Al-Mustaqbal University College

Department of Anesthesia techniques

Third Stage

Lecture 3

Asst. Lect. Raghda M. Alshemari

2021-2022

3. Reducing sample size

• Using random sampling

• Let's assume you are dealing with 2 million observations.

• This creates a problem – whenever you run a procedure, it takes too much time, the computer crashes and/or runs out of disk space.

• To avoid this problem, you may want to pick only 100,000 observations, chosen randomly, from the data set.

3. Reducing sample size

• Go to DATA/SELECT CASES.

• Select the option "**Random Sample of Cases**" by clicking on the round button to the left of it.

• Click on the button "**Sample**."

• Select the option "**Approximately**" by clicking on the round button to the left of it.

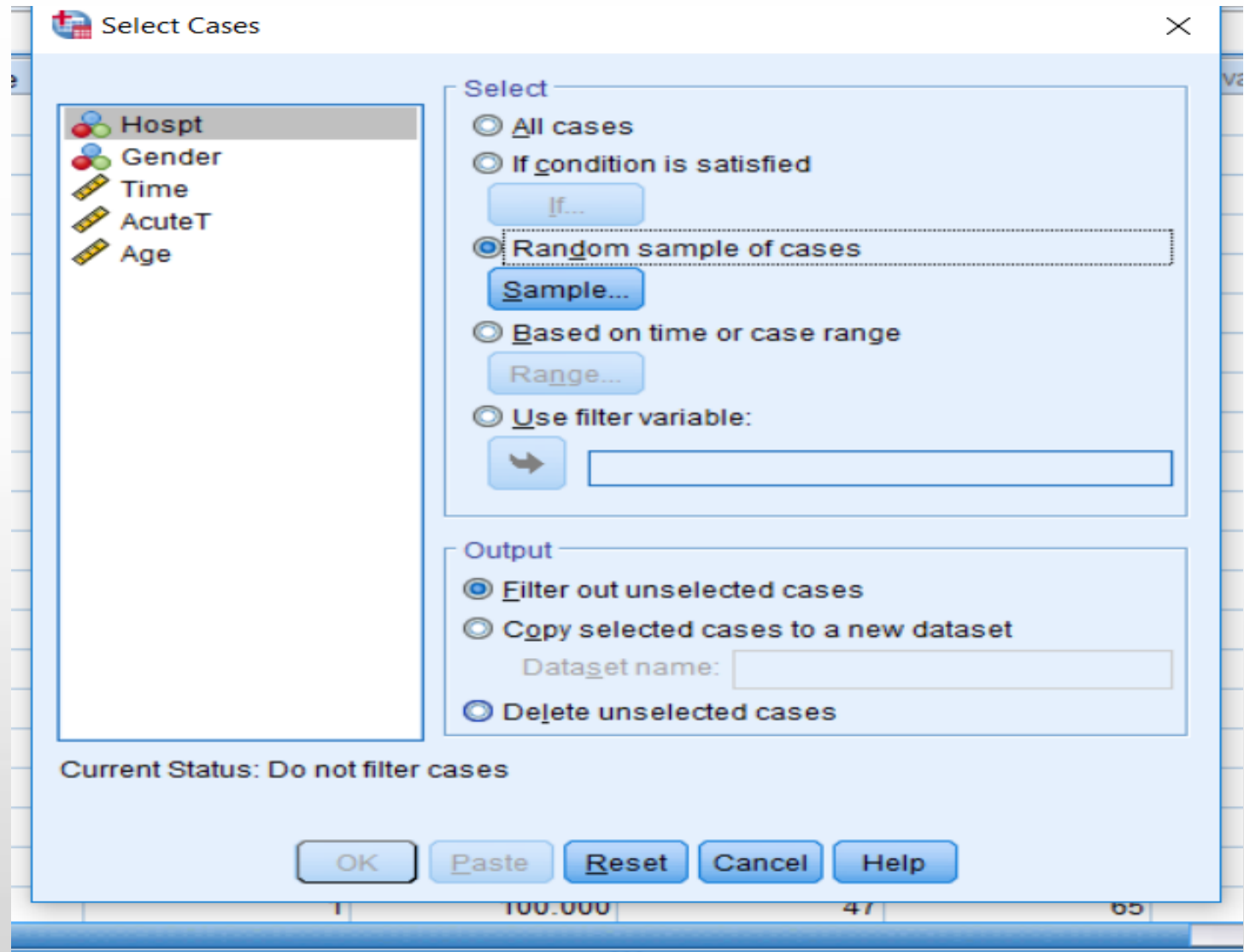• Type in the size of the new sample relative to the size of the entire data set.

3. Reducing sample size

- **Example**

- The relative size is **5%** of the entire data - SPSS will randomly select **100,000** cases from the original data set of 2 million.

3. Reducing sample size

4. Filtering data

- It will often be the case that you will want to select a Sub-set of the data according to **certain criteria**.

- For example, let's assume you want to run procedures on only those cases in which **education level is over 6.**

4. Filtering data

- In effect, you want to temporarily **"hide"** cases in which education level **is 6 or lower**, run your analysis, then have those cases back in your data set.

- Similarly, you can study the statistical attributes of **females only**, **adult females** only, **adult females with high school** or greater education only, etc.
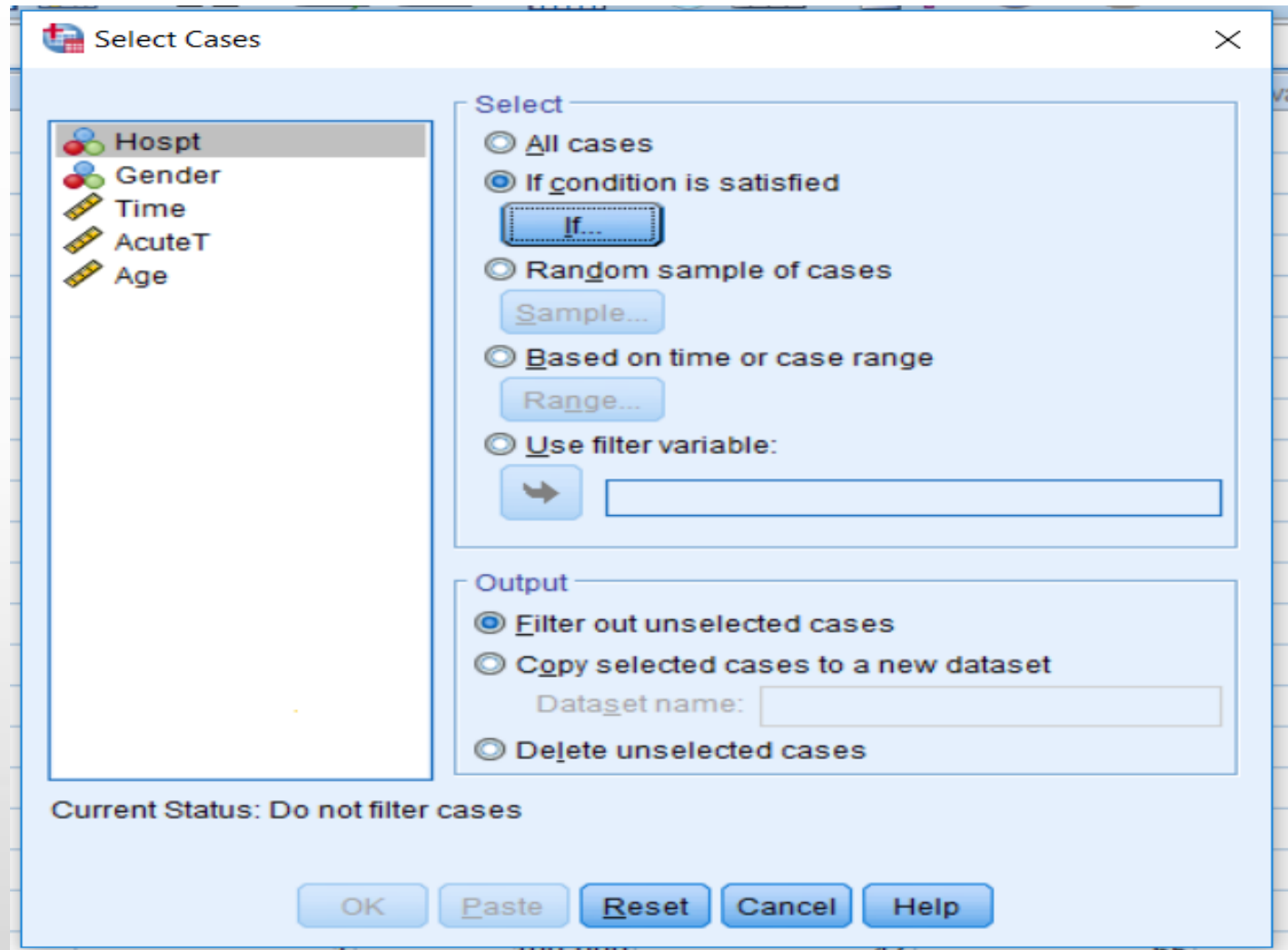
4. Filtering data

- A simple filter

- Suppose you want to run an analysis on only those cases in which the respondent's education level is greater than 6.

- Go to DATA/ SELECT CASE

- When the dialog box opens, click on "If condition is satisfied."

- Click on the button "If."

4. Filtering data

• Such a condition must have variable names. These can be moved from the box on the left (area "1").

• Area "3" has some functions that can be used for creating complex conditions. Area "4" has two buttons you will use often in filtering:

"&" and "|" (for "or"). As you read this section, the purpose and role of each of these areas will become apparent.

4. Filtering data

Thank You For Listening