Descriptive Statistics in SPSS (Quiz)

Descriptive statistics for categorical variables practical questions

In this practical you will explore the distribution of two more categorical variables from the PISA dataset. The first, INFNUCL, is an alternative indicator of the student's environmental awareness in which they were asked to rate how informed they felt about the issue of *nuclear waste*. The second, PARINTSCH, measures parental involvement rather than parental educational status, and gives the students" rating of agreement with the statement "My parents are interested in my school activities".

Use the Frequencies window in SPSS to answer the following questions about the distribution of the variable **INFNUCL**, How informed about nuclear waste:

- Question: How many occurrences are there of Know something, could explain in general?
- Question: What percentage of the dataset is in the category Never heard of this?
- Question: What percentage of valid data (i.e. without missing values) is in category Never heard of this?
- Question: What percentage of values are in the first 2 categories combined?

Use the Frequencies window in SPSS to answer the following questions about the distribution of the variable **PARINTSCH**, Parents are interested in school activities:

- Question: How many occurrences are there of Disagree?
- Question: What percentage of the dataset is in the category Agree?
- Question: What percentage of valid data (i.e. without missing values) is in category Disagree?
- Question: What percentage of values are in the first 3 categories combined?

Use SPSS to draw a bar chart and use it to answer the following questions about the variable **INFNUCL**:

- Question: You will see 4 categories. Approximately how many observations are in category Heard of, but not really able to explain this?

Use SPSS to draw a pie chart and use it to answer the following question about the variable **INFNUCL**:

- Question: Which is the most common category?

Use SPSS to draw a bar chart and use it to answer the following questions about the variable **PARINTSCH**:

- Question: You will see 4 categories. Approximately how many observations are in category Disagree?

Use SPSS to draw a clustered bar chart and use it to answer the following questions:

- Question: Approximately how many observations take values Heard of, but not really able to explain this for **INFNUCL** and Strongly disagree for **PARINTSCH**?
- Question: Approximately how many observations take values Know something, could explain in general for **INFNUCL** and Strongly agree for **PARINTSCH**?
- Question: Approximately how many observations take values Familiar with this, could explain well for **INFNUCL** and Agree for **PARINTSCH**?

Use SPSS to draw a clustered bar chart of percentages of cases and use it to answer the following questions:

- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Disagree take value Know something, could explain in general for **INFNUCL**?
- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Agree take value Never heard of this for **INFNUCL**?
- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Strongly disagree take value Familiar with this, could explain well for **INFNUCL**?

Solutions to Descriptive statistics for categorical variables practical questions

The SPSS instructions for the first question on frequencies are as follows:

- Select **Frequencies** from the **Descriptive Statistics** submenu available from the **Analyze** menu.
- Copy the **How informed about nuclear waste[INFNUCL]** variable into the **Variable(s)** box.
- Click on the **OK** button to produce the tables.

- Question: How many occurrences are there of Know something, could explain in general?
- Question: What percentage of the dataset is in the category Never heard of this?
- Question: What percentage of valid data (i.e. without missing values) is in category Never heard of this?
- Question: What percentage of values are in the first 2 categories combined?

Solutions: The output from SPSS is as follows:

Statistics

How informed about nuclear waste

N	Valid	4803
	Missing	391

How informed about nuclear waste

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Never heard of this	191	3.7	4.0	4.0
	Heard of, but not really able to explain this	1572	30.3	32.7	36.7
	Know something, could explain in general	1920	37.0	40.0	76.7
	Familiar with this, could explain well	1120	21.6	23.3	100.0
	Total	4803	92.5	100.0	
Missing	System	391	7.5		
Total		5194	100.0		

The answers are as follows:

- There are 1920 occurrences of Know something, could explain in general.
- 3.7 percent of the dataset is in category Never heard of this.
- 4.0 percent of valid data is in category Never heard of this.
- 36.7 percent of values are in the first 2 categories combined.

The SPSS instructions for the second question on frequencies are as follows:

- Select **Frequencies** from the **Descriptive Statistics** submenu available from the **Analyze** menu.
- Remove the **How informed about nuclear waste[INFNUCL]** variable into the **Variable(s)** box.
- Copy the **Parents are interested in school activities[PARINTSCH]** variable into the **Variable(s)** box.
- Click on the **OK** button to produce the tables.

- Question: How many occurrences are there of Disagree?
- Question: What percentage of the dataset is in the category Agree?
- Question: What percentage of valid data (i.e. without missing values) is in category Disagree?
- Question: What percentage of values are in the first 3 categories combined?

Solutions: The output from SPSS is as follows:

Statistics

Parents are interested in school activities

N	Valid	5052
	Missing	142

Parents are interested in school activities

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Strongly disagree	118	2.3	2.3	2.3
	Disagree	190	3.7	3.8	6.1
	Agree	2150	41.4	42.6	48.7
	Strongly agree	2594	49.9	51.3	100.0
	Total	5052	97.3	100.0	
Missing	System	142	2.7		
Total		5194	100.0		

The answers are as follows:

- There are 190 occurrences of Disagree.
- 41.4 percent of the dataset is in category Agree.
- 3.8 percent of valid data is in category Disagree.
- 48.7 percent of values are in the first 3 categories combined.

The SPSS instructions for the first bar chart are as follows:

- Select **Bar...** from the **Legacy Dialogs** submenu available from the **Graphs** menu.
- We will use the defaults which are **Simple** and **Summaries for groups of cases** and then click on **Define** .
- Copy the **How informed about nuclear waste[INFNUCL]** variable into the **Category Axis** box.
- Click on the **OK** button to produce the graph.

- Question: You will see 4 categories. Approximately how many observations are in category Heard of, but not really able to explain this?

Solution: The output from SPSS is as follows:
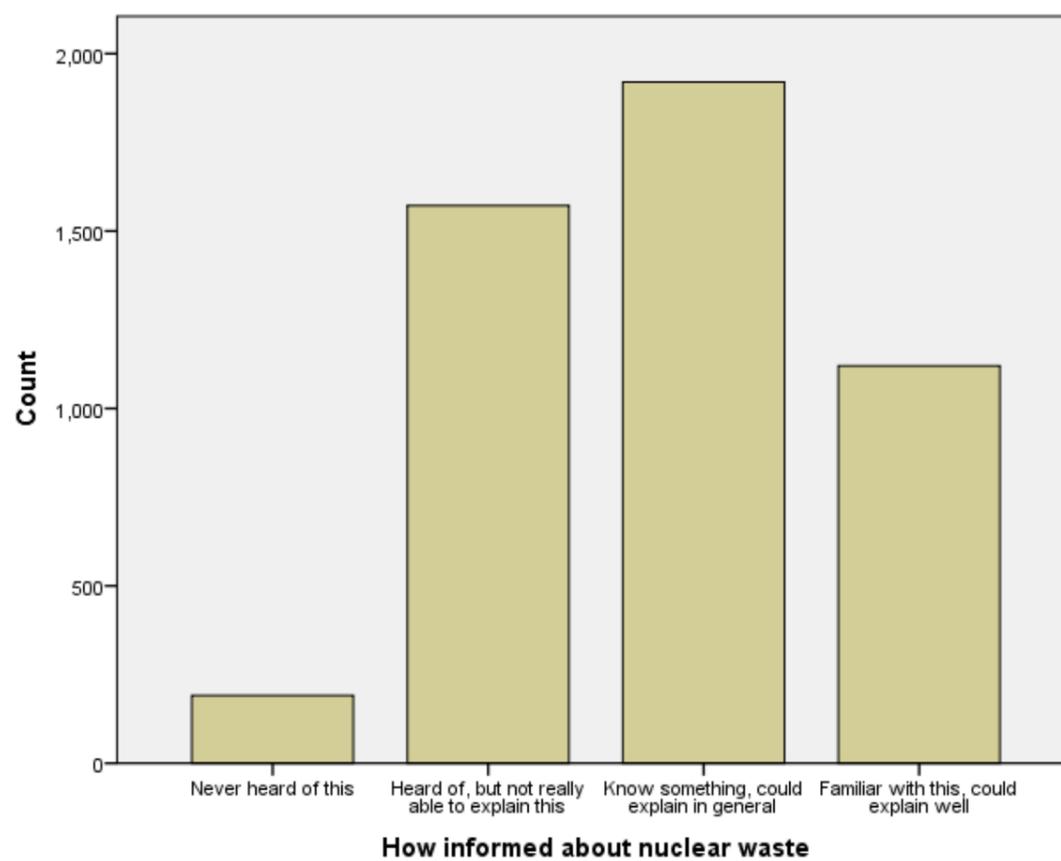
- There are 1572 observations in category Heard of, but not really able to explain this.

The SPSS instructions for the pie chart are as follows:

- Select **Pie...** from the **Legacy Dialogs** submenu available from the **Graphs** menu.
- Keep the choices as **Summaries for groups of cases** before clicking on **Define**.
- Add the **How informed about nuclear waste[INFNUCL]** variable in the **Define Slices by:** box.
- Click on the **OK** button to produce the graph.

- Question: Which is the most common category?

Solution: The output from SPSS is as follows:
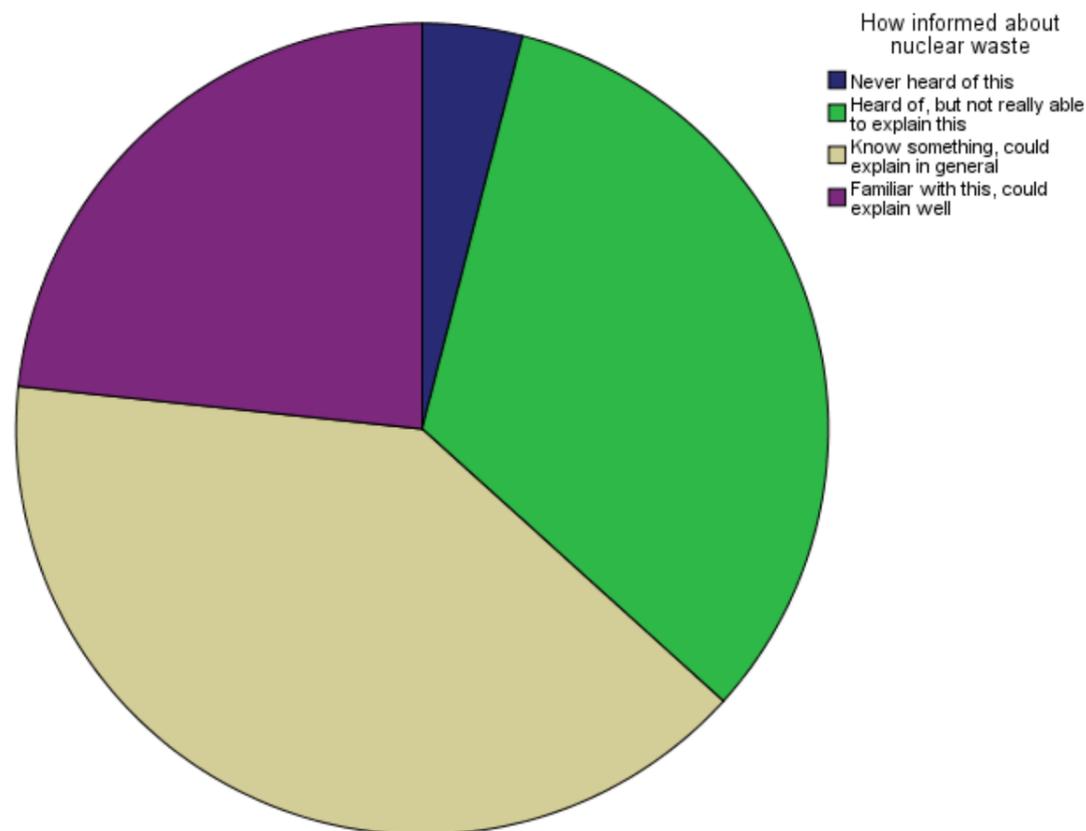
- Know something, could explain in general is the most common category.

The SPSS instructions for the second bar chart are as follows:

- Select again **Bar...** from the **Legacy Dialogs** submenu available from the **Graphs** menu.
- Use the defaults which are **Simple** and **Summaries for groups of cases** and then click on **Define**.
- Remove the **How informed about nuclear waste[INFNUCL]** variable into the **Category Axis** box.
- Copy the **Parents are interested in school activities[PARINTSCH]** variable into the **Category Axis** box.
- Click on the **OK** button to produce the graph as shown below.

- Question: You will see 4 categories. Approximately how many observations are in category Disagree?

Solution: The output from SPSS is as follows:
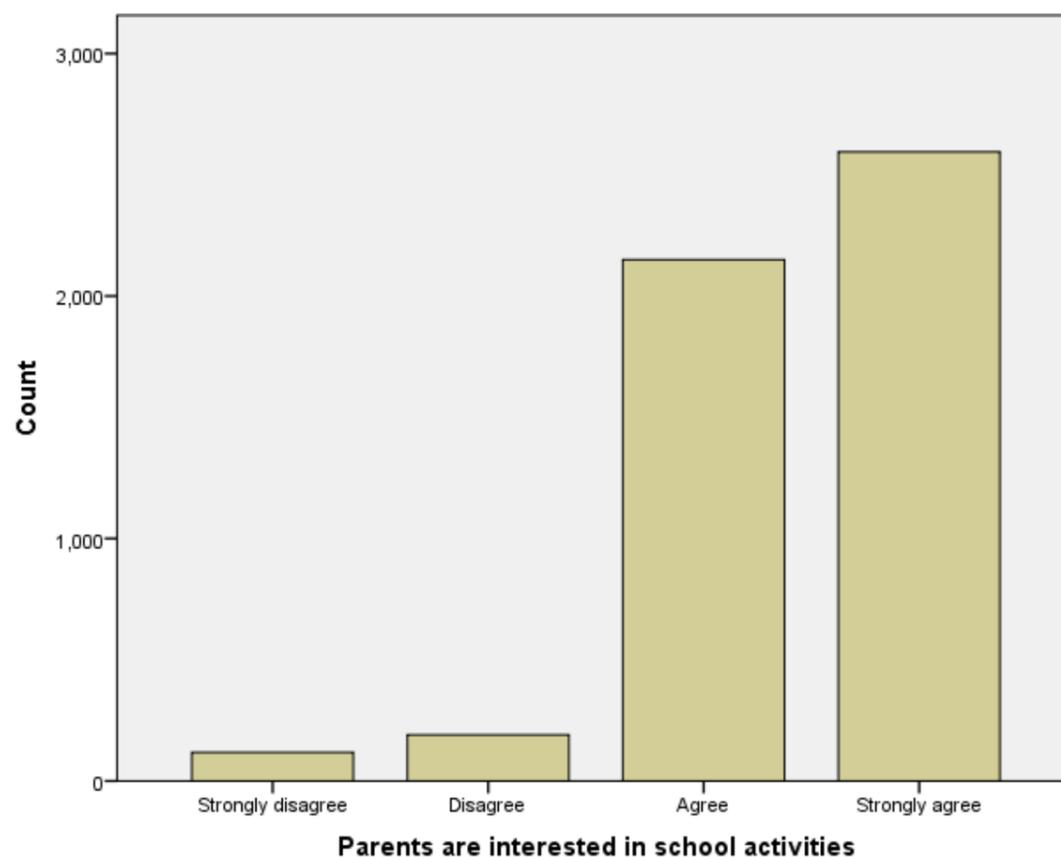
- There are 190 observations in category Disagree.

The SPSS instructions for the clustered bar chart of counts are as follows:

- Select **Bar...** from the **Legacy Dialogs** submenu available from the **Graphs** menu.
- This time we will change from the defaults to **Clustered** and **Summaries for groups of cases** before clicking on **Define** .
- Copy the **How informed about nuclear waste[INFNUCL]** variable into the **Category Axis** box.
- Copy the **Parents are interested in school activities[PARINTSCH]** variable into the **Define Clusters by** box.
- Click on the **OK** button to produce the graph.

- Question: Approximately how many observations take values Heard of, but not really able to explain this for **INFNUCL** and Strongly disagree for **PARINTSCH**?
- Question: Approximately how many observations take values Know something, could explain in general for **INFNUCL** and Strongly agree for **PARINTSCH**?
- Question: Approximately how many observations take values Familiar with this, could explain well for **INFNUCL** and Agree for **PARINTSCH**?

Solutions: The output from SPSS is as follows:
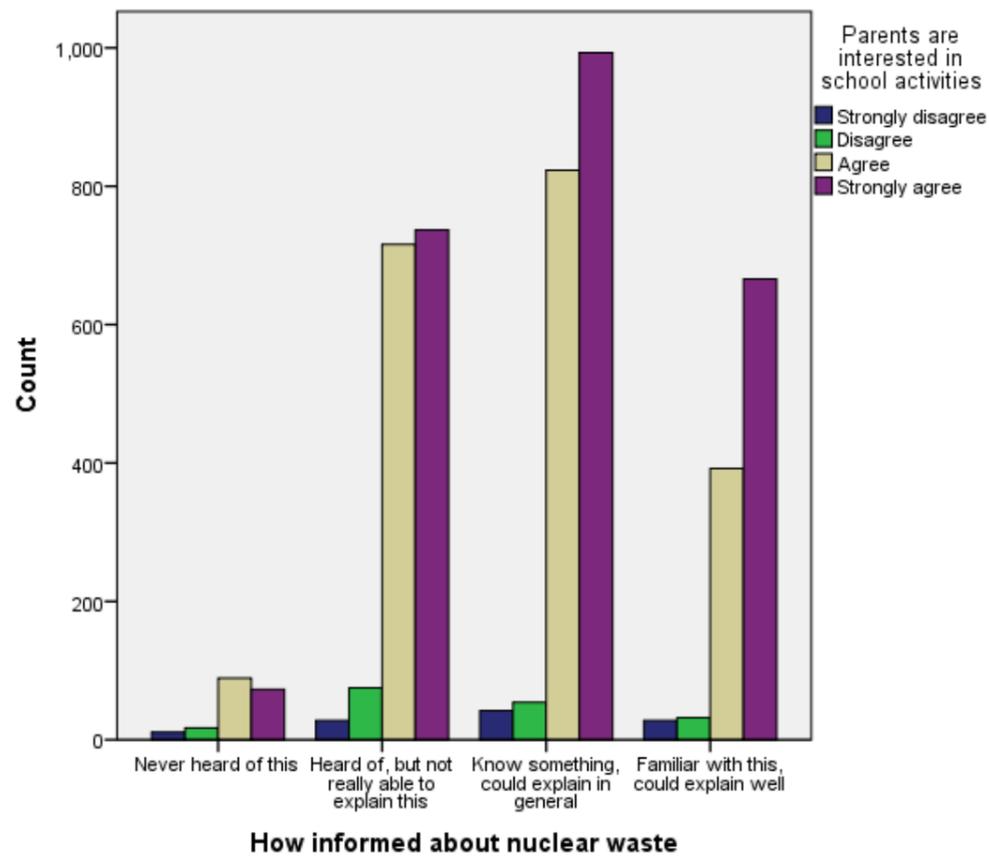
The answers are as follows:

- There are 28 observations taking values Heard of, but not really able to explain this for **INFNUCL** and Strongly disagree for **PARINTSCH**.
- There are 993 observations taking values Know something, could explain in general for **INFNUCL** and Strongly agree for **PARINTSCH**.
- There are 392 observations taking values Familiar with this, could explain well for **INFNUCL** and Agree for **PARINTSCH**.

The SPSS instructions for the clustered bar chart of percentages are as follows:

- Once again select **Bar...** from the **Legacy Dialogs** submenu available from the **Graphs** menu.
- Keep the choices as **Clustered** and **Summaries for groups of cases** before clicking on **Define**.
- Keep the **How informed about nuclear waste[INFNUCL]** variable in the **Category Axis** box.
- Keep the **Parents are interested in school activities[PARINTSCH]** variable in the **Define Clusters by** box.
- Select **% of cases** in the **Bars Represent** choices.
- Click on the **OK** button to produce the graph.

- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Disagree take value Know something, could explain in general for **INFNUCL**?
- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Agree take value Never heard of this for **INFNUCL**?
- Question: What percentage of observations in the sub-group for which **PARINTSCH** takes value Strongly disagree take value Familiar with this, could explain well for **INFNUCL**?

Solutions: The output from SPSS is as follows:
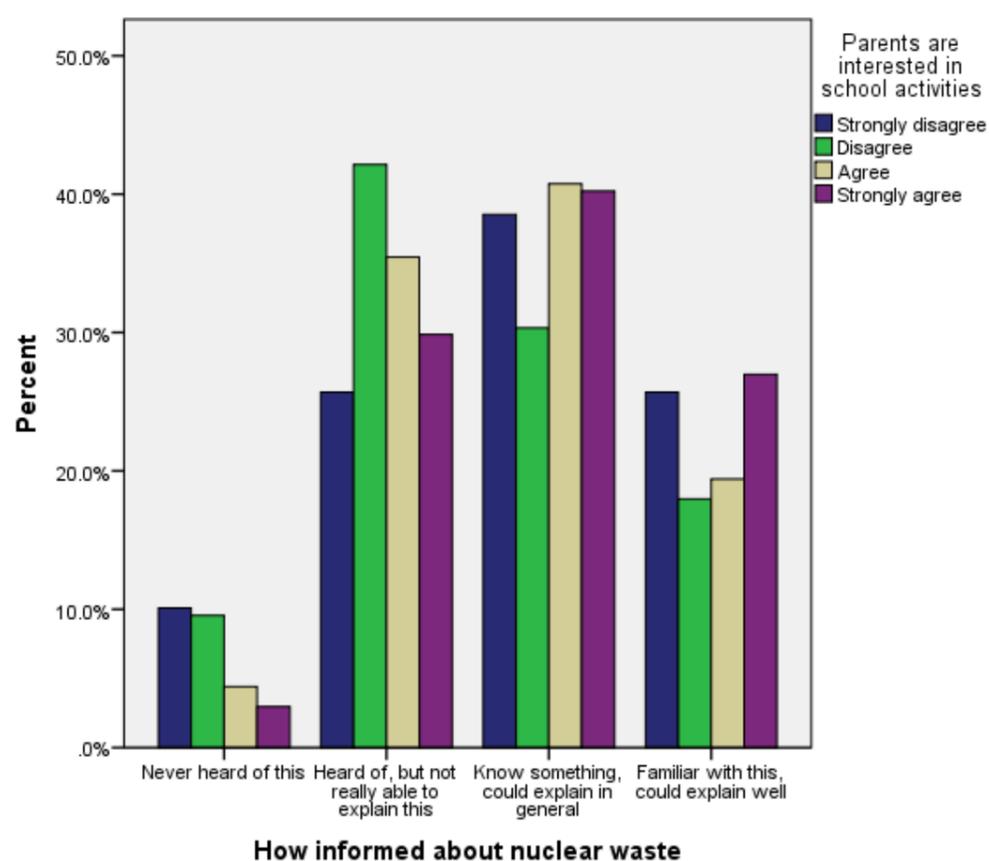
The answers are as follows:

- There are 30.3% percent of observations where **PARINTSCH** takes value Disagree that take value Know something, could explain in general for **INFNUCL**.
- There are 4.4% percent of observations where **PARINTSCH** takes value Agree that take value Never heard of this for **INFNUCL**.
- There are 25.7% percent of observations where **PARINTSCH** takes value Strongly disagree that take value Familiar with this, could explain well for **INFNUCL**.

Only a small minority of students disagreed that their parents were interested in their school activities, but students in this group were more likely than other students to report that they had never heard of nuclear waste.