Output of the show DIRECTORY command

If you have your working directory saved to the folder location, the following code should open the BSA dataset.

```
GET FILE='BSA\UKDA-9005-spss\spss\spss25\bsa2020_archive.sav'.
```

The screenshot shows the IBM SPSS Statistics Data Editor interface with the 'bsa2020_archive.sav' dataset open. The 'Variables View' is displayed, showing a list of 22 variables with their properties.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	serial	Numeric	15	0	Serial number	{-9, Refused...	LO --1	17	Right	Scale	Input
2	QnrVersion	Numeric	8	0	Version (1-6)	{-1, Not appl...	LO --1	8	Right	Scale	Input
3	RespSx2cat	Numeric	3	0	Sex of respond...	{-1, Not appl...	LO --1	12	Right	Nominal	Input
4	RespAgeE	Numeric	8	0	Age last birthda...	{-1, Not appl...	LO --1	10	Right	Nominal	Input
5	MarStat6	Numeric	3	0	Marital status	{-1, Not appl...	LO --1	8	Right	Nominal	Input
6	REconFW01	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
7	REconFW02	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
8	REconFW03	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
9	REconFW04	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
10	REconFW05	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
11	REconFW06	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
12	REconFW07	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
13	REconFW08	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
14	REconFW09	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
15	REconFW10	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
16	REconFW11	Numeric	5	0	Which of these ...	{-1, Not appl...	LO --1	18	Right	Nominal	Input
17	EMPSTAT	Numeric	3	0	The following q...	{-1, Not appl...	LO --1	8	Right	Nominal	Input
18	Employ	Numeric	3	0	How many peo...	{-1, Not appl...	LO --1	8	Right	Nominal	Input
19	Superv	Numeric	3	0	In your job, (do/...	{-1, Not appl...	LO --1	8	Right	Nominal	Input
20	EmpOCC	Numeric	3	0	Which of the fol...	{-1, Not appl...	LO --1	8	Right	Nominal	Input
21	TenureE	Numeric	3	0	Do you own or r...	{-1, Not appl...	LO --1	9	Right	Nominal	Input
22	SupParty	Numeric	3	0	Generally spea...	{-1, Not appl...	LO --1	8	Right	Nominal	Input

Figure 2: BSA dataset in SPSS Variables View

1. Explore the dataset

Start by getting an overall feel for the dataset. Either inspect variables and cases in the data editor or use the code below to produce a summary of all the variables in the dataset.

```
CODEBOOK a11.
```

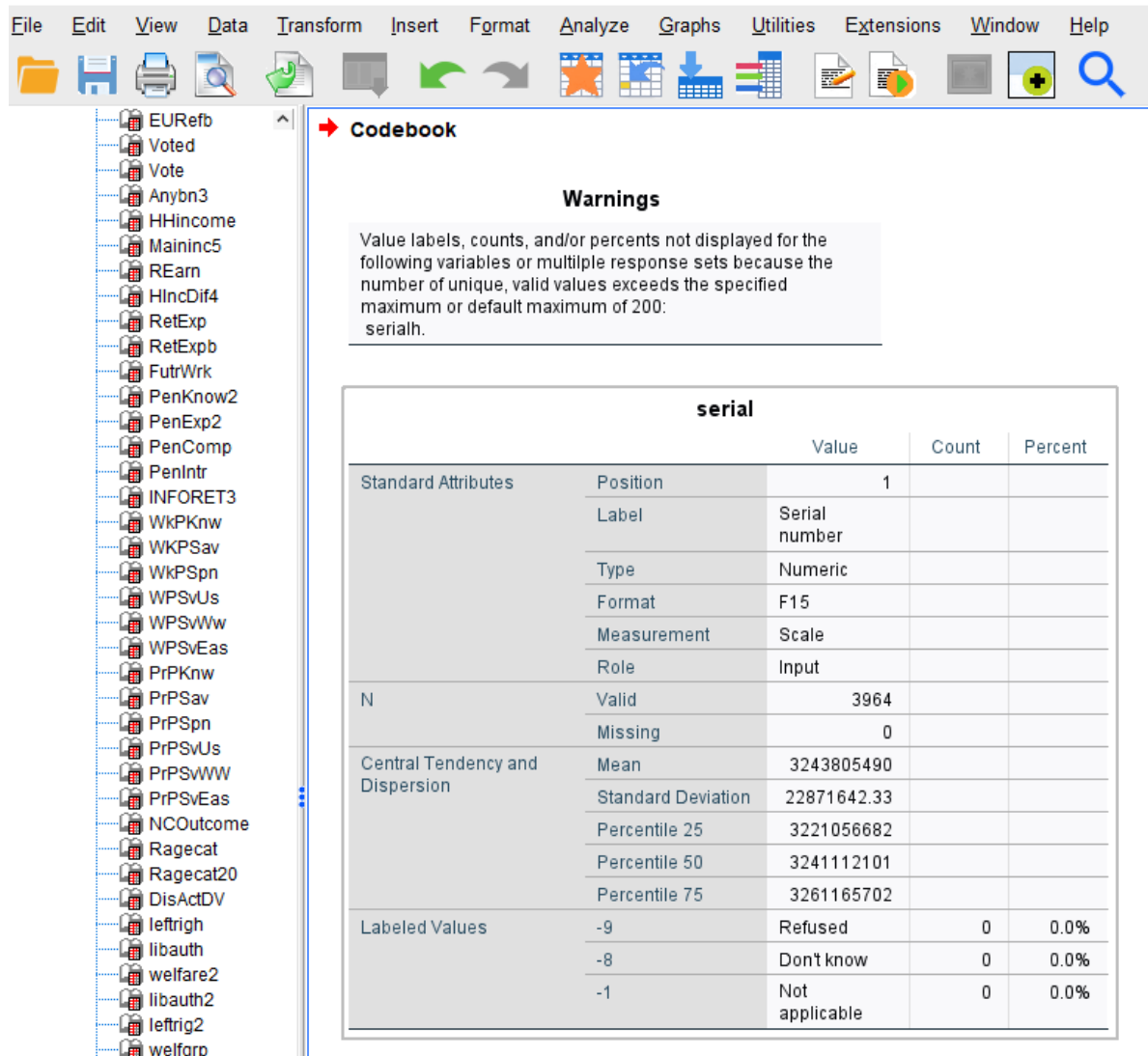


Figure 3: SPSS codebook output for the first variables

Questions

1. What is the overall sample size? 2. How many variables are in the dataset?

Now, focus on the three variables we will use.

CODEBOOK TAXSPEND EUVOTWHO PenExp2.

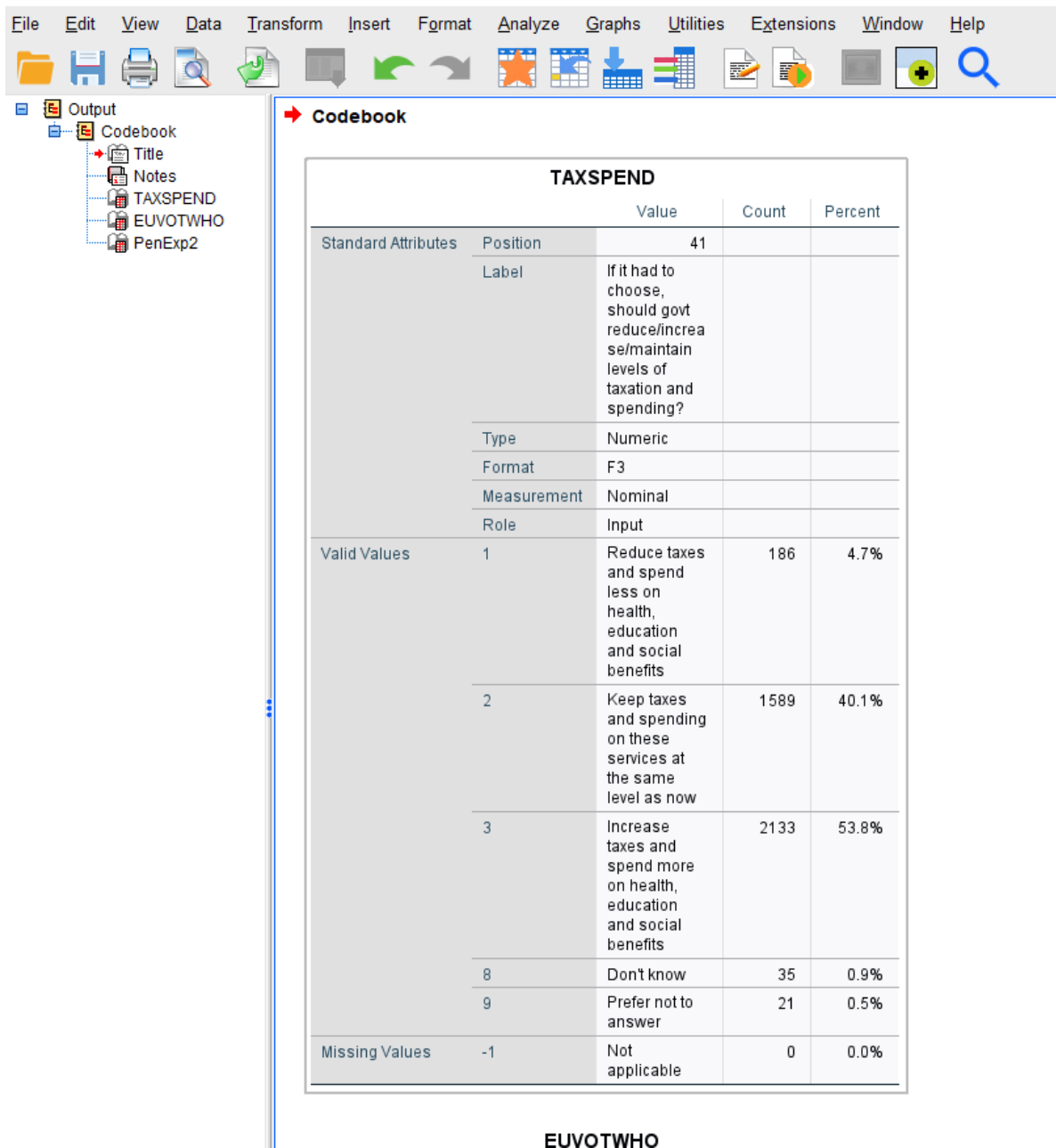


Figure 4: SPSS codebook output for TAXSPEND

Questions 3

What do the variables measure and how?

2. Missing values

Review the frequency tables, examining the not applicable and don't know categories.

Question 4

Why for EUVOTWHO are there so many not applicable? Note, you can use the documentation to check if needed. What does this mean when it comes to interpreting the percentages?

When analysing survey data, it is sometimes convenient to recode item nonresponses such as 'Don't know' and 'Prefer not to say' as system missing so that they do not appear in the results. An example of the syntax required to achieve this with EUVOTWHO and TAXSPEND is provided in the appendix.

Unlike some other surveys, 'Don't knows' and 'Does not apply' were not removed when weights were computed in the BSA. As a result, analyses using weights (ie when planning to use the data to make inference about the British population) need to retain these observations, otherwise estimated results might be incorrect.

3. Compare unweighted and weighted frequencies

Let's examine the weighted responses.

WEIGHT Off.

*This line is probably not unnecessary as we have not applied a weight yet; it has been included.
FREQUENCIES VARIABLES=TAXSPEND EUVOTWHO

/BARChart PERCENT

/ORDER=ANALYSIS.

*Here, we use the FREQUENCIES command for the categorical variables and the EXAMINE command for the continuous variables.
EXAMINE VARIABLES=PenExp2

/PLOT HISTOGRAM

/STATISTICS DESCRIPTIVES

/MISSING LISTWISE

/NOTOTAL.

Frequencies

		Statistics	
		If it had to choose, should govt reduce/increase/maintain levels of taxation and spending?	Did you vote to 'remain a member of the EU' or to 'leave the EU'?
N	Valid	2963	2246
	Missing	1025	1742

Frequency Table

If it had to choose, should govt reduce/increase/maintain levels of taxation and spending?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Reduce taxes and spend less on health, education and social benefits	118	3.0	4.0	4.0
	Keep taxes and spending on these services at the same level as now	928	23.3	31.3	35.3
	Increase taxes and spend more on health, education and social benefits	1805	45.3	60.9	96.2
	(None)	80	2.0	2.7	98.9
	Don't know	31	.8	1.0	100.0
	Refusal	1	.0	.0	100.0
	Total	2963	74.3	100.0	
Missing	skip, version off route	1025	25.7		
Total		3988	100.0		

Did you vote to 'remain a member of the EU' or to 'leave the EU'?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Remain a member of the European Union	1163	29.2	51.8	51.8
	Leave the European Union	1046	26.2	46.6	98.4
	Prefer not to say	13	.3	.6	98.9
	Don't remember	7	.2	.3	99.2
	DK/Refusal	17	.4	.8	100.0
	Total	2246	56.3	100.0	
Missing	skip, version off route	1002	25.1		
	Item not applicable	740	18.6		

Figure 5: SPSS output for Frequency distribution of TAXSPEND and EUVOTWHO

What is the (unweighted) percent who say they voted remain in the EU referendum? The answer is about 58 percent of those who voted in the referendum say they voted to remain. This figure seems a bit high (though people do not always report accurately).

Let's add the weight.

*The weight is added by the command below. It will remain on for all subsequent analyses.
WEIGHT BY BSA20_wt_new.

```
FREQUENCIES VARIABLES=TAXSPEND EUVOTWHO
```

```
  /ORDER=ANALYSIS.
```

```
EXAMINE VARIABLES=PenExp2
```

```
  /PLOT HISTOGRAM
```

```
  /STATISTICS DESCRIPTIVES
```

```
  /CINTERVAL 95
```

```
  /MISSING LISTWISE
```

```
  /NOTOTAL.
```

*To stop weighting the data you can use the following command.

```
WEIGHT off.
```


Descriptives				
		Statistic	Std. Error	
How much do you think someone who reaches State Pension age today would receive in pounds per week?	Mean	1292.65	58.185	
	95% Confidence Interval for Mean	Lower Bound	1178.56	
		Upper Bound	1406.73	
	5% Trimmed Mean	878.12		
	Median	160.00		
	Variance	9777451.114		
	Std. Deviation	3126.892		
	Minimum	0		
	Maximum	9999		
	Range	9999		
	Interquartile Range	80		
	Skewness	2.421	.046	
	Kurtosis	3.877	.091	

How much do you think someone who reaches State Pension age today would receive in pounds per week?

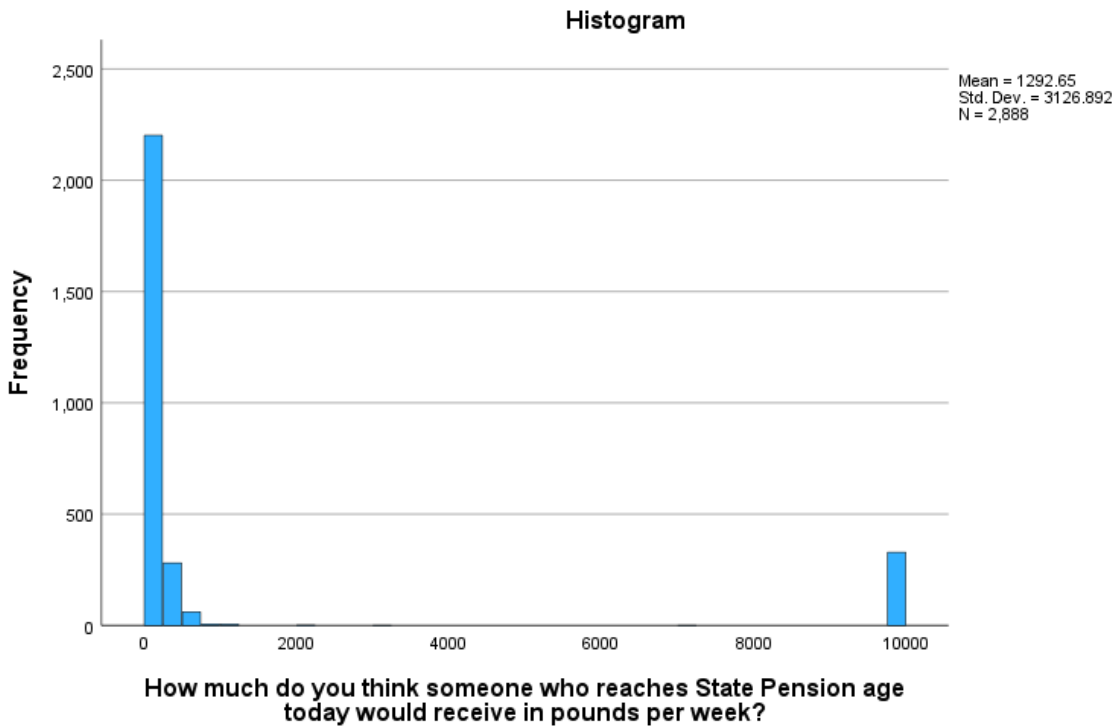


Figure 6: SPSS output for Frequency distribution of TAXSPEND and EUVOTWHO

Now, what proportion say they voted remain in the EU referendum? It is about 54 percent,

lower than the unweighted proportion and closer to the actual referendum results.

4. Confidence intervals

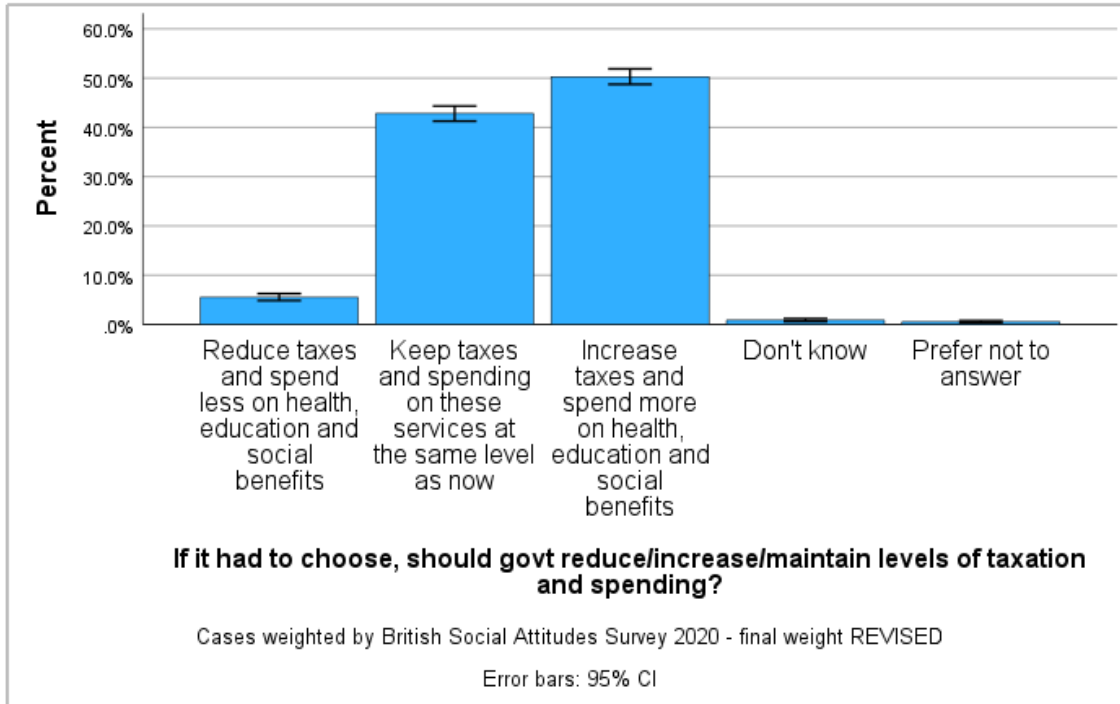
Add confidence intervals to the bar charts and mean to indicate uncertainty due to sampling error.

```
WEIGHT BY BSA20_wt_new.  
GRAPH  
  /BAR(SIMPLE)=PCT BY TAXSPEND  
  /INTERVAL CI(95.0).
```

```
GRAPH  
  /BAR(SIMPLE)=PCT BY EUVOTWHO  
  /INTERVAL CI(95.0).
```

```
EXAMINE VARIABLES=PenExp2  
  /PLOT NONE  
  /STATISTICS DESCRIPTIVES  
  /CINTERVAL 95  
  /MISSING LISTWISE  
  /NOTOTAL.
```

Graph



Graph

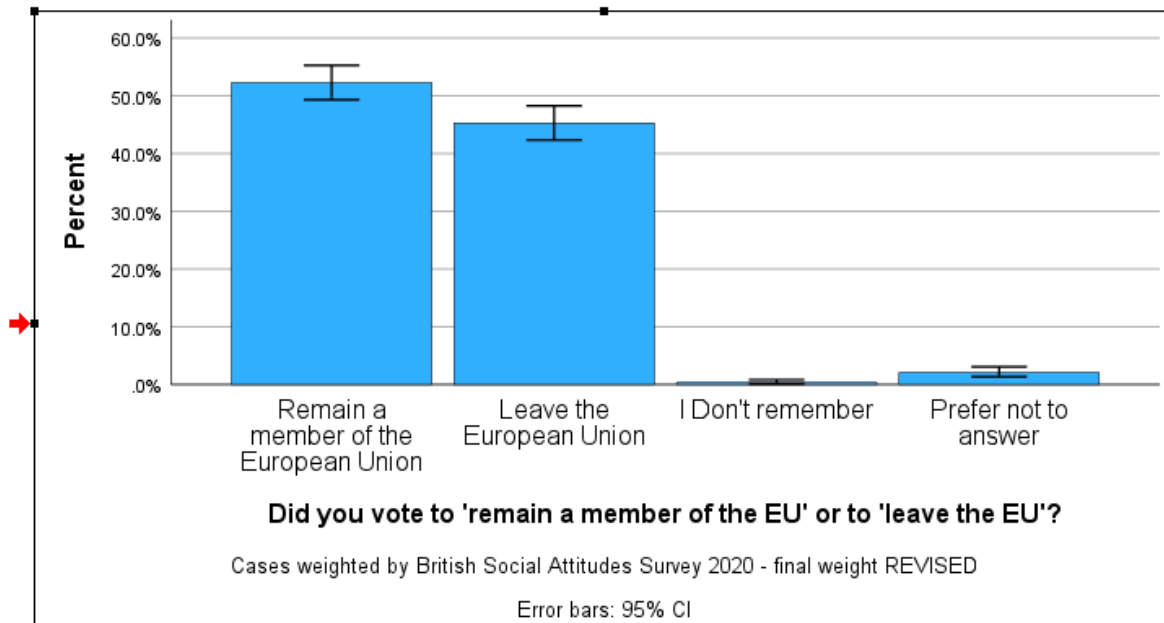


Figure 7: SPSS output for GRAPH BAR of TAXSPEND and EUVOTWHO

Question 5

What proportion think government should increase taxes and spend more on health, education and social benefits?

Question 6

How much do people think they will get at state pension age?

Additional question

Select two variables that interest you and examine their distribution.

Answers

1. There are 3964 cases in the dataset.
2. The total number of variables is 213.
3. TAXSPEND records responses to the questions of whether government should reduce/increase/maintain levels of taxation and spending? There are three possible responses to the question. EUVOTWHO records responses to the question 'Did you vote to 'remain a member of the EU' or to 'leave the EU'?' The responses are Remain or Leave . PenExp2 contains responses to the question 'How much do you think someone who reaches State Pension age today would receive in pounds per week?' Responses are numeric.
4. There are two reasons for the many 'not applicable'.
 - Routing: the question is only asked to those who said yes to a previous question (EURefV2).
 - Versions 5 and 6 - The BSA uses a split sample and the question is only asked in Versions 5 and 6.
5. About 50% say the government should increase taxes and spend more.
6. The amount people think they will get at state pension age varies between £0 and £7000, with an average in the region between £170 and £184

Appendix: recoding nonresponses as system missing

The code below provides an example of how to recode missing values including 'don't knows' and 'prefer not to say' into system missing.

The SPSS syntax below includes the command, the variables and the relevant missing values in (). Note, you can set missing values more than 1 at a time if they have the same missing value pattern.

```
COMPUTE EUVOTWHO_m=EUWOTWHO.
```

```
COMPUTE TAXSPEND_m=TAXSPEND.
```

```
COMPUTE PenExp2_m=PenExp2.
```

```
MISSING VALUES PenExp2_m (-1, 9998, 9999).
```

```
MISSING VALUES TAXSPEND_m (-1, 8, 9).
```

```
MISSING VALUES EUVOTWHO_m (-1, 3 THRU 9).
```