# MENGAPLIKASIKAN SPSS

Aryan Eka Prastya Nugraha

2018

# S P S S

When SPSS, Inc., an IBM Company, was conceived in 1968, it stood for ***Statistical Package for Social Sciences***. Since the company's purchase by IBM in 2009, IBM has decided to simply use the name SPSS to describe its core product of predictive analytics. IBM describes predictive analytics as tools that ***help connect data to effective action by drawing reliable conclusions about current conditions and future events***.

SPSS is an integrated system of computer programs designed for the **analysis of social sciences data**. It is one of the most popular of the many statistical packages currently available for statistical analysis. Its popularity stems from the fact that the program:

- Allows for a great deal of flexibility in the format of data
- Provides the user with a comprehensive set of procedures for data transformation and fle manipulation
- Offers the researcher a large number of statistical analyses commonly used in social sciences

# C A S E

Suppose a survey has been conducted to investigate the extent to which This agree with increases in government spending in the three areas of defense, social security, and childcare services.

Survey Questionnaire

a. Gender      1. _____ Male      2. _____ Female

b. Age _____ (in years)

c. The following three statements relate to increases in government spending in the areas of defense, social security, and childcare services. Please consider these three statements carefully and then decide your level of agreement with the government's decision to increase spending. Please indicate your level of agreement by circling the number on each six-point scale.

i. Increased spending on defense

| 1_____ | 2_____ | 3_____ | 4_____ | 5_____ | 6_____ |
|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Barely Disagree | Barely Agree | Moderately Agree | Strongly Agree |

ii. Increased spending on social security

| 1_____ | 2_____ | 3_____ | 4_____ | 5_____ | 6_____ |
|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Barely Disagree | Barely Agree | Moderately Agree | Strongly Agree |

iii. Increased spending on childcare services

| 1_____ | 2_____ | 3_____ | 4_____ | 5_____ | 6_____ |
|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Barely Disagree | Barely Agree | Moderately Agree | Strongly Agree |

Prior to data entry, it will be useful to prepare a codebook that contains the names of the variables in the questionnaire, their corresponding SPSS variable names, and their coding instructions. An important purpose of such a codebook is **to allow the researcher to keep track of all the variables in the survey questionnaire and the way they are defned in the SPSS data file**.

## Codebook

| Variable | SPSS Variable Name | Code |
|---|---|---|
| Gender | Gender | 1 = male<br>2 = female |
| Age | Age | Age in years |
| Defense | Defense | 1 = Strongly Disagree<br>2 = Moderately Disagree<br>3 = Barely Disagree<br>4 = Barely Agree<br>5 = Moderately Agree<br>6 = Strongly Agree |
| Social Security | Social | 1 = Strongly Disagree<br>2 = Moderately Disagree<br>3 = Barely Disagree<br>4 = Barely Agree<br>5 = Moderately Agree<br>6 = Strongly Agree |
| Childcare Services | Child | 1 = Strongly Disagree<br>2 = Moderately Disagree<br>3 = Barely Disagree<br>4 = Barely Agree<br>5 = Moderately Agree<br>6 = Strongly Agree |

# DATA

## Raw Data

| Gender | Age | Defense | Social | Child |
|--------|-----|---------|--------|-------|
| 1 | 24 | 4 | 2 | 1 |
| 1 | 18 | 5 | 1 | 4 |
| 2 | 33 | 2 | 5 | 6 |
| 1 | 29 | 5 | 3 | 4 |
| 2 | 26 | 3 | 5 | 5 |
| 2 | 19 | 2 | 5 | 2 |
| 1 | 36 | 4 | 4 | 3 |
| 2 | 34 | 3 | 6 | 6 |
| 1 | 20 | 3 | 5 | 1 |
| 2 | 21 | 2 | 5 | 3 |
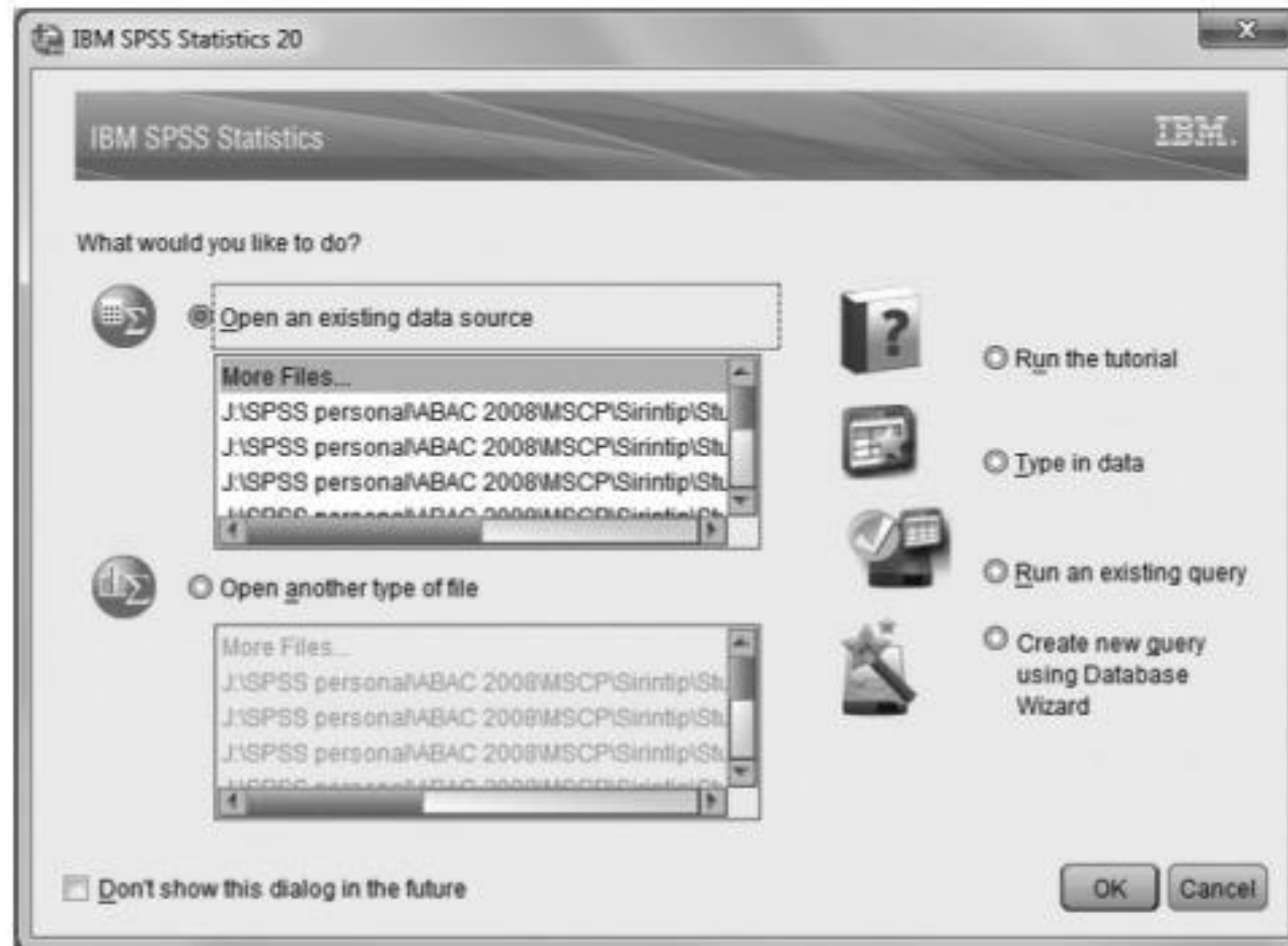
Since the purpose of the present exercise is to create a new data file, close this window by clicking [Cancel]. The following **Untitled1 [DataSet0]— IBM SPSS Statistics Data Editor** screen will then be displayed.
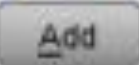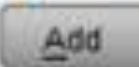
Prior to data entry, the variables in the data set must be named and defned. In the **Untitled1 [DataSet0]—IBM SPSS Statistics Data Editor** screen, the names of the variables are listed down the side (under the **Name** column), with their characteristics listed along the top (Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure).

The codebook presented in Table will serve as a guide in naming and defning the variables. For example, the frst variable is named GENDER and is coded 1 = male and 2 = female. Thus, in the frst cell under **Name** in the **Data Editor** screen, type in the name GENDER. To assign the coded values (1 = male, 2 = female) to this variable, click the corresponding cell under **Values** in the **Data Editor** screen. Click the shaded area to open the following **Value Labels** window.

In order to define the code for male respondents, type **1** in the **Value:** cell, and in the **Label:** cell, type **Male**. Next, click [Add] to complete the coding for the male respondents. For female respondents, type **2** in the **Value:** cell, and in the **Label:** cell, type **Female**. Next, click [Add] to complete the coding for the female respondents. The completed **Value Labels** window is presented below.
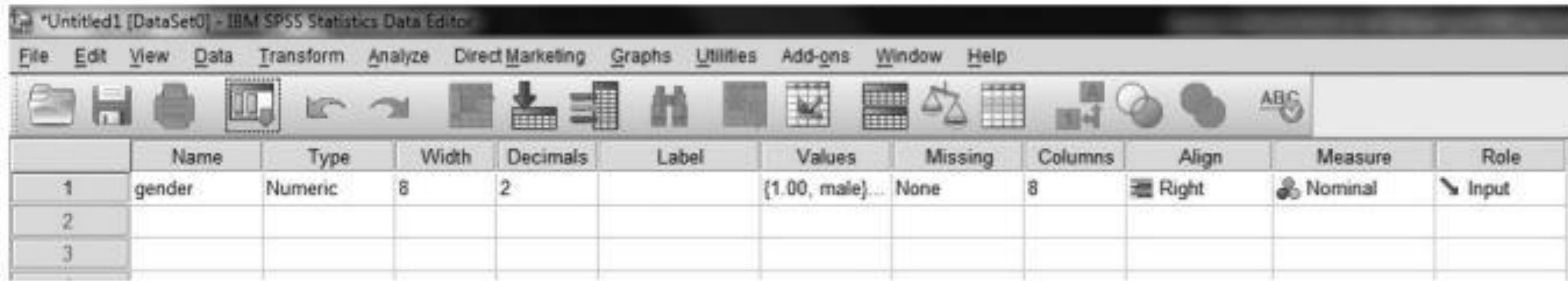
Repeat the above coding procedure for the rest of variables in the codebook. Please note that the AGE variable is a *continuous* variable and therefore has no coded values.
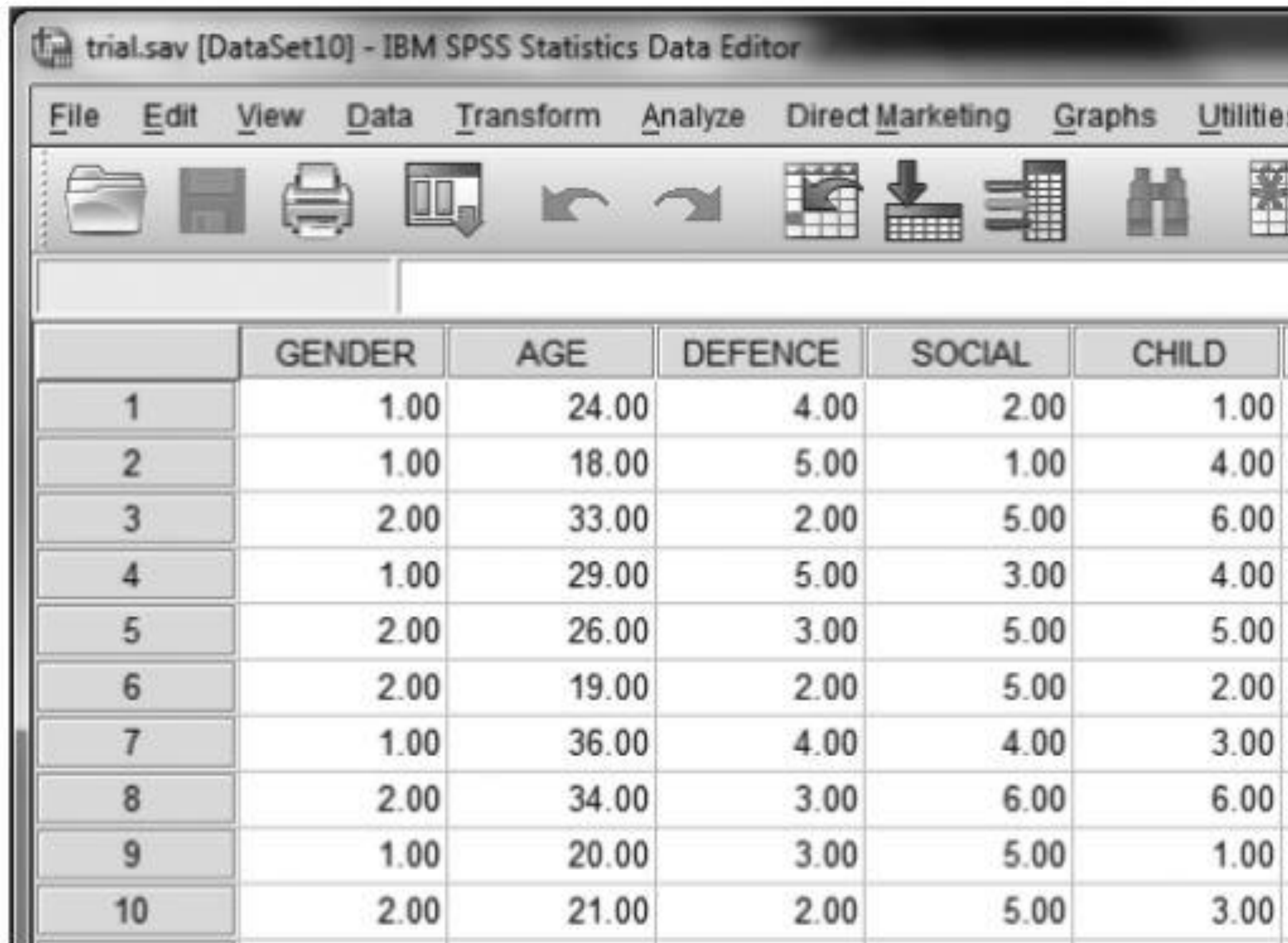
# Data Entry



trial.sav [DataSet10] – IBM SPSS Statistics Data Editor

File   Edit   View   Data   Transform   Analyze   Direct Marketing   Graphs   Utilities

|    | GENDER | AGE   | DEFENCE | SOCIAL | CHILD |
|----|--------|-------|---------|--------|-------|
| 1  | 1.00   | 24.00 | 4.00    | 2.00   | 1.00  |
| 2  | 1.00   | 18.00 | 5.00    | 1.00   | 4.00  |
| 3  | 2.00   | 33.00 | 2.00    | 5.00   | 6.00  |
| 4  | 1.00   | 29.00 | 5.00    | 3.00   | 4.00  |
| 5  | 2.00   | 26.00 | 3.00    | 5.00   | 5.00  |
| 6  | 2.00   | 19.00 | 2.00    | 5.00   | 2.00  |
| 7  | 1.00   | 36.00 | 4.00    | 4.00   | 3.00  |
| 8  | 2.00   | 34.00 | 3.00    | 6.00   | 6.00  |
| 9  | 1.00   | 20.00 | 3.00    | 5.00   | 1.00  |
| 10 | 2.00   | 21.00 | 2.00    | 5.00   | 3.00  |

# S P S S   A n a l y s i s

1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Frequencies**. The following **Frequencies** window will open.
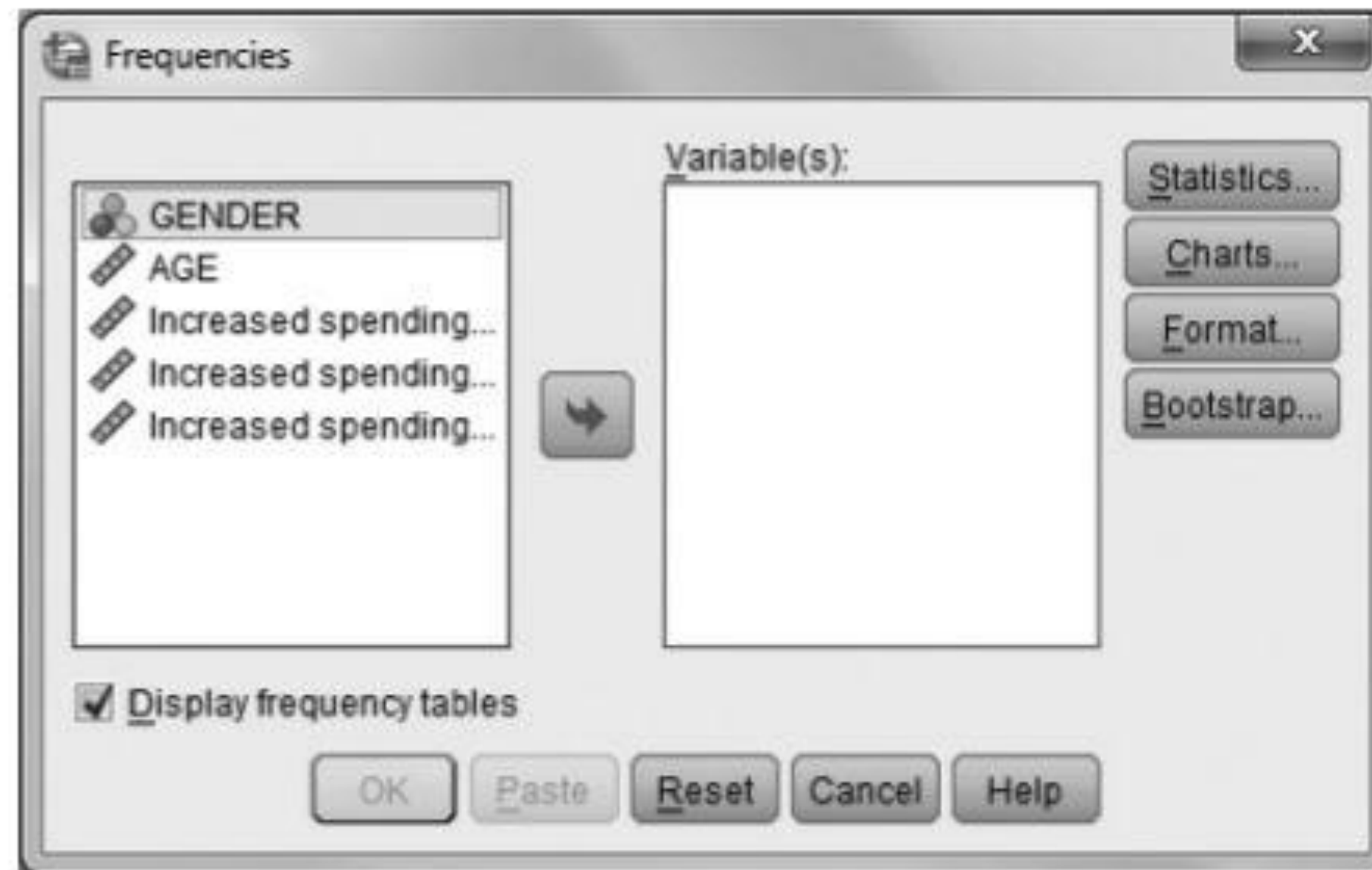
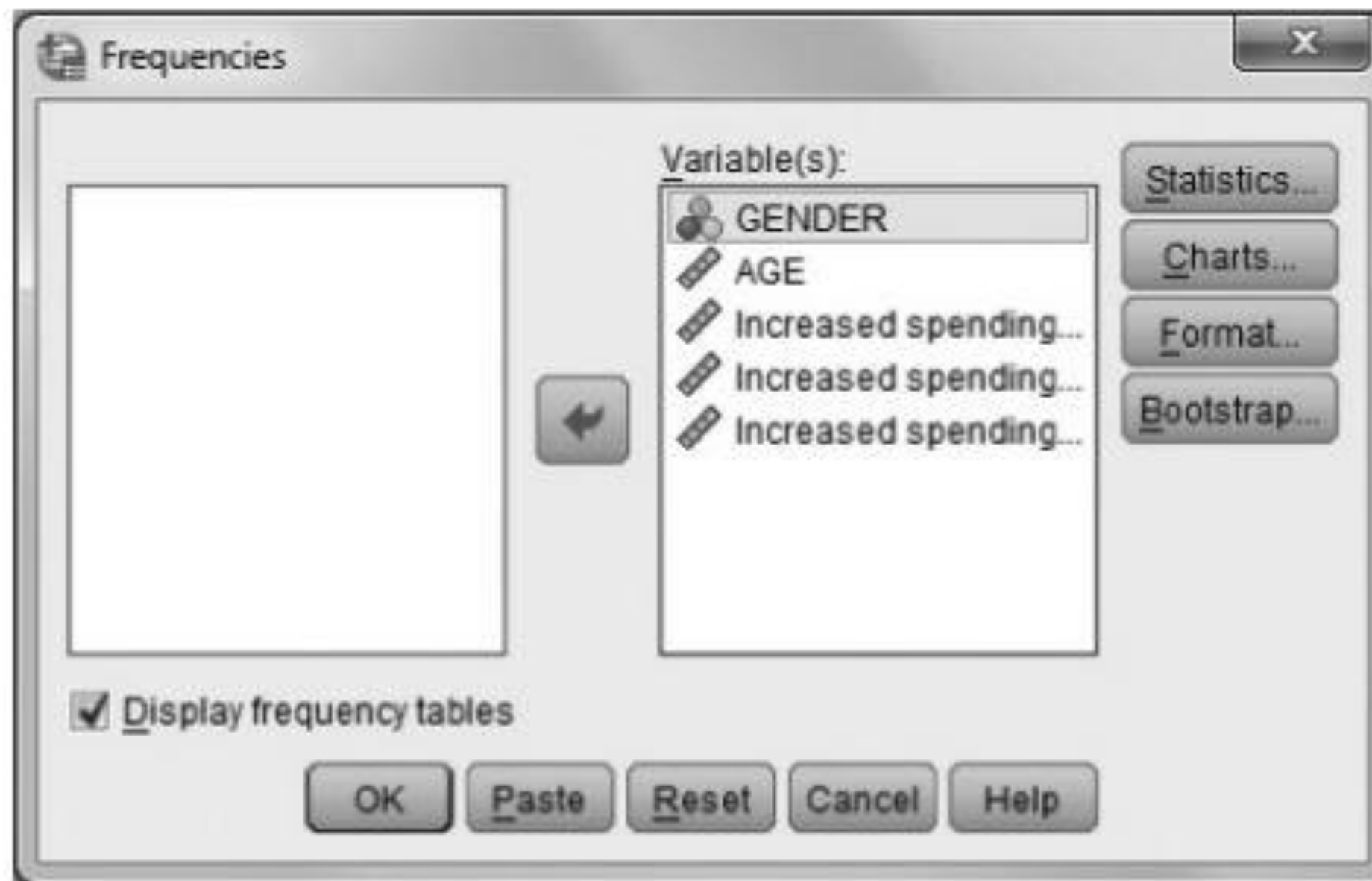2. In the left-hand field containing the study's five variables, click (highlight) these variables, and then click ➡ to transfer the selected variables to the **Variable(s):** field.

3. Click [Statistics] to open the **Frequencies: Statistics** window below. Suppose the researcher is only interested in obtaining statistics for the **Mean**, **Median**, **Mode**, and **Standard Deviation** for the five variables. In the **Frequencies: Statistics** window, check the fields related to these statistics.

Next click [Continue].
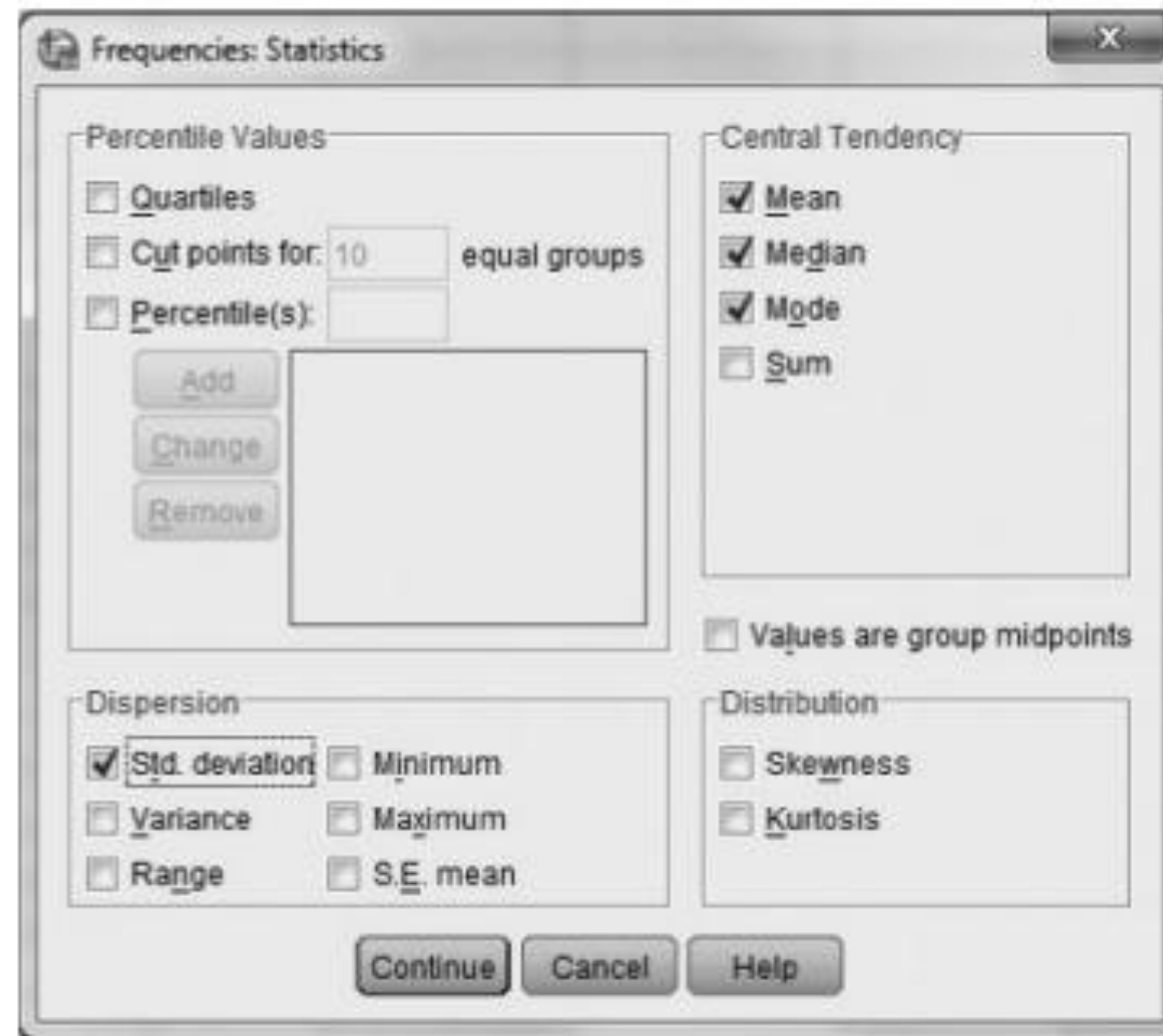
When the **Frequencies** window opens, run the analysis by clicking [OK]. See Table 2.4 for the results.
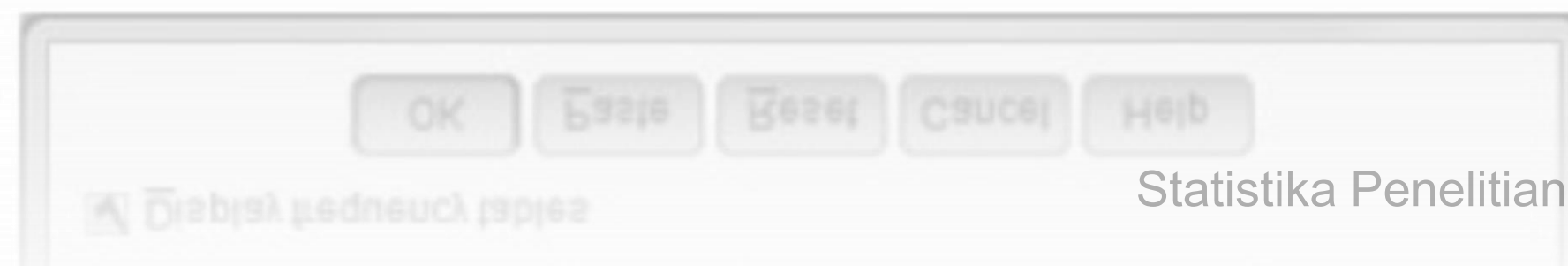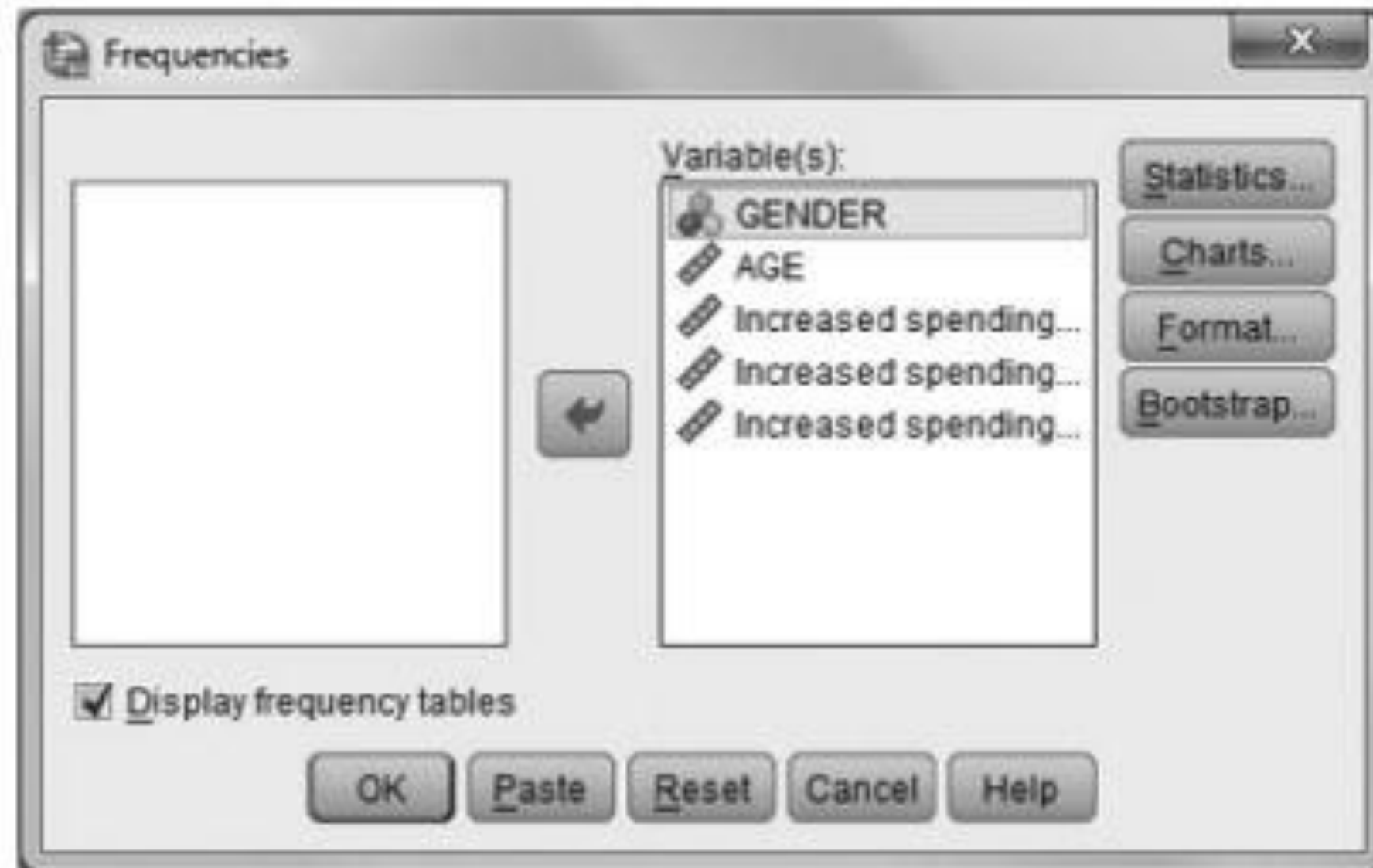
# SPSS Output

## Frequencies Output

### Frequencies

#### Statistics

| | | Gender | Age | Increased Spending on Defence | Increased Spending on Social Security | Increased Spending on Childcare Services |
|---|---|---|---|---|---|---|
| N | Valid | 10 | 10 | 10 | 10 | 10 |
| | Missing | 0 | 0 | 0 | 0 | 0 |
| Mean | | 1.5000 | 26.0000 | 3.3000 | 4.1000 | 3.5000 |
| Median | | 1.5000 | 25.0000 | 3.0000 | 5.0000 | 3.5000 |
| Mode | | 1.00[a] | 18.00[a] | 2.00[a] | 5.00 | 1.00[a] |
| Std. Deviation | | .52705 | 6.66667 | 1.15950 | 1.59513 | 1.84089 |

[a] Multiple modes exist. The smallest value is shown.

## Frequency Table

### Gender

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | Male | 5 | 50.0 | 50.0 | 50.0 |
| | Female | 5 | 50.0 | 50.0 | 100.0 |
| | Total | 10 | 100.0 | 100.0 | |

| | | Age | | | |
|---|---|---|---|---|---|
| | | Frequency | Percent | Valid Percent | Cumulative Percent |
| Valid | 18.00 | 1 | 10.0 | 10.0 | 10.0 |
| | 19.00 | 1 | 10.0 | 10.0 | 20.0 |
| | 20.00 | 1 | 10.0 | 10.0 | 30.0 |
| | 21.00 | 1 | 10.0 | 10.0 | 40.0 |
| | 24.00 | 1 | 10.0 | 10.0 | 50.0 |
| | 26.00 | 1 | 10.0 | 10.0 | 60.0 |
| | 29.00 | 1 | 10.0 | 10.0 | 70.0 |
| | 33.00 | 1 | 10.0 | 10.0 | 80.0 |
| | 34.00 | 1 | 10.0 | 10.0 | 90.0 |
| | 36.00 | 1 | 10.0 | 10.0 | 100.0 |
| | Total | 10 | 100.0 | 100.0 | |

*(Continued)*

## Frequencies Output

### Increased Spending on Defense

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | moderately disagree | 3 | 30.0 | 30.0 | 30.0 |
| | barely disagree | 3 | 30.0 | 30.0 | 60.0 |
| | barely agree | 2 | 20.0 | 20.0 | 80.0 |
| | moderately agree | 2 | 20.0 | 20.0 | 100.0 |
| | Total | 10 | 100.0 | 100.0 | |

## Increased Spending on Social Security

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | strongly disagree | 1 | 10.0 | 10.0 | 10.0 |
| | moderately disagree | 1 | 10.0 | 10.0 | 20.0 |
| | barely disagree | 1 | 10.0 | 10.0 | 30.0 |
| | barely agree | 1 | 10.0 | 10.0 | 40.0 |
| | moderately agree | 5 | 50.0 | 50.0 | 90.0 |
| | strongly agree | 1 | 10.0 | 10.0 | 100.0 |
| | Total | 10 | 100.0 | 100.0 | |

## Increased Spending on Childcare Services

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | strongly disagree | 2 | 20.0 | 20.0 | 20.0 |
| | moderately disagree | 1 | 10.0 | 10.0 | 30.0 |
| | barely disagree | 2 | 20.0 | 20.0 | 50.0 |
| | barely agree | 2 | 20.0 | 20.0 | 70.0 |
| | moderately agree | 1 | 10.0 | 10.0 | 80.0 |
| | strongly agree | 2 | 20.0 | 20.0 | 100.0 |
| | Total | 10 | 100.0 | 100.0 | |

# Results and Interpretation

The **Statistics** table presents the requested mean, median, mode, and standard deviation (SD) statistics for the fve variables. The **Gender** variable is a nominal (categorical) variable and as such, its mean, median, and standard deviation statistics are not meaningful. The remaining four variables of **Age**, **Defense**, **Social**, and **Child** are measured at least at the ordinal level (i.e., they are continuous variables), and as such their mean, median, and standard deviation statistics can be interpreted.

The results presented in the **Statistics** table show that the 10 respondents in the survey have a mean age of 26 years and a median age of 25 years. Since there is no one age that occurs more frequently than others, SPSS presents the lowest age value of 18 as the mode.

For the three variables of "support for increased spending" on defense, social security, and childcare services, the results show that support for increased spending for social security is the highest (mean = 4.10; median = 5.00), followed by childcare services (mean = 3.50; median = 3.50), and defense (mean = 3.30; median = 3.00). The results also show that the variables of Defense and Child have multiple modes, and as such, SPSS has presented their lowest values (defense: mode = 2.00; child: mode = 1.00). The Social variable has a single mode of 5.00.

For the Age variable, the standard deviation shows that its average deviation (dispersion) from the mean is 6.66 years. For the Defense, Social, and Child variables, the results show that support for increased spending on childcare services has the largest average variation (SD = 1.84) from its mean score.

The standard deviation scores for support for increased spending for defense (SD = 1.59) and social security (SD = 1.59) are similar.

The **frequency** table presents the breakdown of the frequency distributions for the fve variables (Gender, Age, Defense, Social, Child). For each variable, the frequency table presents (1) the **frequency** of occurrence for each value within that variable, (2) the frequency for each value expressed as a **percentage** of the total sample, (3) the **valid percentage** for each value, controlling for missing cases, and (4) the **cumulative percentage** for each succeeding value within that variable.

For example, the Frequency table for the Gender variable shows that there are five males and five females in the sample, and that these two groups represent 50% each of the total sample.

Since there are no missing cases, the valid percentage values are identical to the percentage values. *If there are missing cases, then the valid percentage values should be interpreted*.

The cumulative percentage presents the percentage of scores falling at or below each score. Thus, for the sample of 10 respondents, the fve males in the sample represent 50% of the sample, and the additional fve females represent a cumulative percentage of 100%.

The frequency tables for the Age, Defense, Social, and Child variables are interpreted in exactly the same way.

# R E F L E K S I

1. Informasi penting hari ini

2. Manfaat penting dari informasi penting hari ini

3. Tindak lanjut yang dapat saudara lakukan

# THANK YOU!

ANY QUESTIONS?