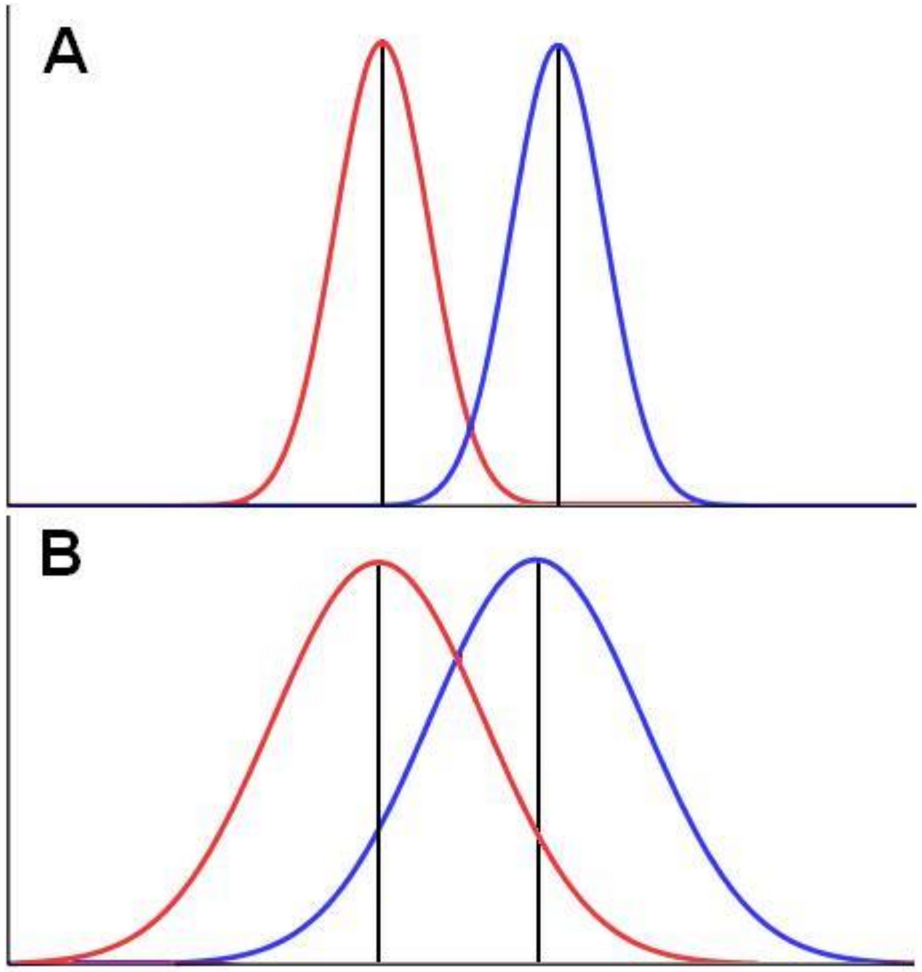
- http://www.stefanokaburu.com/r-workshop.html

Basic Statistics: SPSS

Dr. Colin Dubreuil

Comparing two groups

- A common activity in most fields of research is comparing 2 (or more) groups
- Many of our research questions are based around this concept
  - Do males and females of my study species differ in size?
  - Do males and females of my study species differ in terms of activity budget?
  - Do unaltered, natural bodies of water host a wider variety of species than anthropogenically modified bodies of water?
  - Do birds living near sources of anthropogenic noise produce songs that differ in structure as compared to birds living further away from anthropogenic noise?
  - Do animals produce more stress hormones before, or after a hurricane?



How do we ‘know’ if there is a difference between groups?

- We could compare the distributions visually
- We could calculate the means (or some other measure of central tendency) and see if they are the same

The issue

- The issue comes down to this: If I graph my data, or compare means, how do I know when a difference is different enough? Let's say I:
  - Count the number of water foul species living in 7 'natural' bodies of water
  - Count the number of water foul species living in 7 anthropogenically modified bodies of water

- Now, Calculate the means for each 'type' of habitat. I find that:
  - Natural bodies of water have (on average) 22.6 species of waterfowl (SD = 2.5)
  - Modified bodies of water have on average 20.1 species (SD = 6.1)

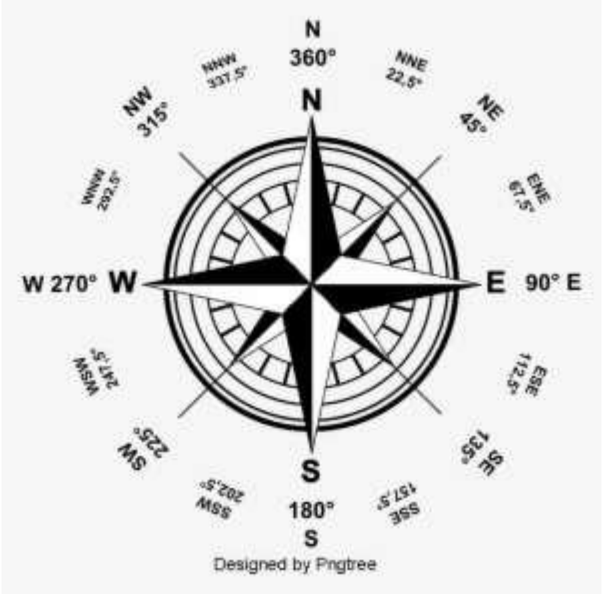
  - **Is that different enough? Can I say my data suggest there is a real difference?**

The issue

- The issue is that when we sample, there is a random aspect to it:
- Our sample statistics (mean of our sample, mode, etc…) are very likely to match the true population parameter perfectly
- You know that 10 flips of a coins should theoretically give you 5 heads, and 5 tails…
  - But it often doesn't because of the randomness of sampling

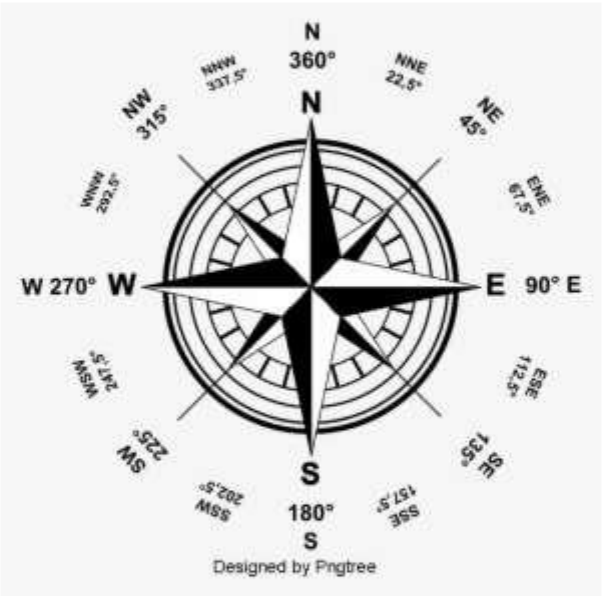
The issue

- **Take this silly question:**
  - Are people who sit on the East side of graduate seminars generally taller than people who sit on the West side?



When is ‘different’ different enough?

- **Take this silly question:**
  - Are people who sit on the ‘east’ side of graduate seminars generally taller than people who sit on the ‘west’ side?



- To answer this question, my best course of action would be to:
  - Divide people in this room into ‘east sitters’, and ‘west sitters’
  - Take hight measurements of people from each group
  - Perhaps I could graph this data
  - Take the mean (average) hight of people in each group
  - If I get different average values, have I proved anything?

The problem

- Likely not…
- In truth, I am likely to find a difference every time I do a comparison of 2 (or more) groups like this… Even if there isn’t a ‘real’ difference

- I personally don’t believe there is ANY relationship between hight and whether you tend to sit on the ‘east side’ of a room vs the ‘west side’…
- But If I broke you all into those 2 groups, measured everyone in each group, and took the average, it’s actually EXTREAMLY unlikely I would get the exact same average height for the 2 groups

A solution

- So, to determine if 2 groups are ‘different enough’ for us to consider them different, we tend to relay on **inferential statistical tests**

- Today, we will look at some **basic statistical tests** that will allow us to address the issue of: “Are these groups different enough to conclude there is a real world difference?”
  - The two sample T-Test
  - The non-parametric equivalent
  - The paired samples T-test
  - The non-parametric equivalent
  - Chi-squared test

The independent samples T-Test

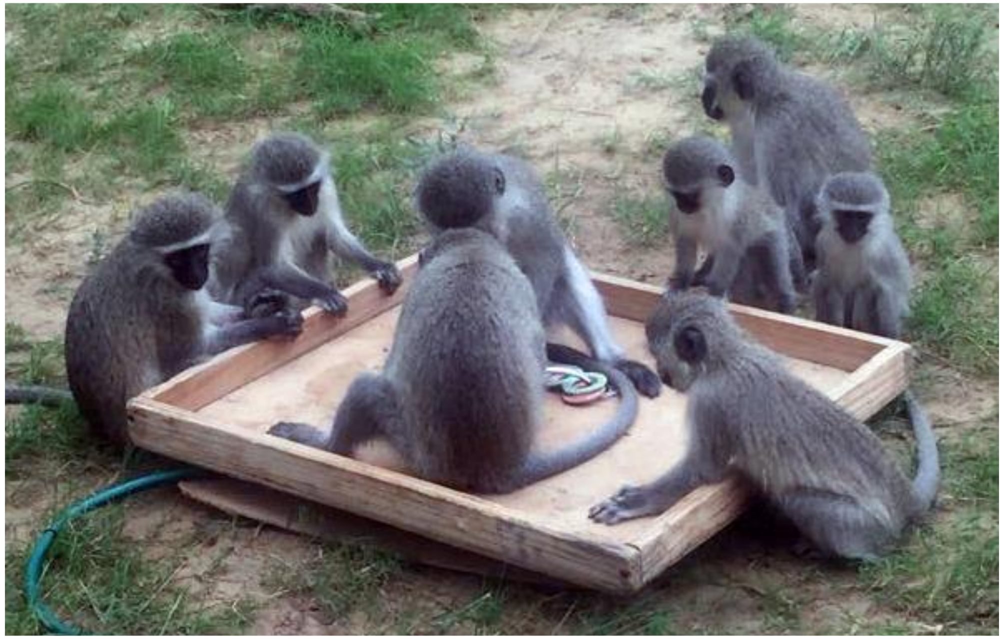
- Inferential statistic used to determine if there is a significant difference between the means (the averages) of two groups.

An Example: Vervet body size dimorphism

- You are working in South Africa with a group of vervet monkeys
- Vervet monkeys live in multimale-multifemale groups
- You get the impression in your first few weeks in the field that adult male vervet monkeys are somewhat larger than females, but…
  - Based on visual inspection alone, you can't be certain
  - No one would EVER let you publish a paper that reveals there are sex differences in body size in vervets without some statistical evidence that this is true!

Vervet body size dimorphism

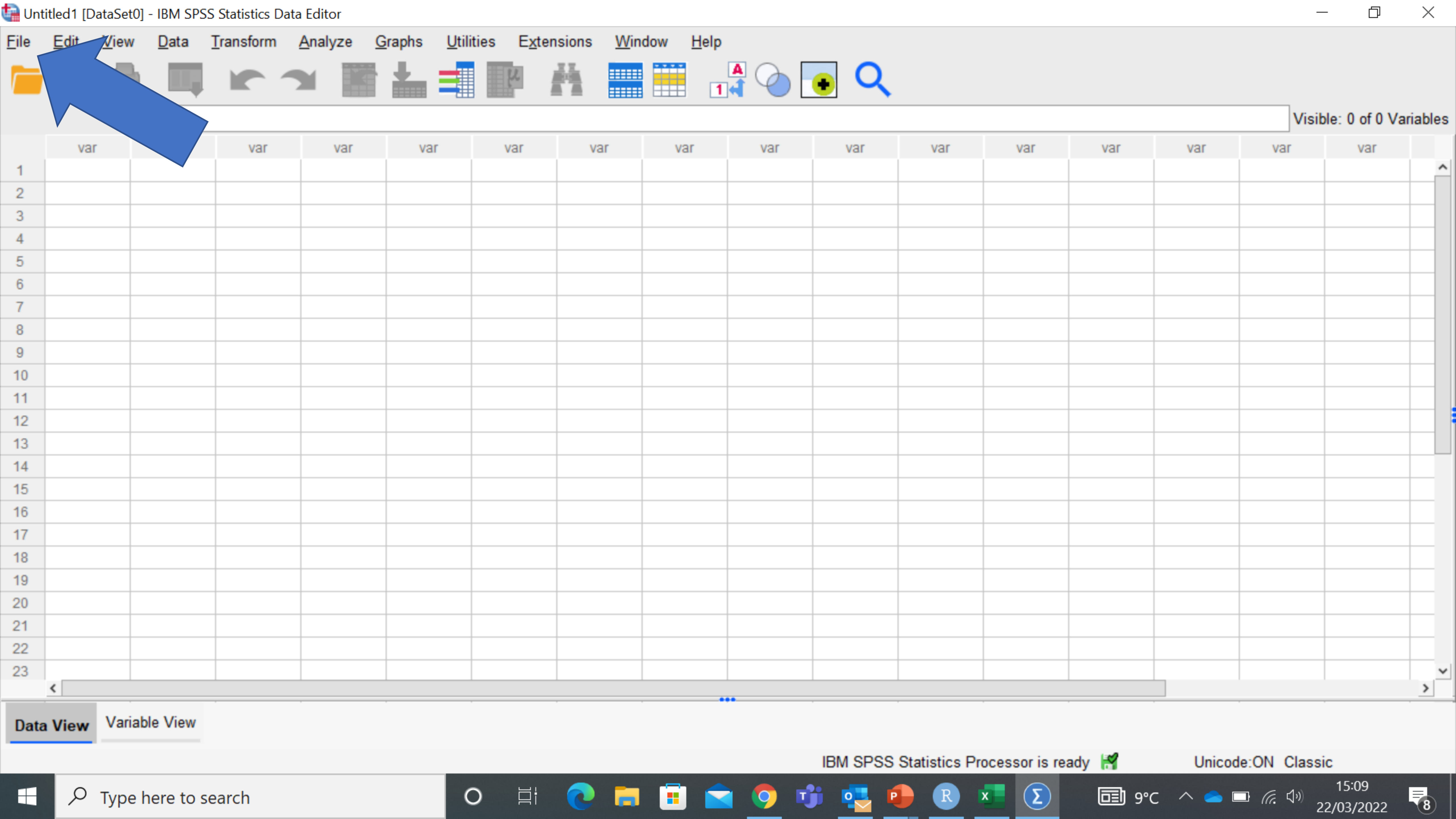
- You devise a plan to bring a scale into the field, bait it with honey, and get weights for individual monkeys
- After getting ethics approval, you move forward!

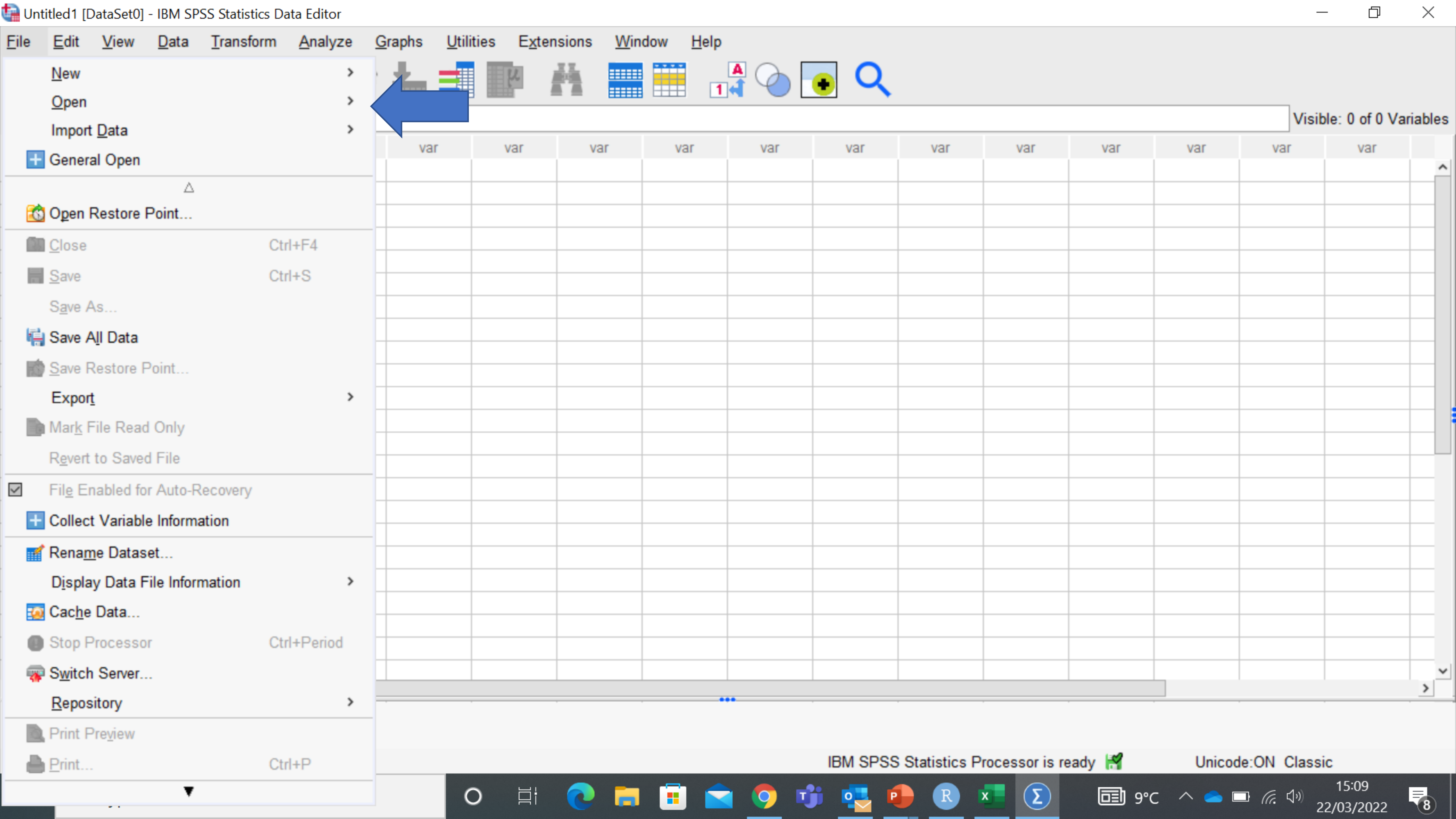


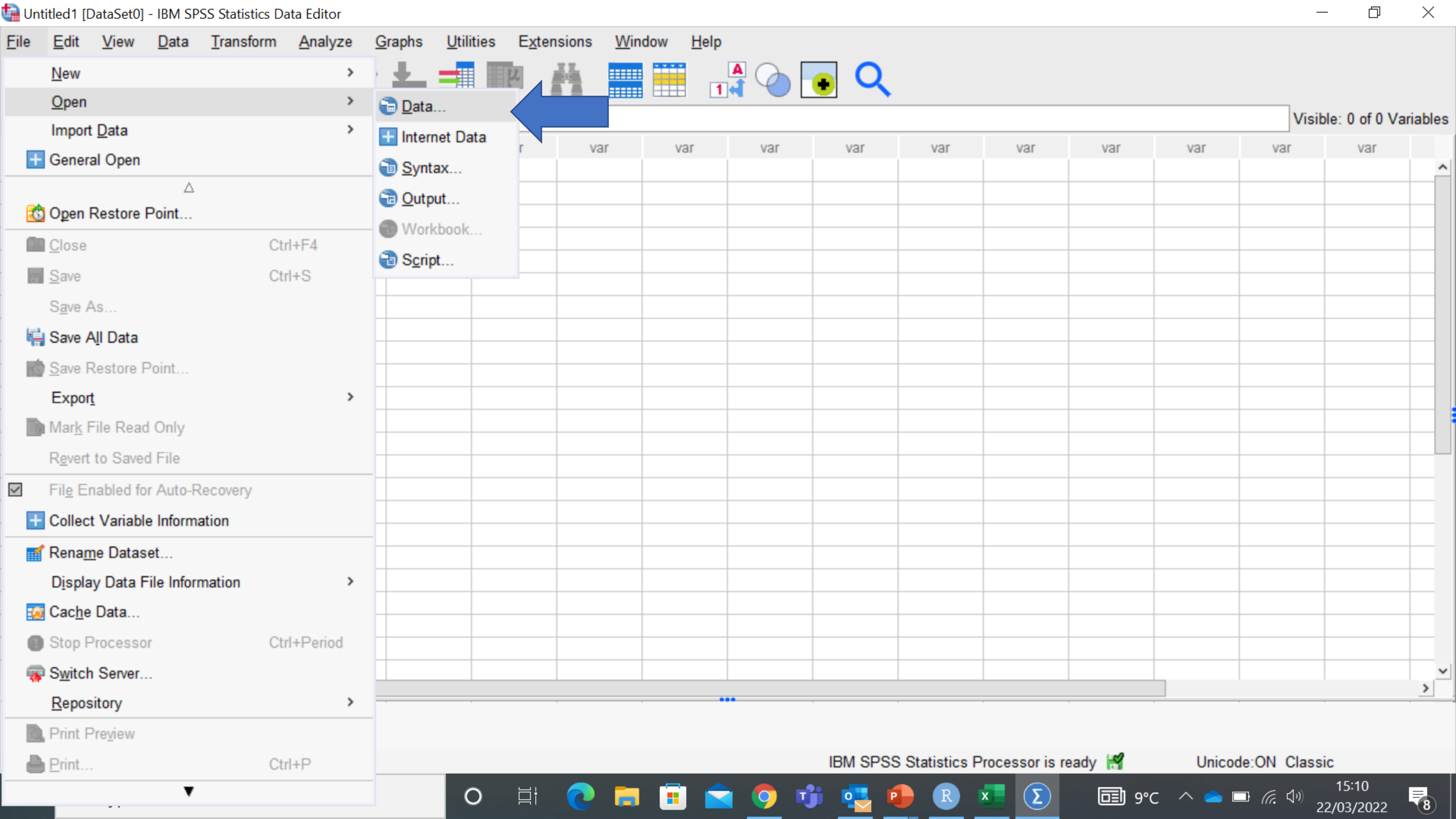
What a great field season!

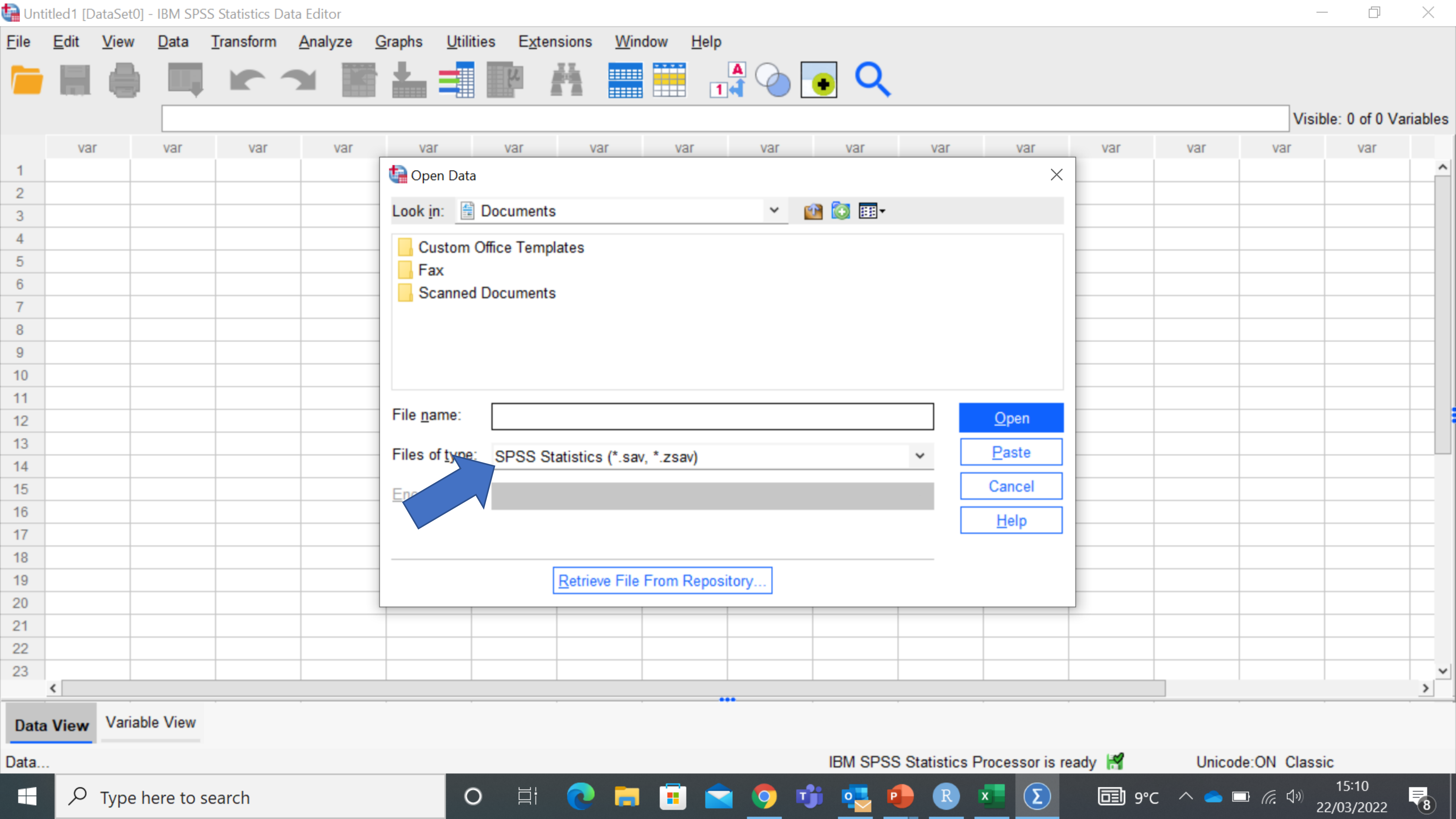
- You now have data!
- Probably the first thing you should do is take a look at the data

- In SPSS, please open ‘vervetweight.xlsx’

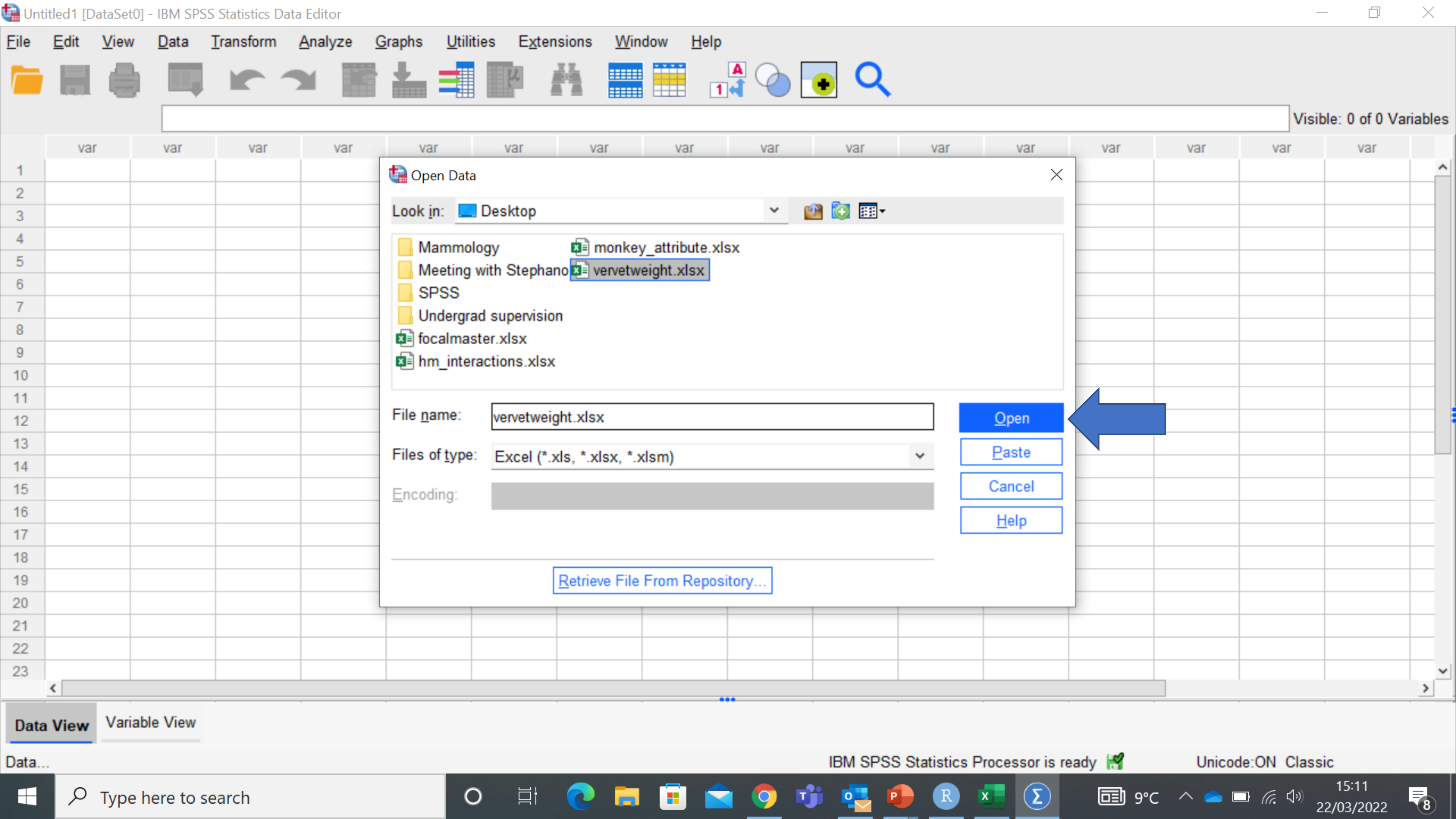

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 0 of 0 Variables
var
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
9°C
15:09
22/03/2022

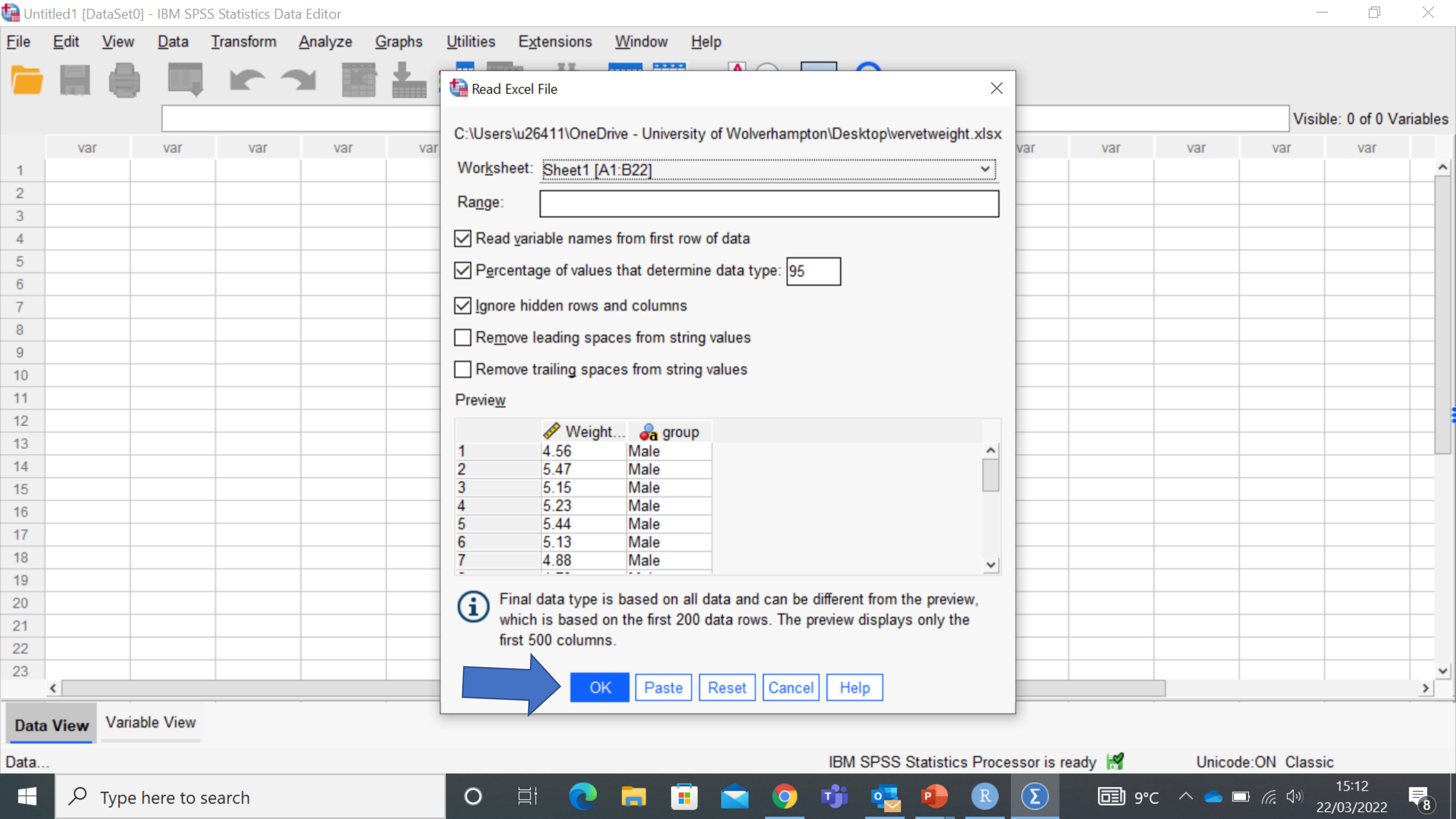

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
New
Open
Import Data
General Open
Open Restore Point...
Close Ctrl+F4
Save Ctrl+S
Save As...
Save All Data
Save Restore Point...
Export
Mark File Read Only
Revert to Saved File
File Enabled for Auto-Recovery
Collect Variable Information
Rename Dataset...
Display Data File Information
Cache Data...
Stop Processor Ctrl+Period
Switch Server...
Repository
Print Preview
Print... Ctrl+P
Visible: 0 of 0 Variables
var
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
9°C
15:09
22/03/2022


Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
New
Open
Import Data
General Open
Open Restore Point...
Close Ctrl+F4
Save Ctrl+S
Save As...
Save All Data
Save Restore Point...
Export
Mark File Read Only
Revert to Saved File
File Enabled for Auto-Recovery
Collect Variable Information
Rename Dataset...
Display Data File Information
Cache Data...
Stop Processor Ctrl+Period
Switch Server...
Repository
Print Preview
Print... Ctrl+P
Data...
Internet Data
Syntax...
Output...
Workbook...
Script...
Visible: 0 of 0 Variables
var
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
9°C
15:10
22/03/2022



Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 0 of 0 Variables
Open Data
Look in: Documents
Custom Office Templates
Fax
Scanned Documents
File name:
Files of type: SPSS Statistics (*.sav, *.zsav)
Open
Paste
Cancel
Help
Retrieve File From Repository...
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
15:10
22/03/2022


Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 0 of 0 Variables
var
Open Data
Look in:
Desktop
Mammology
Meeting with Stephano
SPSS
Undergrad supervision
focalmaster.xlsx
hm_interactions.xlsx
monkey_attribute.xlsx
vervetweight.xlsx
File name:
vervetweight.xlsx
Files of type:
Excel (*.xls, *.xlsx, *.xlsm)
Encoding:
Open
Paste
Cancel
Help
Retrieve File From Repository...
Data View
Variable View
Data...
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
9°C
15:11
22/03/2022


Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 0 of 0 Variables
Read Excel File
C:\Users\u26411\OneDrive - University of Wolverhampton\Desktop\vervetweight.xlsx
Worksheet: Sheet1 [A1:B22]
Range:
Read variable names from first row of data
Percentage of values that determine data type: 95
Ignore hidden rows and columns
Remove leading spaces from string values
Remove trailing spaces from string values
Preview
Weight... group
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
Final data type is based on all data and can be different from the preview, which is based on the first 200 data rows. The preview displays only the first 500 columns.
OK Paste Reset Cancel Help
Data View Variable View
Data...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:12
22/03/2022

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

	Weight_kg	group
1	4.56	Male
2	5.47	Male
3	5.15	Male
4	5.23	Male
5	5.44	Male
6	5.13	Male
7	4.88	Male
8	4.79	Male
9	4.98	Male
10	5.01	Male
11	4.56	Female
12	4.76	Female
13	4.34	Female
14	3.97	Female
15	5.15	Female
16	4.45	Female
17	4.14	Female
18	4.55	Female
19	4.78	Female
20	4.84	Female
21	4.55	Female

Column: ‘group’
- 10 Males
- 11 Females

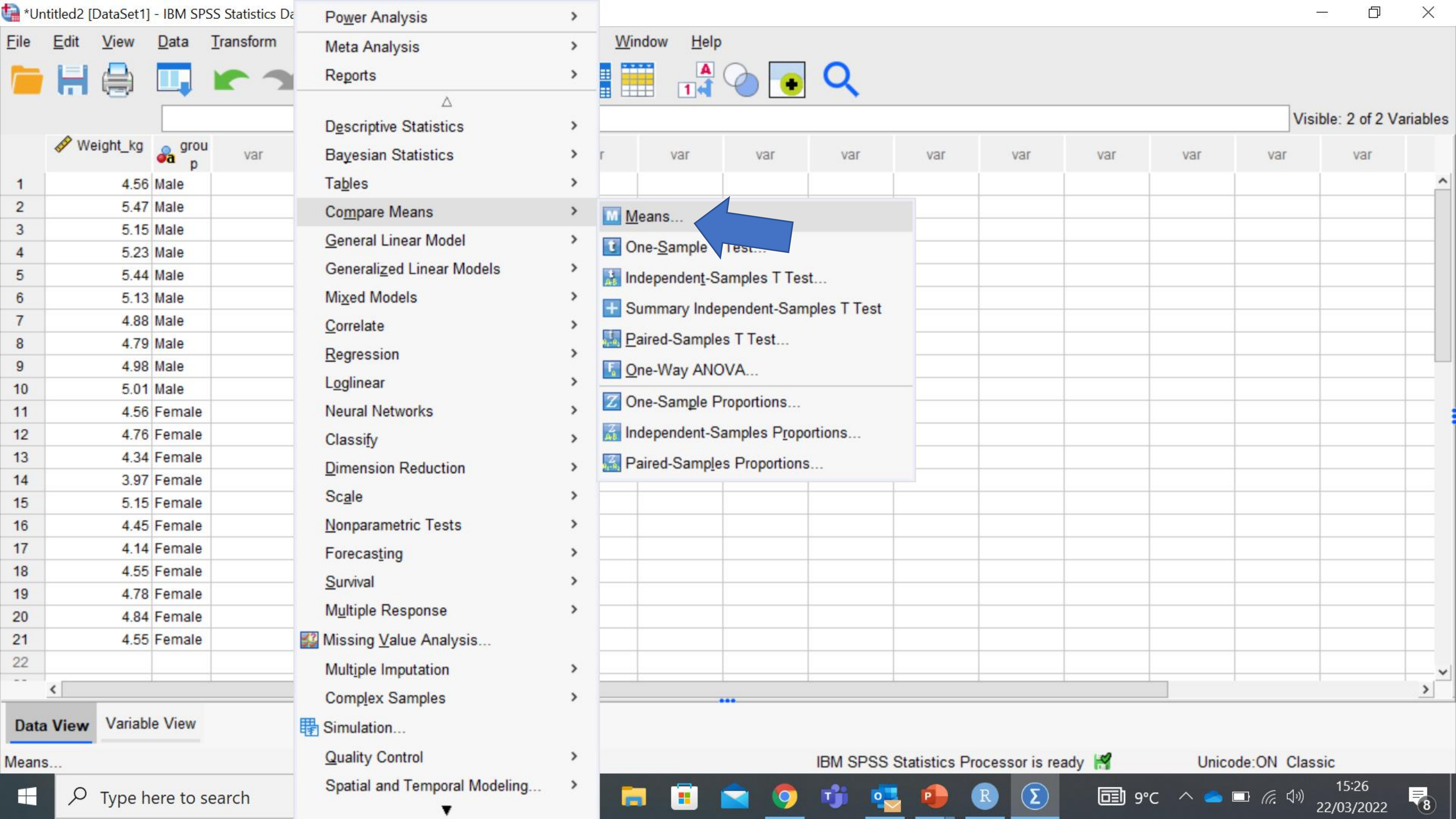
Column: Weight_kg
- Weights of different individuals

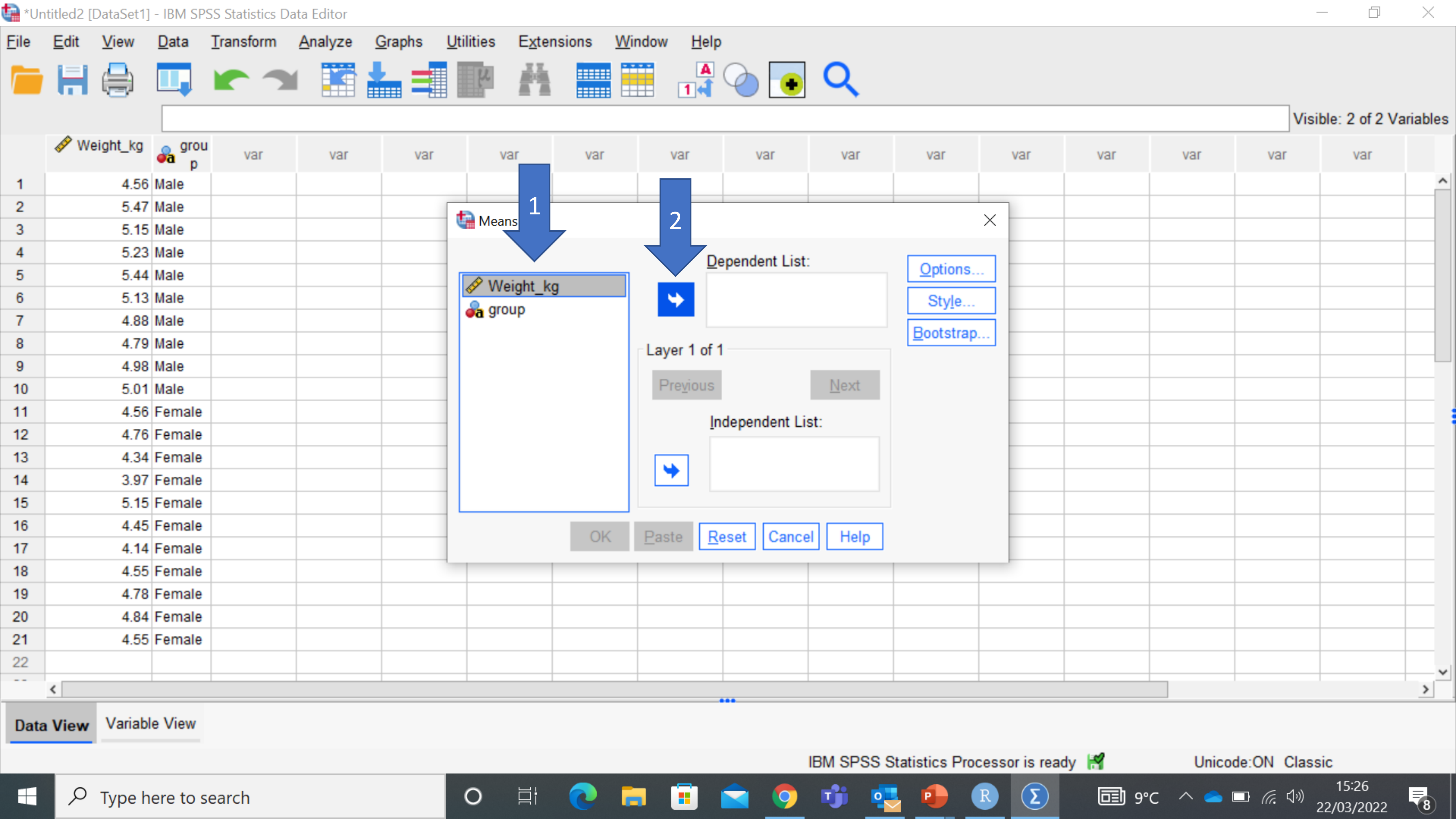
	Weight_kg	group
1	4.56	Male
2	5.47	Male
3	5.15	Male
4	5.23	Male
5	5.44	Male
6	5.13	Male
7	4.88	Male
8	4.79	Male
9	4.98	Male
10	5.01	Male
11	4.56	Female
12	4.76	Female
13	4.34	Female
14	3.97	Female
15	5.15	Female
16	4.45	Female
17	4.14	Female
18	4.55	Female
19	4.78	Female
20	4.84	Female
21	4.55	Female

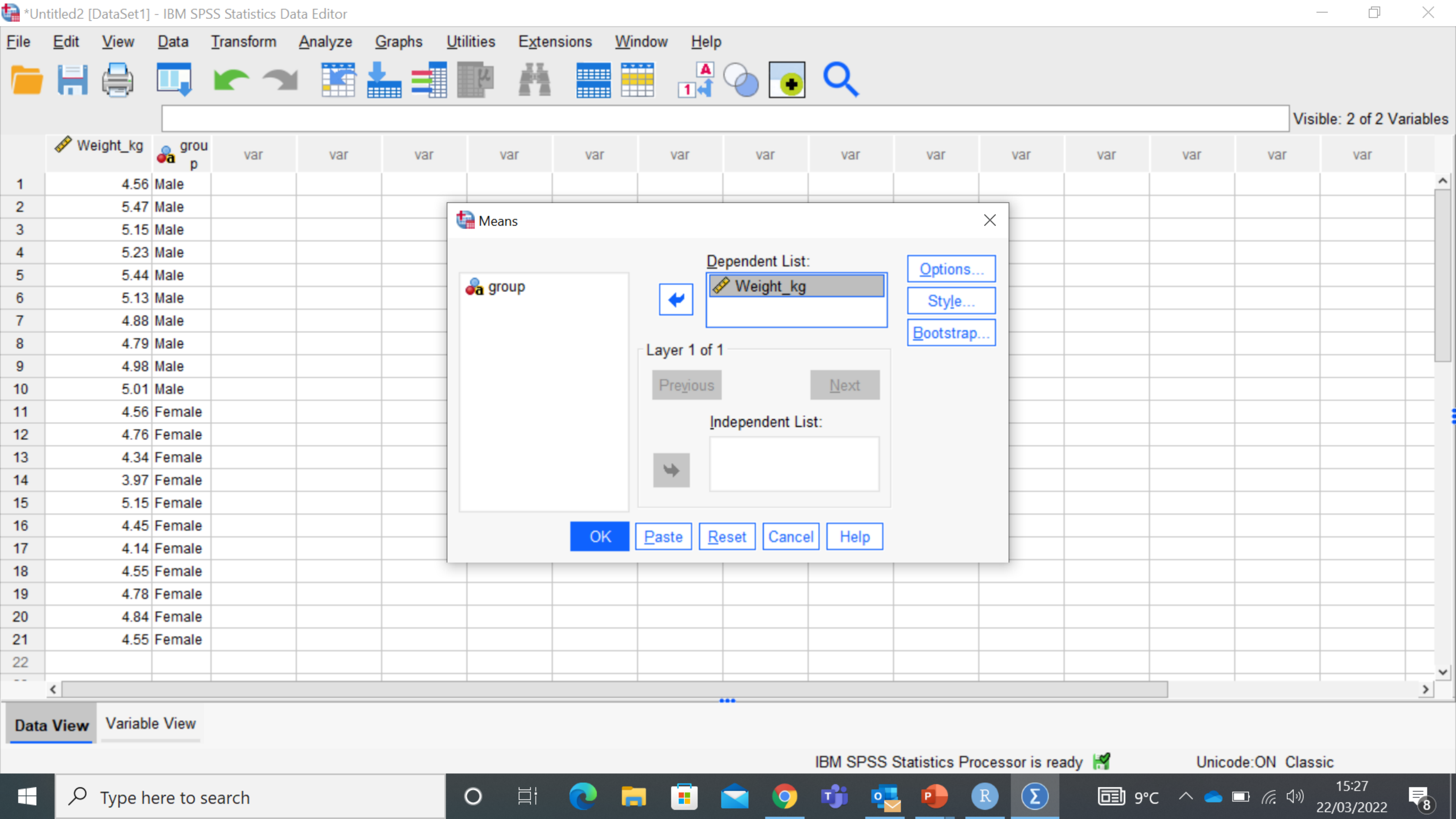
Lets briefly explore our data!

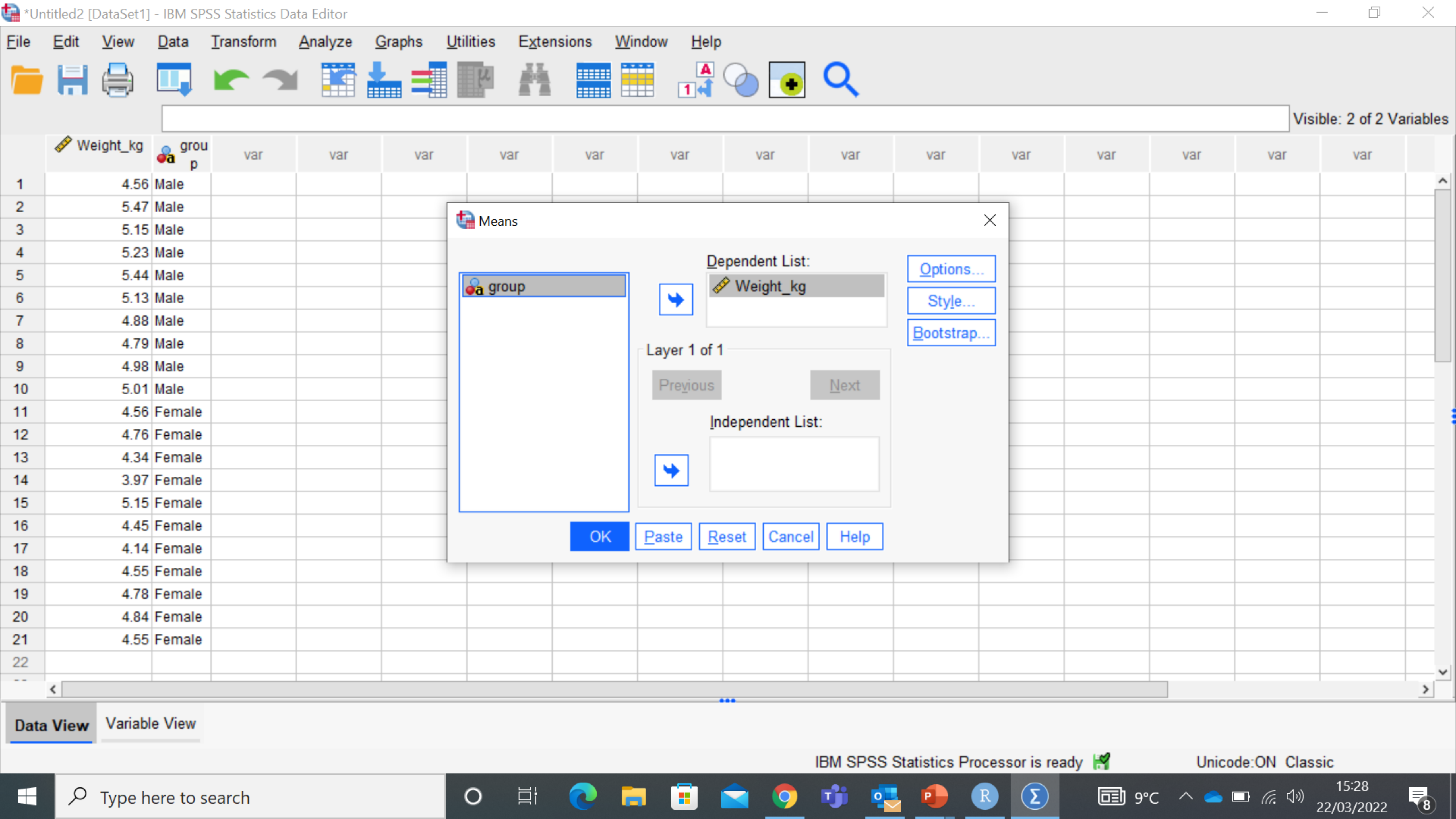
Perhaps the first thing we can look into is the mean value of our groups…

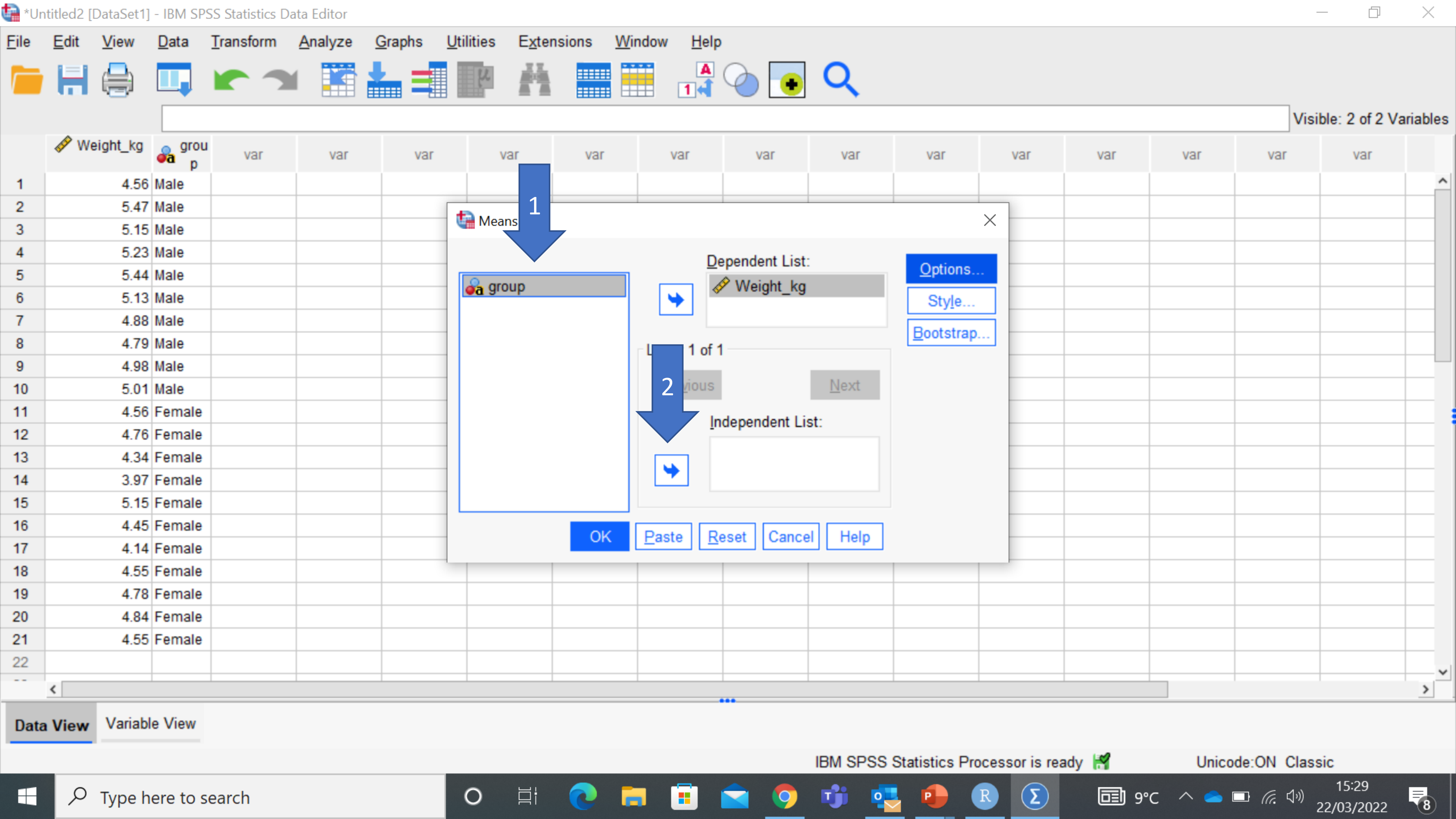
Did males and females have a different weight on average?

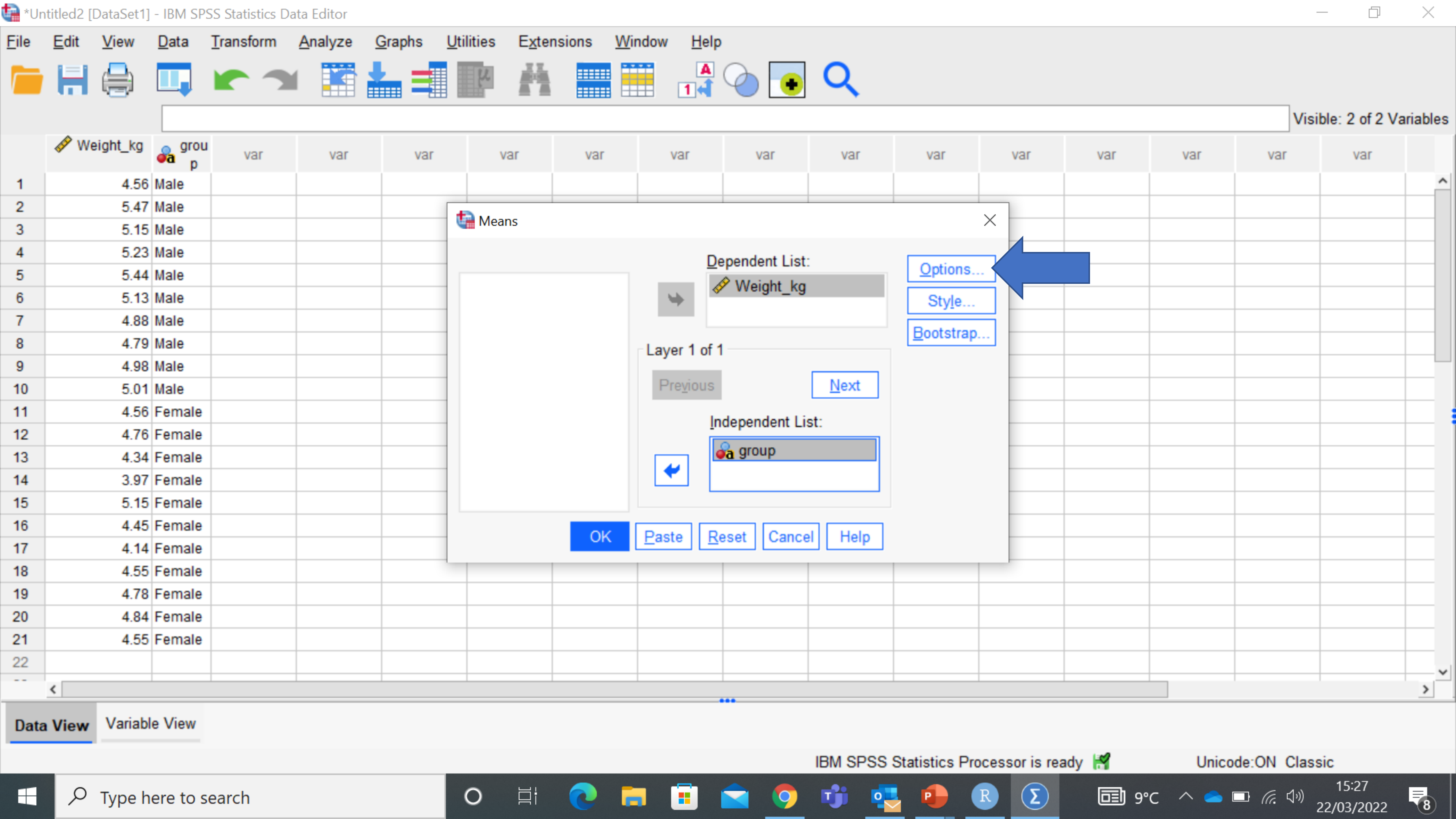

*Untitled2 [DataSet1] - IBM SPSS Statistics Da
File Edit View Data Transform
Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Data View
Variable View
Means...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:26
22/03/2022

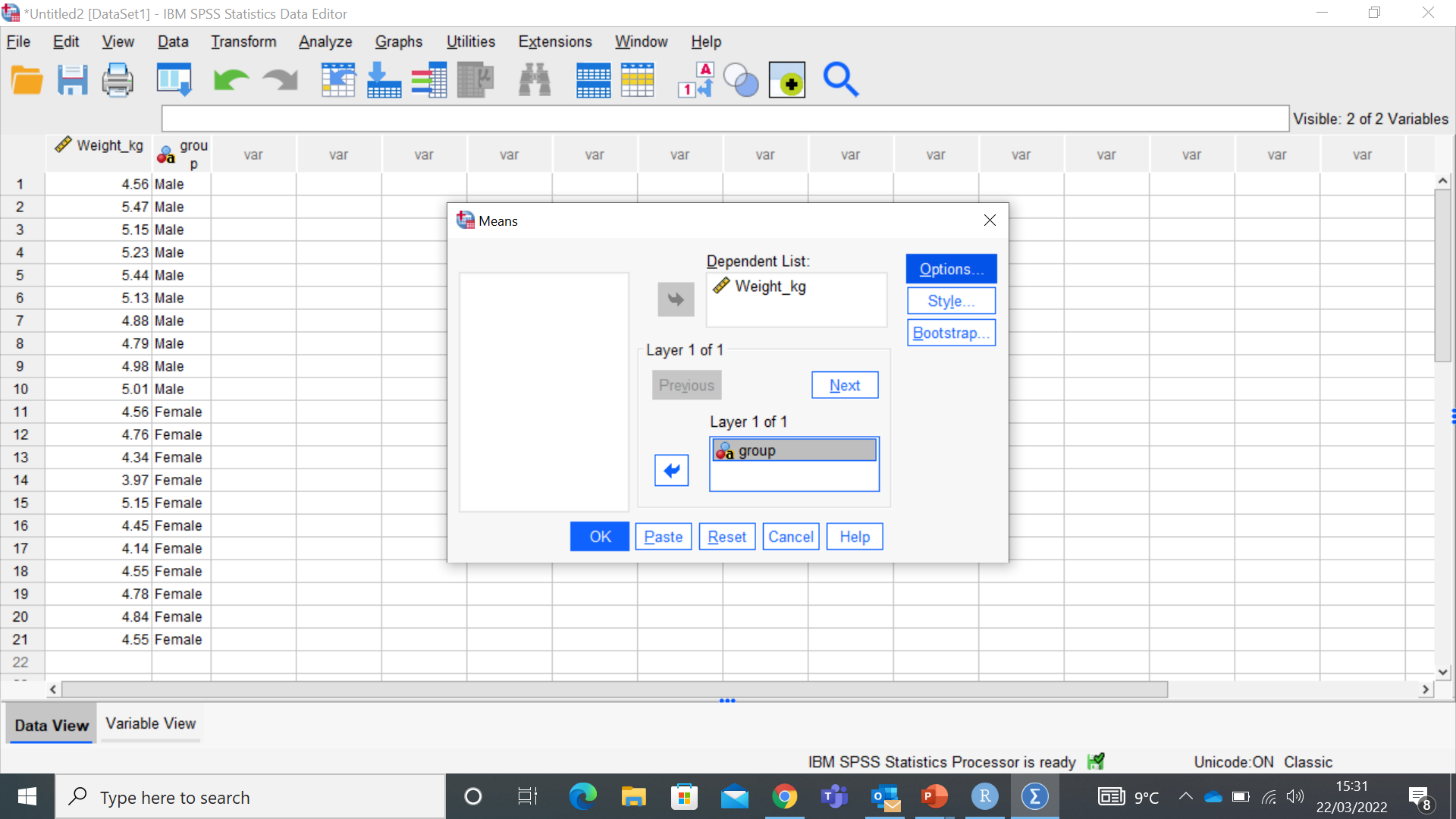

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
1
2
Weight_kg
group
Dependent List:
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Independent List:
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:26
22/03/2022

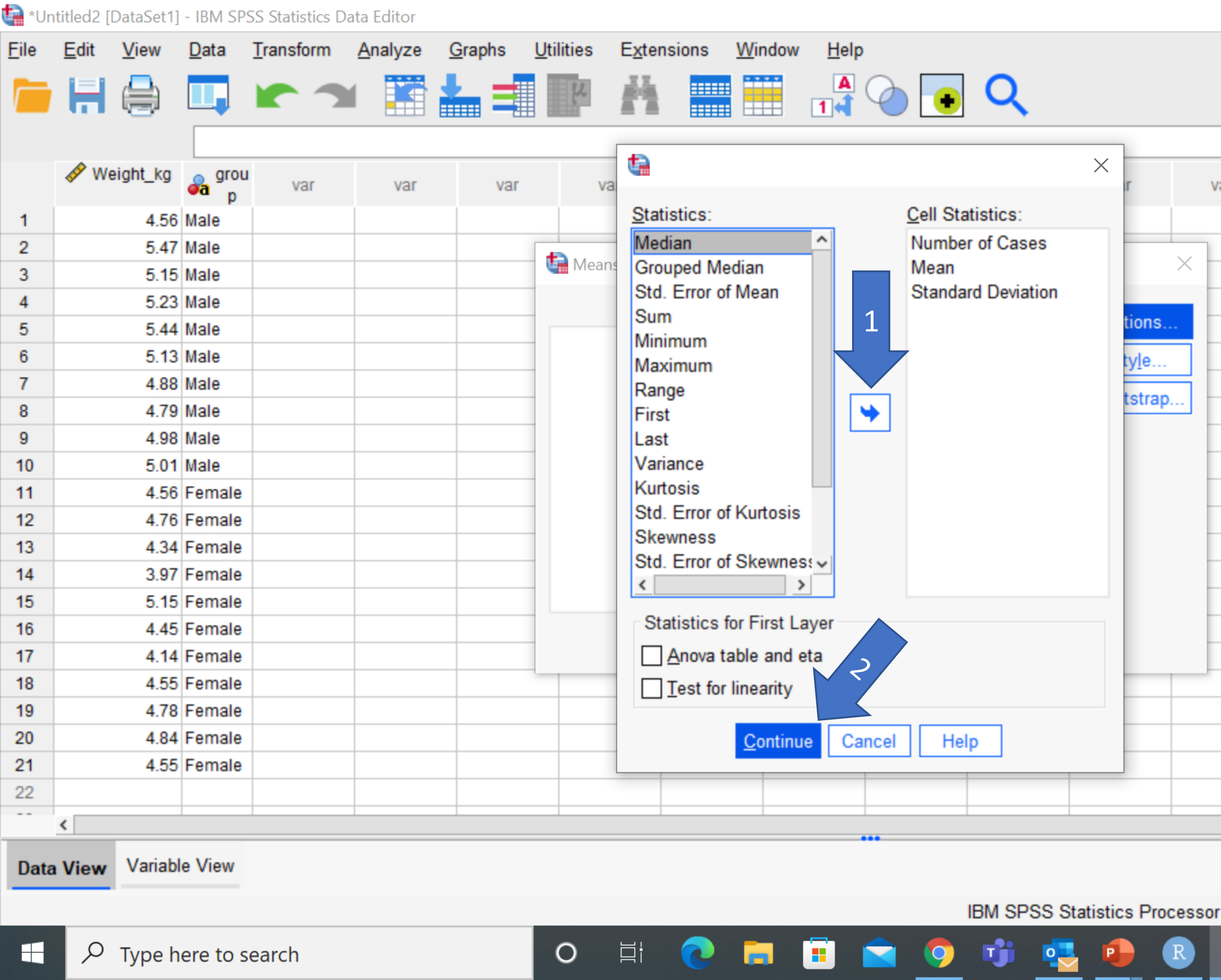

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
group
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Independent List:
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:27
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg group var var var var var var var var var var var var var var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
group
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Independent List:
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:28
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
Means
1
group
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
1 of 1
2
Next
Independent List:
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:29
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Independent List:
group
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:27
22/03/2022

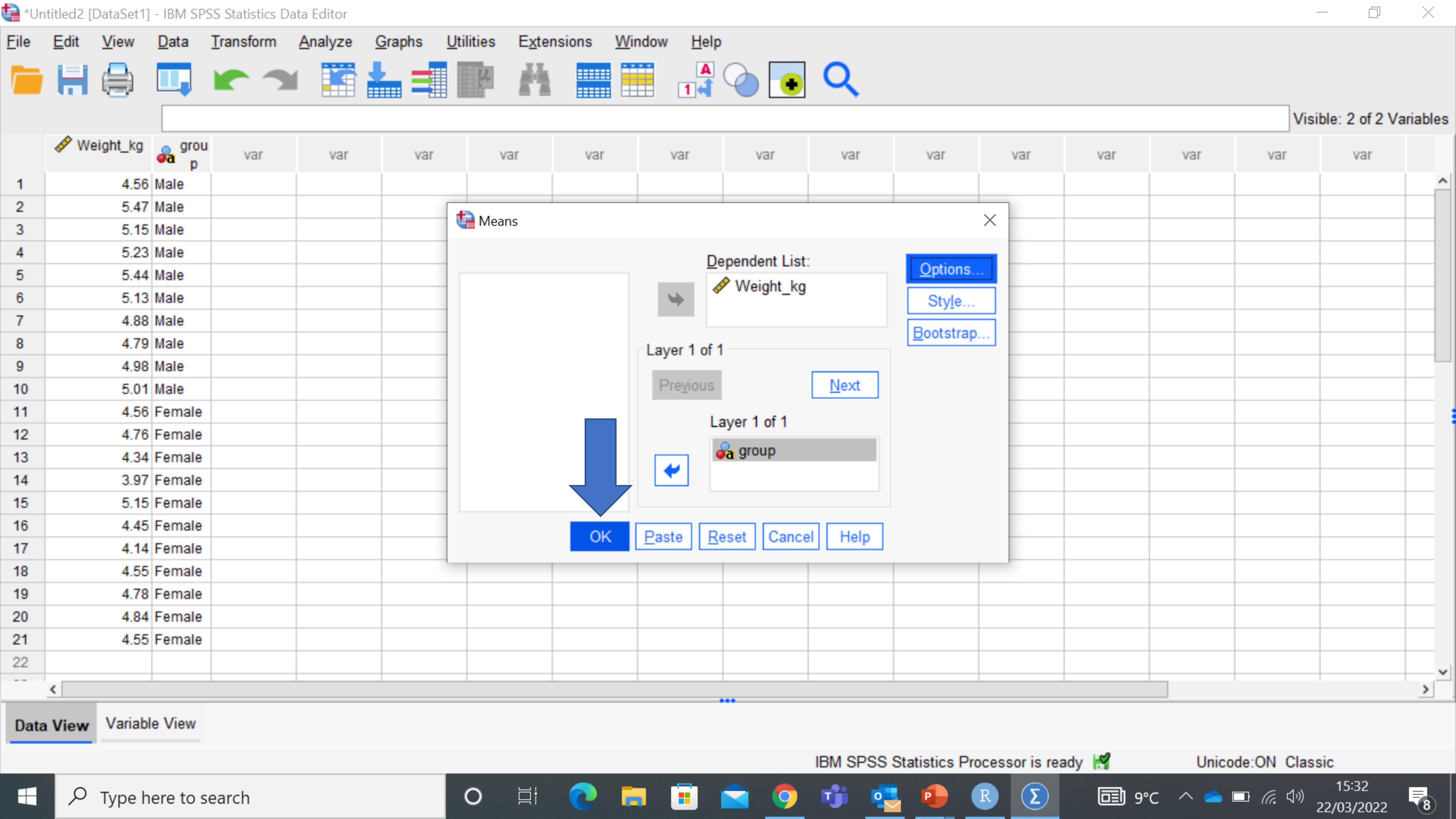

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
group
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:31
22/03/2022



Select the measurements you are interested in from the 'Statistics' box

Move them over to the 'Cell statistics' box with the arrow button

When you are done, click continue


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Means
Dependent List:
Weight_kg
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
group
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
15:32
22/03/2022

*Output5 [Document5] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - Means
    - Title
    - Notes
    - Case Processing Summary
    - Report

Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Weight_kg * group	21	100.0%	0	0.0%	21	100.0%

Report

Weight_kg

group	N	Mean	Std. Deviation
Female	11	4.5536	.33161
Male	10	5.0640	.28254
Total	21	4.7967	.39890

Double-click to activate

Notice here, we have counts by group

- We have 11 females in our sample
- We have 10 males in our sample
- A total of 21 individuals in our sample!

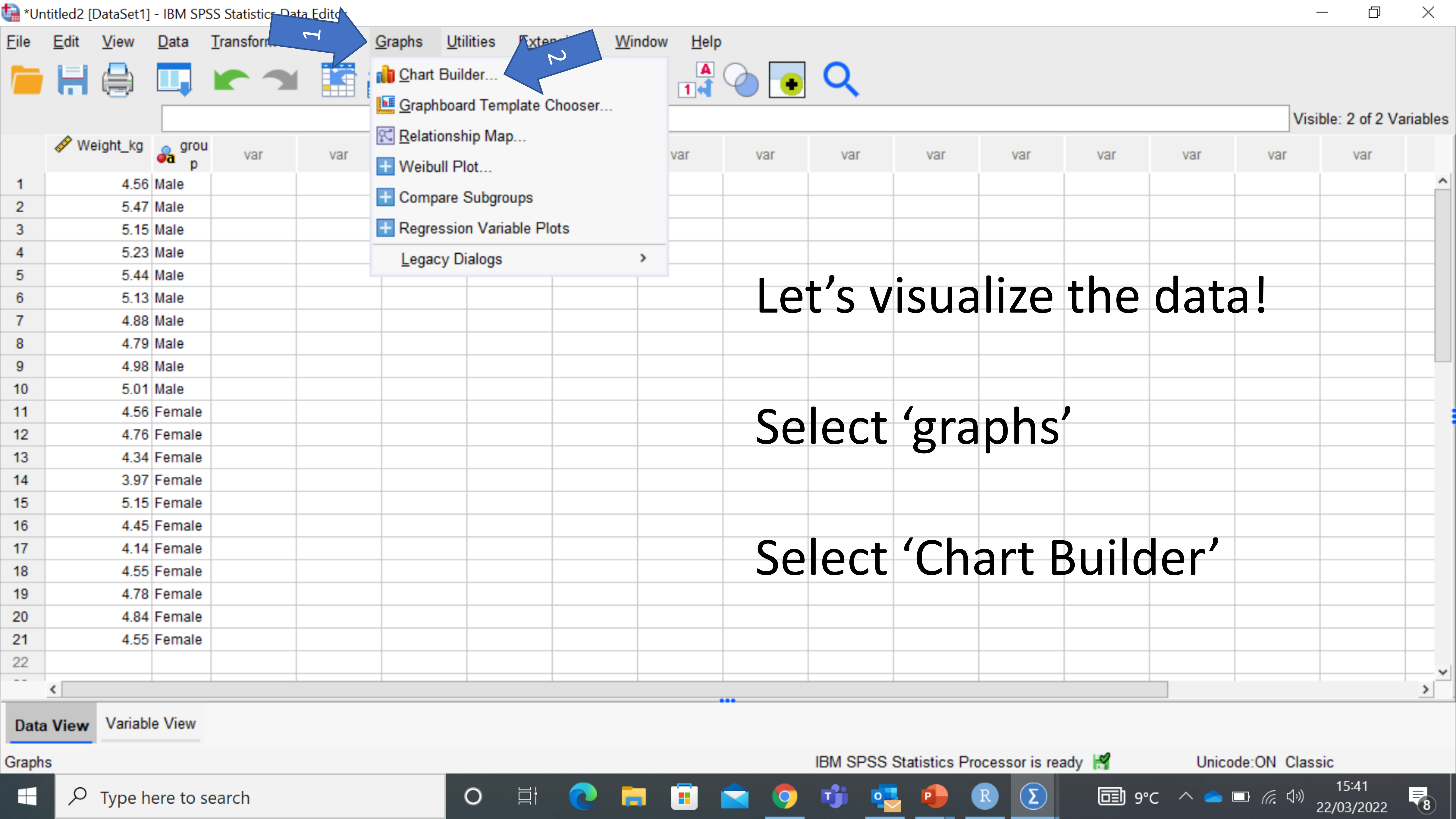
We also have mean weight by group:

- Females have an average weight of 4.55kg
- Males have an average weight of 5.06kg
- The average weight of all monkeys is 4.80kg

- Based on this, I looks like males *might* be heavier than males… But there is more we can do to investigate this properly!

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

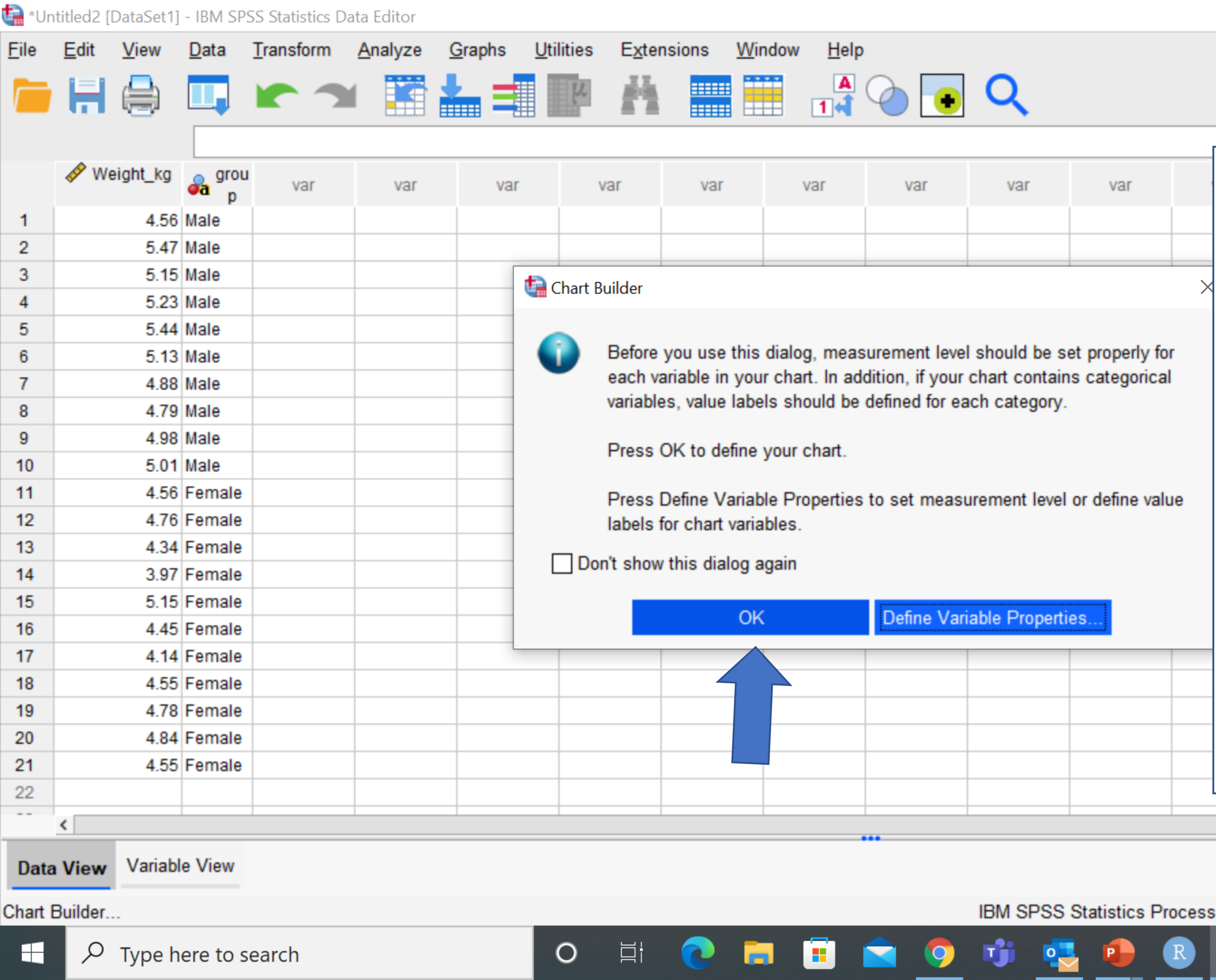
Type here to search   9°C   15:32 22/03/2022



Let’s visualize the data!

Select ‘graphs’

Select ‘Chart Builder’

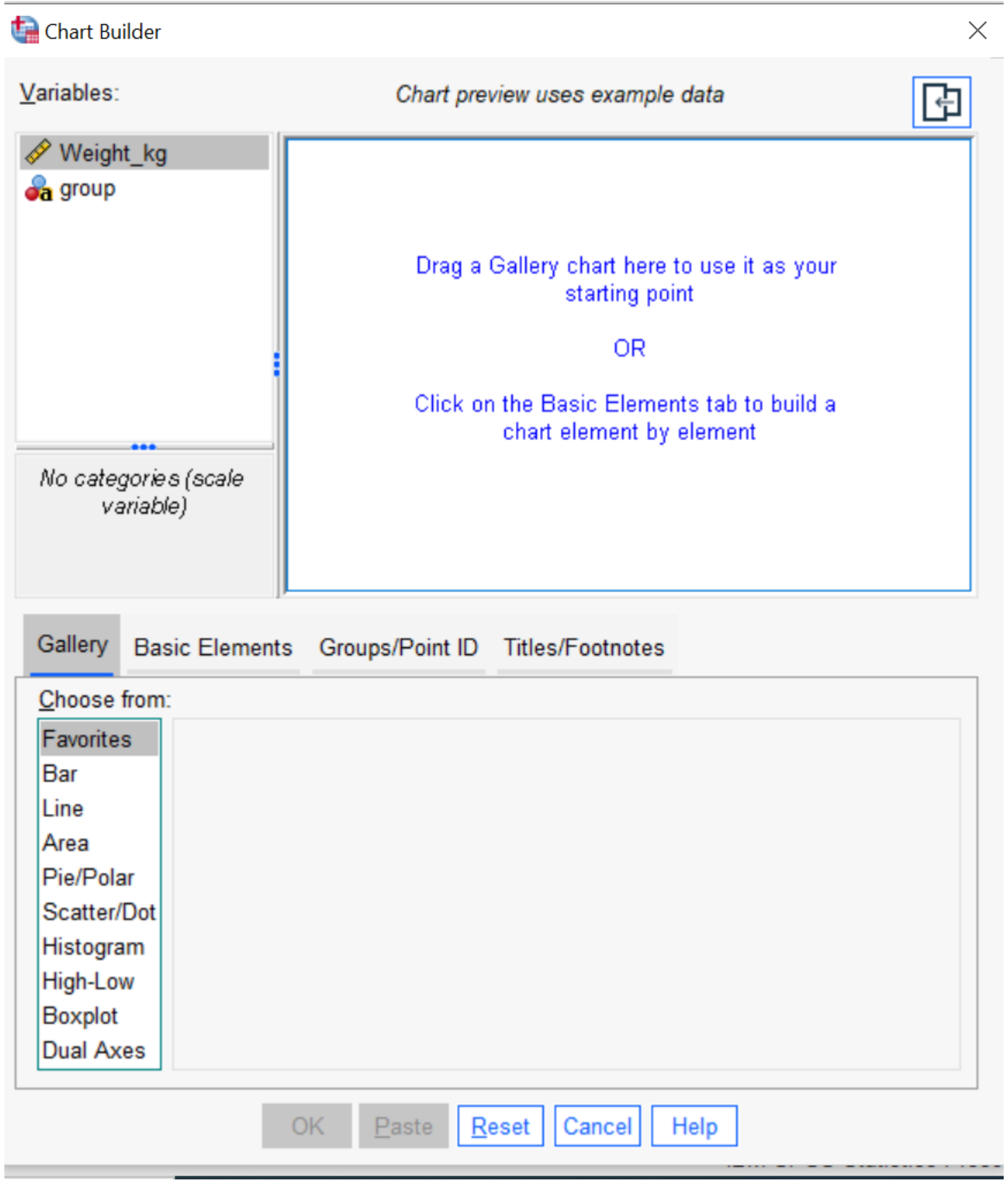


This warning dialogue may come up

It is telling you to be sure your data are set to the correct 'measurement level'

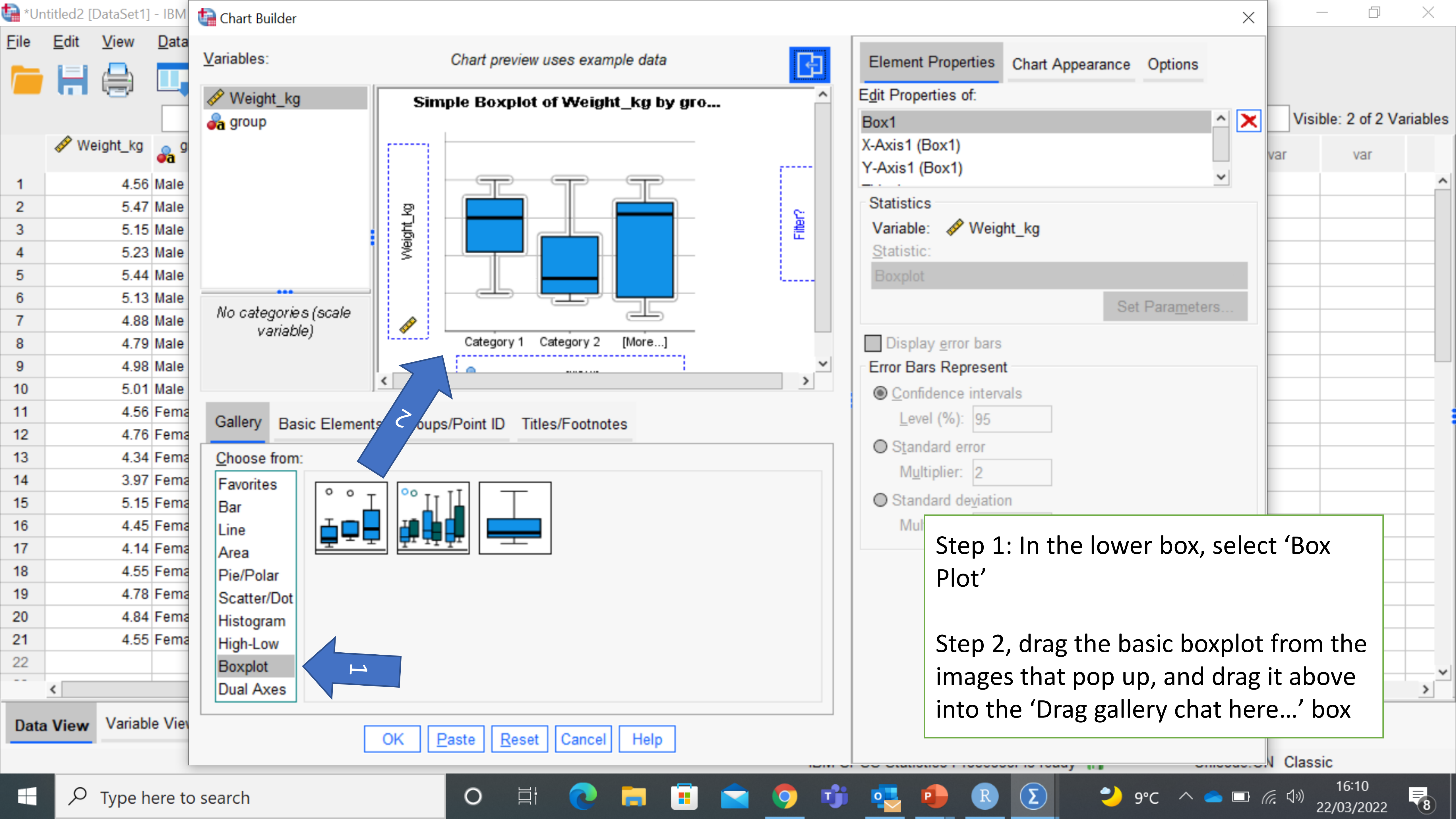
Remember that sex is a categorical variable. In SPSS this is coded as 'Nominal'

Weight is continuous
In SPSS, this is coded as 'Scale'



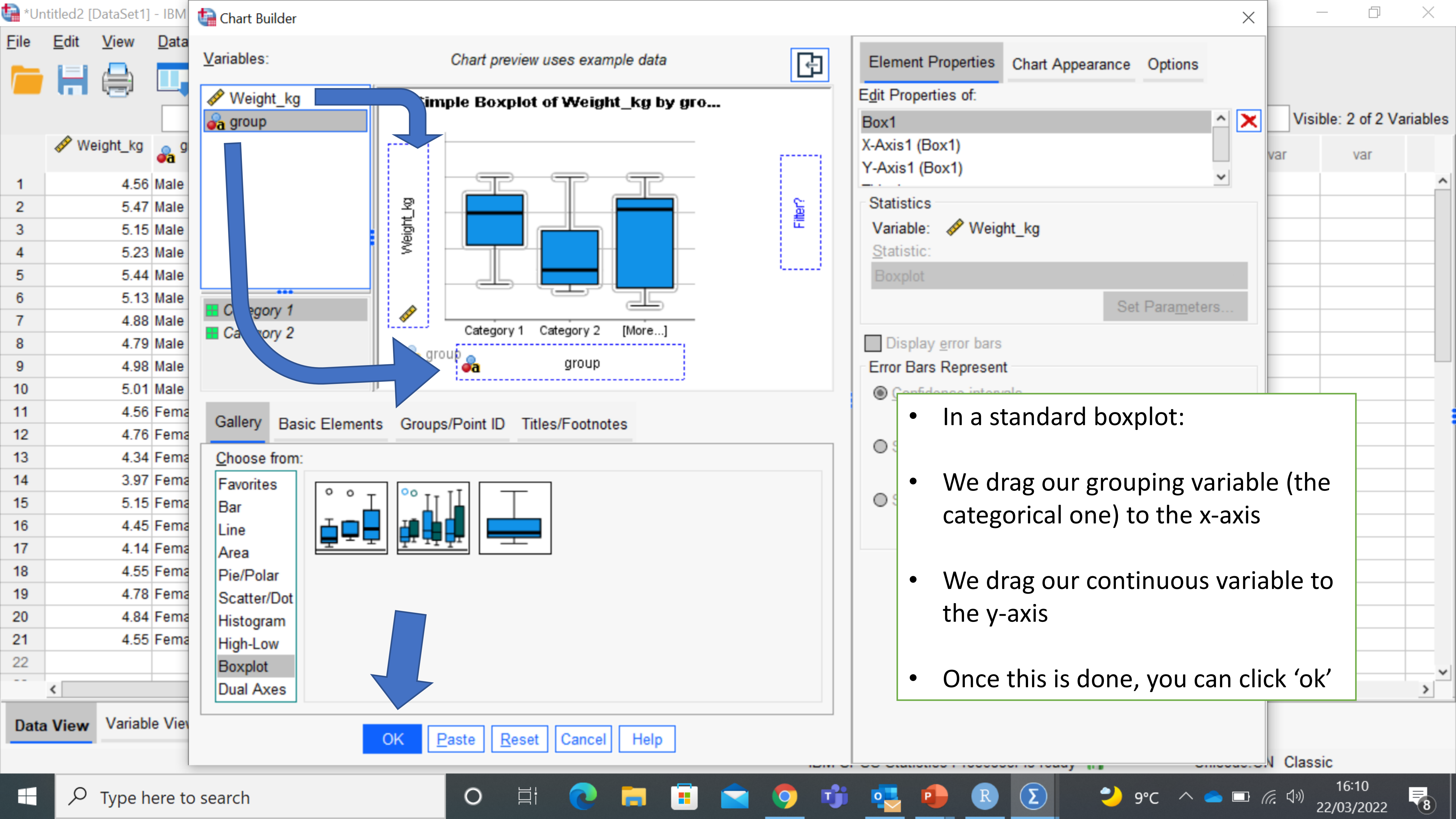
- This should bring up a window that looks something like this…

- We want to compare the value of a continuous variable across different categories
- A box plot is our best friend for this!

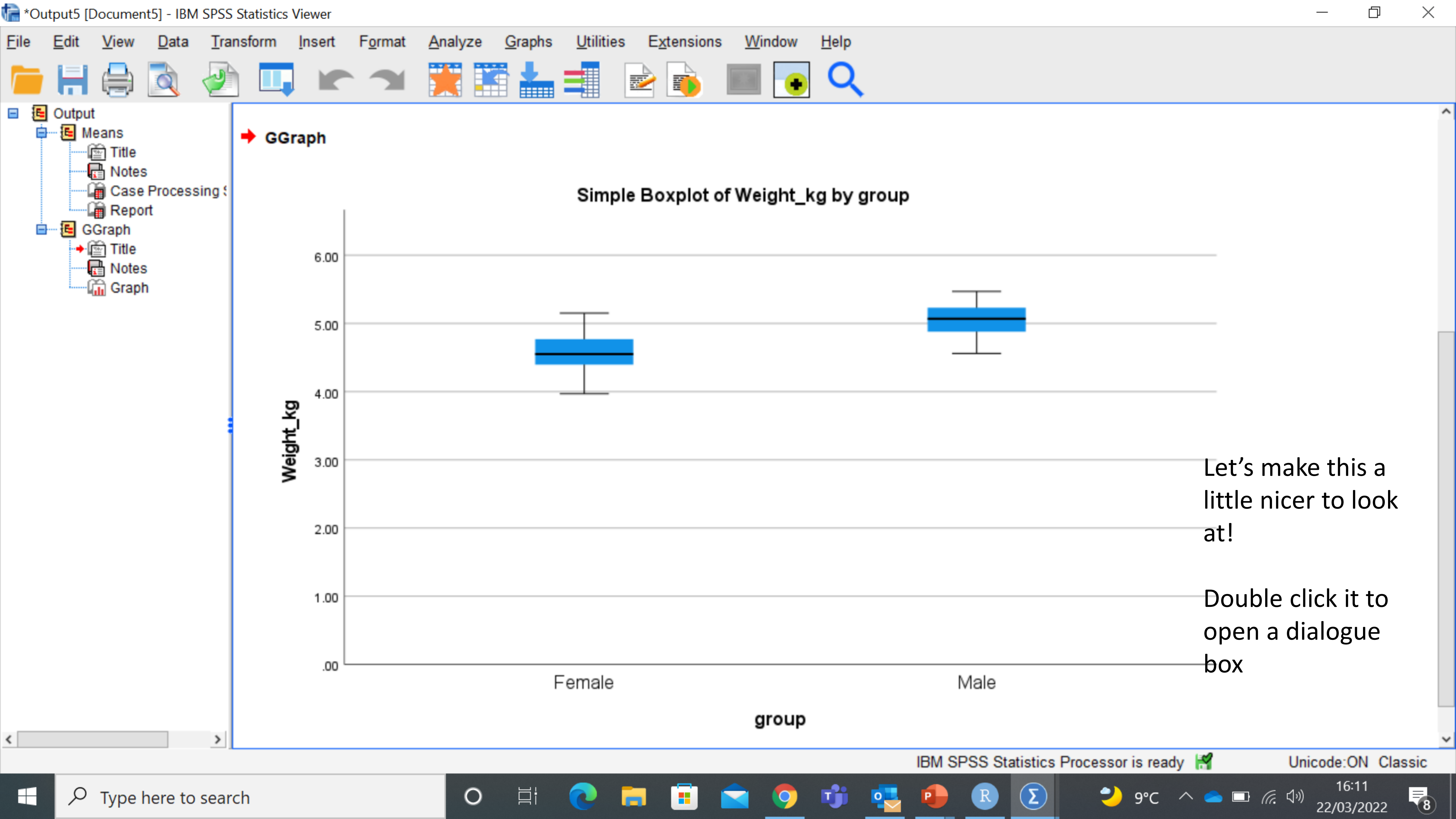


Step 1: In the lower box, select ‘Box Plot’

Step 2, drag the basic boxplot from the images that pop up, and drag it above into the ‘Drag gallery chat here…’ box

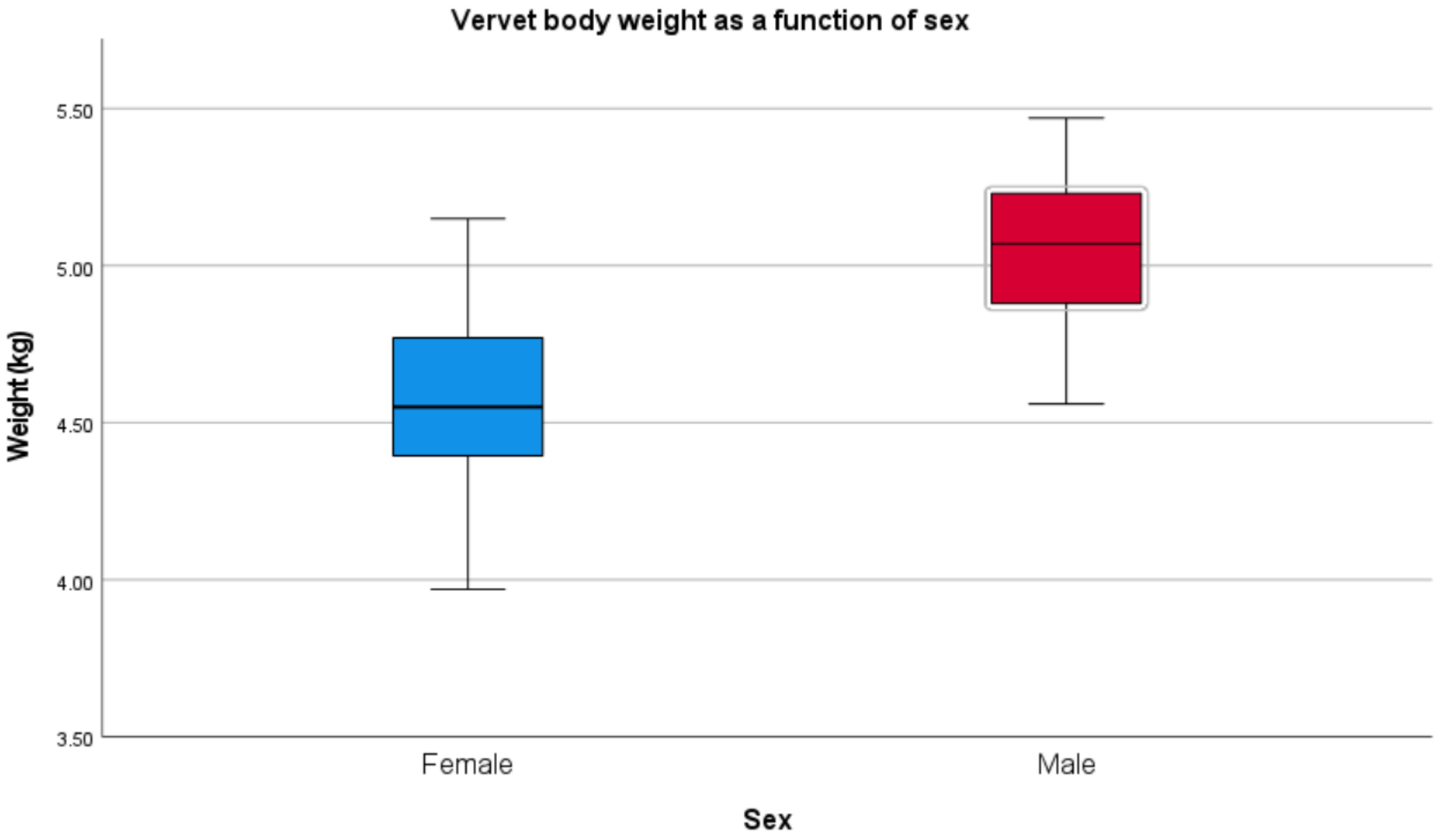


- In a standard boxplot:
- We drag our grouping variable (the categorical one) to the x-axis
- We drag our continuous variable to the y-axis
- Once this is done, you can click ‘ok’



Let’s make this a little nicer to look at!

Double click it to open a dialogue box



- That's better

- Hey, this figure seems to line up nicely with our inference from the means:
  - Males seem to be heavier than females!

- Ok. So we have different mean values for the weight of males vs female vervet monkeys
- Our figure seems to back up the idea that males are heavier than females

- But is there a statistically significant difference between the males and females?
  - For this, we will explore the **independent samples T Test**

Independent samples T Test

- A test used to see if there is a significant difference between the **means** of **two groups**
  - **You can start thinking about T Tests when you have 2 groups**
  - **The variable that represents the ‘groups’ is by definition a categorical variable**
  - **You have a second variable that you think might vary by group, and that variable is continuous**

- **The null hypothesis:** the 2 population (group) means are equal

- **Alternative hypothesis:** the 2 population (group) means are different

IMPORTANT: Assumptions

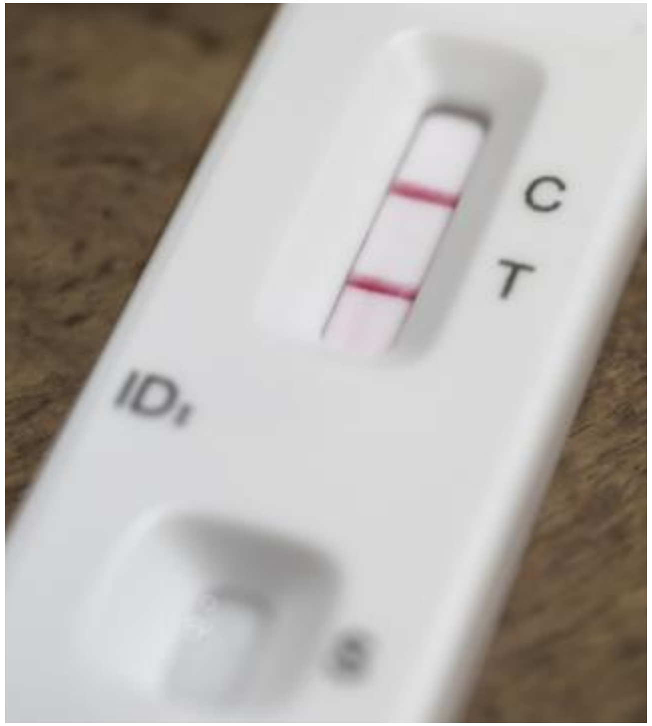
- Any statistical test is a procedure that can tell us something about out data, provided certain **assumptions** are met…
- Results from a test cannot be interpreted properly if basic assumptions are not met!

- Let’s take/use a COVID rapid antigen test as a metaphore

IMPORTANT: Assumptions

- If you showed me a COVID test you brought from home, and it is displaying a positive result, I would believe you had COVID, as long as some ASSUMPTIONS were met
  - Did YOU take the test, or did you swab our brother?
  - Did you take the test recently, or was that test taken 3 weeks ago?
  - Did you follow the instructions on the box?
- Assuming these ASSUMPTIONS are met, I would infer from a positive COVID test that you have COVID…
- Otherwise, that test doesn't mean much!

- The same logic holds true for any statistical test!



Assumptions of an Two Samples T Test

- Independence of observations
- Normality
- No significant Outliers
- Homogeneity of Variance

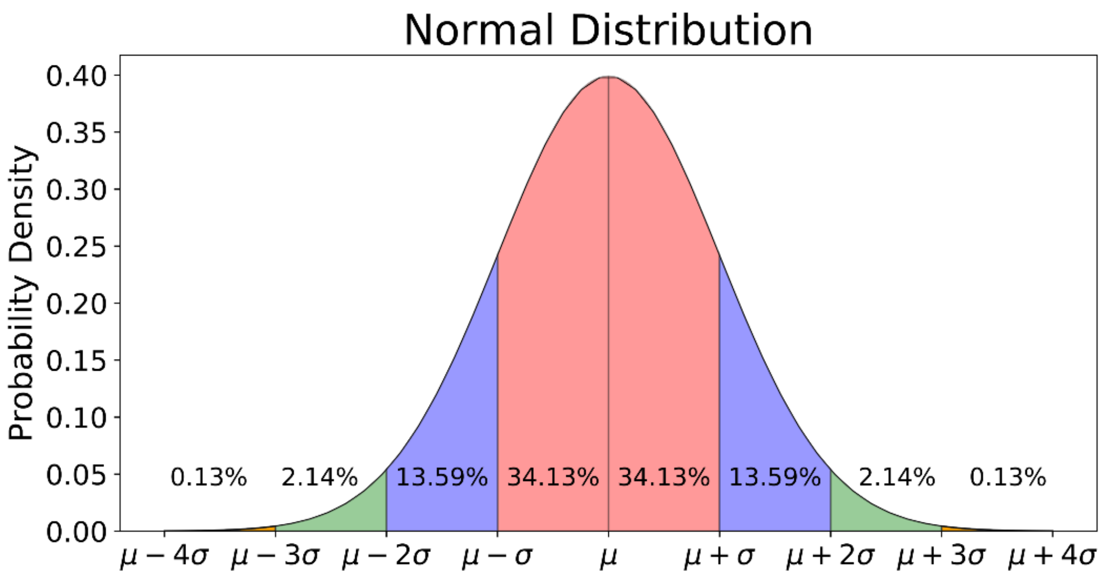
Assessing assumptions

- Independence
  - Here, you are really assessing the **research design**
  - Was the data sampled randomly?
  - Can you consider the data points independent of one another?

- Possible independence issues to consider:
  - If you took all your "female" measurements from a group of related females
  - Did you accidentally collect male weights before a drought, and then females during the drought?
- This assumption needs to be considered thoroughly!
  - If the assumption is not met, **there are other, more appropriate statistics you can use**!

Normality + Outliers

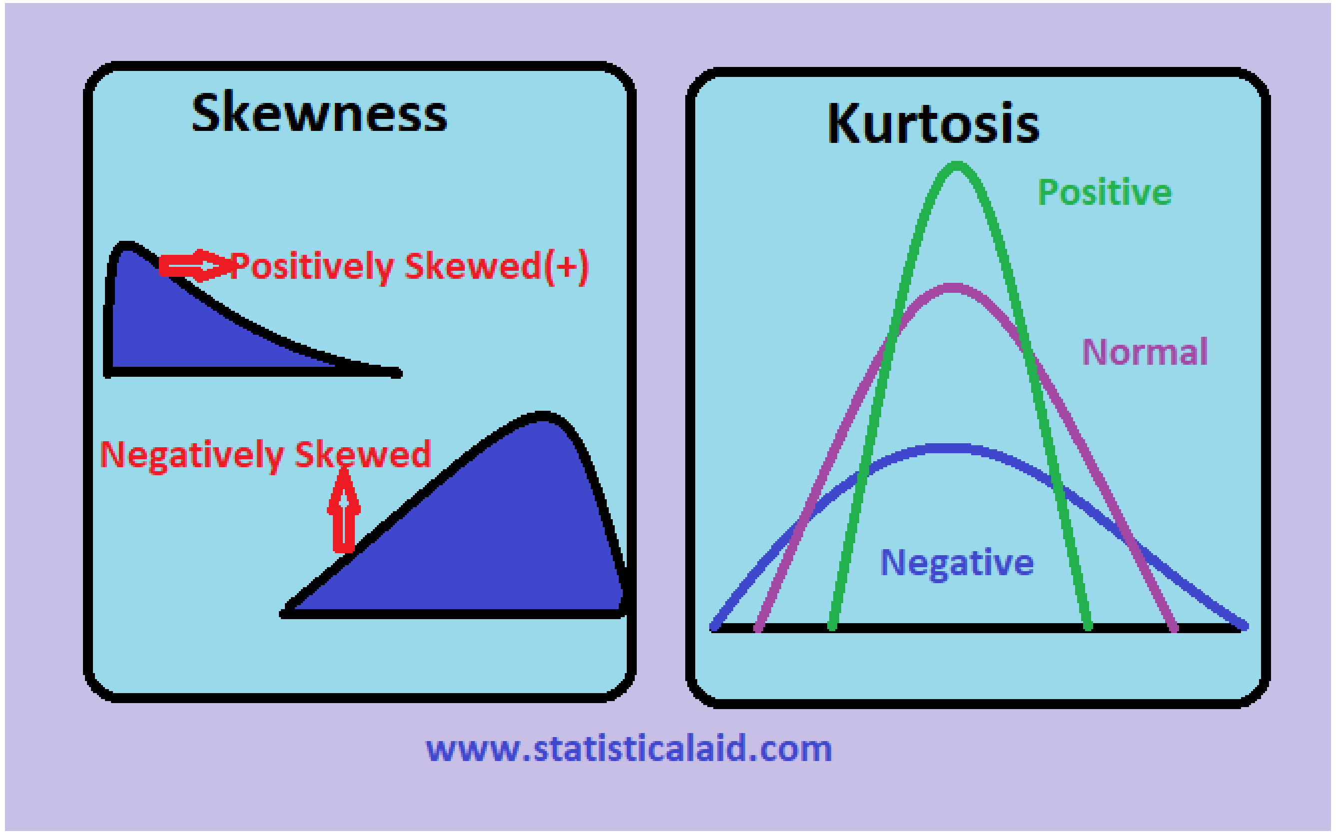
- The distribution of the data **in both your groups** should be 'normal'
- There are some specific mathematical descriptions of what makes a **normal distribution**, but for our purposes today:
  - The Normal distribution (AKA Gaussian distribution), is a distribution where data is spread symmetrical around the mean
- The bulk of the data is closer to the mean, and tapers off as you move away from the mean.
- The standard 'bell curve'.



~68% of the data should be with 1 SD from the mean
~95% of the data should be with 2 SD of the mean
~99.7% of the data should be within 3 SD of the mean

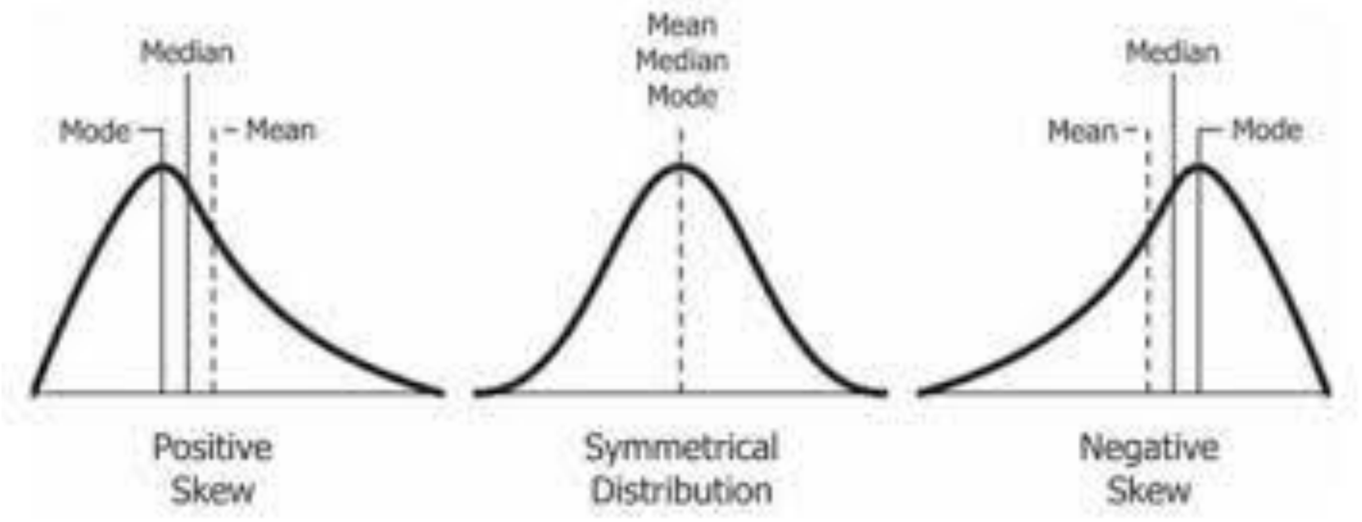
Normality + Outliers

- There are some very specific criteria that must be met for a ‘bell curve’ to be considered normally distributed
- One way of checking for normality is to look at **skewness** and **kurtosis**



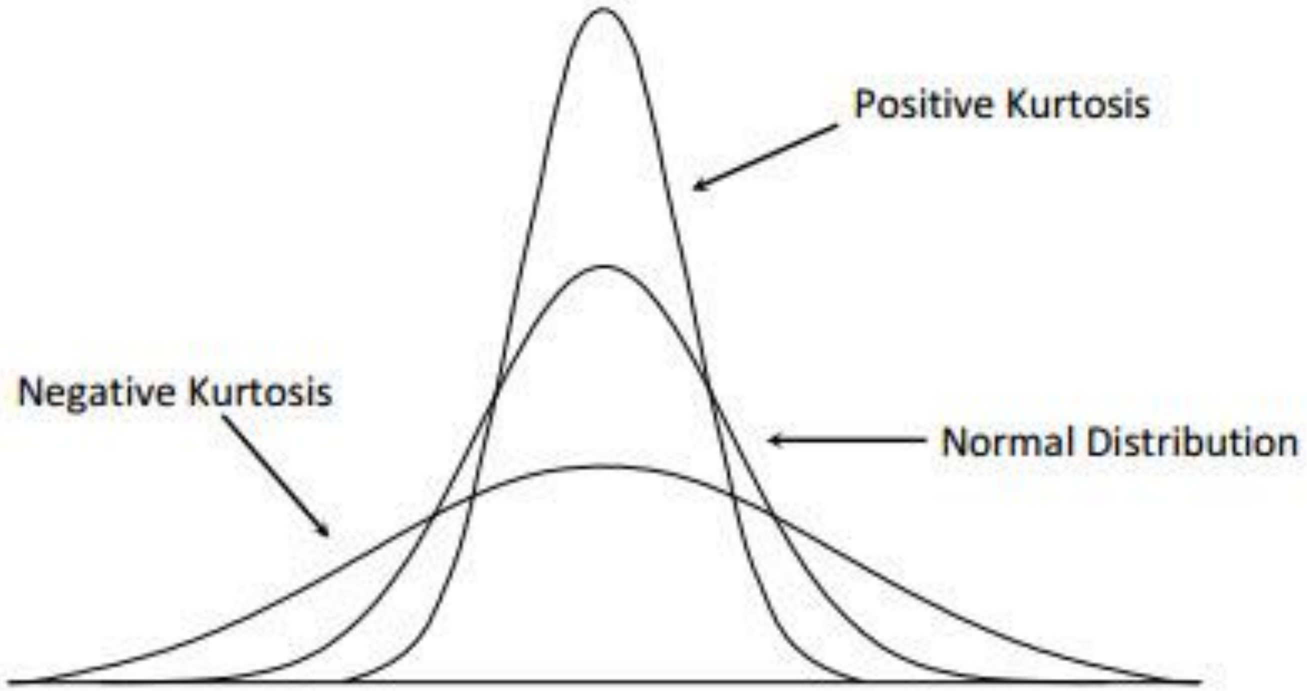
Skewness

- **Skewness**: a measure of symmetry
  - **Positively skewed** distributions have long tails to the right (positive side) of the distribution
  - **Negatively skewed** distributions have long tails to the left (negative side) of the distribution
- Normal distributions should have low (close to zero) levels of skewness



Kurtosis

- **Kurtosis**: How ‘fat’ are the tails of your distribution
- Normal distributions have kurtosis values close to zero
  - If your data have **positive kurtosis values**, sharper peak and heavier tails compared to a normal distribution
  - **Negative kurtosis values** suggest flatter peak and thinner tails compared to a normal distribution



Assessing Normality: Skewness and Kurtosis

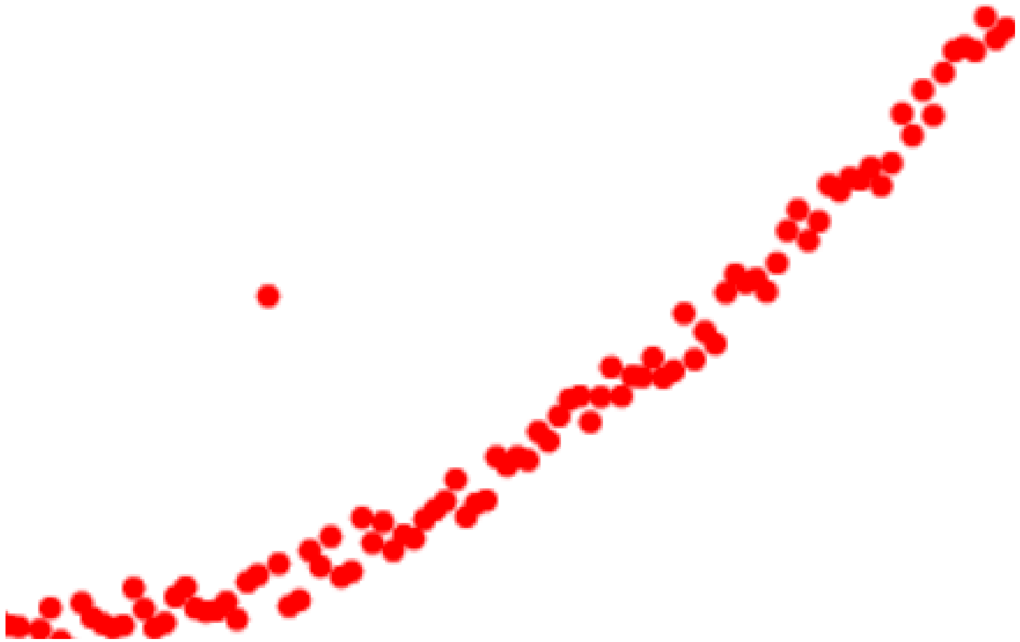
- Without getting too math-y: As a rule of thumb, you are looking for values of kurtosis and skewness that are **between -2 and +2 (zero would be optimum)**
- Normality can also be assessed using a statistical test: **Shapiro-Wilk**
- We also assess normality using **qqplots**.

- But wait… what about outliers?!

Normality + **Outliers**

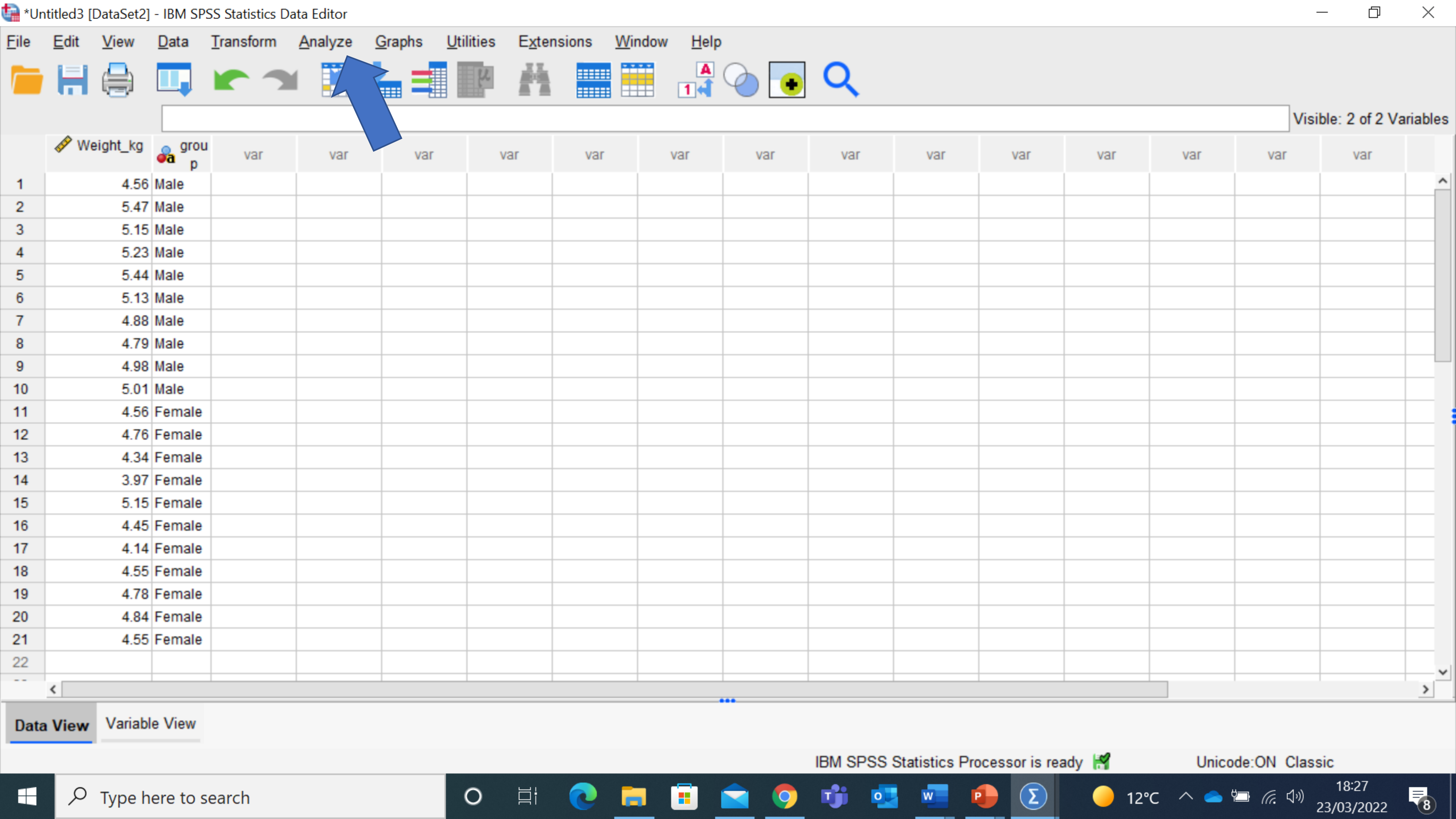
- Simply put, an outlier is an extremely high or extremely low data point relative to the rest of your data set

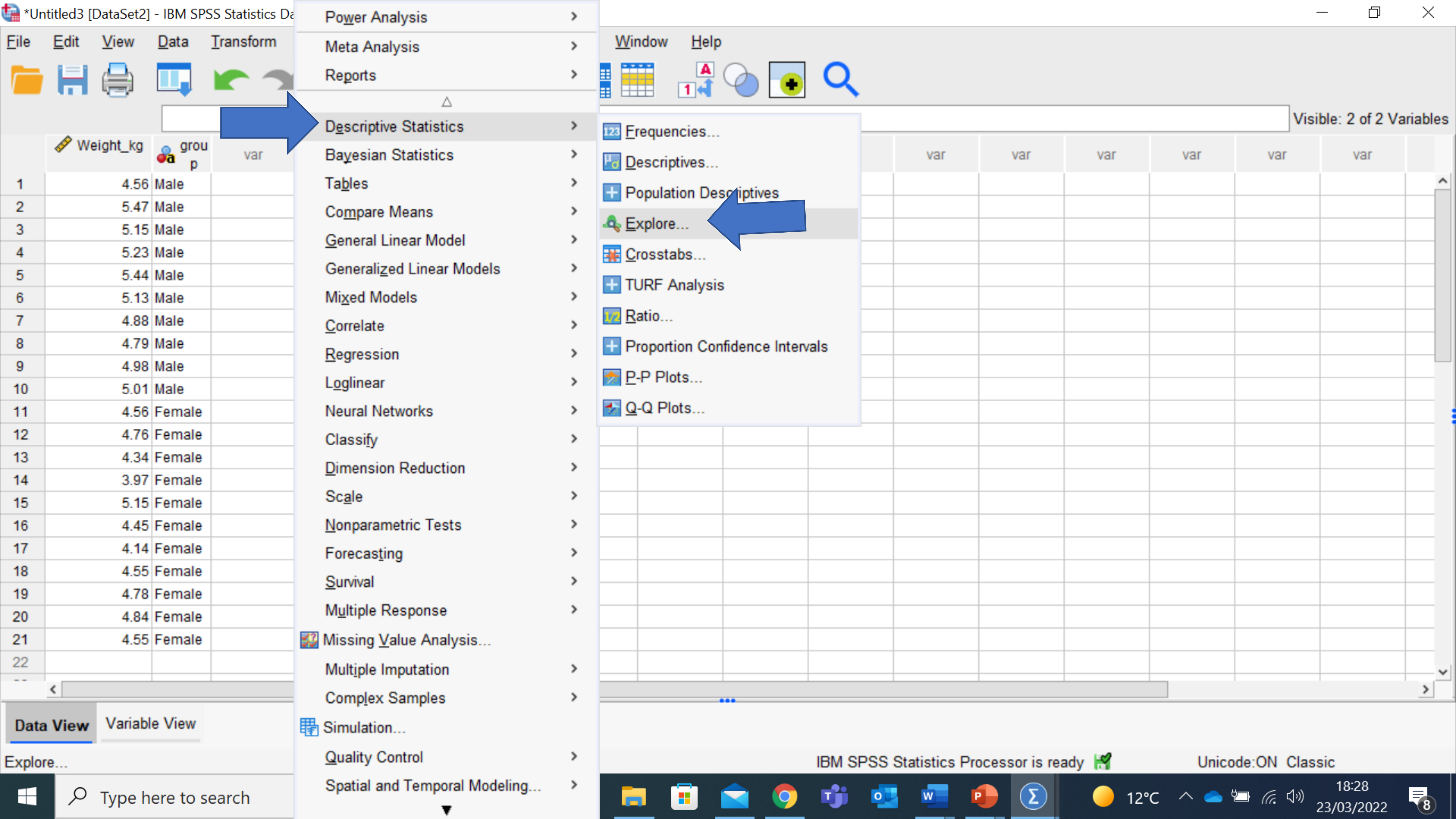
- Outliers are extreme values that stand out greatly from the overall pattern of values in a dataset or graph.

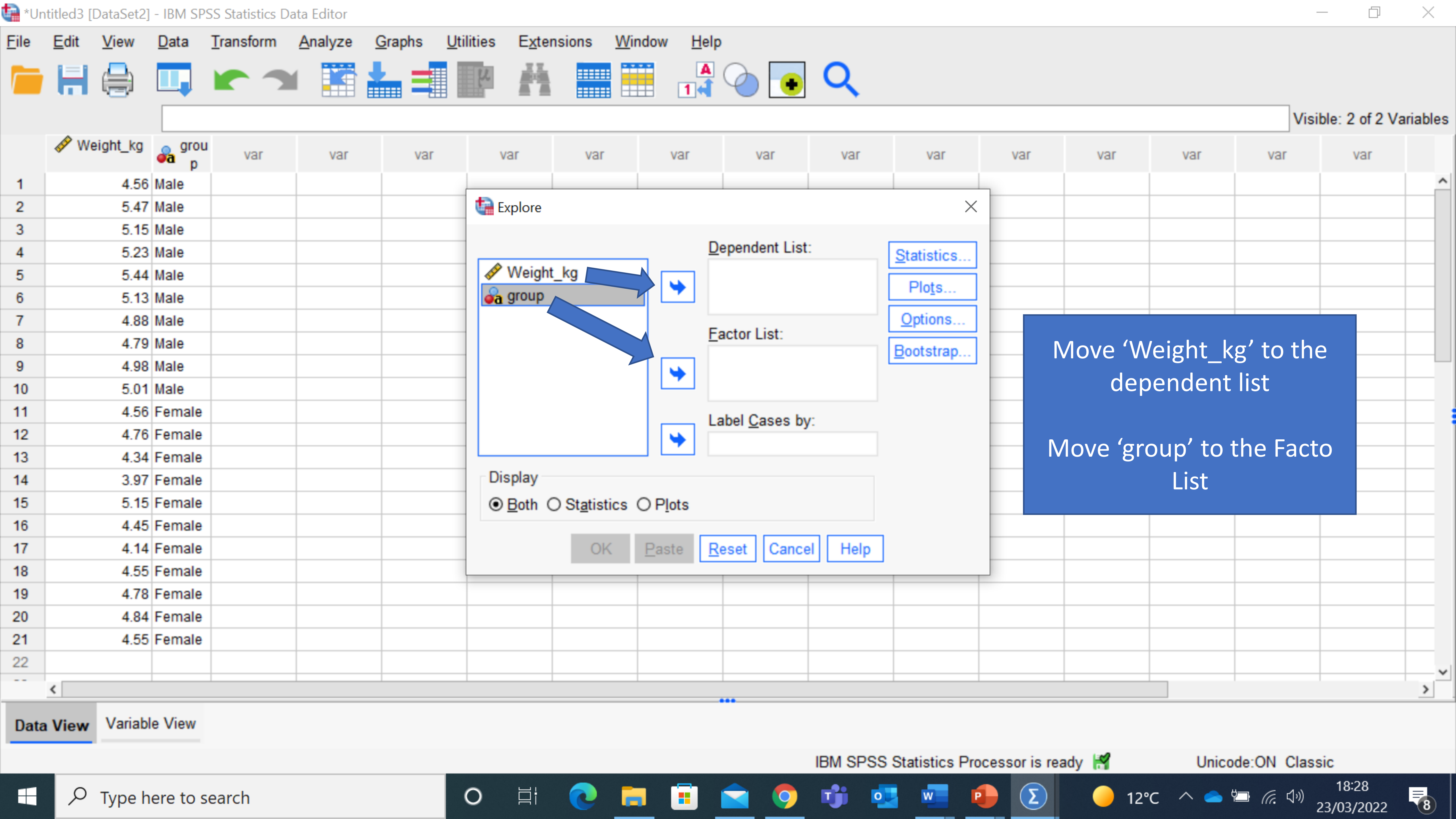


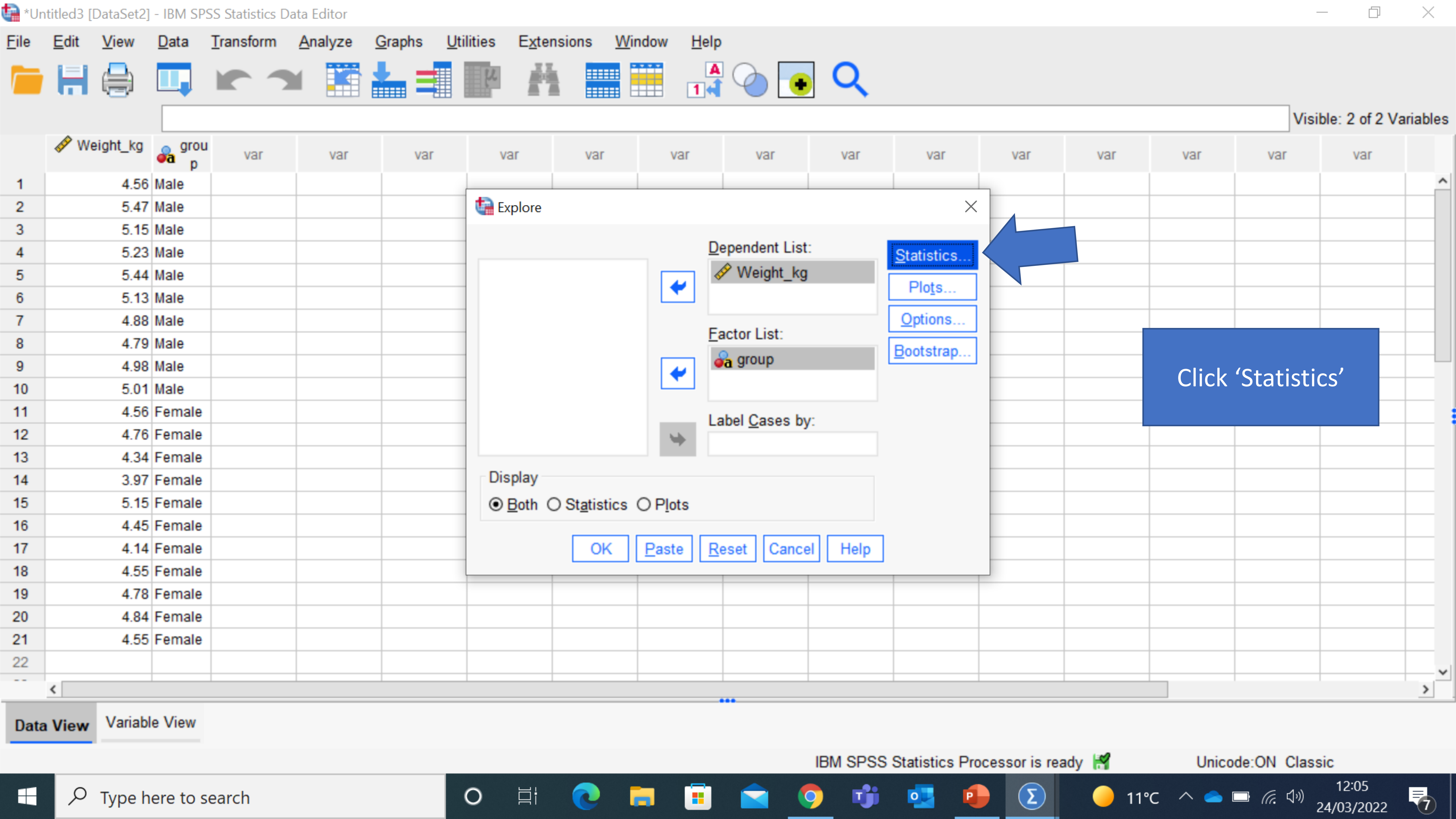
Assessing Normality + Outliers in SPSS = Easy!

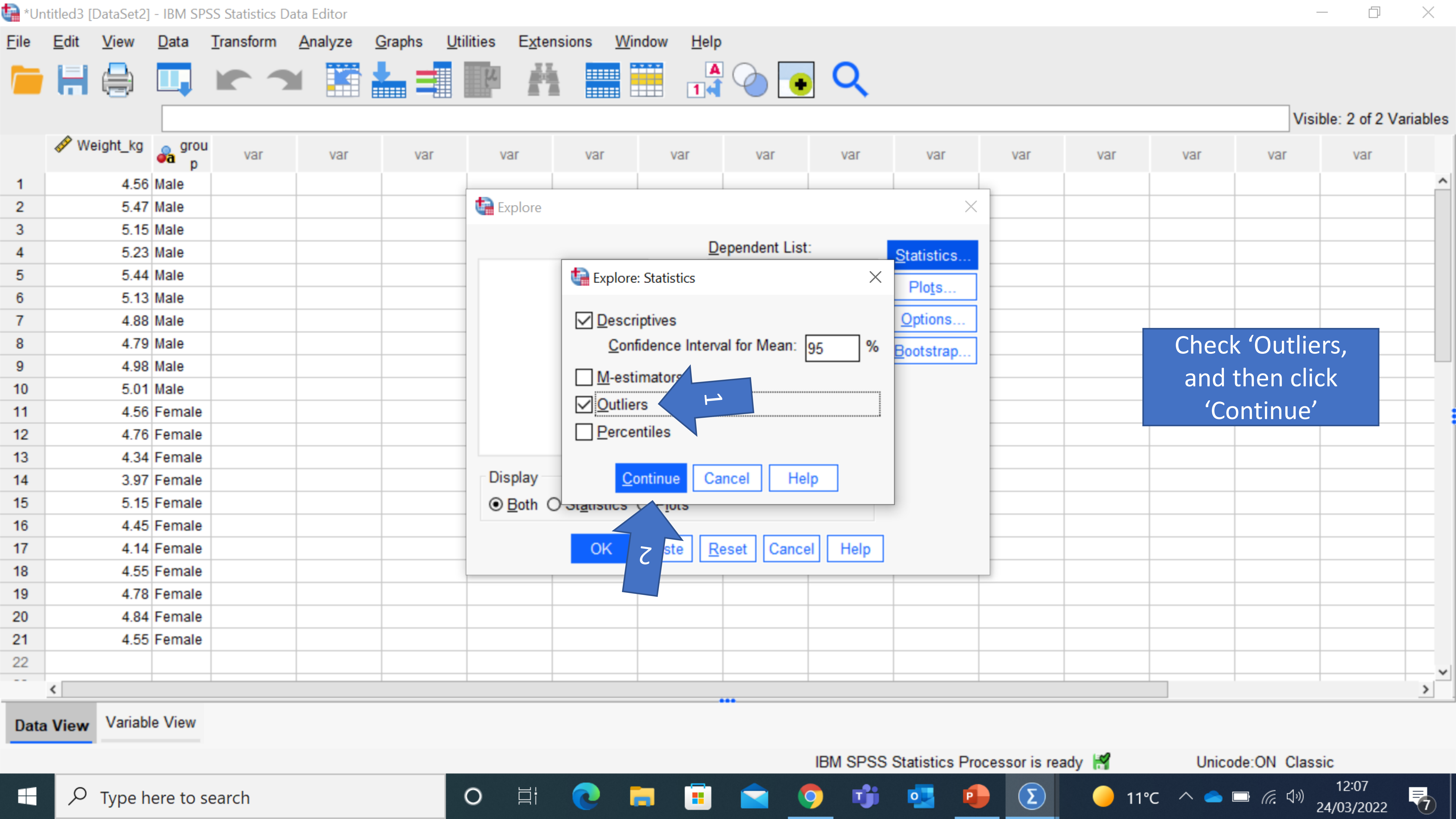
- Great news! SPSS can help us look at Both Normality AND Outliers in one output!

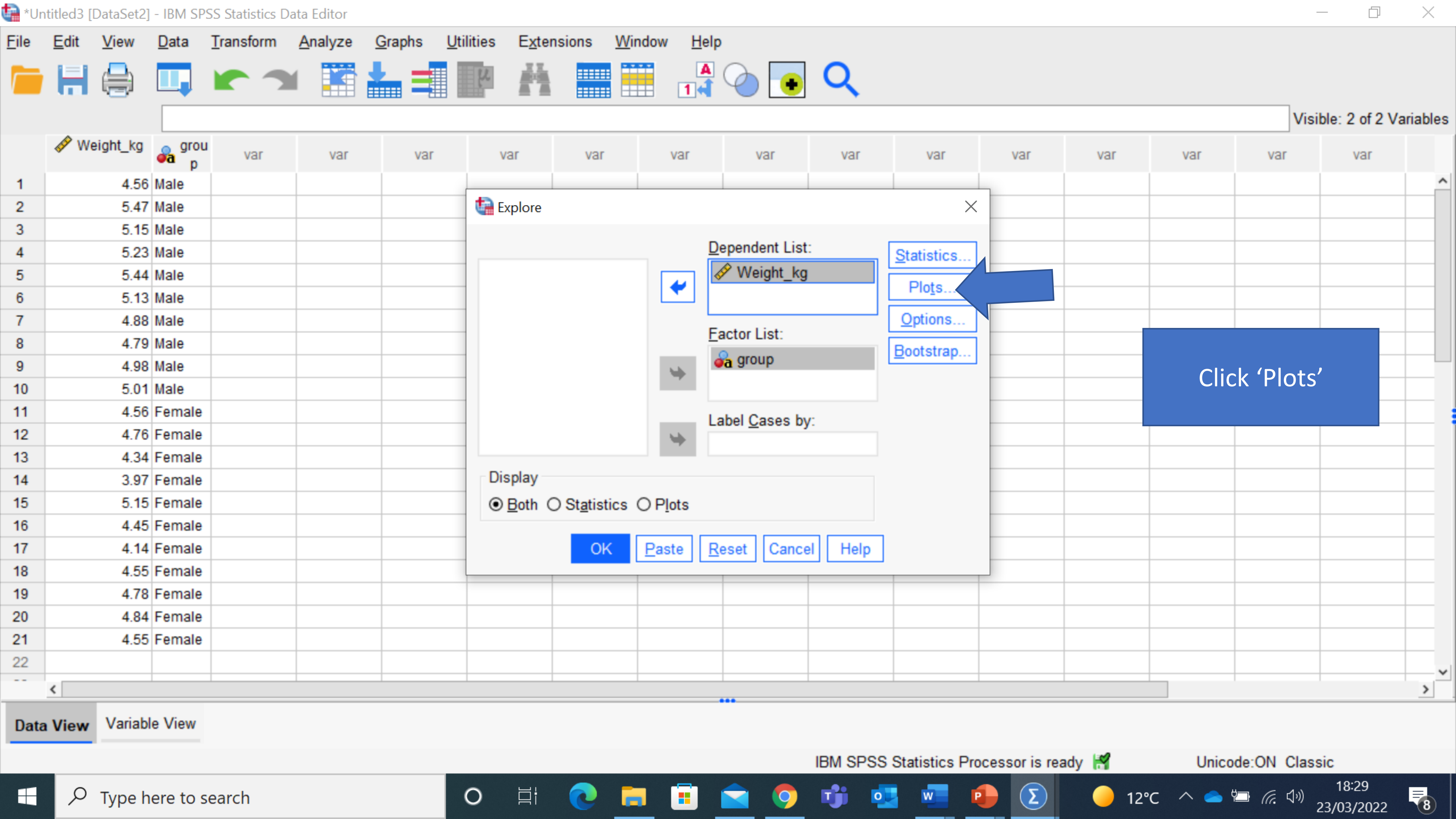

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
18:27
23/03/2022

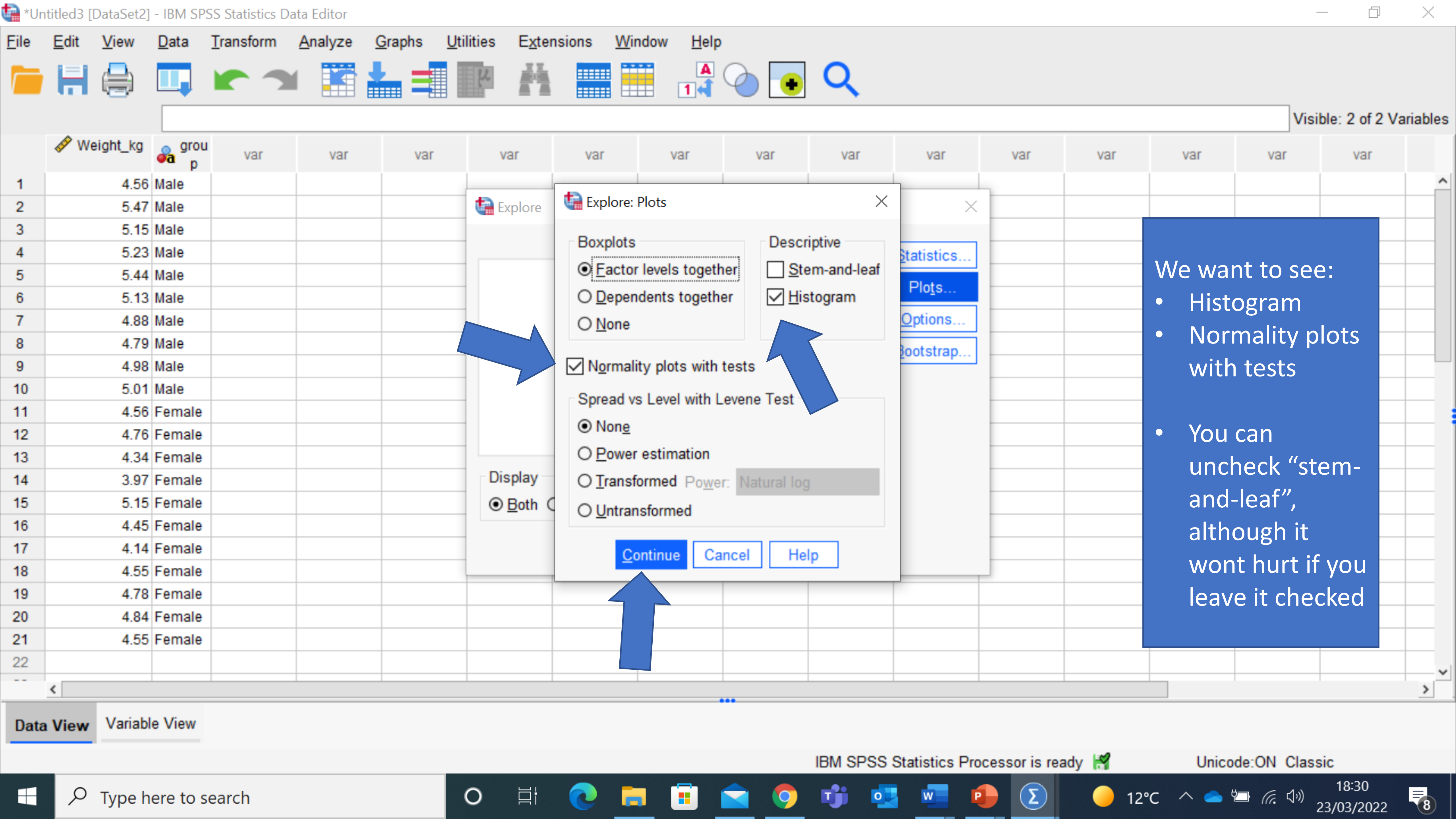

*Untitled3 [DataSet2] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Frequencies...
Descriptives...
Population Descriptives
Explore...
Crosstabs...
TURF Analysis
Ratio...
Proportion Confidence Intervals
P-P Plots...
Q-Q Plots...
Data View
Variable View
Explore...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
18:28
23/03/2022

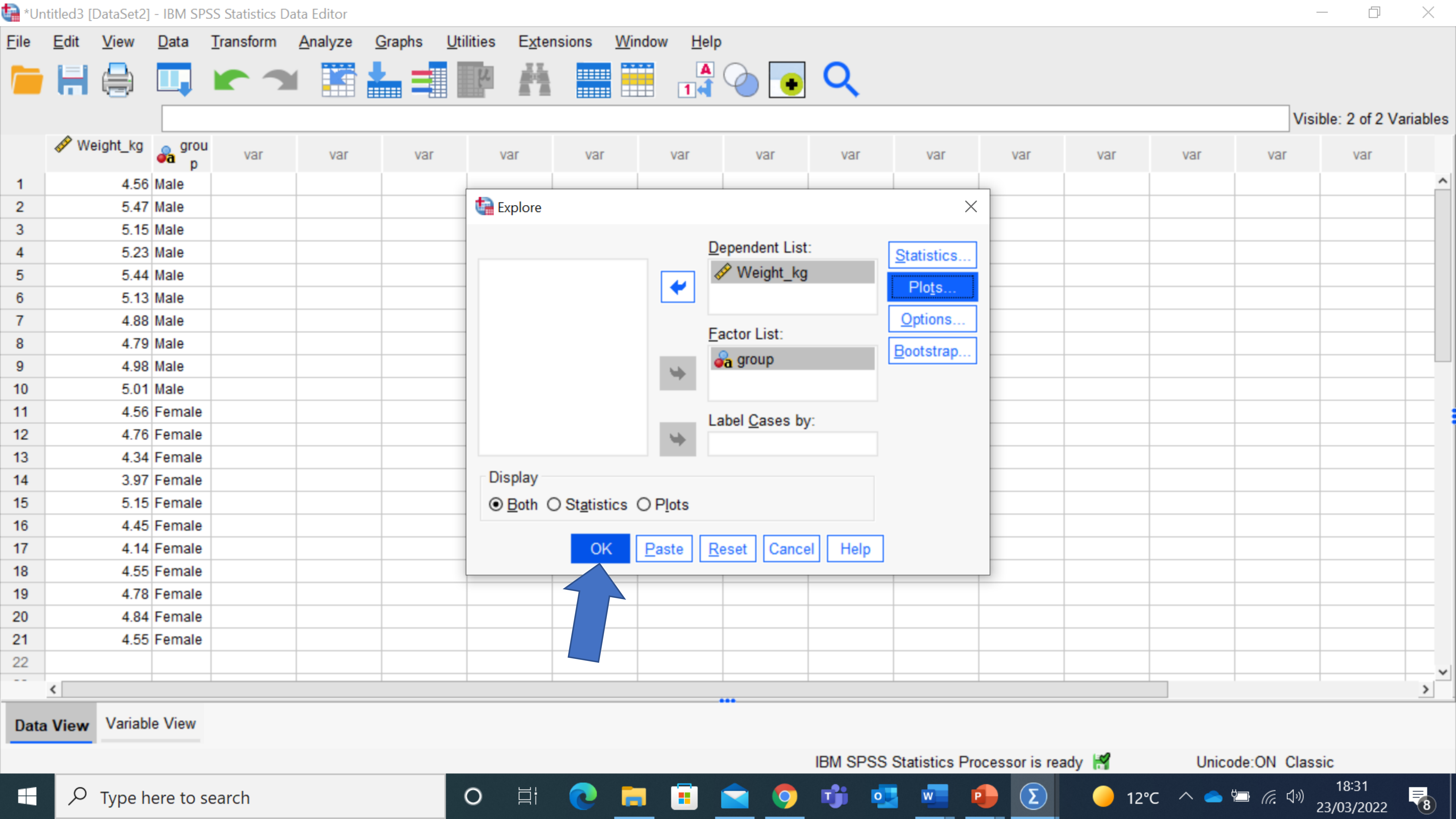

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
Explore
Dependent List:
Factor List:
Label Cases by:
Weight_kg
group
Statistics...
Plots...
Options...
Bootstrap...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help
Move 'Weight_kg' to the dependent list
Move 'group' to the Facto List
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
18:28
23/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Explore
Dependent List:
Weight_kg
Factor List:
group
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help
Click 'Statistics'
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
12:05
24/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
Explore
Dependent List:
Statistics...
Plots...
Options...
Bootstrap...
Explore: Statistics
Descriptives
Confidence Interval for Mean: 95 %
M-estimators
Outliers
Percentiles
Continue
Cancel
Help
Display
Both
OK
Reset
Cancel
Help
1
2
Check ‘Outliers, and then click ‘Continue’
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
12:07
24/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
Explore
Dependent List:
Weight_kg
Factor List:
group
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Display
Both Statistics Plots
OK Paste Reset Cancel Help
Click 'Plots'
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
Explore
Explore: Plots
Boxplots
Factor levels together
Dependents together
None
Descriptive
Stem-and-leaf
Histogram
Normality plots with tests
Spread vs Level with Levene Test
None
Power estimation
Transformed Power: Natural log
Untransformed
Continue
Cancel
Help
Statistics...
Plots...
Options...
Bootstrap...
Display
Both
We want to see:
• Histogram
• Normality plots with tests
• You can uncheck "stem-and-leaf", although it wont hurt if you leave it checked


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
var
Explore
Dependent List:
Weight_kg
Factor List:
group
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Display
Both Statistics Plots
OK Paste Reset Cancel Help
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
18:31
23/03/2022

	group			Statistic	Std. Error
Weight_kg	Female	Mean		4.5536	.09998
		95% Confidence Interval for Mean	Lower Bound	4.3309	
			Upper Bound	4.7764	
		5% Trimmed Mean		4.5529	
		Median		4.5500	
		Variance		.110	
		Std. Deviation		.33161	
		Minimum		3.97	
		Maximum		5.15	
		Range		1.18	
		Interquartile Range		.44	
		Skewness		-.082	.661
		Kurtosis		.157	1.279
	Male	Mean		5.0640	.[illegible]35
		95% Confidence Interval for Mean	Lower Bound	4.8619	
			Upper Bound	5.2661	
		5% Trimmed Mean		5.0694	
		Median		5.0700	
		Variance		.080	
		Std. Deviation		.28254	
		Minimum		4.56	
		Maximum		5.47	
		Range		.91	
		Interquartile Range		.43	
		Skewness		-.185	.687
		Kurtosis		-.251	1.334

For this test, we need to check for normality and outliers in BOTH our groups:

ie., in Females AND in males **separately**

I see here that for both sexes, skewness and kurtosis are between 2 and -2 for both males and females!

Great!

	Variance	.080	
	Std. Deviation	.28254	
	Minimum	4.56	
	Maximum	5.47	
	Range	.91	
	Interquartile Range	.43	
	Skewness	-.185	.687
	Kurtosis	-.251	1.334

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	group	Statistic	df	Sig.	Statistic	df	Sig.
Weight_kg	Female	.132	11	.200*	.979	11	.959
	Male	.108	10	.200*	.974	10	.927

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Weight_kg

Histograms

Histogram

for group= Female

The **Shapiro-Wilk test** is a test for normality

When this test is significant (ie., the p value is BELOW 0.05), it means our distribution is statistically DIFFERENT than normal

We want a bigger value here, and we have it for both males AND females!!! Good news!

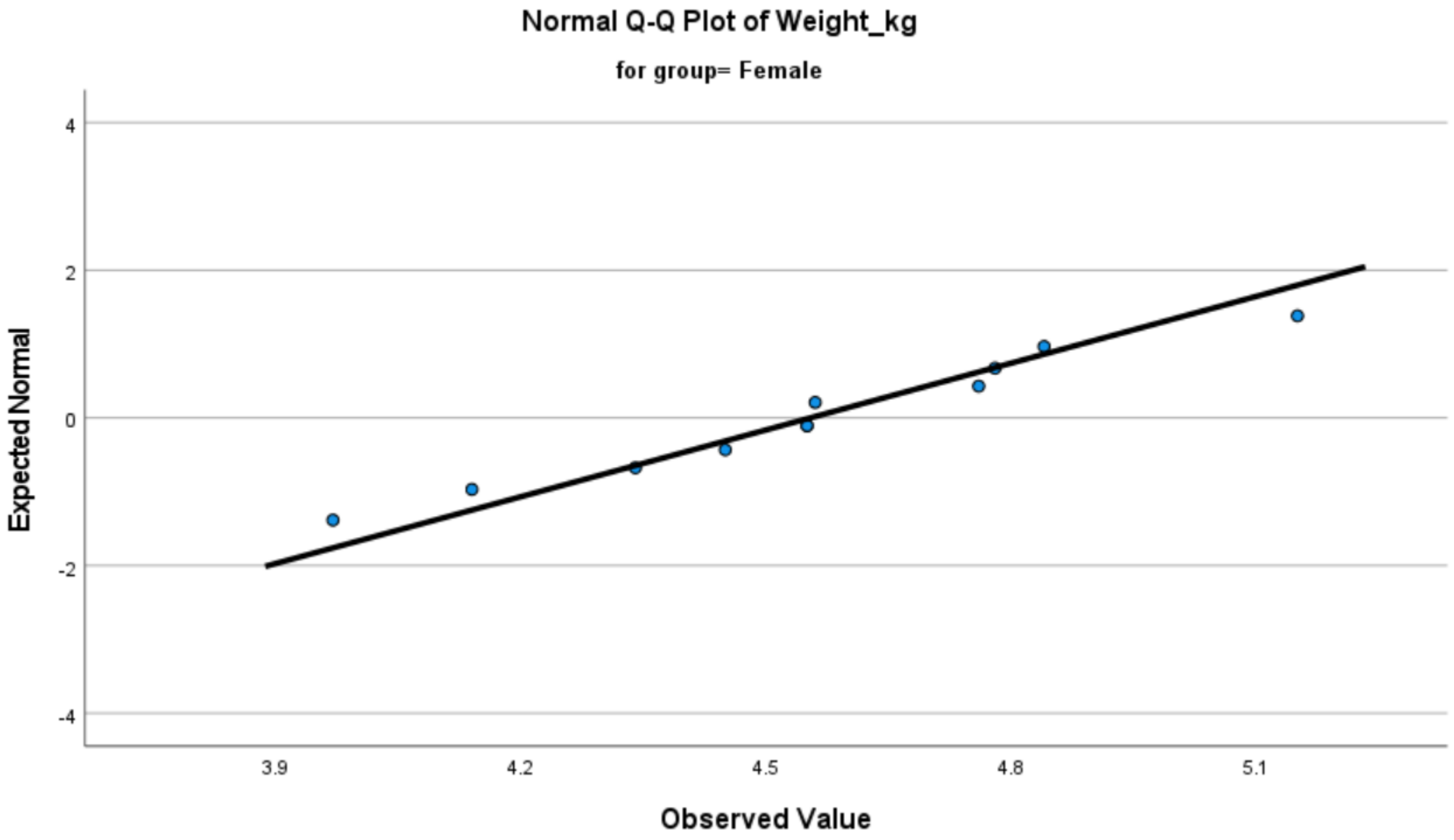
Small sample size: Makes it easier to get a non-significant result. Our sample is small, so we want to look at more than this

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - Explore
    - Title
    - Notes
    - Active Dataset
    - group
      - Title
      - Case Process
      - Descriptives
      - Tests of Norm
      - Weight_kg
        - Title
        - Histograr
          - Title
          - grou|
          - grou|
        - Normal Q
          - Title
          - grou|
          - grou|
        - Detrende
          - Title
          - grou|
          - grou|
        - Boxplot

Normal Q-Q Plots



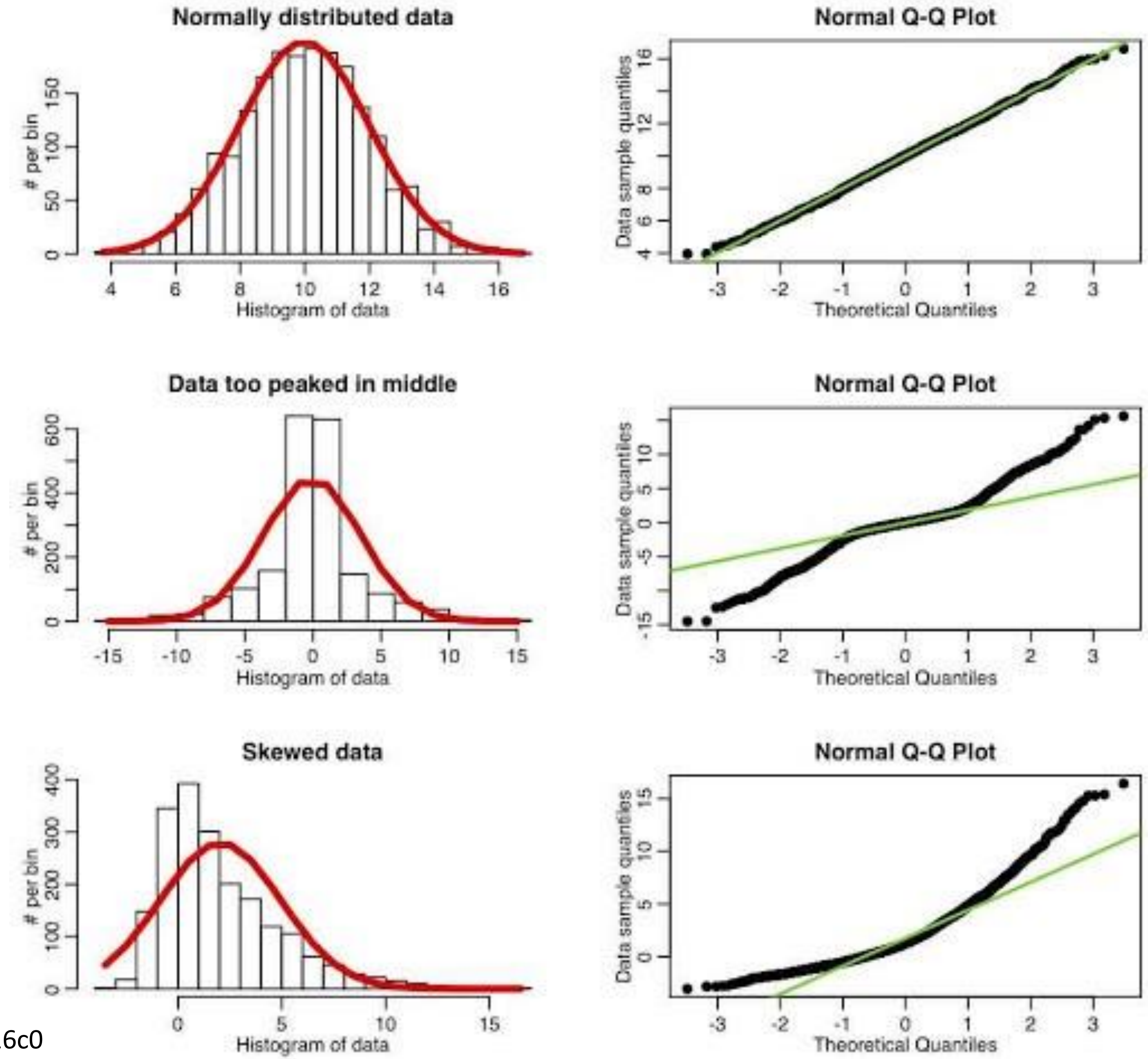
Double-click to activate

IBM S Processor is ready Unicode:ON Classic

Type here to search 12°C 19:05 23/03/2022

qqplots

- If data are normally distributed, you should see the points generally follow the straight line on the plot
- Having one or two points being off the line is **fine**
- **A trend**, or **pattern** in the points away from the line indicates a lack of normality



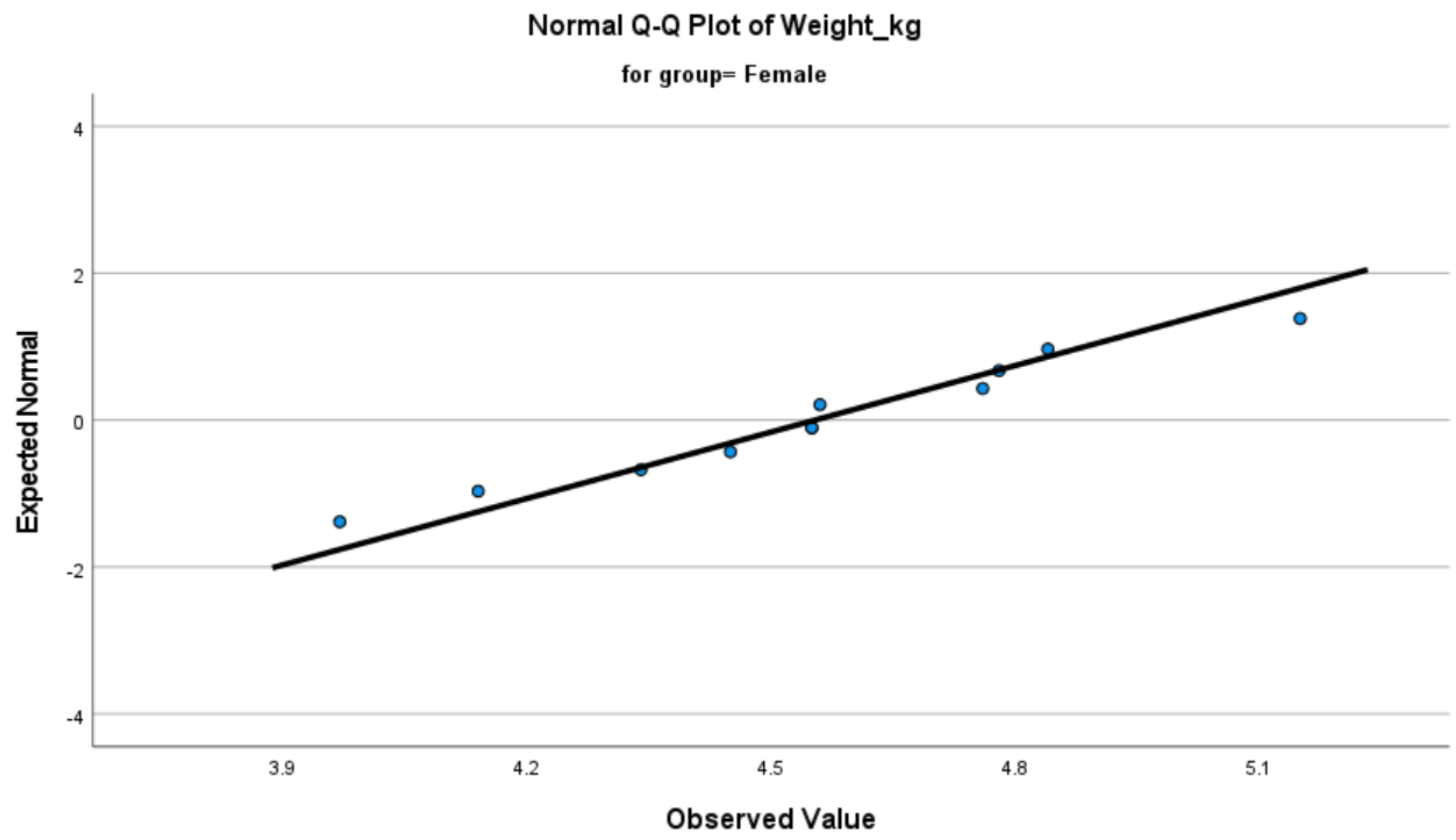
https://towardsdatascience.com/q-q-plots-explained-5aa8495426c0

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - Explore
    - Title
    - Notes
    - Active Dataset
    - group
      - Title
      - Case Process
      - Descriptives
      - Tests of Norm
      - Weight_kg
        - Title
        - Histograr
          - Title
          - grou
          - grou
        - Normal Q
          - Title
          - grou
          - grou
        - Detrende
          - Title
          - grou
          - grou
        - Boxplot

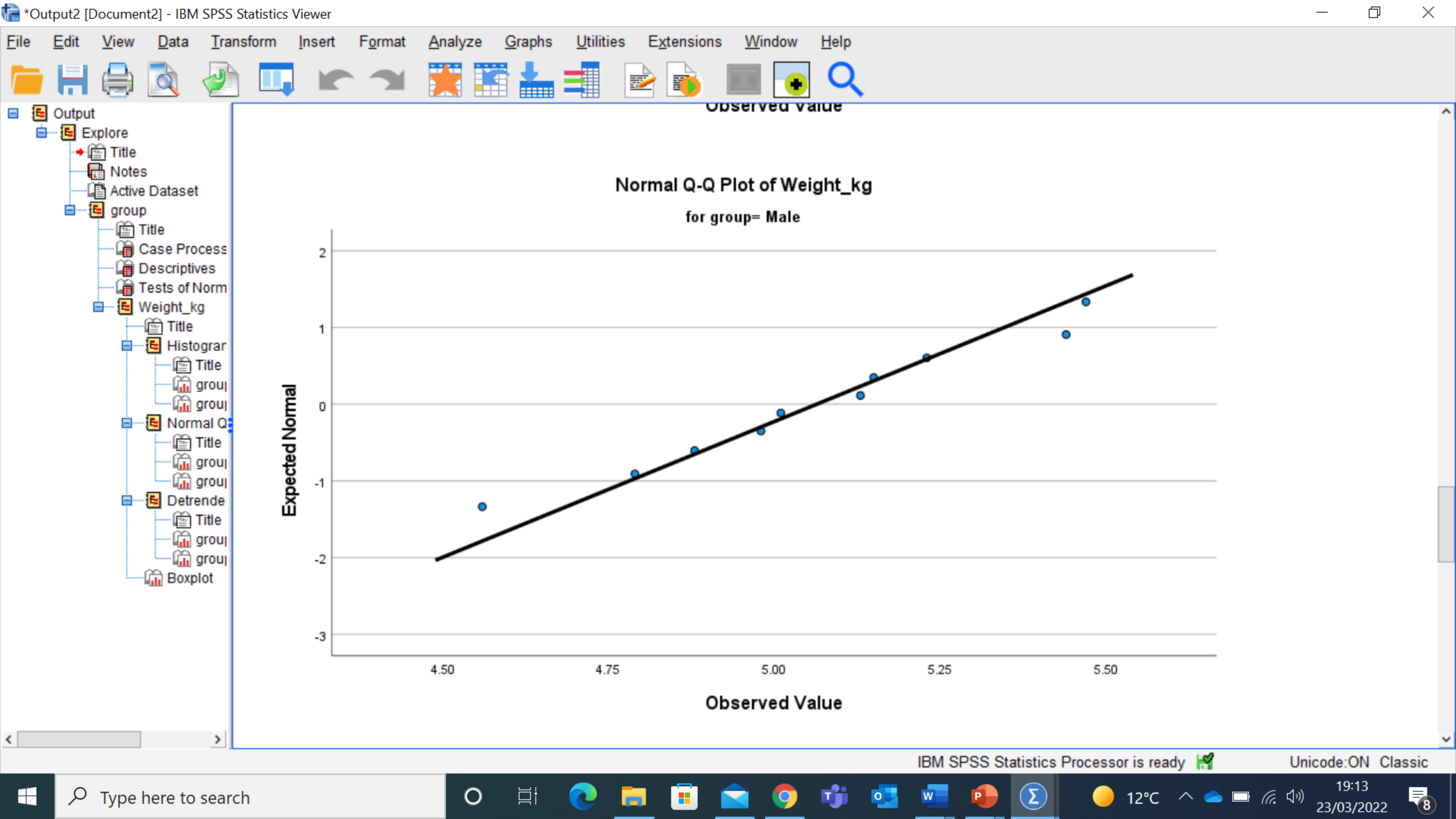
Normal Q-Q Plots

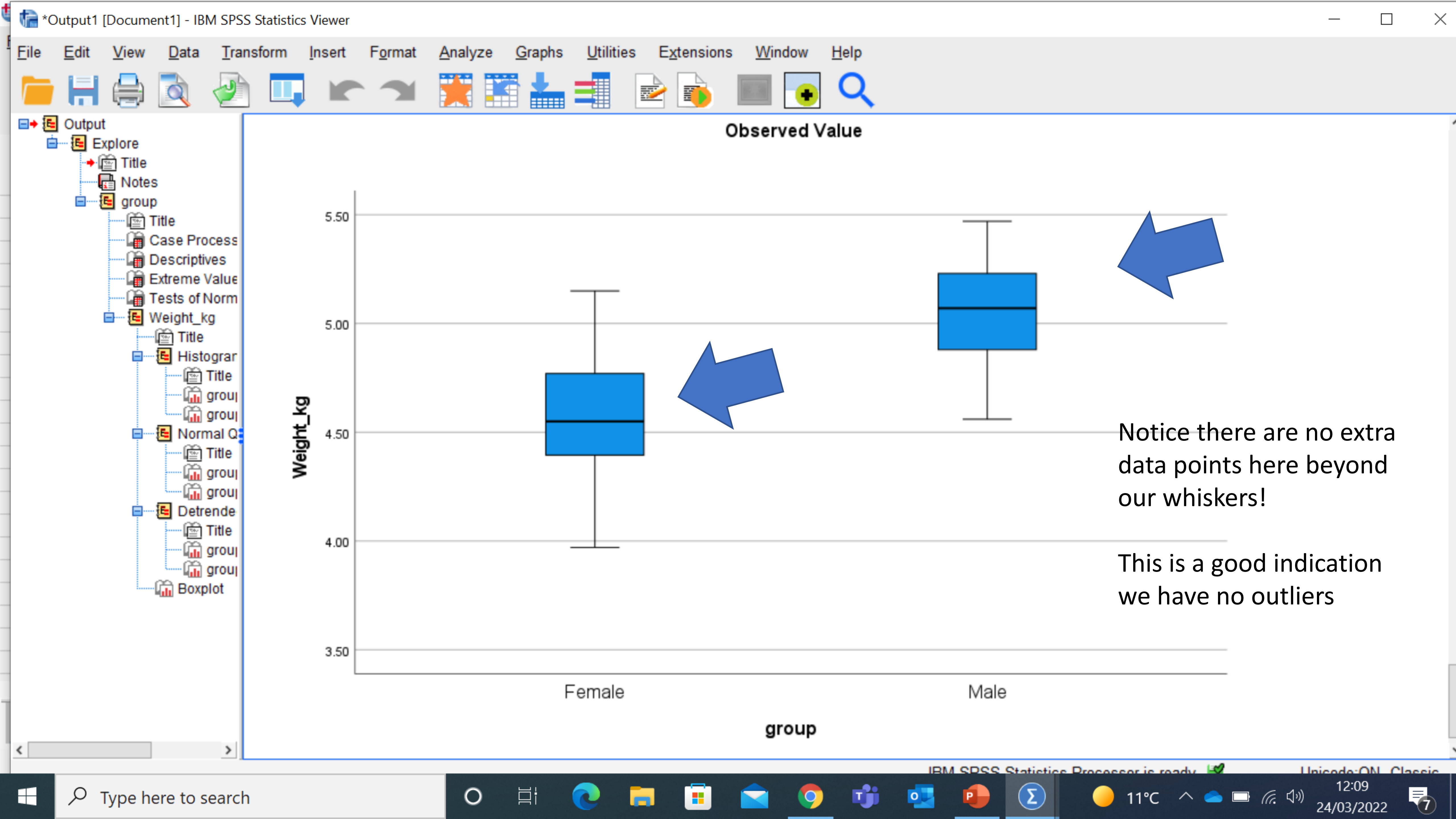


Double-click to activate

IBM S... Processor is ready Unicode:ON Classic

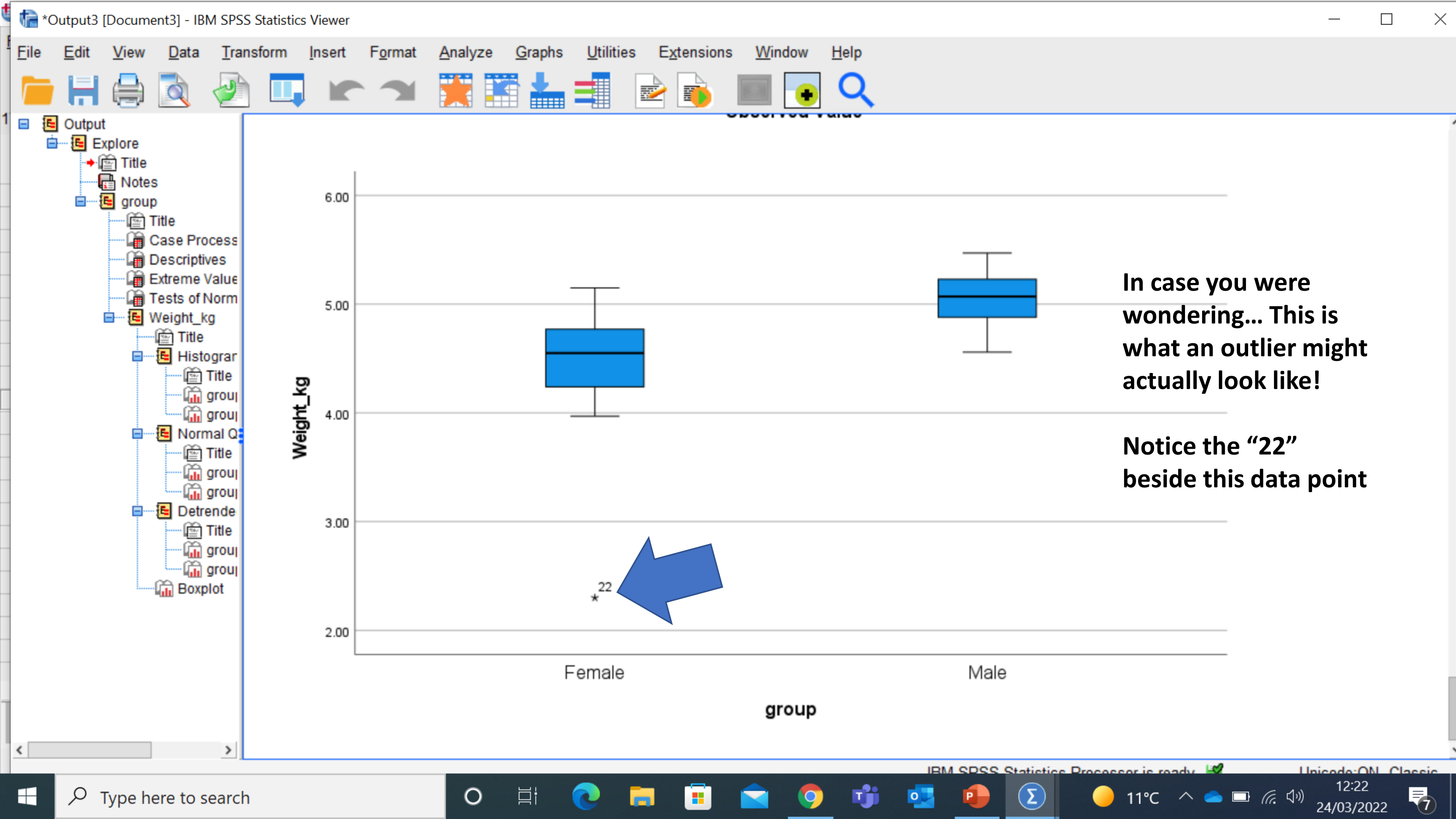
Type here to search 12°C 19:05 23/03/2022


*Output2 [Document2] - IBM SPSS Statistics Viewer
File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help
Output
Explore
Title
Notes
Active Dataset
group
Title
Case Process
Descriptives
Tests of Norm
Weight_kg
Title
Histograr
Title
grou|
grou|
Normal Q
Title
grou|
grou|
Detrende
Title
grou|
grou|
Boxplot
Normal Q-Q Plot of Weight_kg
for group= Male
Expected Normal
2
1
0
-1
-2
-3
4.50
4.75
5.00
5.25
5.50
Observed Value
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
19:13
23/03/2022



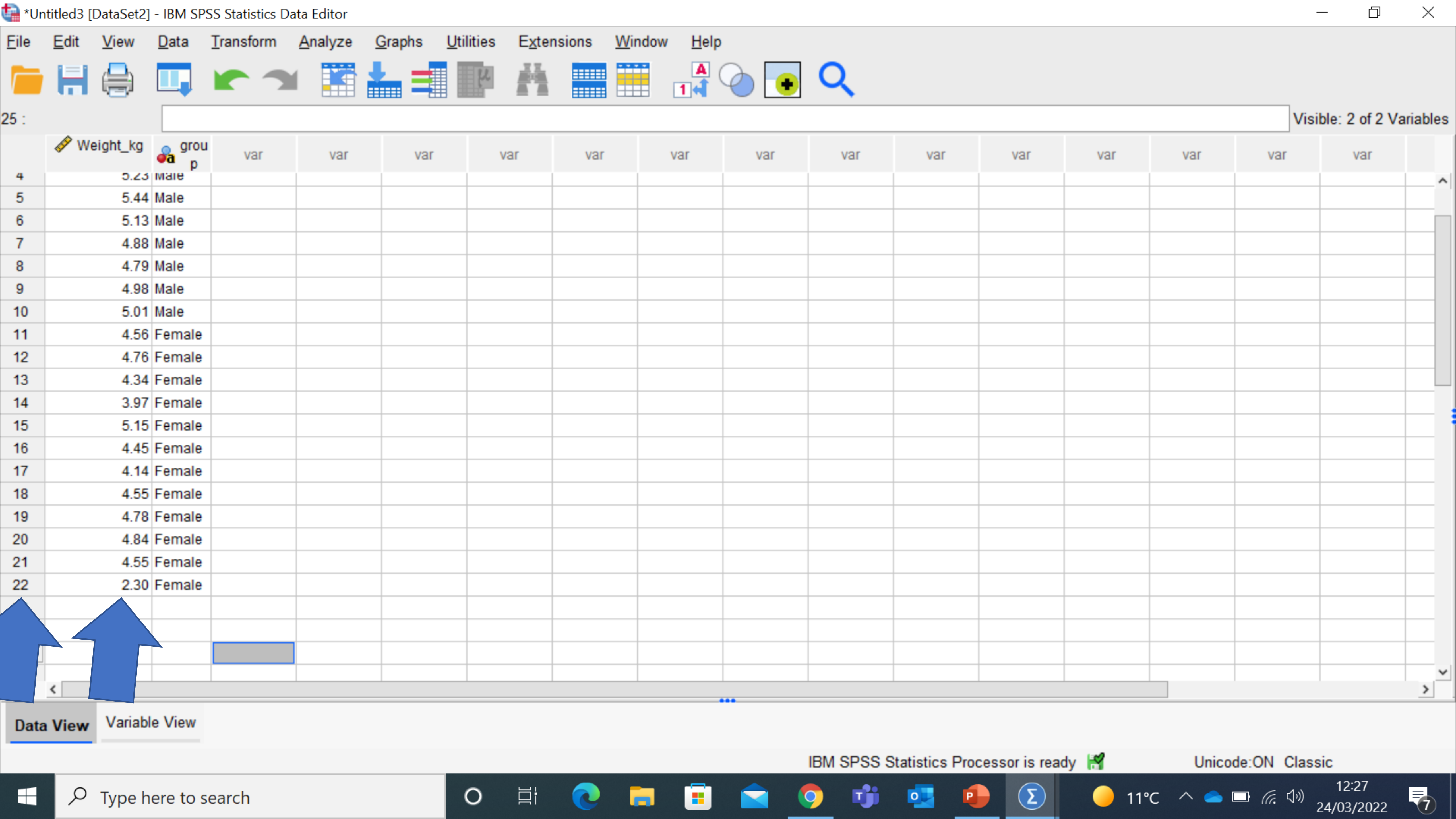
Notice there are no extra data points here beyond our whiskers!

This is a good indication we have no outliers



In case you were wondering… This is what an outlier might actually look like!

Notice the “22” beside this data point


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
25 :
Visible: 2 of 2 Variables
Weight_kg
group
var
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22 2.30 Female
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
11°C
12:27
24/03/2022

*Output3 [Document3] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Extreme Values

	group			Case Number	Value
Weight_kg	Female	Highest	1	15	5.15
			2	20	4.84
			3	19	4.78
			4	12	4.76
			5	11	4.56
		Lowest	1	22	2.30
			2	14	3.97
			3	17	4.14
			4	13	4.34
			5	16	4.45
	Male	Highest	1	2	5.47
			2	5	5.44
			3	4	5.23
			4	3	5.15
			5	6	5.13
		Lowest	1	1	4.56
			2	8	4.79
			3	7	4.88
			4	9	4.98
			5	10	5.01

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	group	Statistic	df	Sig.	Statistic	df	Sig.

Homogeneity of variance

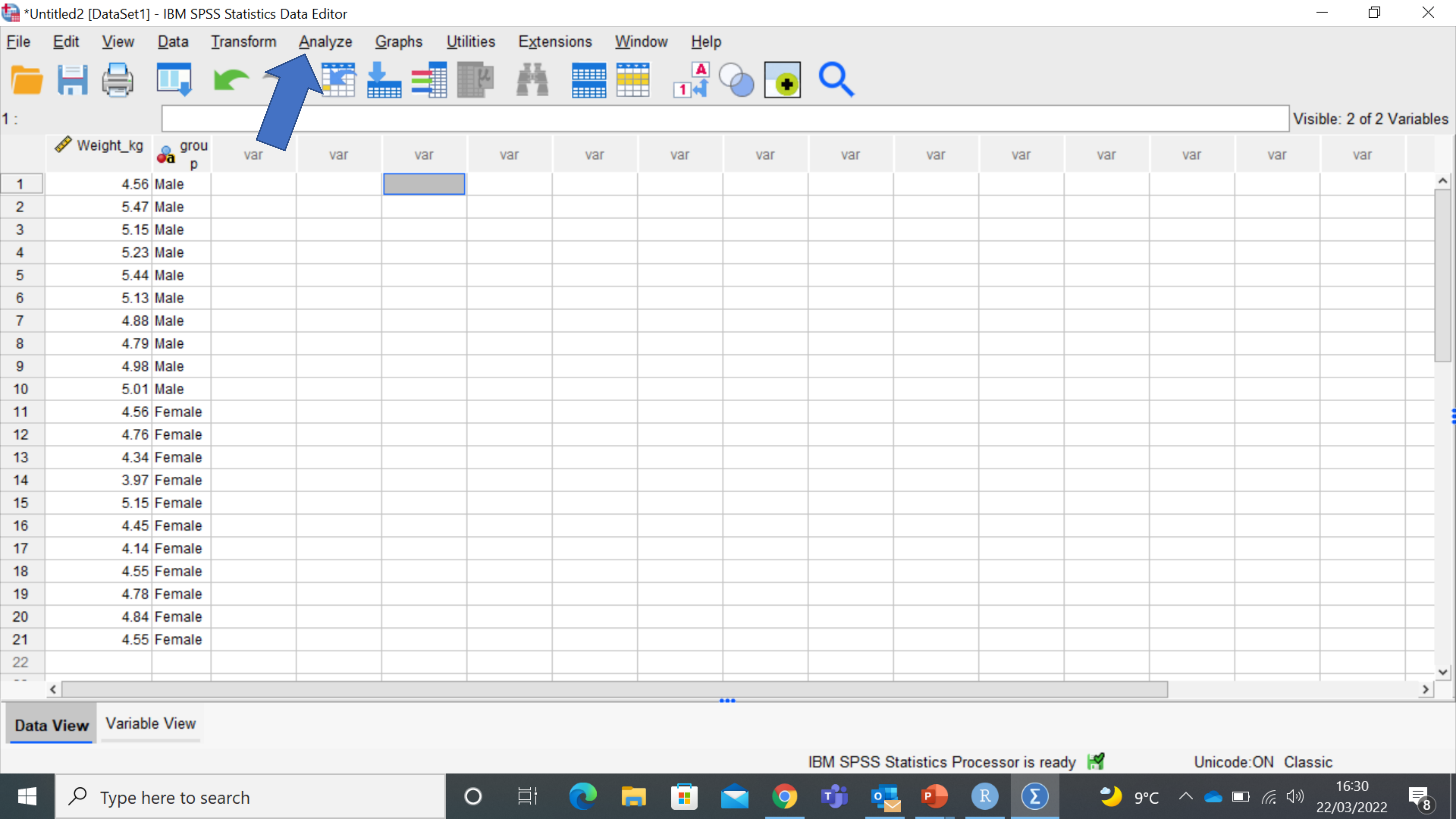
- Is the distribution of the data around the mean approximately equal in each group?

- Good news, SPSS does a really nice thing for us here, and we don’t have to test this assumption before we run the test…
- It will check for us during the test, and it gives us a solution if this assumption isn’t met! ☺

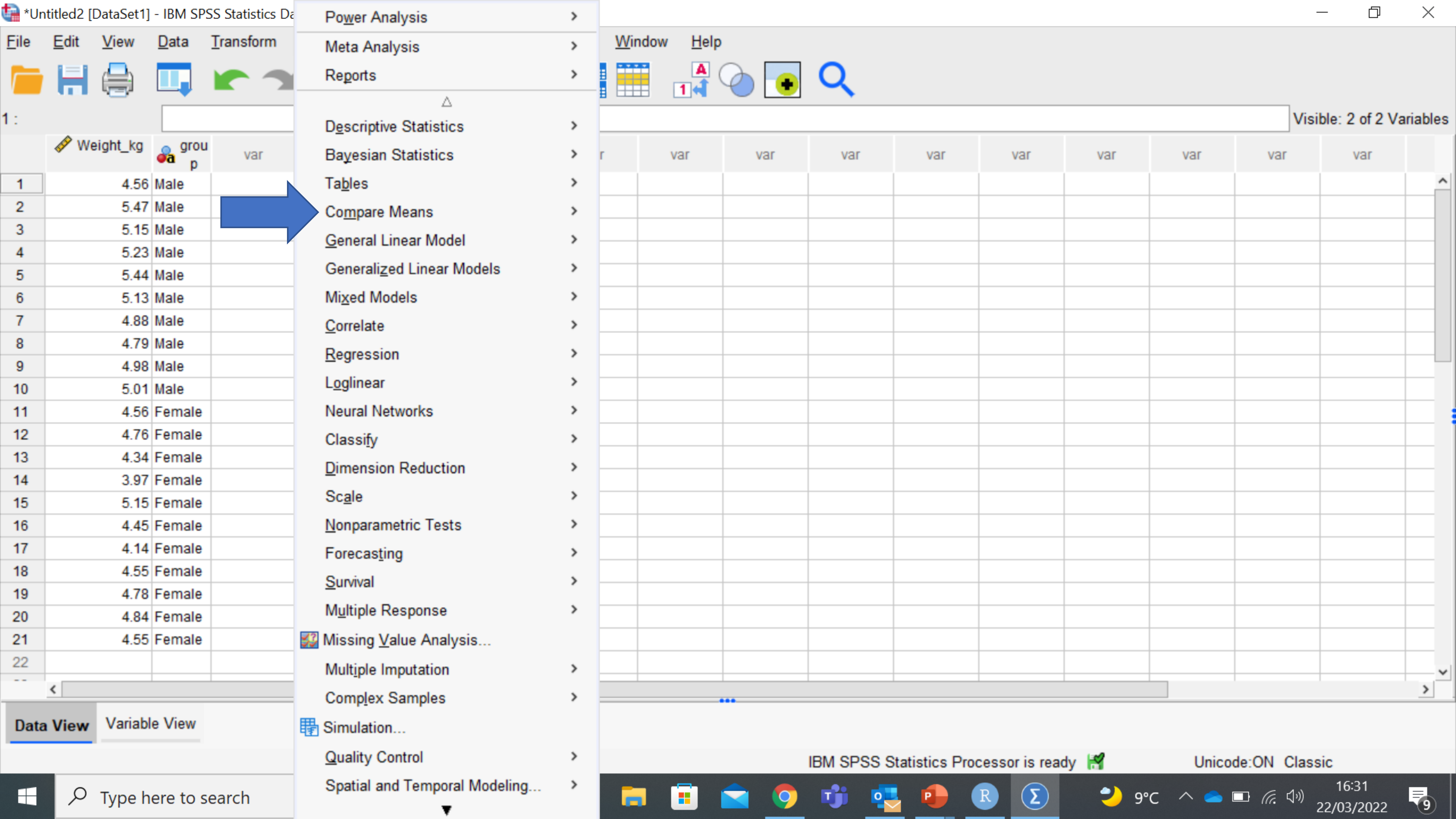
Let’s do this thing!

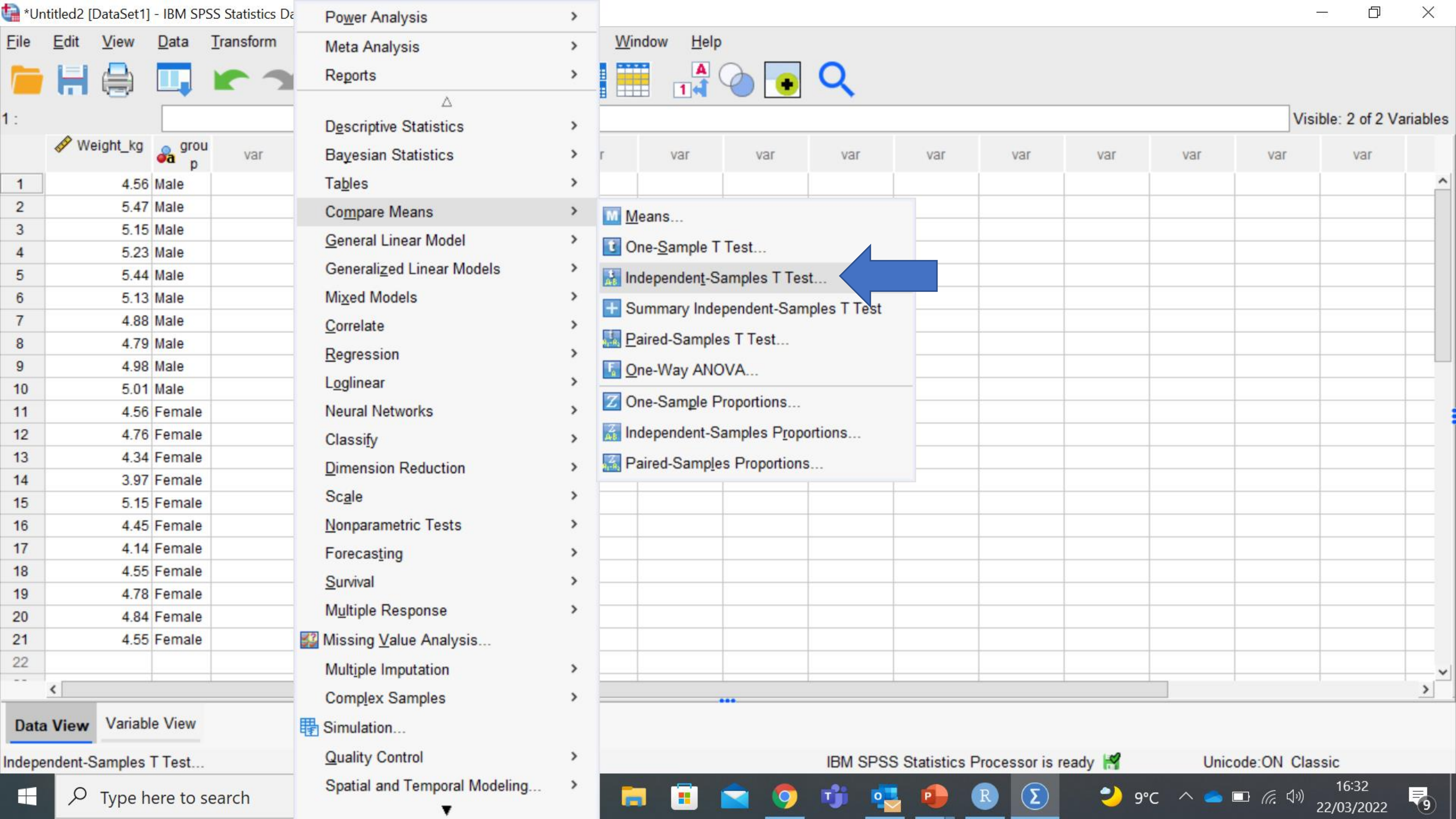
- Okay!
- Our data were collected independently!
- Our data appear normal!
- We don’t have any outliers!
- Colin said we don’t have to worry about the equality of variances yet…
- Can we please run a T Test!?!?!?

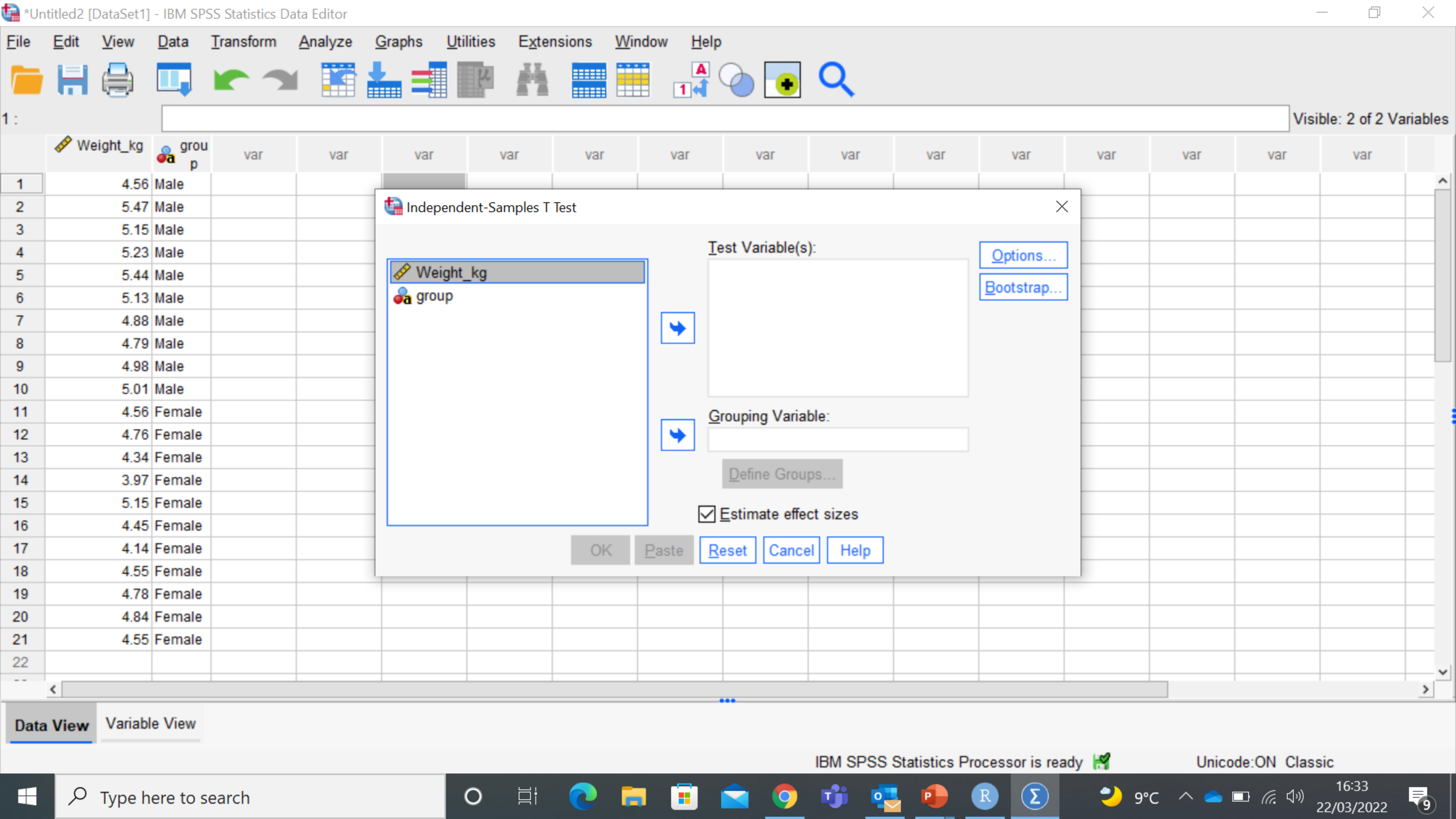
Yes!

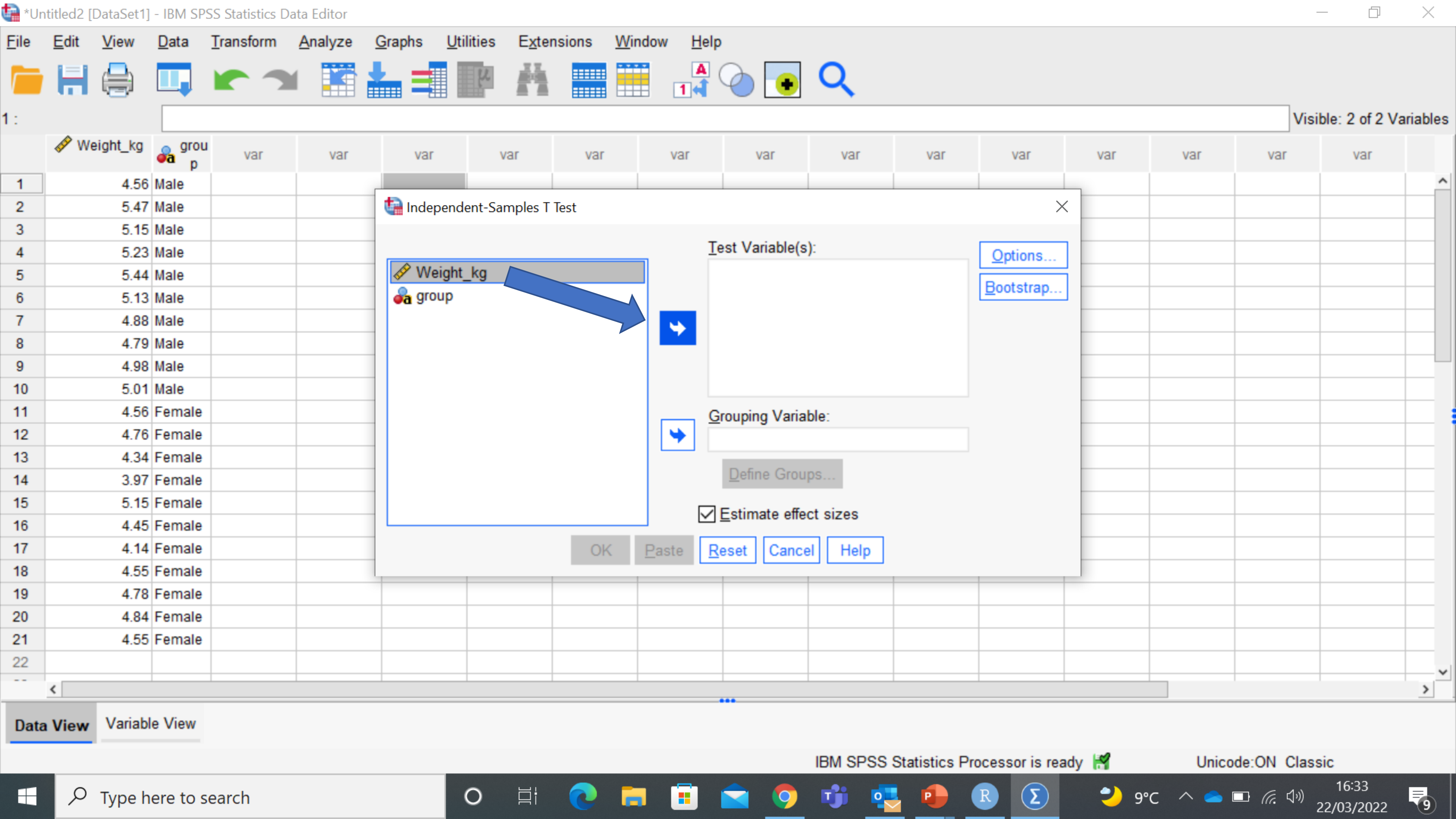


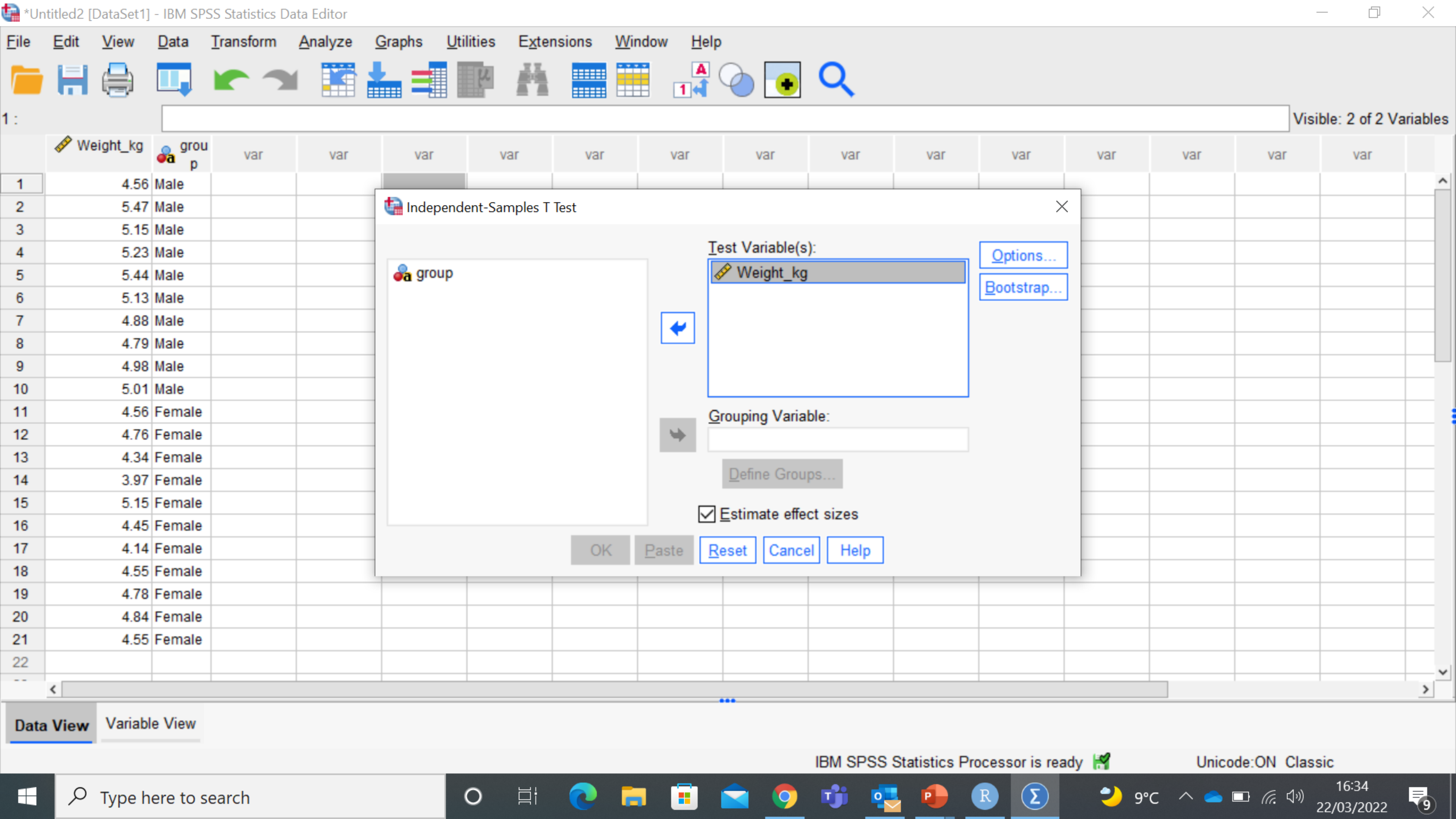
	Weight_kg	group
1	4.56	Male
2	5.47	Male
3	5.15	Male
4	5.23	Male
5	5.44	Male
6	5.13	Male
7	4.88	Male
8	4.79	Male
9	4.98	Male
10	5.01	Male
11	4.56	Female
12	4.76	Female
13	4.34	Female
14	3.97	Female
15	5.15	Female
16	4.45	Female
17	4.14	Female
18	4.55	Female
19	4.78	Female
20	4.84	Female
21	4.55	Female

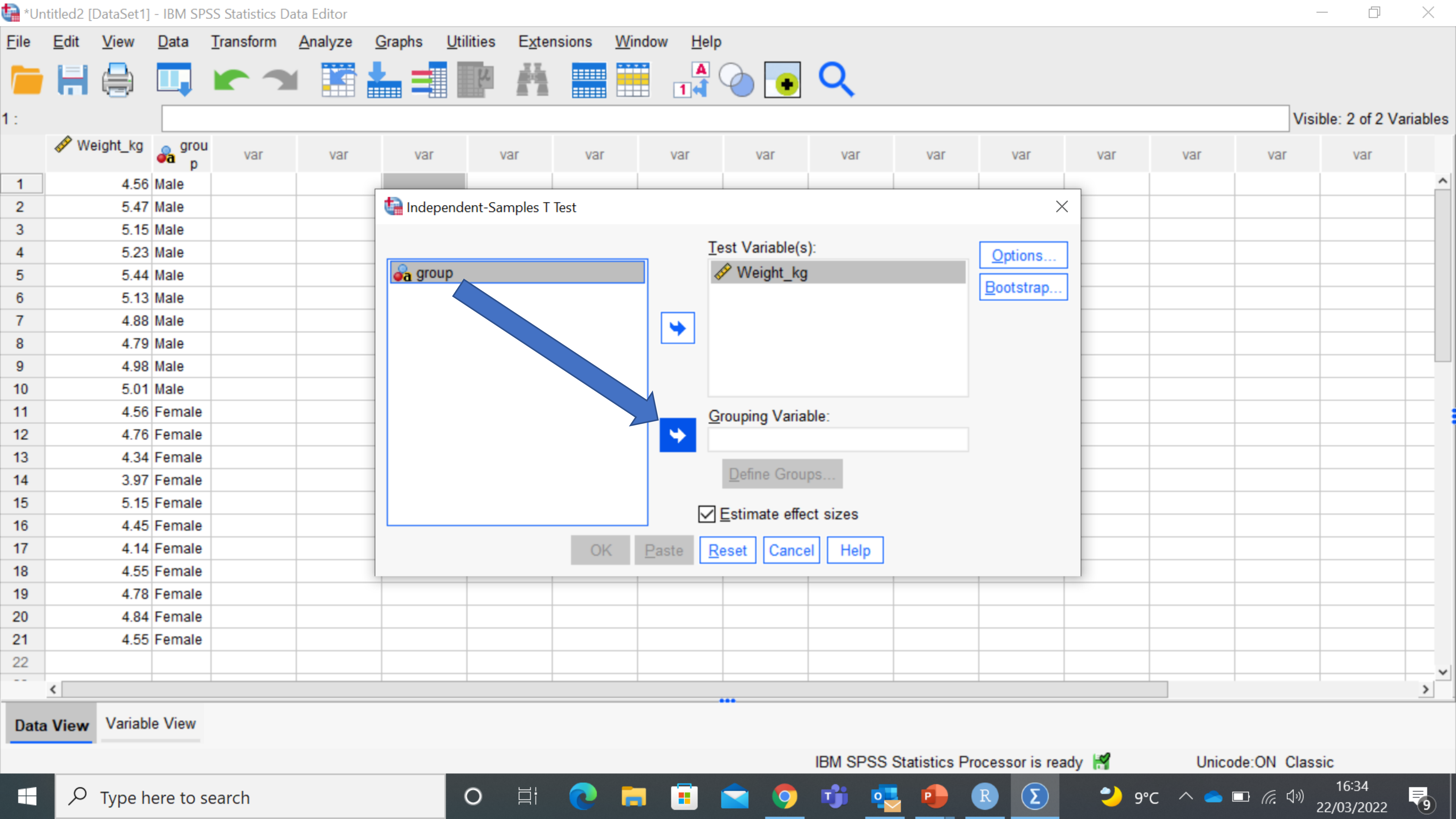

*Untitled2 [DataSet1] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
9°C
16:31
22/03/2022

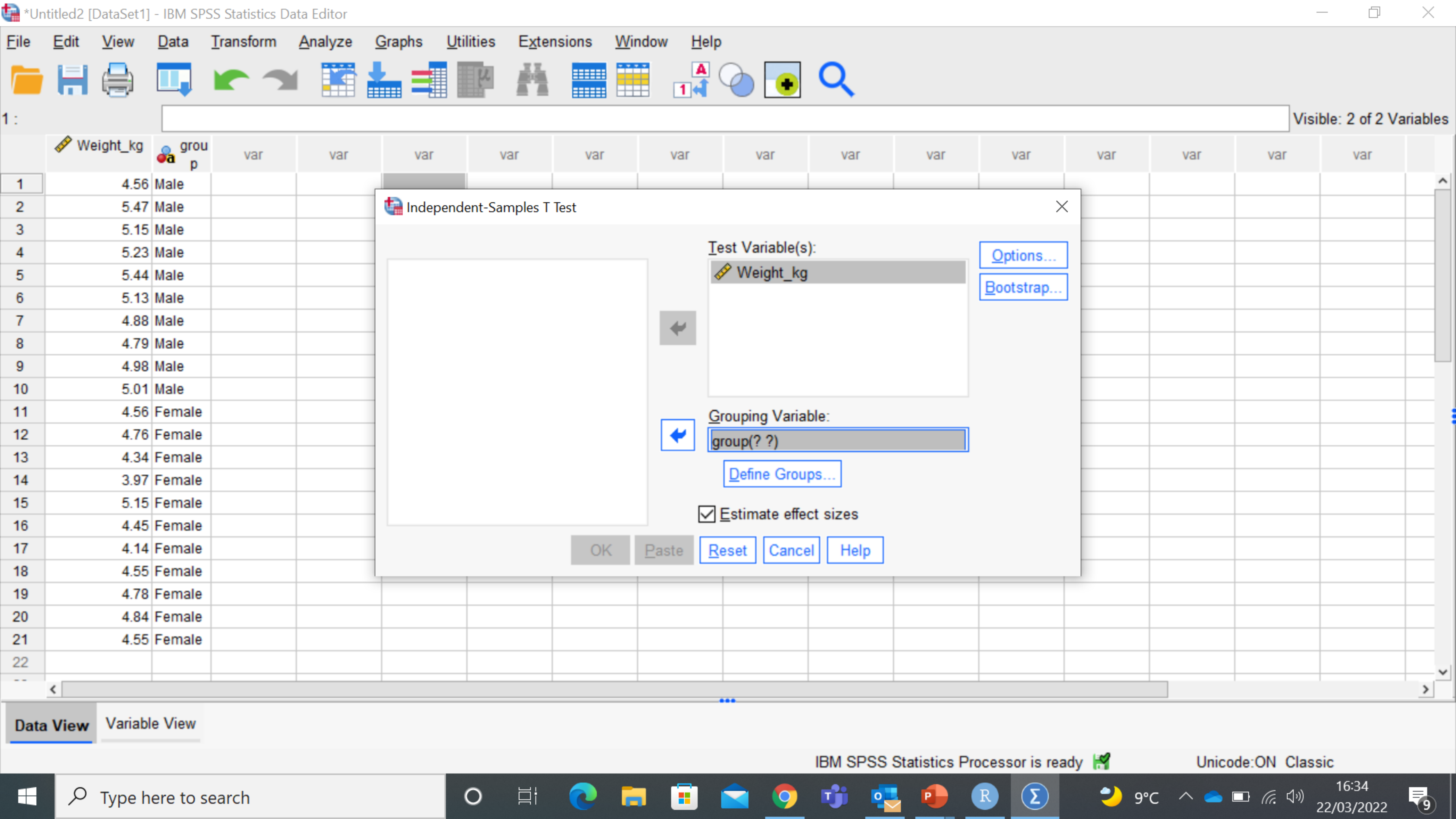

*Untitled2 [DataSet1] - IBM SPSS Statistics Da
File Edit View Data Transform
Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
Data View
Variable View
Independent-Samples T Test...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:32
22/03/2022

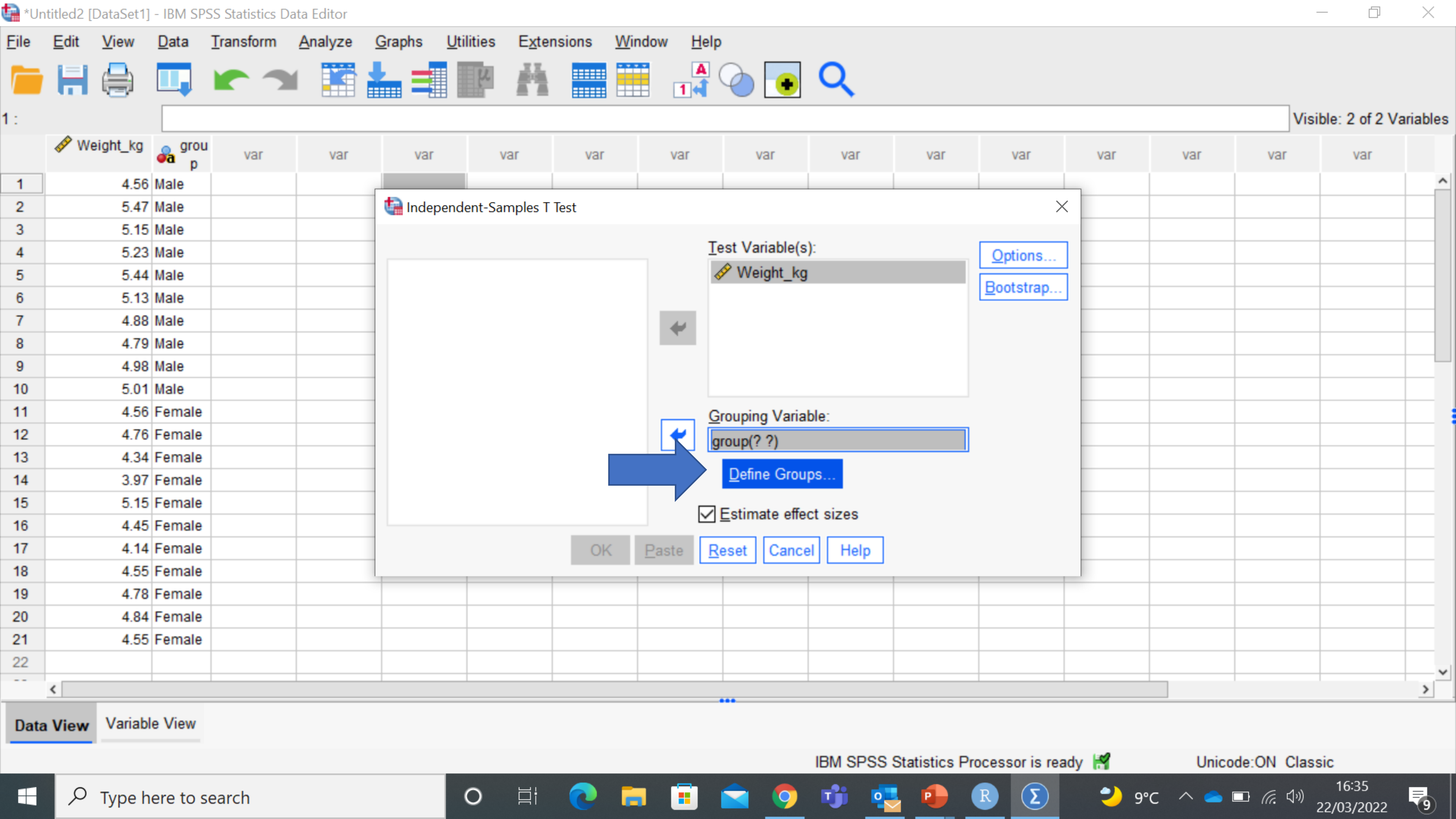

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
Weight_kg
group
Test Variable(s):
Options...
Bootstrap...
Grouping Variable:
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:33
22/03/2022

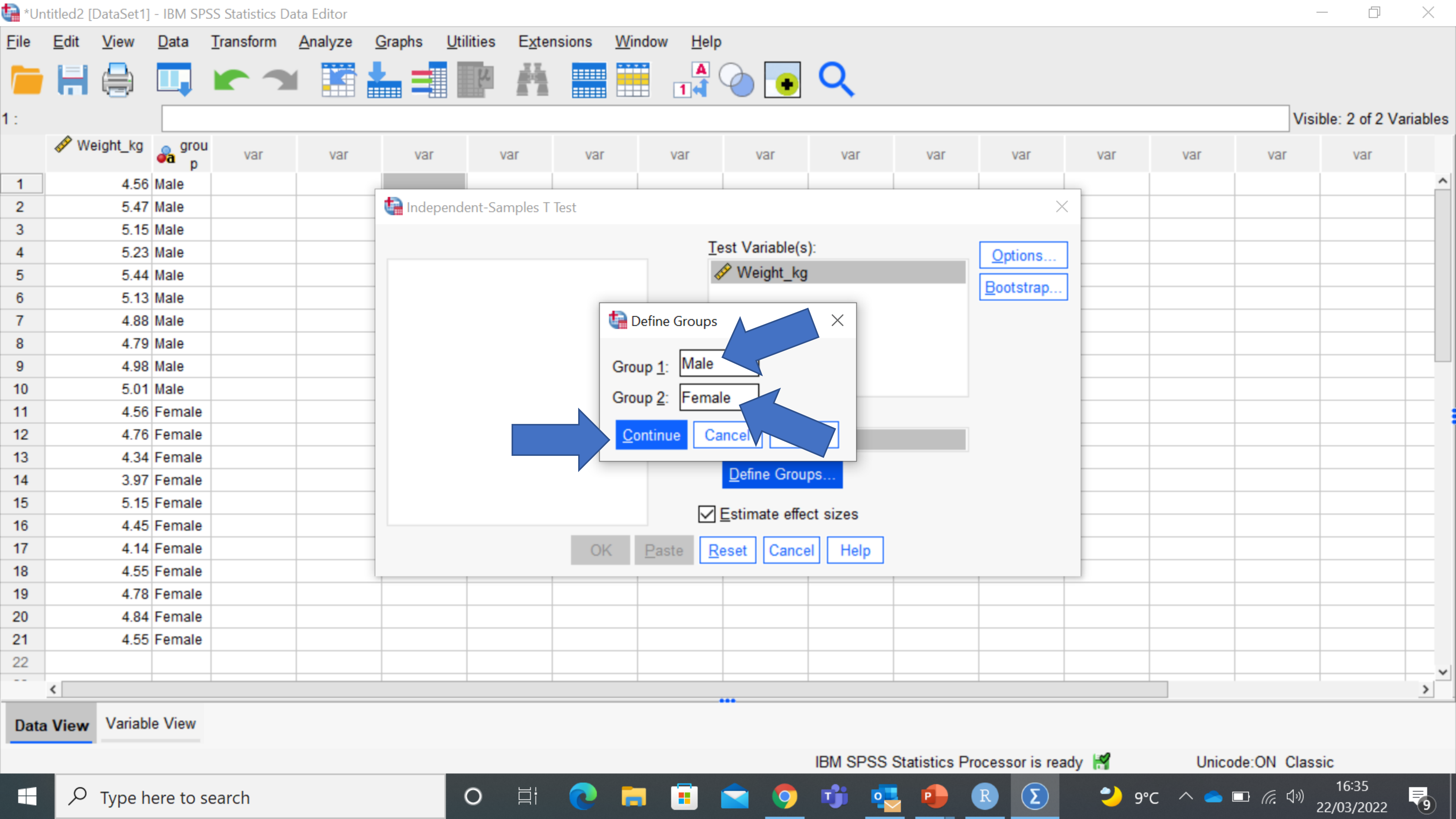

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg group
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
Independent-Samples T Test
Weight_kg
group
Test Variable(s):
Options...
Bootstrap...
Grouping Variable:
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:33
22/03/2022

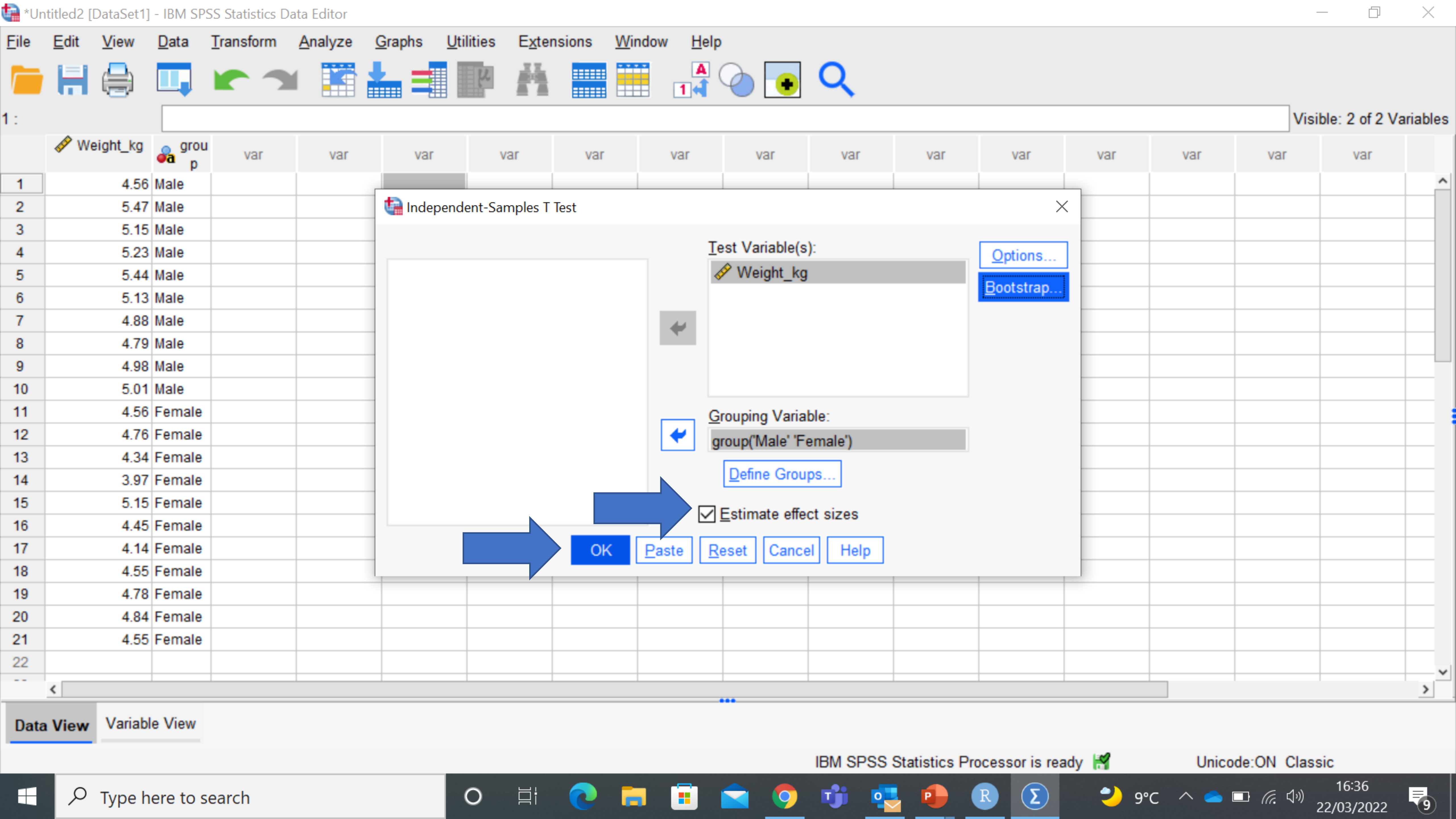

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg group var var var var var var var var var var var var var var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
group
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Grouping Variable:
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:34
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
group
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Grouping Variable:
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:34
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg group var var var var var var var var var var var var var var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Grouping Variable:
group(? ?)
Define Groups...
Estimate effect sizes
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:34
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Grouping Variable:
group(? ?)
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:35
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
1 :
Visible: 2 of 2 Variables
Weight_kg
group
var
1 4.56 Male
2 5.47 Male
3 5.15 Male
4 5.23 Male
5 5.44 Male
6 5.13 Male
7 4.88 Male
8 4.79 Male
9 4.98 Male
10 5.01 Male
11 4.56 Female
12 4.76 Female
13 4.34 Female
14 3.97 Female
15 5.15 Female
16 4.45 Female
17 4.14 Female
18 4.55 Female
19 4.78 Female
20 4.84 Female
21 4.55 Female
22
Independent-Samples T Test
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Define Groups
Group 1: Male
Group 2: Female
Continue
Cancel
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:35
22/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 2 of 2 Variables
Weight_kg
group
Independent-Samples T Test
Test Variable(s):
Weight_kg
Options...
Bootstrap...
Grouping Variable:
group('Male' 'Female')
Define Groups...
Estimate effect sizes
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
9°C
16:36
22/03/2022

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

ut
T-Test
Title
Notes
Group St
Independ
Independ

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval	
				Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 9°C 16:37 22/03/2022

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Descriptive stats by group:

- Number of cases – 10 males, and 11 females
- Means – same as earlier
- Standard deviation/errors for each mean

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Levene's test for equality of variance:

- That other assumption: Is the variance around the mean for each group approximately equal?
- If this statistic is **not significant (0.05 or higher)**, then you can assume the variance is equal

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

SO: Since our Levene's test gave us a p value that was greater than (or equal to) 0.05, we can look at the first row of this table for our results ☺

Easy!

If this wasn't the case, you'd simply look at the second row: "Equal variance not assumed"
welch's t-test

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

- *t* is the computed test statistic (t statistic)
- *df* is the degrees of freedom

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Output6 [Document6] - IBM SPSS Statistics Viewer

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

- *Your p value!* Generally, this is the number people get the most excited about:
- Within the scientific community, we generally accept p values that are <0.05 as significant:
- In this case, out p value is <0.05, so we would say we can reject the null hypothesis that the mean weights in our 2 groups are equal

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

- *Mean Difference* is the difference between the sample means, i.e. $x_1 - x_2$; it also corresponds to the numerator of the test statistic for that test

- *Std. Error Difference* is the standard error of the mean difference estimate; it also corresponds to the denominator of the test statistic for that test

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Effect size:

When using T-tests, we often also want to report the size of the effect we are reporting

p values may tell you whether or not there is an effect, but…

…p values DO NOT tell us anything about effect size

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Output6 [Document6] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

ut
T-Test
Title
Notes
Group St
Independ
Independ

Group Statistics

	group	N	Mean	Std. Deviation	Std. Error Mean
Weight_kg	Male	10	5.0640	.28254	.08935
	Female	11	4.5536	.33161	.09998

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
Weight_kg	Equal variances assumed	.059	.811	3.776	19	<.001	.001	.51036	.13516	.22747	.79325
	Equal variances not assumed			3.806	18.933	<.001	.001	.51036	.13409	.22965	.79108

Independent Samples Effect Sizes

		Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Weight_kg	Cohen's d	.30934	1.650	.632	2.636
	Hedges' correction	.32226	1.584	.606	2.531
	Glass's delta	.33161	1.539	.431	2.600

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the sample standard deviation of the control group.

Effect size:

For effect size, we use Cohen's d: a **standardized** mean difference

General rule:

- small effect (d = 0.2 and up)
- medium effect (d = 0.5 and up)
- large effect (d = 0.8 and up)

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 9°C 16:37 22/03/2022

So what am I reporting?!?!

Take a look at this, and see where I got these numbers from our output!

- General formula:
- *t*(degrees of freedom) = the *t* statistic, *p* = *p* value, d = cohen's d.

- The mean weight of females in our sample was 4.55kg (SD = 0.33), whereas the mean weight in males was 5.06kg (SD = 0.28). An independent samples T Test showed that the difference was statistically significant, t(19), 3.776, p = 0.001, d = 1.65.

Great! That wasn't so bad (I hope!)

- But what happens when we can't meet all the assumptions?

- Well, the assumption of equal variance is easy to deal with: welch's t-test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means					C	
			A	B					95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Mile time	Equal variances assumed	102.98	.000	13.475	390	.000	0:02:14	0:00:10	0:01:55	0:02:34
	Equal variances not assumed			15.047	315.846	.000	0:02:14	0:00:08	0:01:57	0:02:32

Image from: https://libguides.library.kent.edu/spss/independentttest

What about if our data isn't **normally distributed**?

- This is a much bigger issue.

- Lets take a look at an example…

- Load the data: "genderagg.xlsx"
    - Data containing the number of aggressive interactions started by male and female vervet monkeys over the span of a month

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	Sex	Sex_num	aggress	var	var	var	var	var	var	var	var	var	var	var	var	var
1	Female	1	19													
2	Female	1	18													
3	Female	1	9													
4	Female	1	17													
5	Female	1	8													
6	Female	1	7													
7	Female	1	16													
8	Female	1	19													
9	Female	1	20													
10	Female	1	9													
11	Female	1	11													
12	Female	1	18													
13	Male	2	16													
14	Male	2	5													
15	Male	2	15													
16	Male	2	2													
17	Male	2	14													
18	Male	2	15													
19	Male	2	4													
20	Male	2	7													
21	Male	2	15													
22	Male	2	6													
23	Male	2	7													

Take some time to look at the **average number of aggressive interactions** by males and females

Please notice here that Sex_num represents males and females…

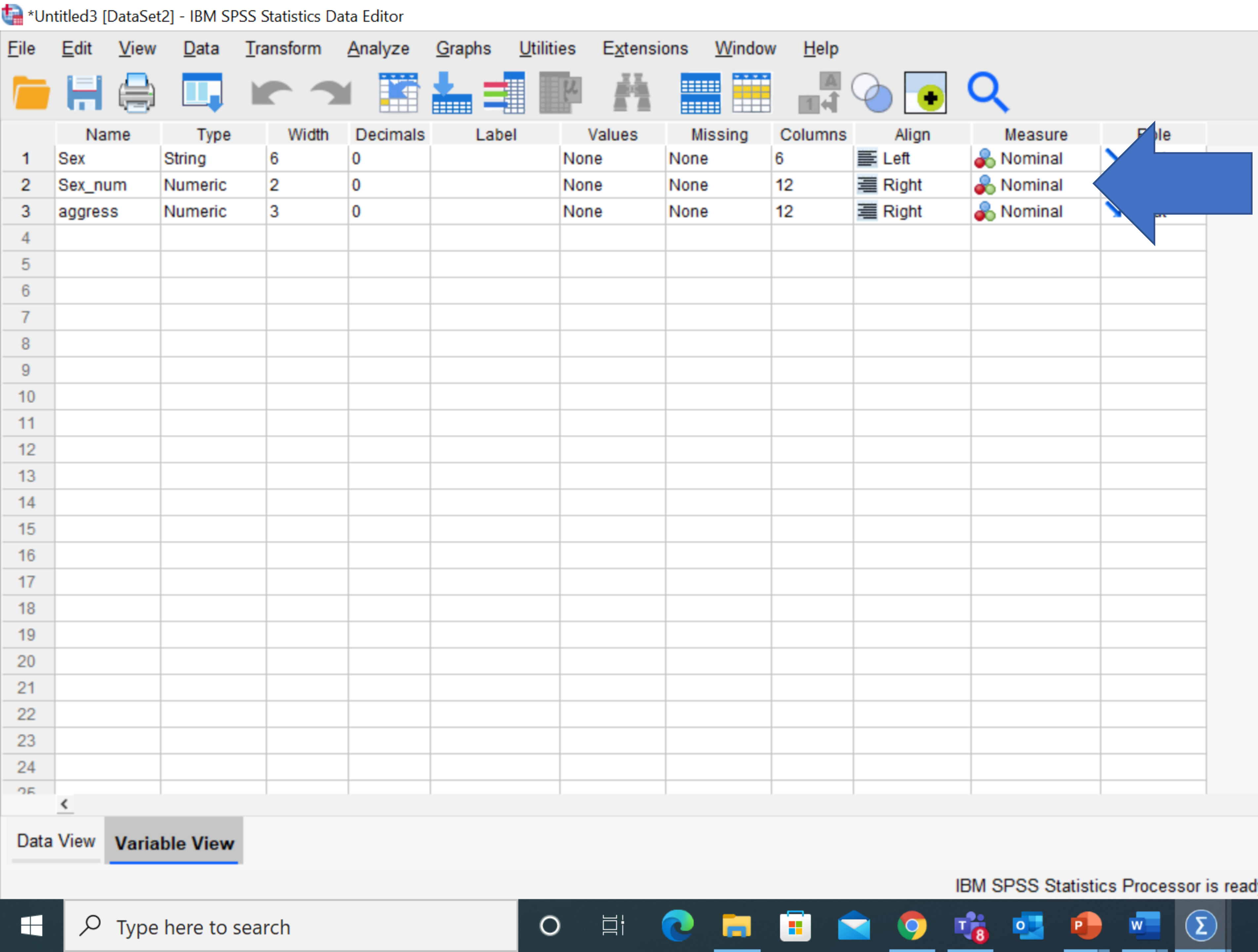
This is important!

Check out your variable view

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 11°C 20:02 24/03/2022



Today, SPSS was on our size, and it automatically realized that the 1's and 2's for Sex_num were categories…

BUT

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor

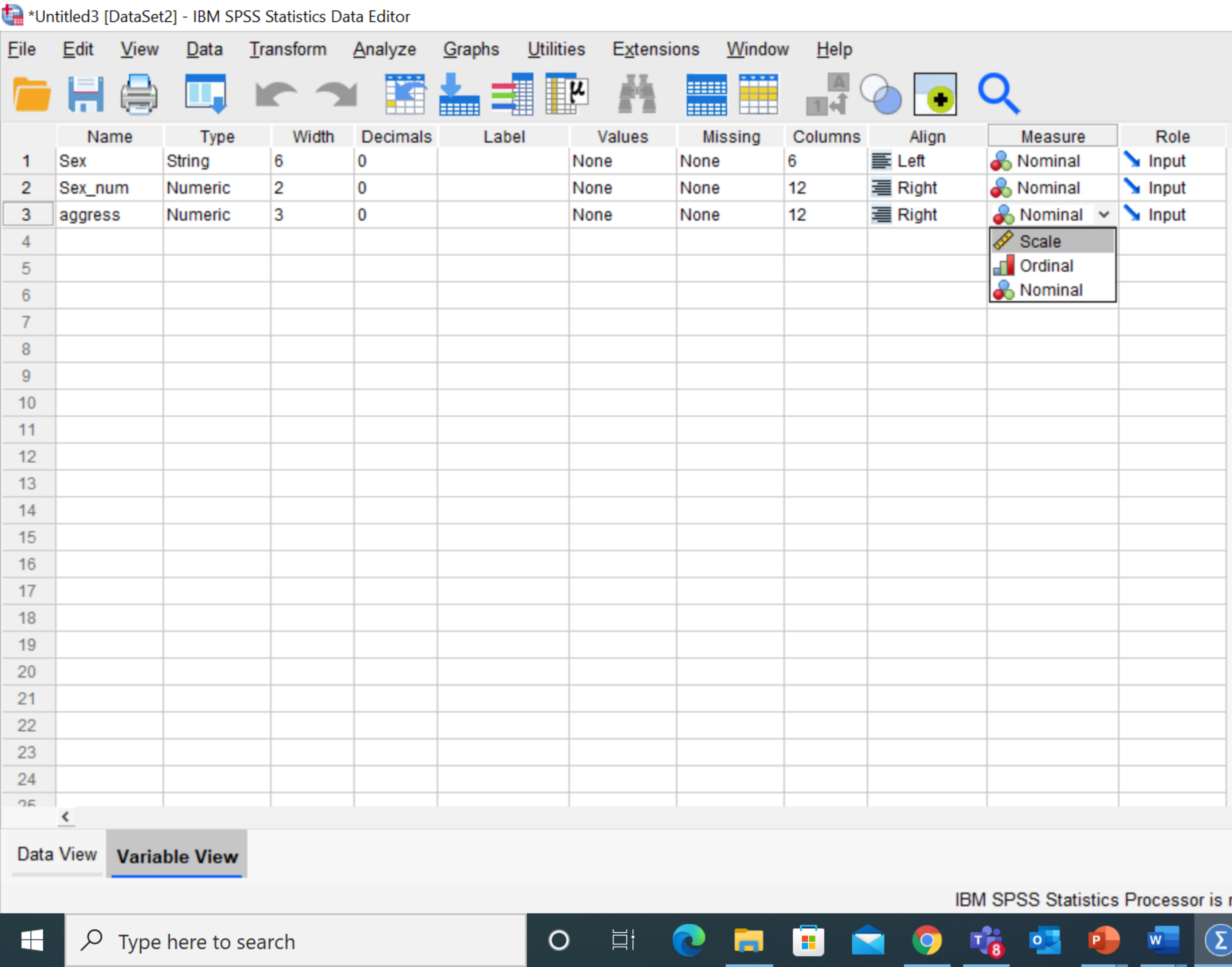
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Sex	String	6	0		None	None	6	Left	Nominal	Input
2	Sex_num	Numeric	2	0		None	None	12	Right	Nominal	
3	aggress	Numeric	3	0		None	None	12	Right	Nominal	

It got the measurement type wrong for ‘aggress’!

This is a COUNT of aggressive interactions, and it is continuous

Data View | Variable View

IBM SPSS Statistics Processor is ready   Unicode:ON   Classic



Click that box and be sure to change it to a 'Scale' variable!

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	Sex	Sex_num	aggress	var	var	var	var	var	var	var	var	var	var	var	var	var
1	Female	1	19													
2	Female	1	18													
3	Female	1	9													
4	Female	1	17													
5	Female	1	8													
6	Female	1	7													
7	Female	1	16													
8	Female	1	19													
9	Female	1	20													
10	Female	1	9													
11	Female	1	11													
12	Female	1	18													
13	Male	2	16													
14	Male	2	5													
15	Male	2	15													
16	Male	2	2													
17	Male	2	14													
18	Male	2	15													
19	Male	2	4													
20	Male	2	7													
21	Male	2	15													
22	Male	2	6													
23	Male	2	7													

You should try to make a **boxplot** to get a feel for the data

Go for it! Help each other out!!!

While you're at it, do you remember how to check out the **mean**?

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 11°C 20:02 24/03/2022

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	Sex	Sex_num	aggress
1	Female	1	19
2	Female	1	18
3	Female	1	9
4	Female	1	17
5	Female	1	8
6	Female	1	7
7	Female	1	16
8	Female	1	19
9	Female	1	20
10	Female	1	9
11	Female	1	11
12	Female	1	18
13	Male	2	16
14	Male	2	5
15	Male	2	15
16	Male	2	2
17	Male	2	14
18	Male	2	15
19	Male	2	4
20	Male	2	7
21	Male	2	15
22	Male	2	6
23	Male	2	7

Now: This looks like a perfect data set to analyse using a paired T Test, but is it normally distributed?

Let's check!

Try it out!

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	Sex	Sex_num	aggress
1	Female	1	19
2	Female	1	18
3	Female	1	9
4	Female	1	17
5	Female	1	8
6	Female	1	7
7	Female	1	16
8	Female	1	19
9	Female	1	20
10	Female	1	9
11	Female	1	11
12	Female	1	18
13	Male	2	16
14	Male	2	5
15	Male	2	15
16	Male	2	2
17	Male	2	14
18	Male	2	15
19	Male	2	4
20	Male	2	7
21	Male	2	15
22	Male	2	6
23	Male	2	7

As this data is not normally distributed, we cannot use a T Test

A commonly used non-parametric equivalent of the independent samples T Test is the **Mann–Whitney U test**

the **Mann–Whitney U test** drops the assumption of normality

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Mann–Whitney U test

- A **Mann–Whitney U test** does not compare means of 2 groups
- It instead RANKS the values of the dependant values from lowest to highest

Sex	Sex_num	aggress
Female	1	19
Female	1	18
Female	1	9
Female	1	17
Female	1	8
Female	1	7
Female	1	16
Female	1	19
Female	1	20
Female	1	9
Female	1	11
Female	1	18
Male	2	16
Male	2	5
Male	2	15
Male	2	2
Male	2	14
Male	2	15
Male	2	4
Male	2	7
Male	2	15
Male	2	6
Male	2	7
Male	2	14

Mann–Whitney U test

- It then compares the **mean ranks** of the two groups:
- Null hypothesis: The two samples come from the same population
- Ie., the probability is 50% that a randomly drawn member of the first population will exceed a member of the second population
- Ie., does one group have higher ranking values than the other group overall?

1	Sex	Sex_num	aggress	Rank
2	Male	2	2	1
3	Male	2	4	2
4	Male	2	5	3
5	Male	2	6	4
6	Female	1	7	6
7	Male	2	7	6
8	Male	2	7	6
9	Female	1	8	8
10	Female	1	9	9.5
11	Female	1	9	9.5
12	Female	1	11	11
13	Male	2	14	12.5
14	Male	2	14	12.5
15	Male	2	15	15
16	Male	2	15	15
17	Male	2	15	15
18	Female	1	16	17.5
19	Male	2	16	17.5
20	Female	1	17	19
21	Female	1	18	20.5
22	Female	1	18	20.5
23	Female	1	19	22.5
24	Female	1	19	22.5
25	Female	1	20	24

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

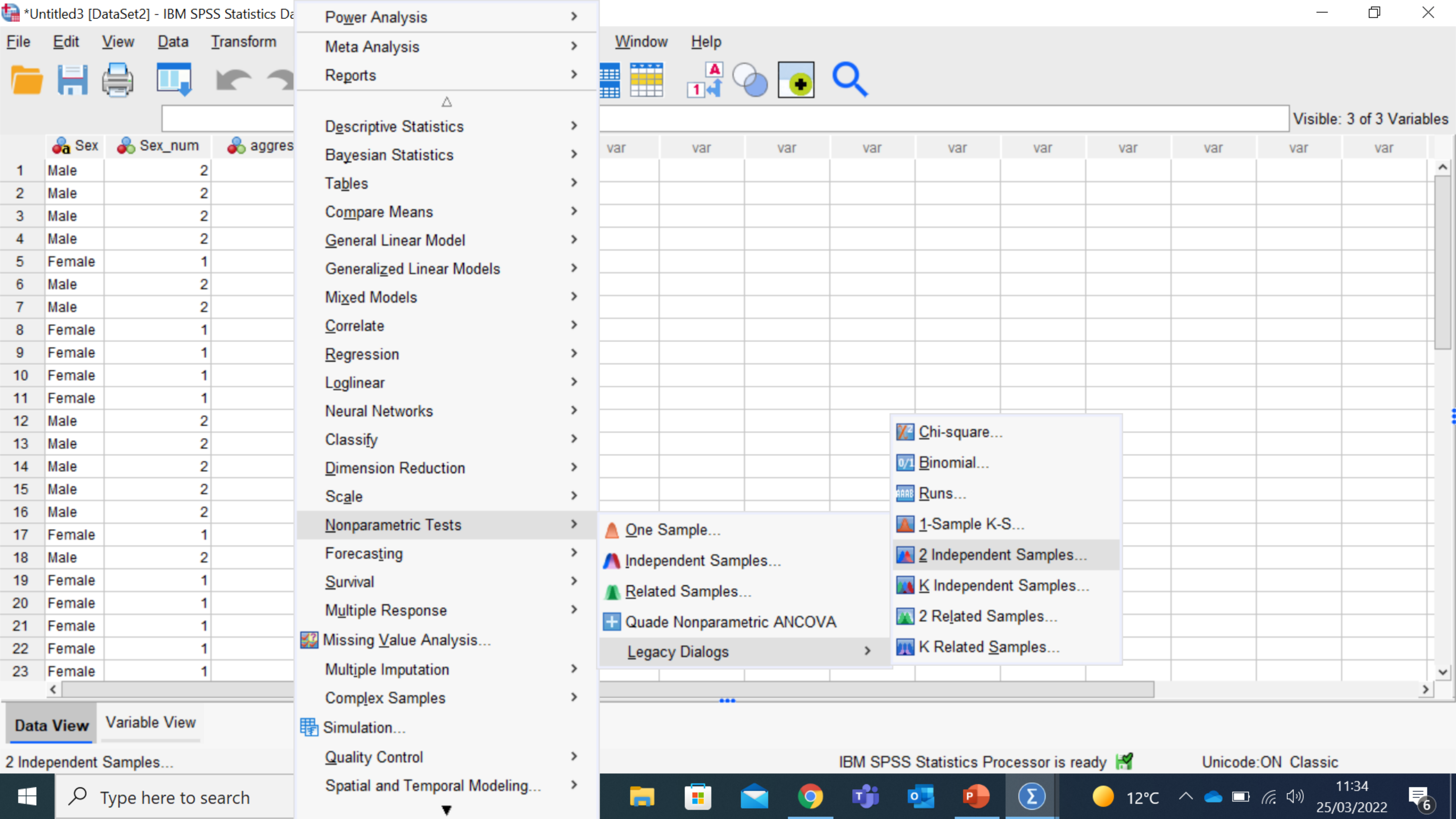
Visible: 3 of 3 Variables

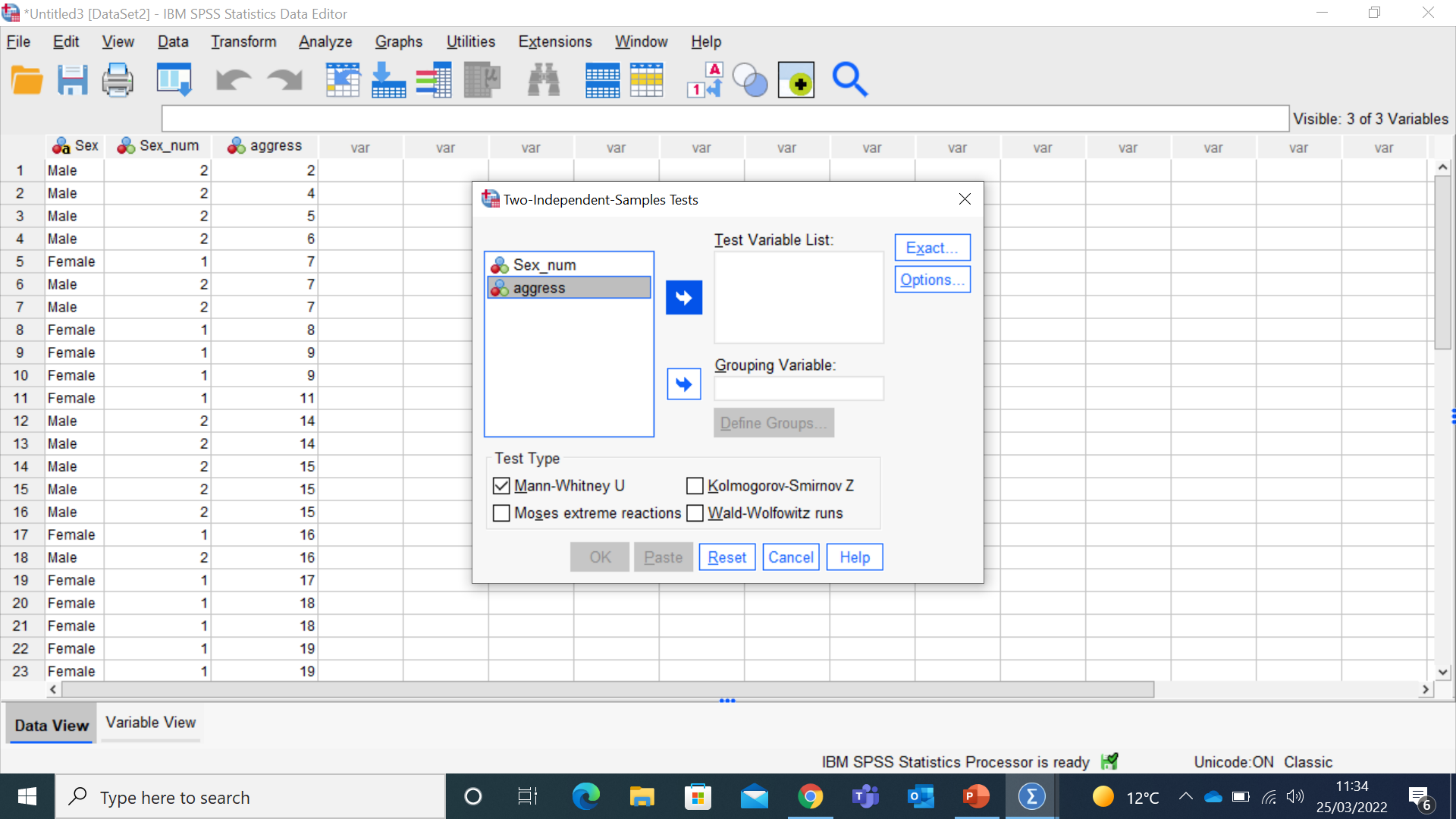
	Sex	Sex_num	aggress	var	var	var	var	var	var	var	var	var	var	var	var	var
1	Male	2	2													
2	Male	2	4													
3	Male	2	5													
4	Male	2	6													
5	Female	1	7													
6	Male	2	7													
7	Male	2	7													
8	Female	1	8													
9	Female	1	9													
10	Female	1	9													
11	Female	1	11													
12	Male	2	14													
13	Male	2	14													
14	Male	2	15													
15	Male	2	15													
16	Male	2	15													
17	Female	1	16													
18	Male	2	16													
19	Female	1	17													
20	Female	1	18													
21	Female	1	18													
22	Female	1	19													
23	Female	1	19													

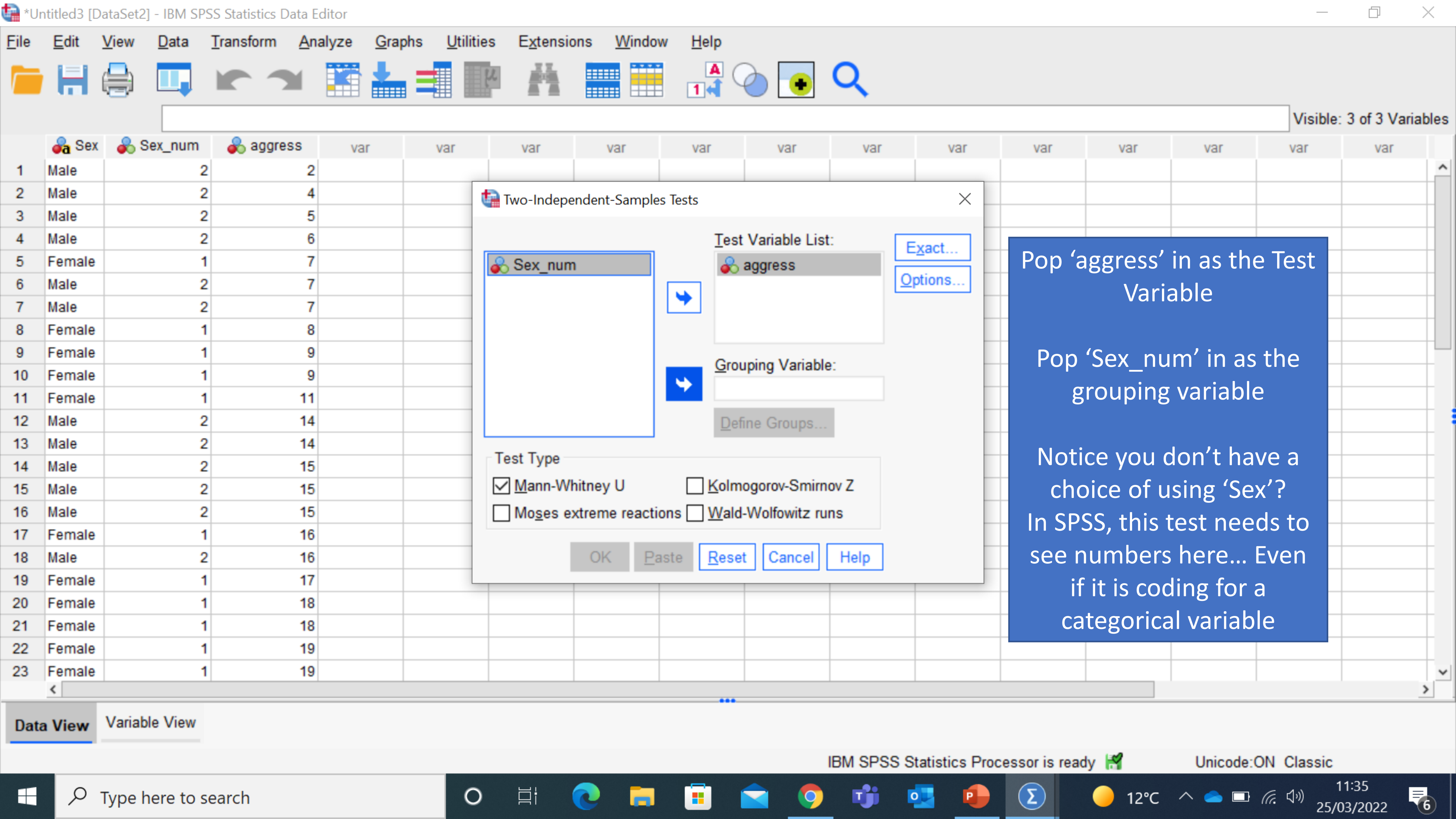
Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 12°C 11:33 25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
Visible: 3 of 3 Variables
Sex
Sex_num
aggres
var
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
One Sample...
Independent Samples...
Related Samples...
Quade Nonparametric ANCOVA
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...
1 Male 2
2 Male 2
3 Male 2
4 Male 2
5 Female 1
6 Male 2
7 Male 2
8 Female 1
9 Female 1
10 Female 1
11 Female 1
12 Male 2
13 Male 2
14 Male 2
15 Male 2
16 Male 2
17 Female 1
18 Male 2
19 Female 1
20 Female 1
21 Female 1
22 Female 1
23 Female 1
Data View
Variable View
2 Independent Samples...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
11:34
25/03/2022

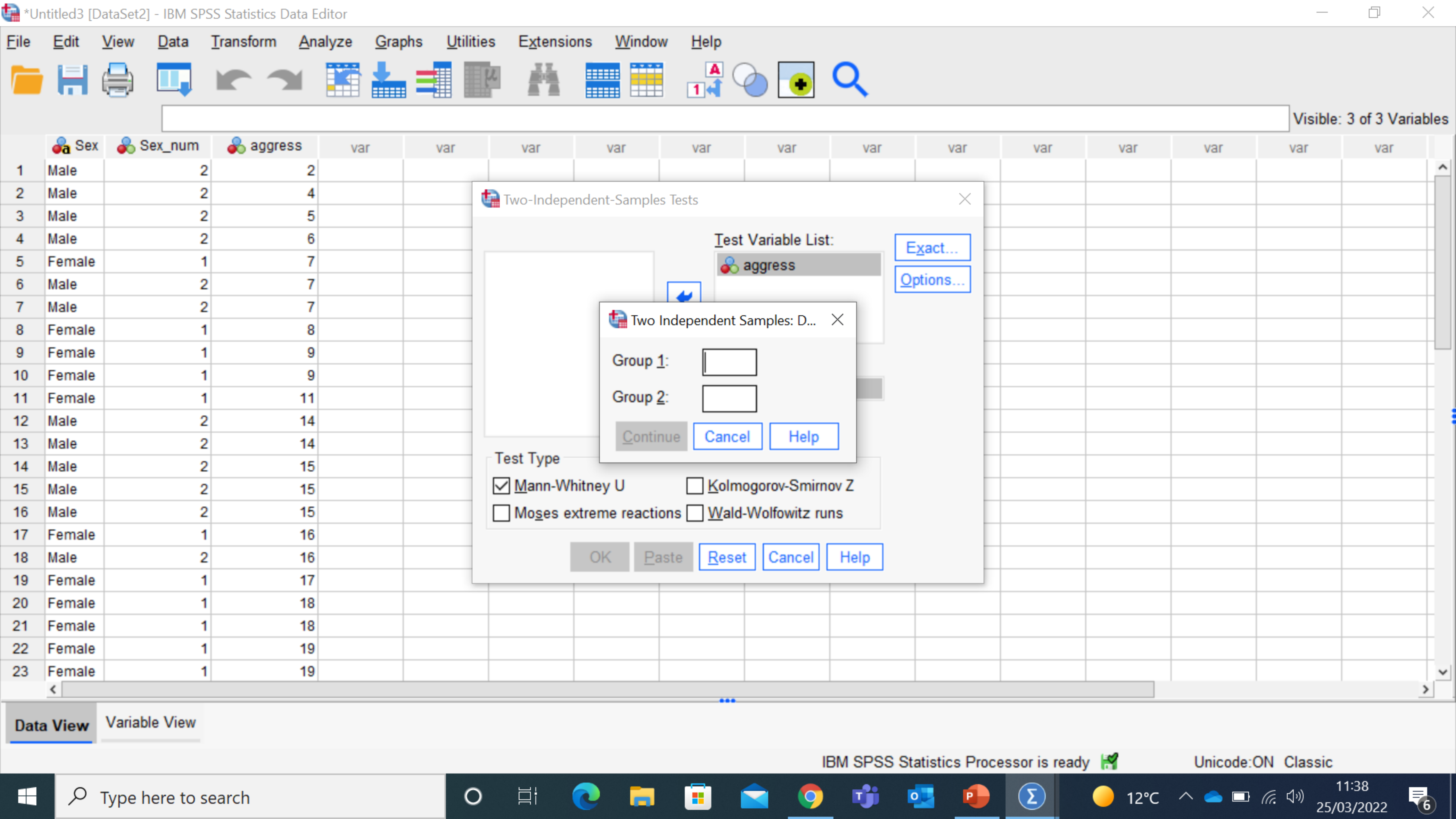

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var var var var var var var var var var var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Two-Independent-Samples Tests
Test Variable List:
Exact...
Options...
Sex_num
aggress
Grouping Variable:
Define Groups...
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
11:34
25/03/2022



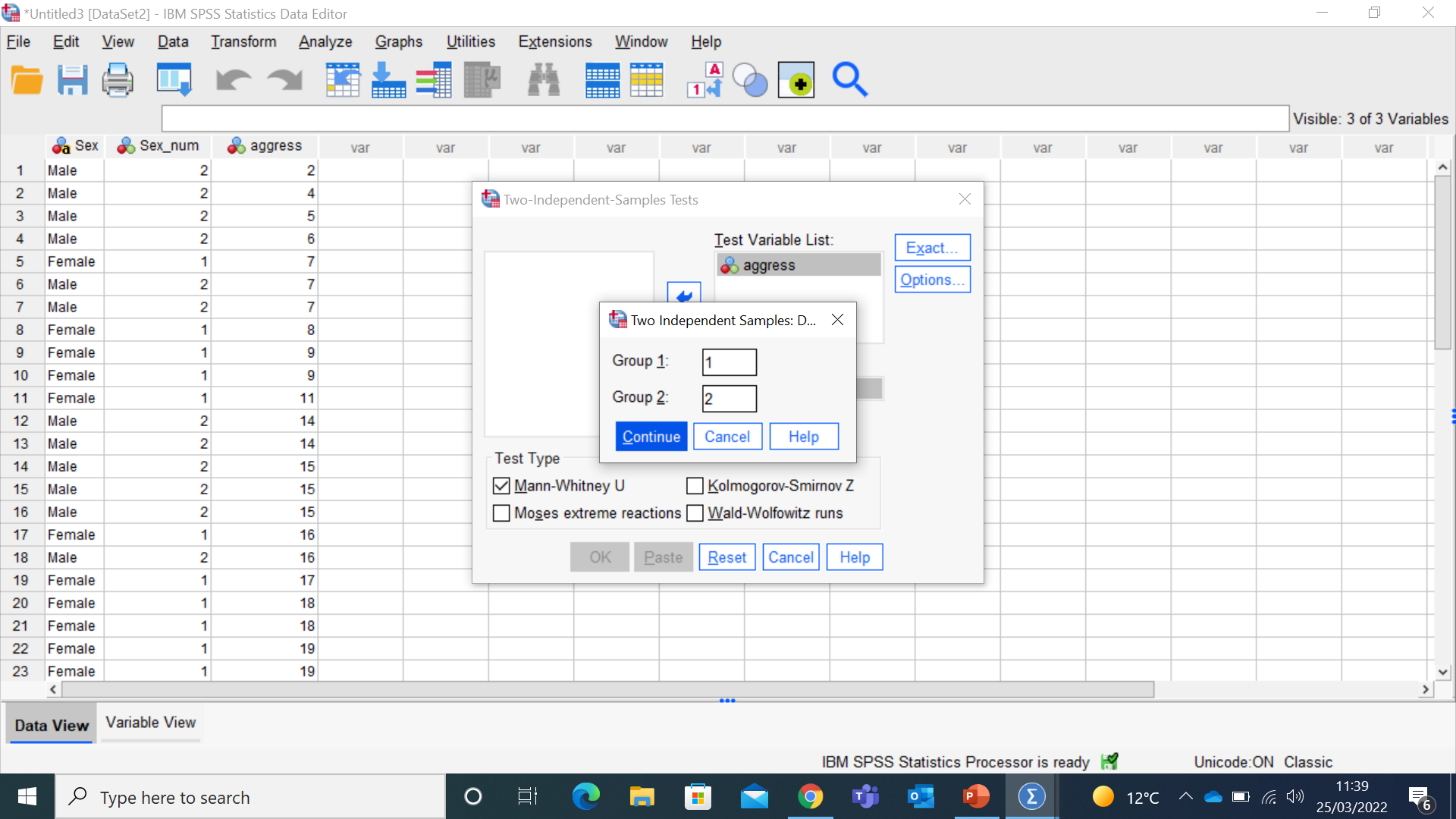
Pop 'aggress' in as the Test Variable

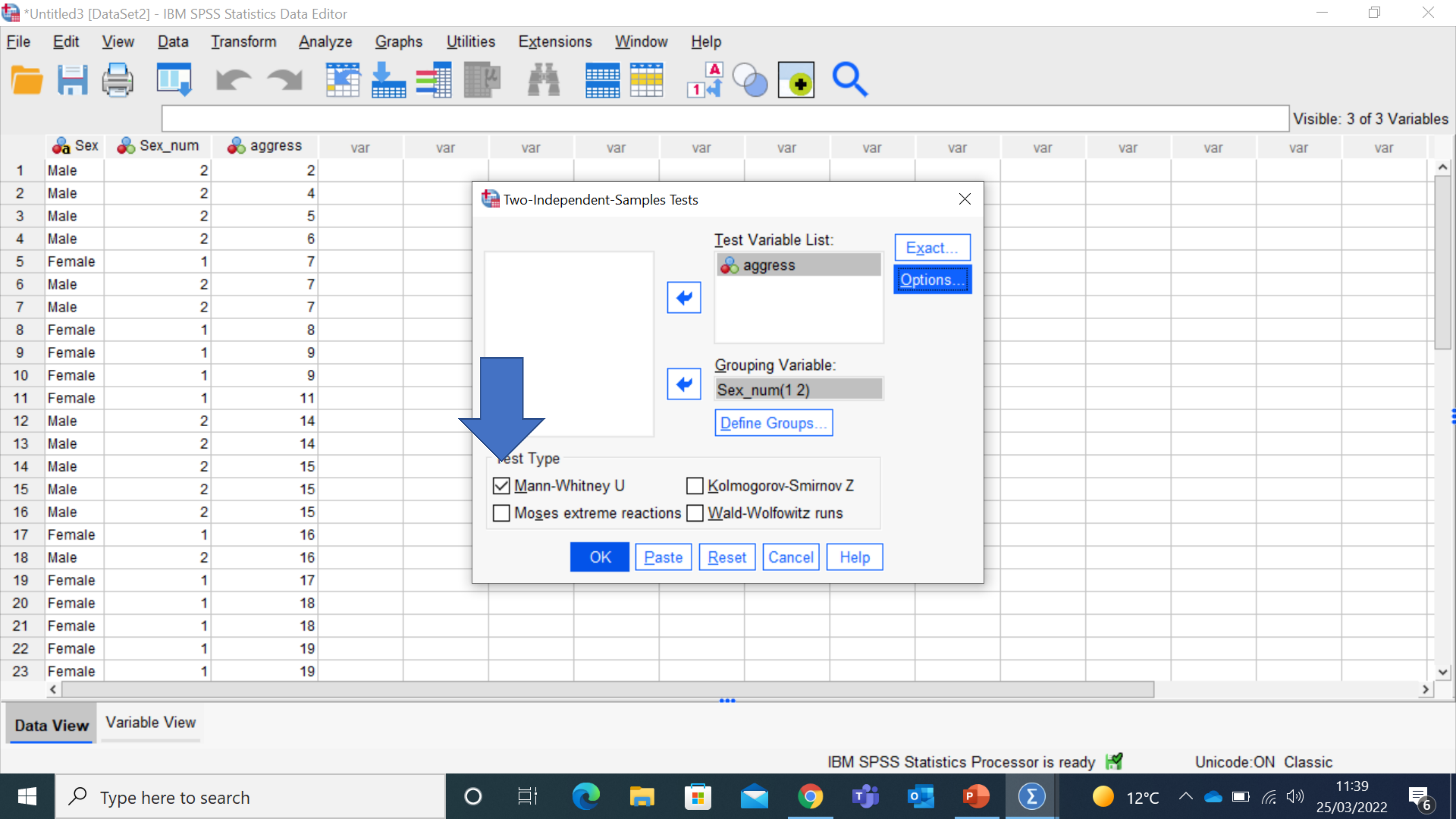
Pop 'Sex_num' in as the grouping variable

Notice you don't have a choice of using 'Sex'?
In SPSS, this test needs to see numbers here… Even if it is coding for a categorical variable



*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 3 of 3 Variables
Sex
Sex_num
aggress
var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Two-Independent-Samples Tests
Test Variable List:
aggress
Exact...
Options...
Two Independent Samples: D...
Group 1:
Group 2:
Continue
Cancel
Help
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
12°C
11:38
25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var var var var var var var var var var var var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Two-Independent-Samples Tests
Test Variable List:
aggress
Exact...
Options...
Two Independent Samples: D...
Group 1: 1
Group 2: 2
Continue
Cancel
Help
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
11:39
25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var
Two-Independent-Samples Tests
Test Variable List:
aggress
Exact...
Options...
Grouping Variable:
Sex_num(1 2)
Define Groups...
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
12°C
11:39
25/03/2022

*Output1 [Document1] - IBM SPSS Statistics Viewer

NPar Tests

[DataSet2]

Mann-Whitney Test

Ranks

	Sex_num	N	Mean Rank	Sum of Ranks
aggress	1	12	15.88	190.50
	2	12	9.13	109.50
	Total	24		

Test Statistics[a]

	aggress
Mann-Whitney U	31.500
Wilcoxon W	109.500
Z	-2.345
Asymp. Sig. (2-tailed)	.019
Exact Sig. [2*(1-tailed Sig.)]	.017[b]

a. Grouping Variable: Sex_num

b. Not corrected for ties.

Our first bit of output is the **ranks** table.

Here, we can see that Sex_num 1 (females) had a higher mean rank than Sex_num 2 (males)

Importantly: If sex had no effect on the number of aggressive interactions an individual took part in, we would expect the mean ranks of males and females to be quite similar genderagg - check • Last Modified: Just now

	A	B	C	D	E	F	G	H	I	J
1	Male	2	2	1			Female	1	7	6
2	Male	2	4	2			Female	1	8	8
3	Male	2	5	3			Female	1	9	9.5
4	Male	2	6	4			Female	1	9	9.5
5	Male	2	7	6			Female	1	11	11
6	Male	2	7	6			Female	1	16	17.5
7	Male	2	14	12.5			Female	1	17	19
8	Male	2	14	12.5			Female	1	18	20.5
9	Male	2	15	15			Female	1	18	20.5
10	Male	2	15	15			Female	1	19	22.5
11	Male	2	15	15			Female	1	19	22.5
12	Male	2	16	17.5			Female	1	20	24
15			Average Rank Males	9.125					Average Rank females	15.875
17			Sum or Ranks Males	109.5					Sum of Ranks Females	190.5

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - NPar Tests
    - Title
    - Notes
    - Active Dataset
    - Mann-Whitney Tes
      - Title
      - Ranks
      - Test Statistics

NPar Tests

[DataSet2]

Mann-Whitney Test

Ranks

	Sex_num	N	Mean Rank	Sum of Ranks
aggress	1	12	15.88	190.50
	2	12	9.13	109.50
	Total	24		

Test Statistics[a]

	aggress
Mann-Whitney U	31.500
Wilcoxon W	109.500
Z	-2.345
Asymp. Sig. (2-tailed)	.019
Exact Sig. [2*(1-tailed Sig.)]	.017[b]

a. Grouping Variable: Sex_num

b. Not corrected for ties.

Mann-Whitney U and **Wilcoxon W** are test statistics; they summarize the difference in mean rank.

We prefer reporting Exact Sig. (2-tailed): the p-value corrected for ties.

Importantly, this p-value is lower than 0.05

This means we can say that there was a significant difference in the mean ranks of our 2 samples!

This suggests a difference between the 2 groups

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

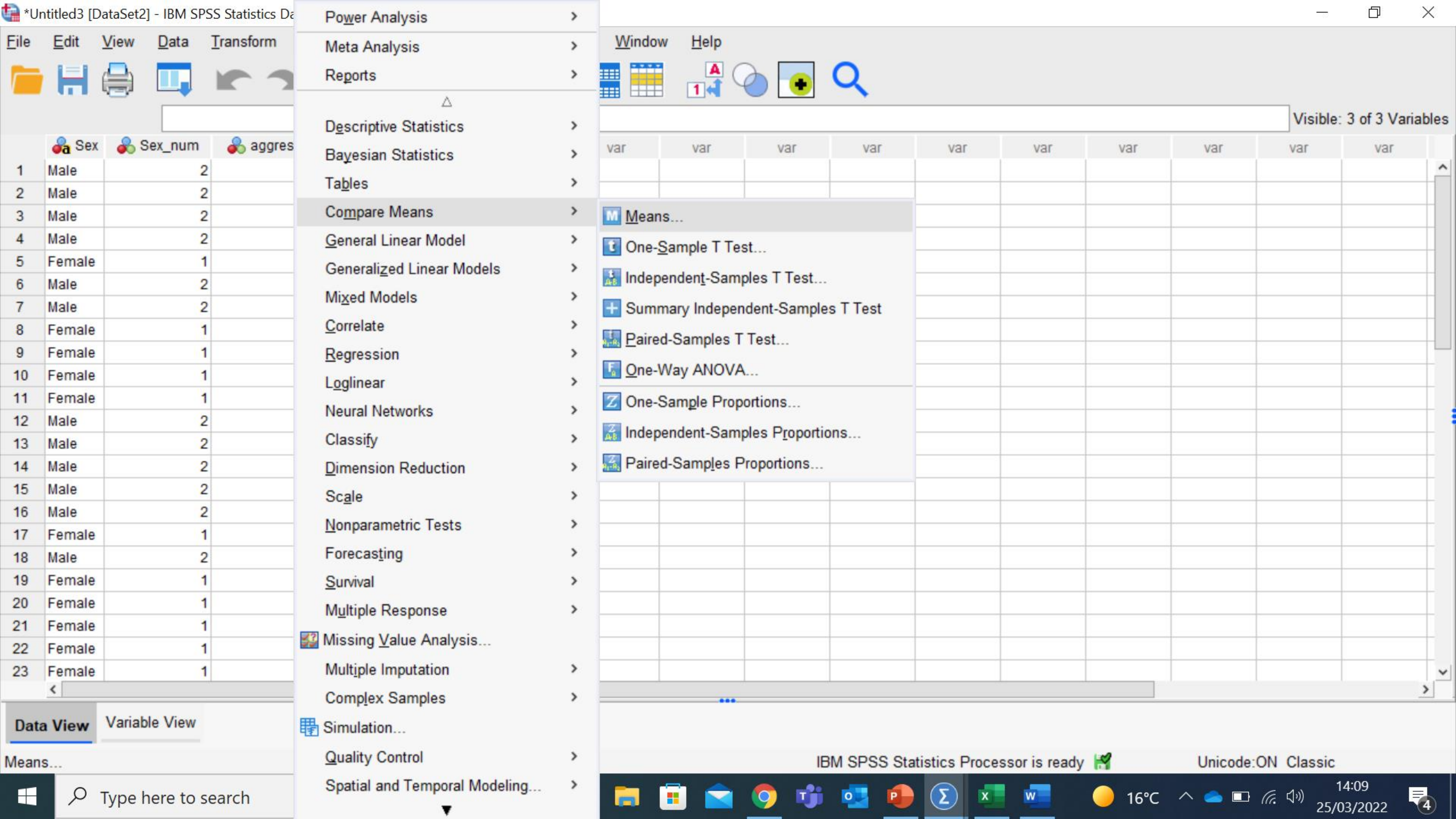
Type here to search   15°C   12:44 25/03/2022

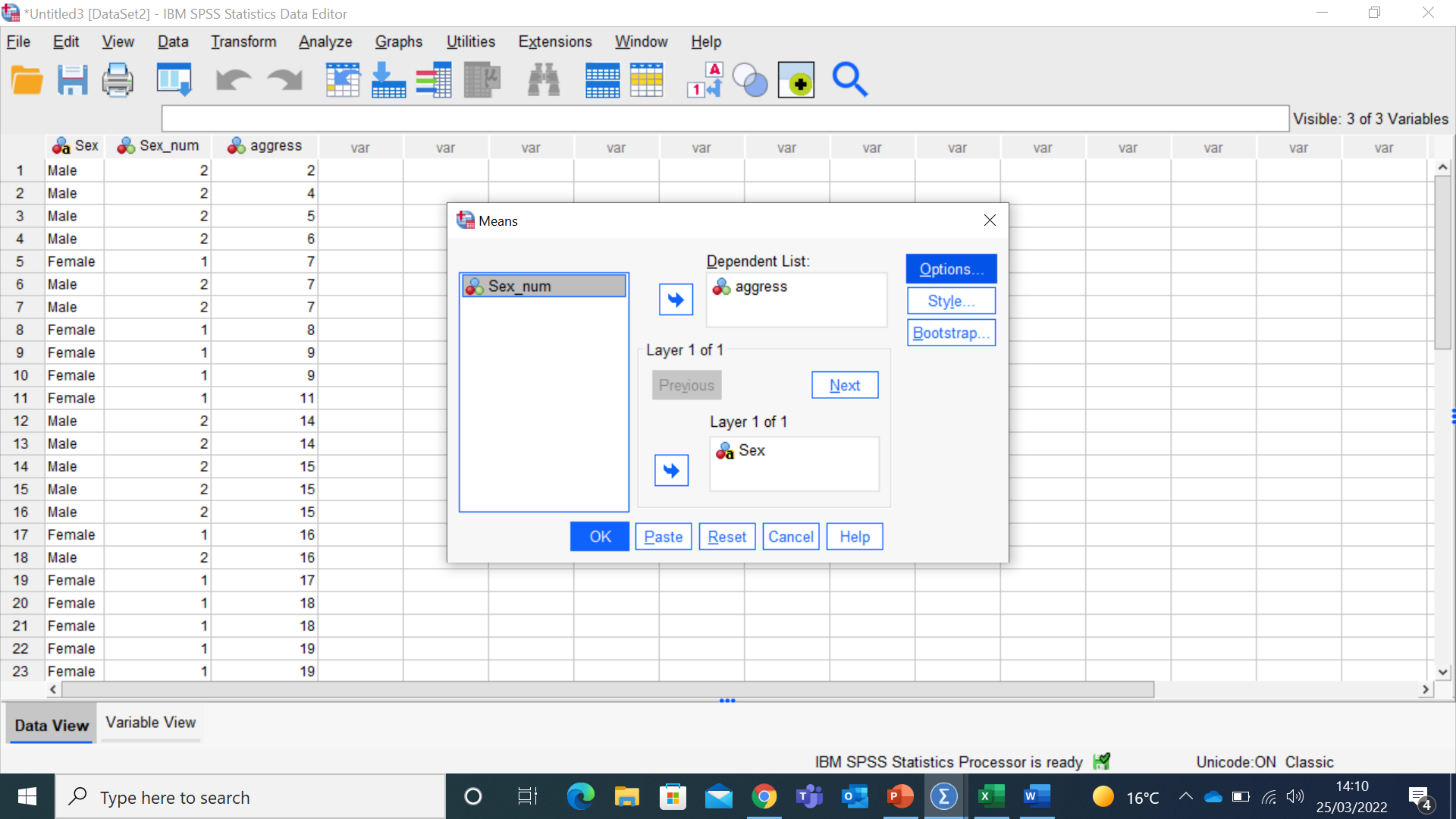
Reporting

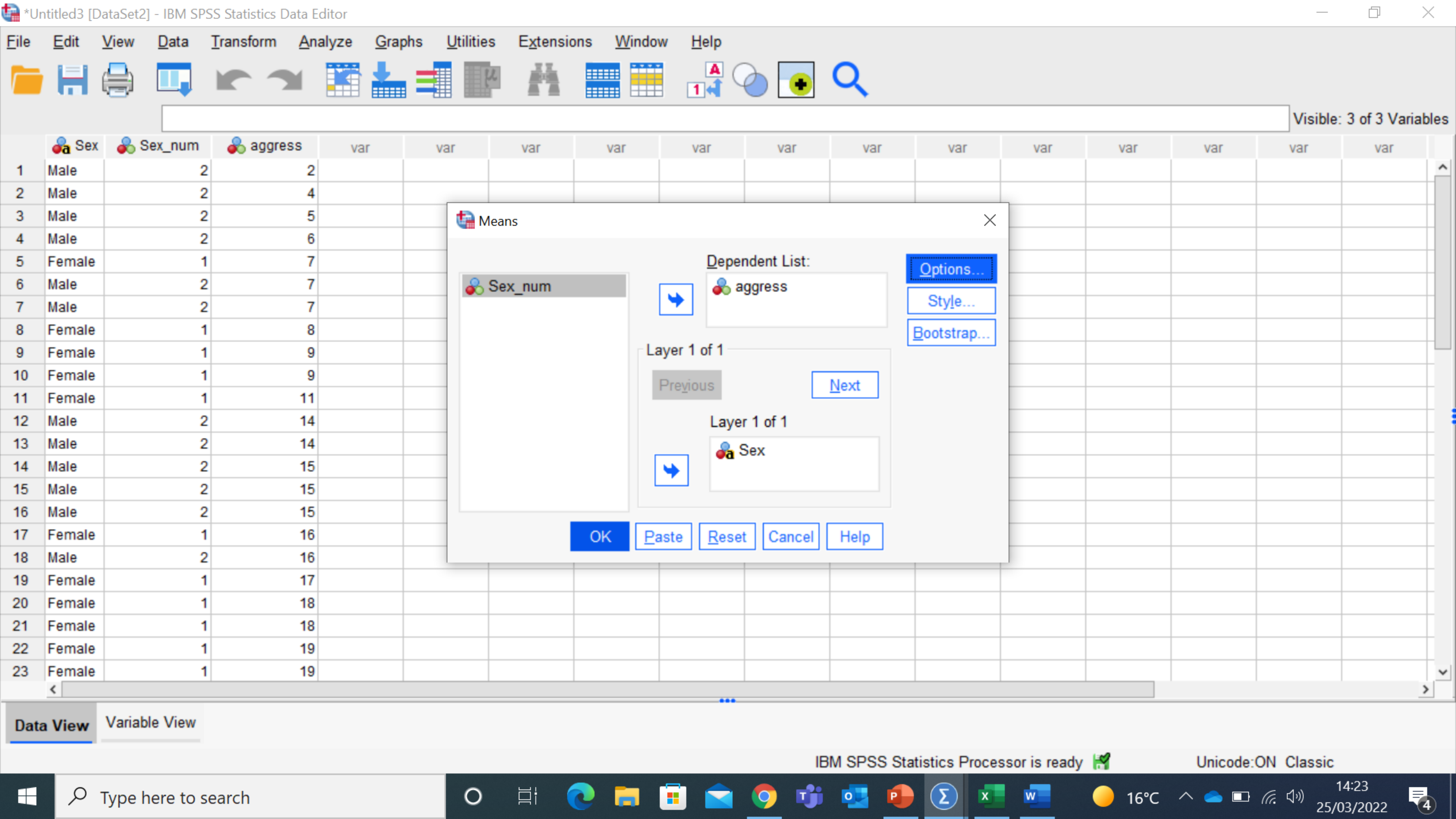
- General formula:
- U = (value of Mann-Whitney U test statistic), p = (p value)
  - Also, you should generally report the medians, and mean rank values for your two groups.

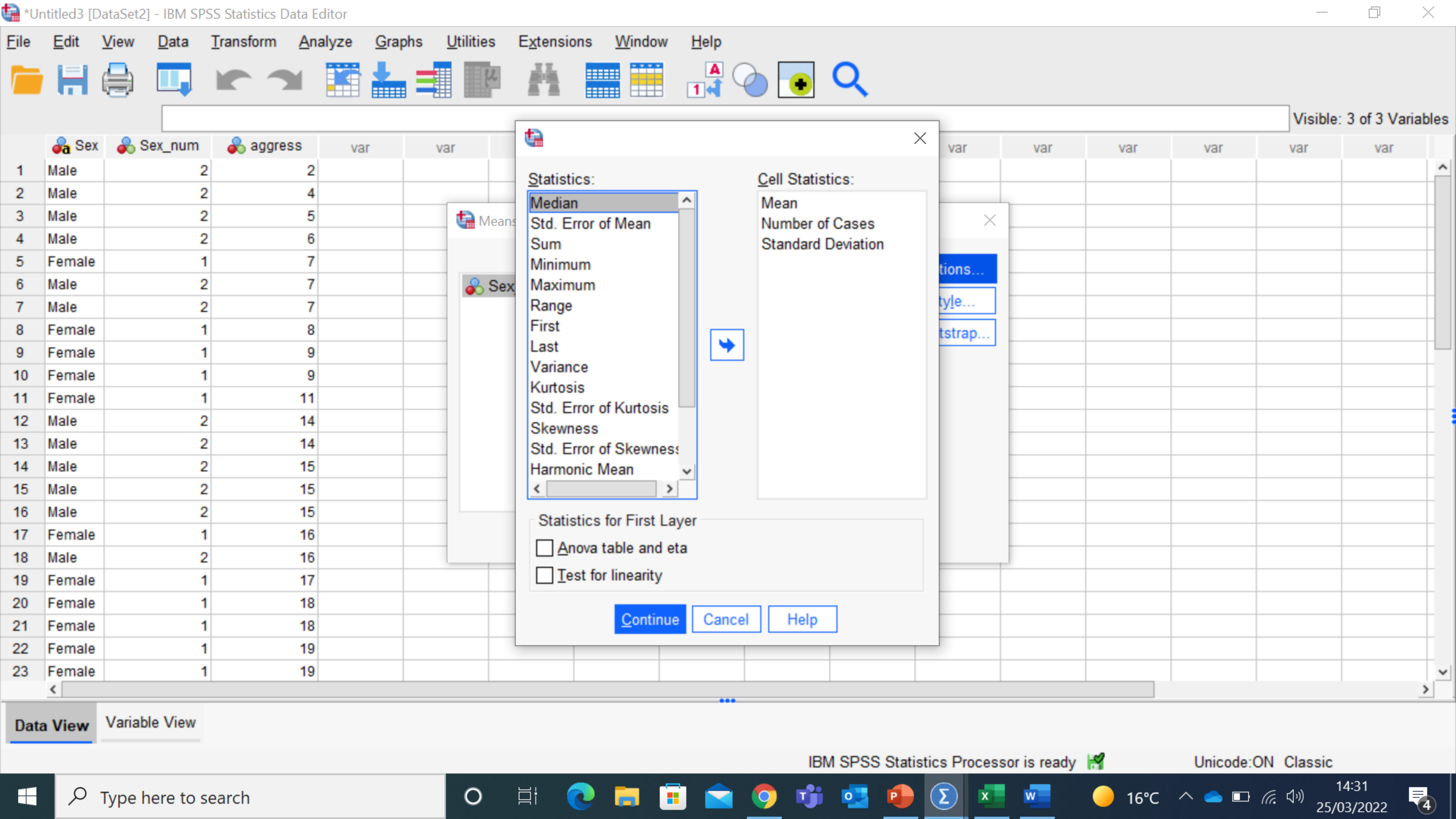
  - A Mann-Whitney test indicated that the number of aggressive interactions that an individual engaged in per day was greater for females (Mdn = ★ , MeanRk = 15.88) than it was for males (Mdn = ★ , MeanRk = 9.13), U = 31.5, p = 0.019.

- Importantly, all this info is available on your SPSS output, except for **the medians**… I can show you were that comes from super quick!

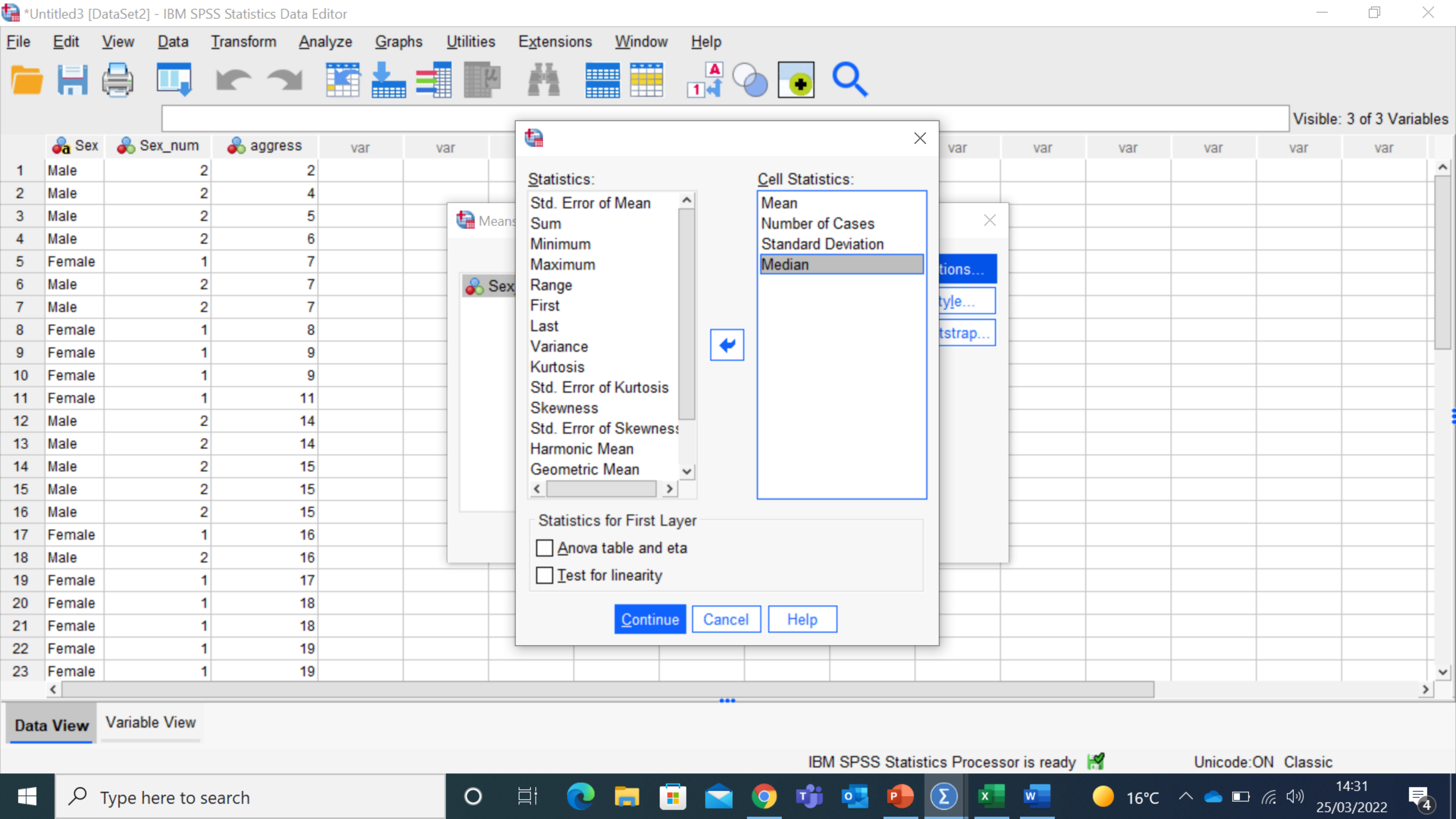

*Untitled3 [DataSet2] - IBM SPSS Statistics Da
File Edit View Data Transform
Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggres
var
1 Male 2
2 Male 2
3 Male 2
4 Male 2
5 Female 1
6 Male 2
7 Male 2
8 Female 1
9 Female 1
10 Female 1
11 Female 1
12 Male 2
13 Male 2
14 Male 2
15 Male 2
16 Male 2
17 Female 1
18 Male 2
19 Female 1
20 Female 1
21 Female 1
22 Female 1
23 Female 1
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
Data View Variable View
Means...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
16°C
14:09
25/03/2022

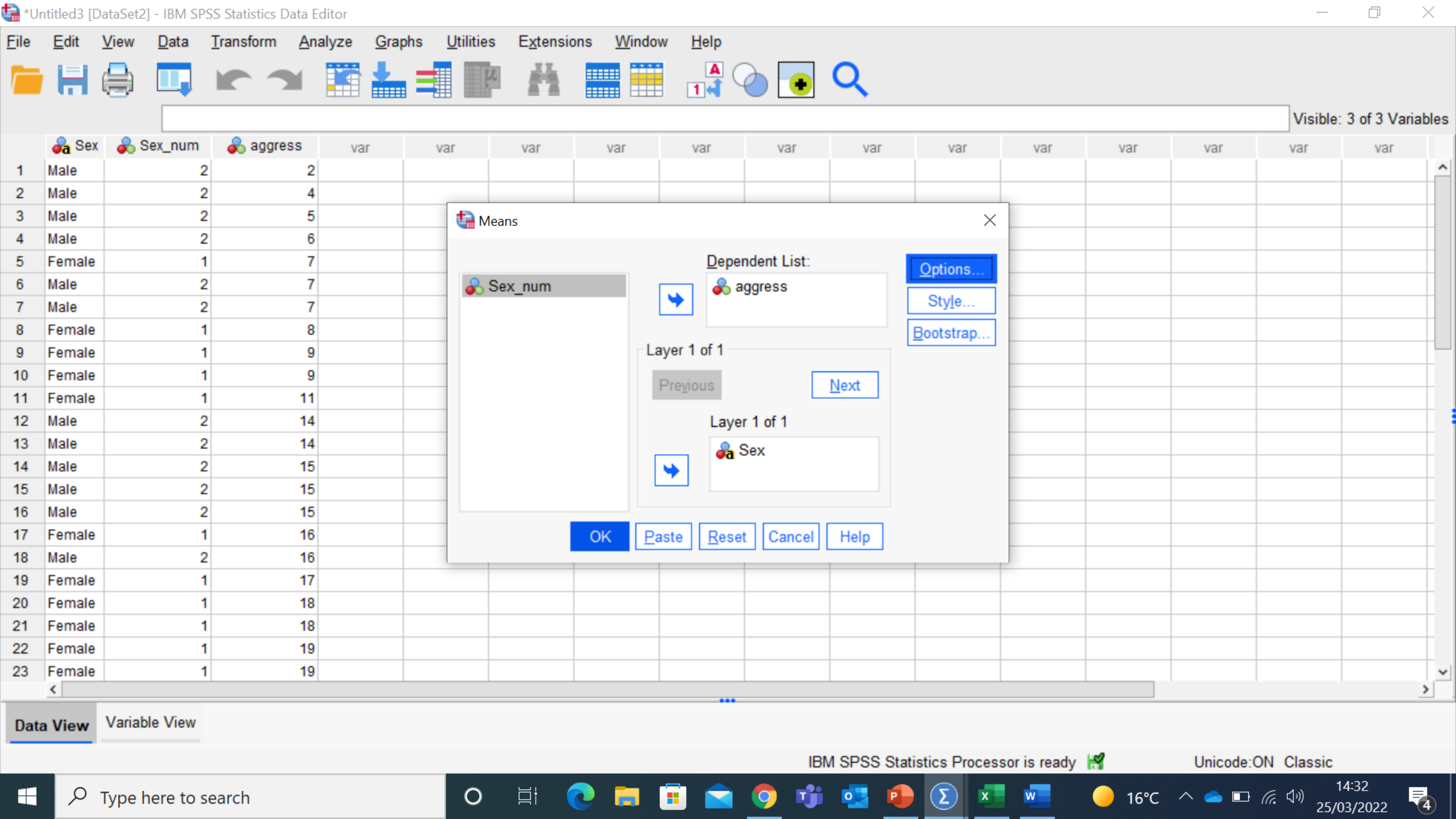

*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var var var var var var var var var var var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Means
Sex_num
Dependent List:
aggress
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
Sex
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
16°C
14:10
25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var var var var var var var var var var var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Means
Sex_num
Dependent List:
aggress
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
Sex
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
16°C
14:23
25/03/2022



*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Means
Statistics:
Median
Std. Error of Mean
Sum
Minimum
Maximum
Range
First
Last
Variance
Kurtosis
Std. Error of Kurtosis
Skewness
Std. Error of Skewness
Harmonic Mean
Cell Statistics:
Mean
Number of Cases
Standard Deviation
Statistics for First Layer
Anova table and eta
Test for linearity
Continue
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
16°C
14:31
25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 3 of 3 Variables
Sex
Sex_num
aggress
var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Means
Sex
tions...
tyle...
tstrap...
Statistics:
Std. Error of Mean
Sum
Minimum
Maximum
Range
First
Last
Variance
Kurtosis
Std. Error of Kurtosis
Skewness
Std. Error of Skewness
Harmonic Mean
Geometric Mean
Cell Statistics:
Mean
Number of Cases
Standard Deviation
Median
Statistics for First Layer
Anova table and eta
Test for linearity
Continue
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
16°C
14:31
25/03/2022


*Untitled3 [DataSet2] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Sex Sex_num aggress var var var var var var var var var var var var var
1 Male 2 2
2 Male 2 4
3 Male 2 5
4 Male 2 6
5 Female 1 7
6 Male 2 7
7 Male 2 7
8 Female 1 8
9 Female 1 9
10 Female 1 9
11 Female 1 11
12 Male 2 14
13 Male 2 14
14 Male 2 15
15 Male 2 15
16 Male 2 15
17 Female 1 16
18 Male 2 16
19 Female 1 17
20 Female 1 18
21 Female 1 18
22 Female 1 19
23 Female 1 19
Means
Sex_num
Dependent List:
aggress
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
Sex
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
16°C
14:32
25/03/2022

*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Means
Title
Notes
Case Processing S
Report

Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
aggress * Sex	24	100.0%	0	0.0%	24	100.0%

Report

aggress

Sex	Mean	N	Std. Deviation	Median
Female	14.25	12	4.993	16.50
Male	10.00	12	5.240	10.50
Total	12.13	24	5.456	14.00

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 16°C 14:32 25/03/2022

Reporting

- General formula:
- U = (value of Mann-Whitney U test statistic), p = (p value)
  - Also, you should generally report the medians, and mean rank values for your two groups.

  - Let's add in the median values!

  - A Mann-Whitney test indicated that the number of aggressive interactions that an individual engaged in per day was greater for females (Mdn = 16.5, MeanRk = 15.88) than it was for males (Mdn = 10.5, MeanRk = 9.13), U = 31.5, p = 0.019.

Paired groups

- In our first 2 analyses, we were comparing 2 groups
- In those tests, each group was made up of different individuals
- The males in group one were different individuals than the females I group 2!

- But what if we measure the SAME individuals twice?
  - Ie., measure the animals once before an experimental treatment
  - Then measure the SAME ANIMALS after the treatment.

- This requires different tests!

Paired samples T Test

- **Paired Samples T Test** compares the means of related group of samples.
  - It is used when you have 2 values measured **from the same individual** – each measurement representing a different time, treatment, etc…
  - Our measurements are therefore "paired"

- Examples:
  - You might want to compare the average weight of 20 mice **before**, and **after** a treatment
    - The data contains 20 sets of values before treatment, and 20 sets of values after treatment - the weight of each mouse was weighed twice (once before treatment, and once after treatment)

  - A measurement taken at two different times (e.g., Weight of animals in the fall compared to the spring)
    - The data contains 15 sets of values from the fall, and 15 from the spring – each animal was captured and weighed in the fall, then **re-captured** and weighed **again** in the winter

The data

- Let’s open the data “mice2.xlsx”

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

	id	before	after
1	1	187.2	429.5
2	2	194.2	404.4
3	3	231.7	405.6
4	4	200.5	397.2
5	5	201.7	377.9
6	6	235.0	445.8
7	7	208.7	408.4
8	8	172.4	337.0
9	9	184.6	414.3
10	10	189.6	380.3

- You are made aware that a new shopping centre with a large food court is moving into your town
- You are concerned, as your favourite species, the **brown rat**, live in your town, and you are afraid that the access to food court garbage will negatively affect their diet, and weight.
- You set out to quickly collect weight data before the shopping food court is built to get some baseline data!

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

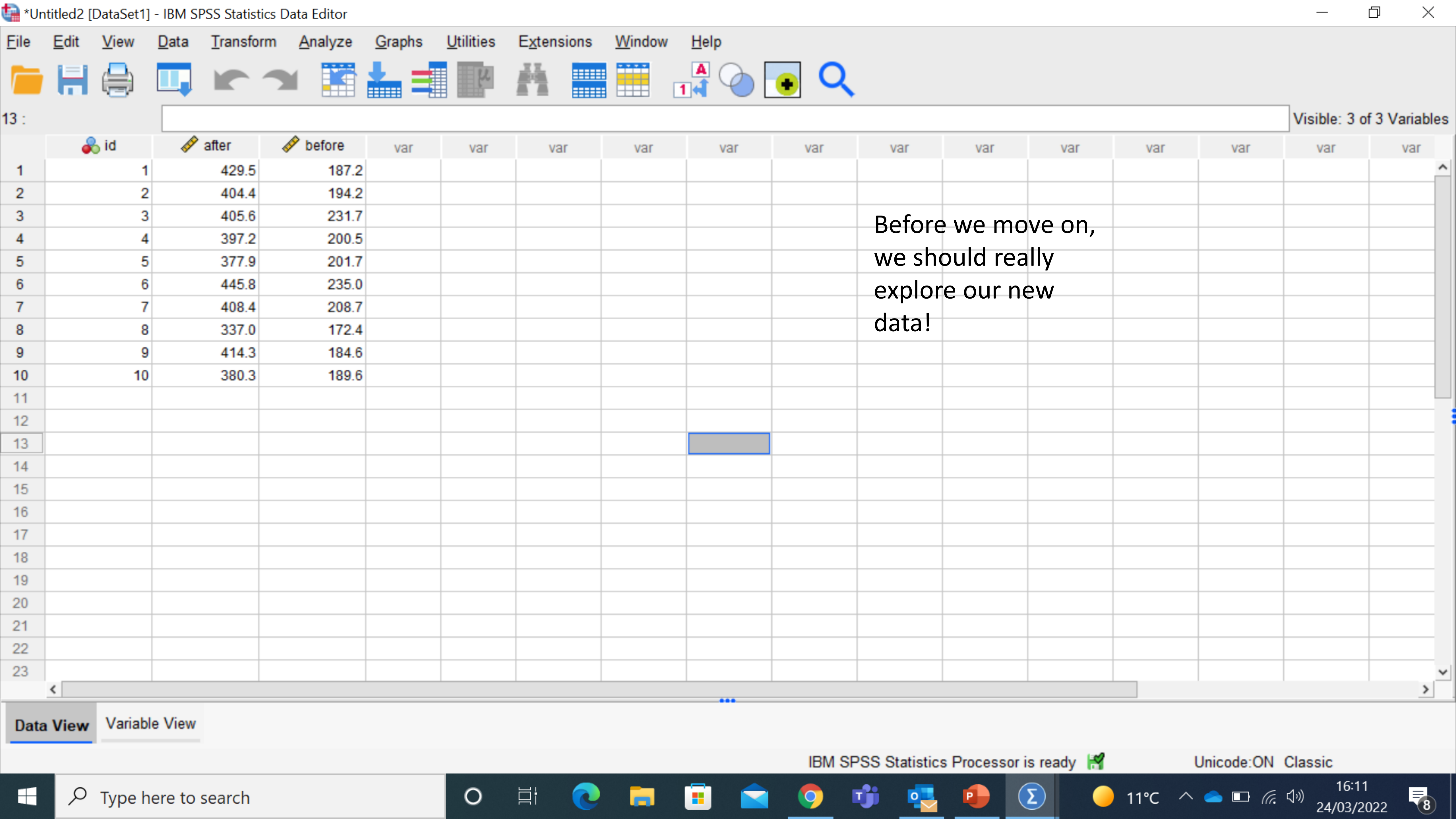
Visible: 3 of 3 Variables

	id	before	after
1	1	187.2	429.5
2	2	194.2	404.4
3	3	231.7	405.6
4	4	200.5	397.2
5	5	201.7	377.9
6	6	235.0	445.8
7	7	208.7	408.4
8	8	172.4	337.0
9	9	184.6	414.3
10	10	189.6	380.3

- Our data represent the weights of 10 individual brown rats
- Individual IDs are indicated by the numbers (1 through 10) under the 'id' column
- The 'before' column represents weight measurements of these individuals **before the food court was operational**
- The individual rats were caught, marked, weighed, and released back into the 'wilds' of your town
- The column marked 'after' represents the weight of those **SAME RATS**, who were recaptured 5 months **AFTER the food court was finished and opened**

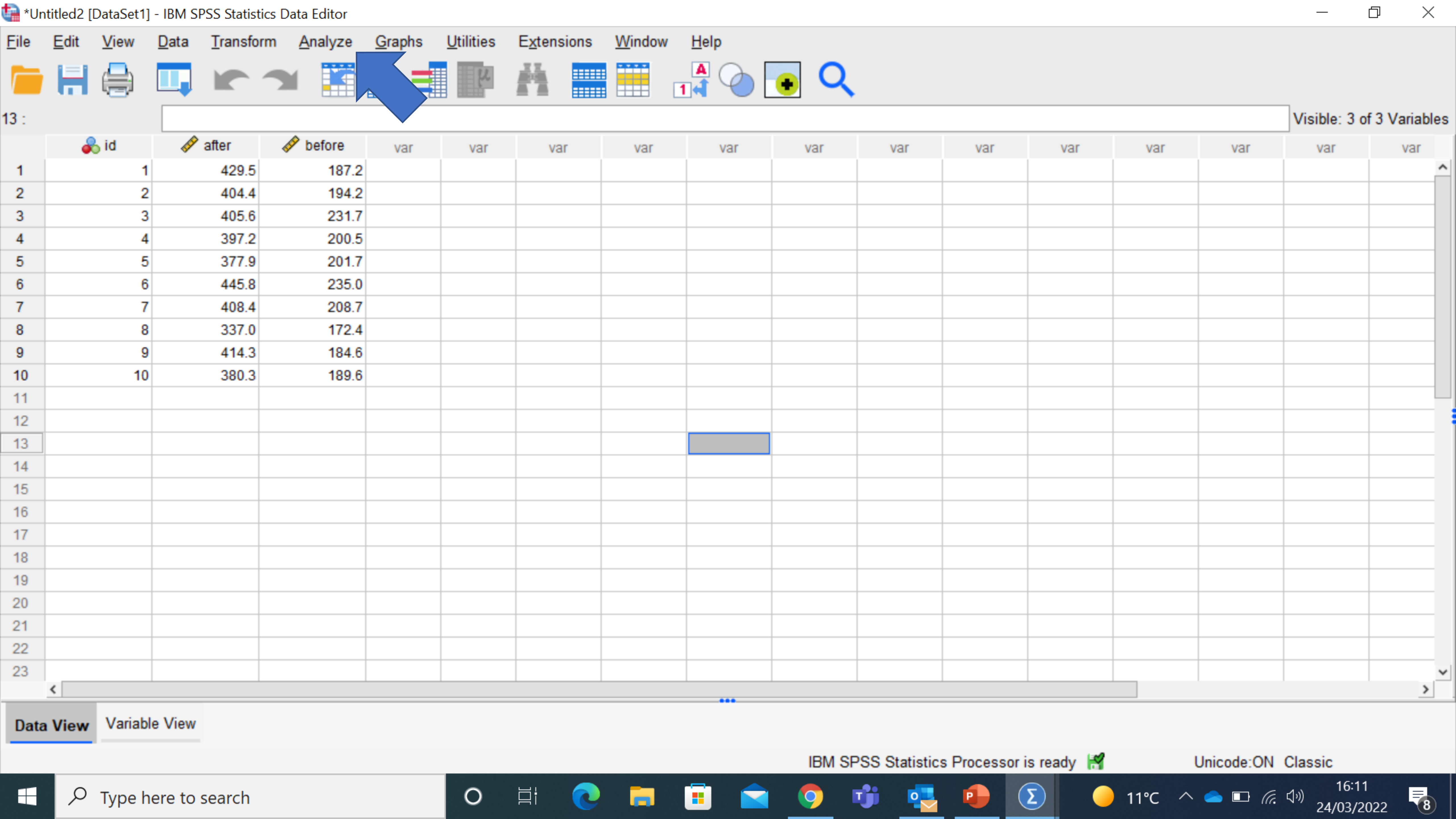
Data View Variable View

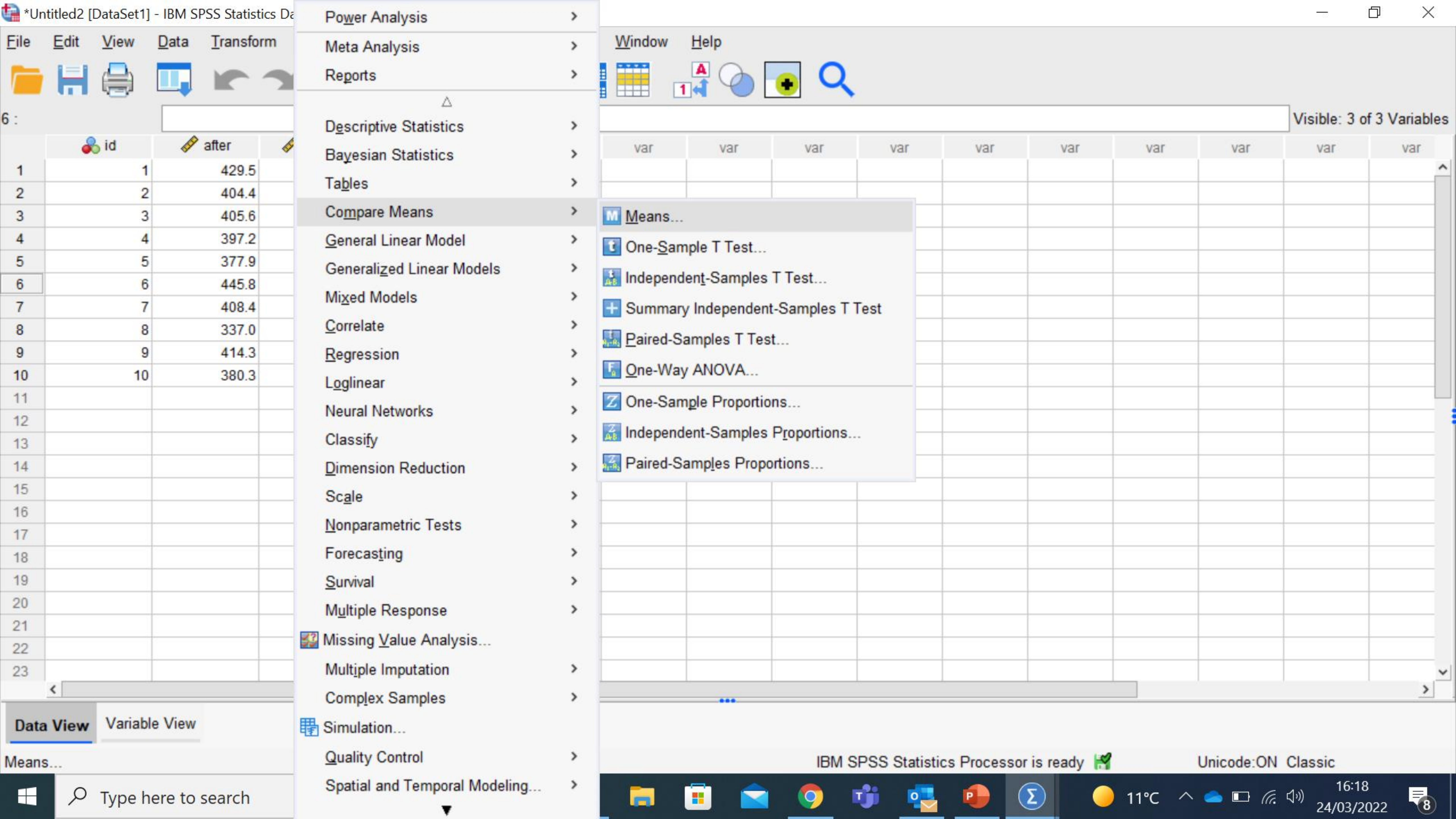
IBM SPSS Statistics Processor is ready Unicode:ON Classic

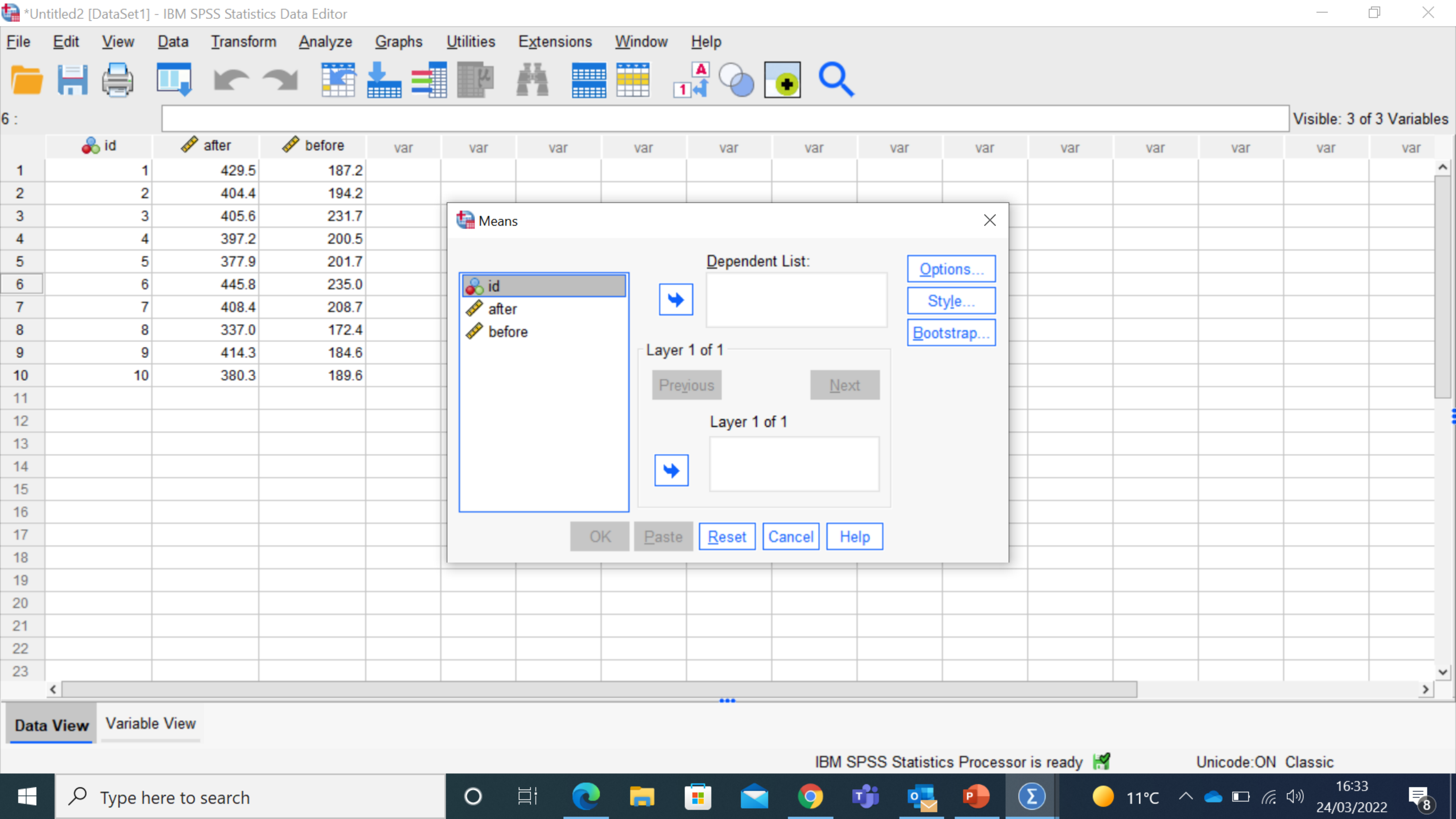


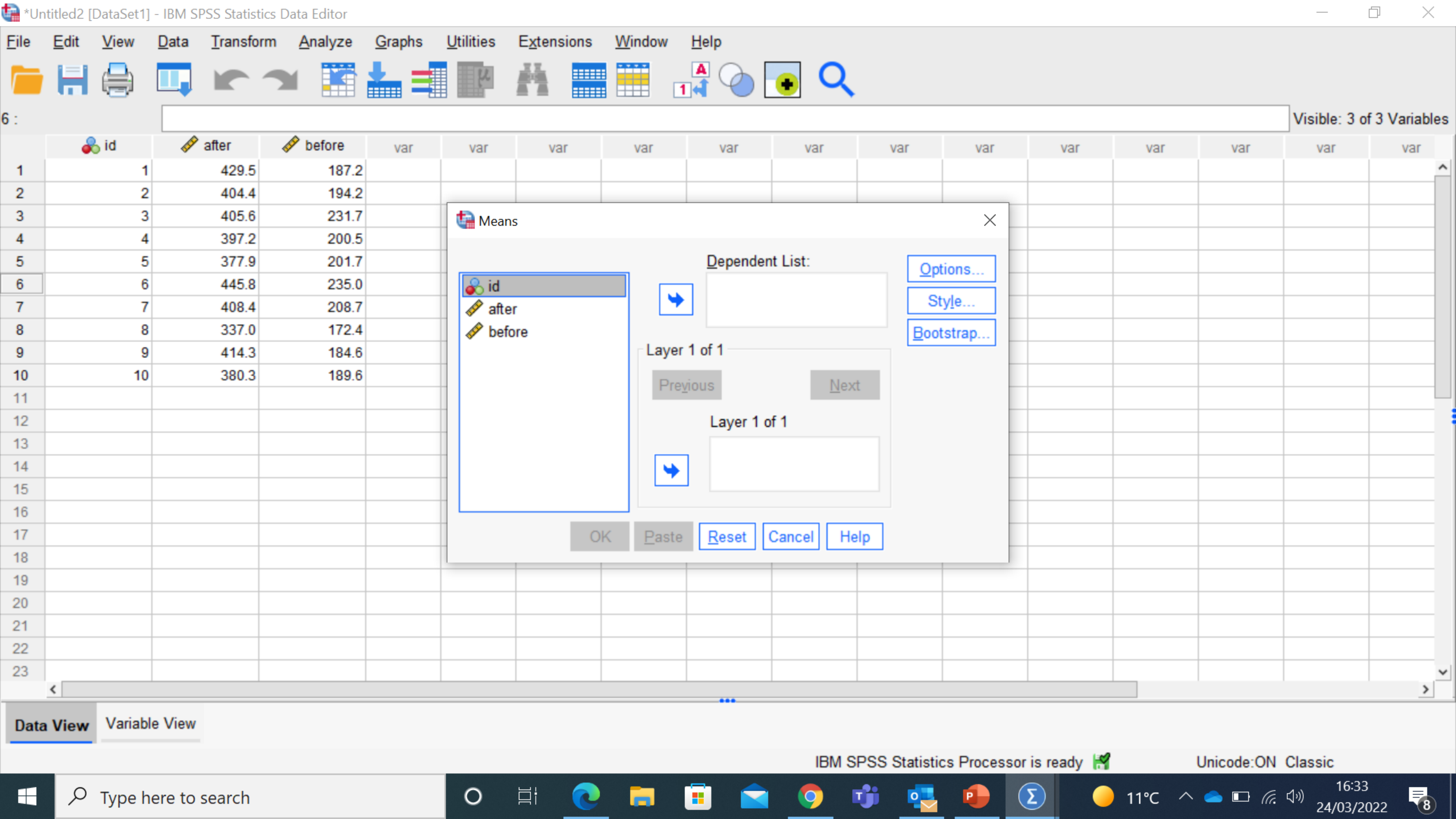
	id	after	before
1	1	429.5	187.2
2	2	404.4	194.2
3	3	405.6	231.7
4	4	397.2	200.5
5	5	377.9	201.7
6	6	445.8	235.0
7	7	408.4	208.7
8	8	337.0	172.4
9	9	414.3	184.6
10	10	380.3	189.6

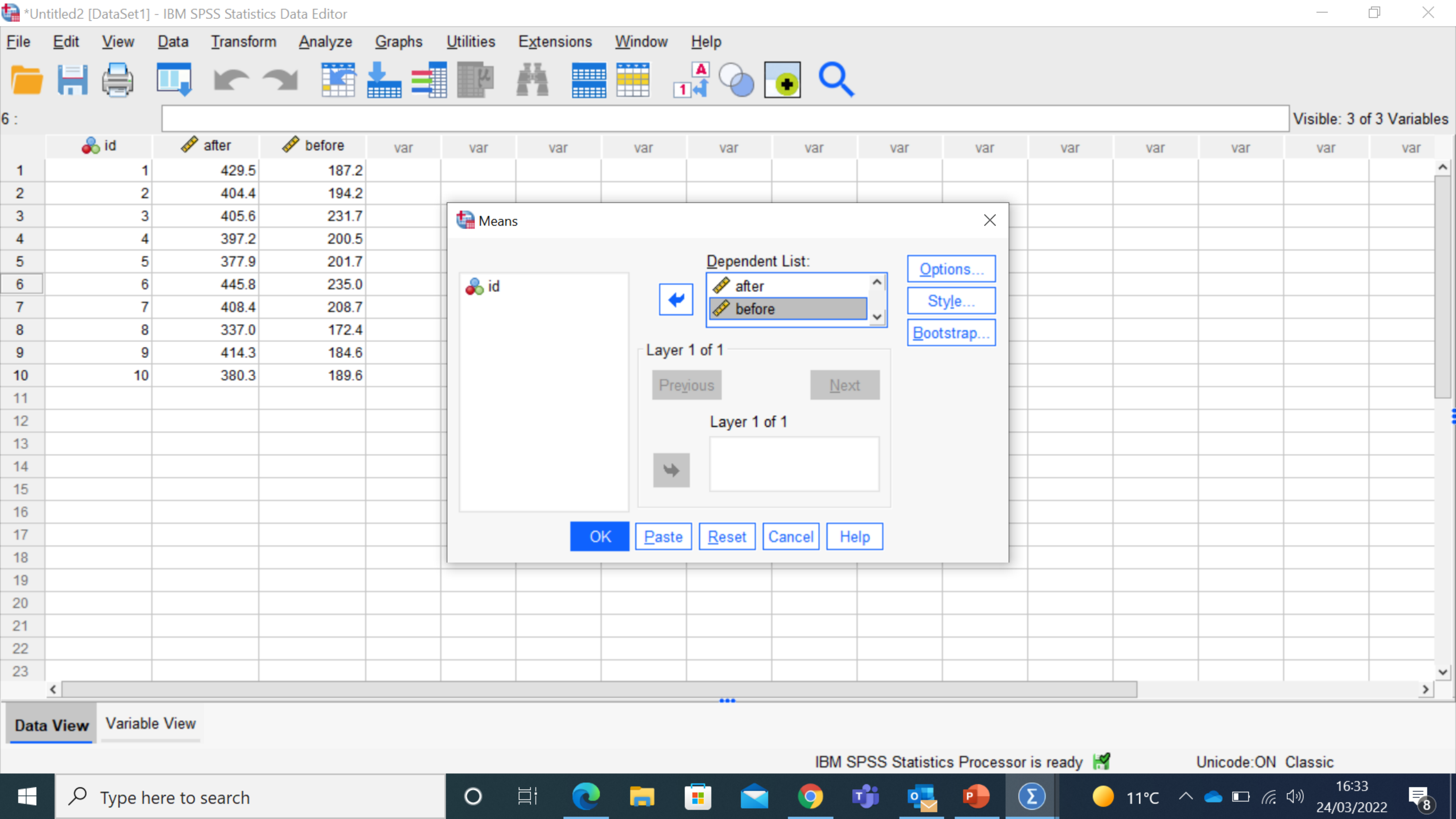
Before we move on, we should really explore our new data!

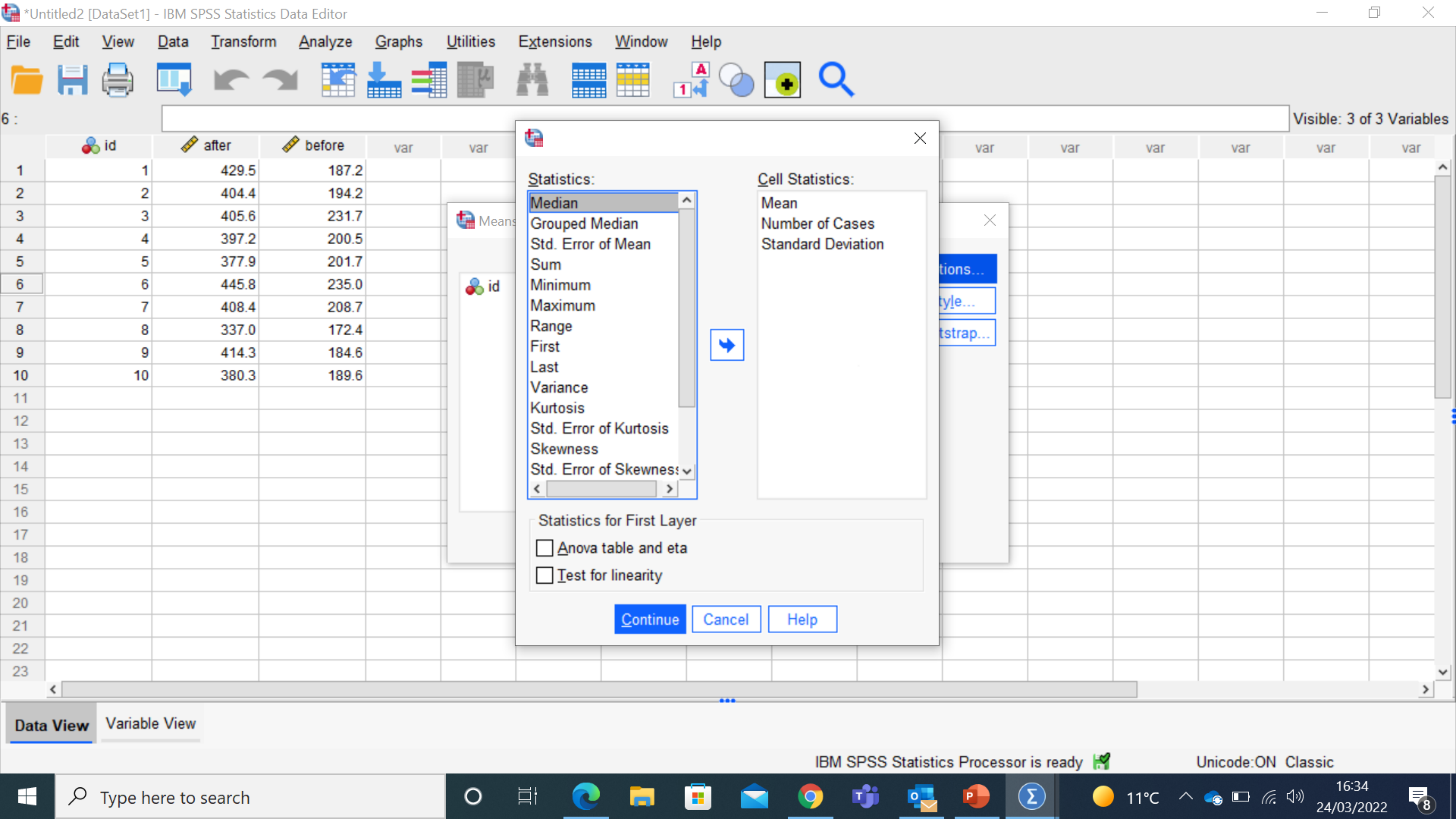

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
13 :
Visible: 3 of 3 Variables
id after before var var var var var var var var var var var var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
11
12
13
14
15
16
17
18
19
20
21
22
23
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:11
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
6 :
Visible: 3 of 3 Variables
id
after
var
1 1 429.5
2 2 404.4
3 3 405.6
4 4 397.2
5 5 377.9
6 6 445.8
7 7 408.4
8 8 337.0
9 9 414.3
10 10 380.3
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
Data View
Variable View
Means...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:18
24/03/2022

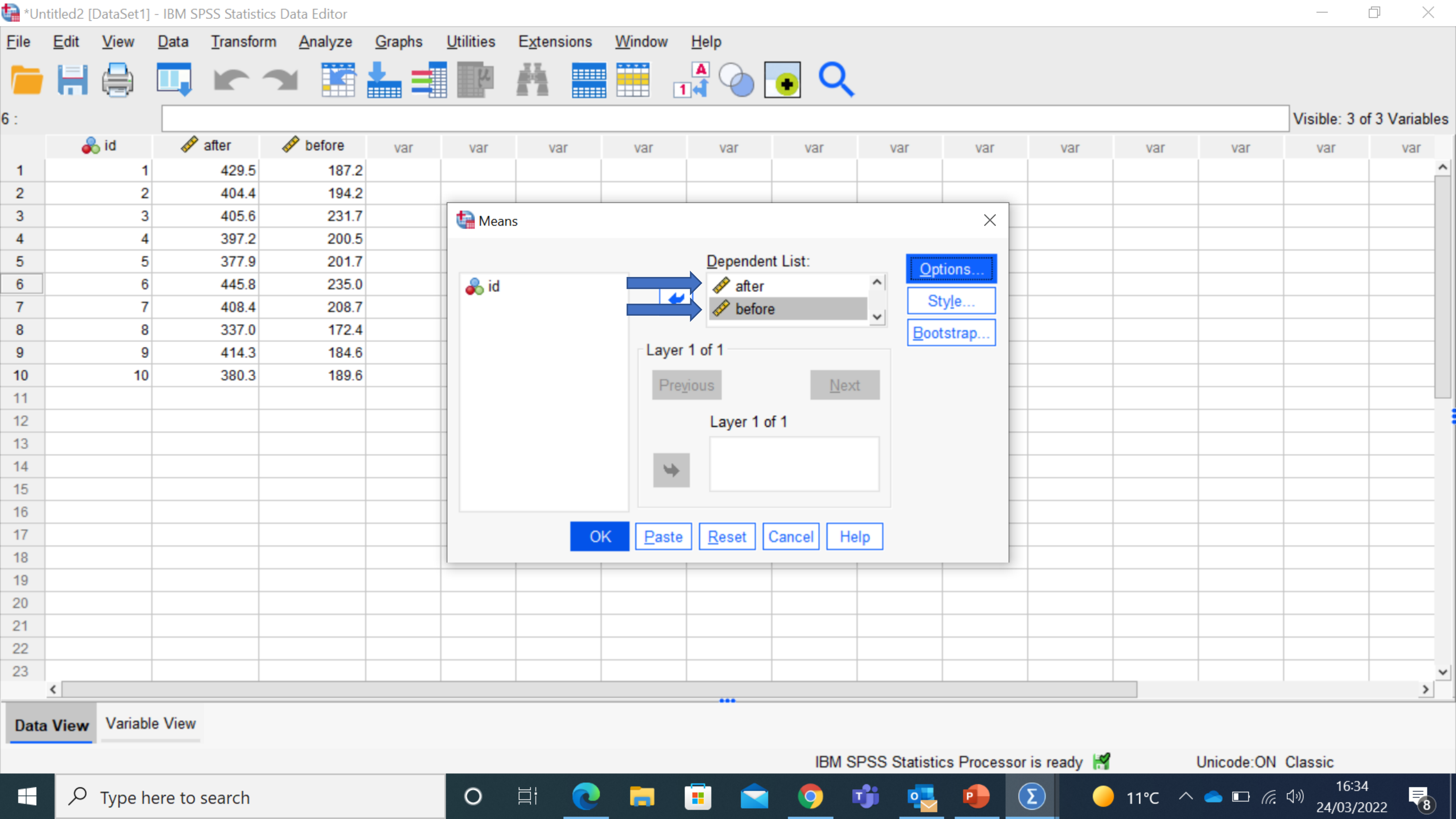

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
6 :
Visible: 3 of 3 Variables
id after before
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
after
before
Dependent List:
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:33
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
6 :
Visible: 3 of 3 Variables
id after before var var var var var var var var var var var var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
after
before
Dependent List:
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:33
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
6 :
Visible: 3 of 3 Variables
id after before var var var var var var var var var var var var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
Dependent List:
after
before
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:33
24/03/2022



*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
6 :
Visible: 3 of 3 Variables
id after before var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
Statistics:
Median
Grouped Median
Std. Error of Mean
Sum
Minimum
Maximum
Range
First
Last
Variance
Kurtosis
Std. Error of Kurtosis
Skewness
Std. Error of Skewness
Cell Statistics:
Mean
Number of Cases
Standard Deviation
Statistics for First Layer
Anova table and eta
Test for linearity
Continue
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:34
24/03/2022



*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id after before var var var var var var var var var var var var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
Dependent List:
after
before
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:34
24/03/2022

Data structure – Wide format

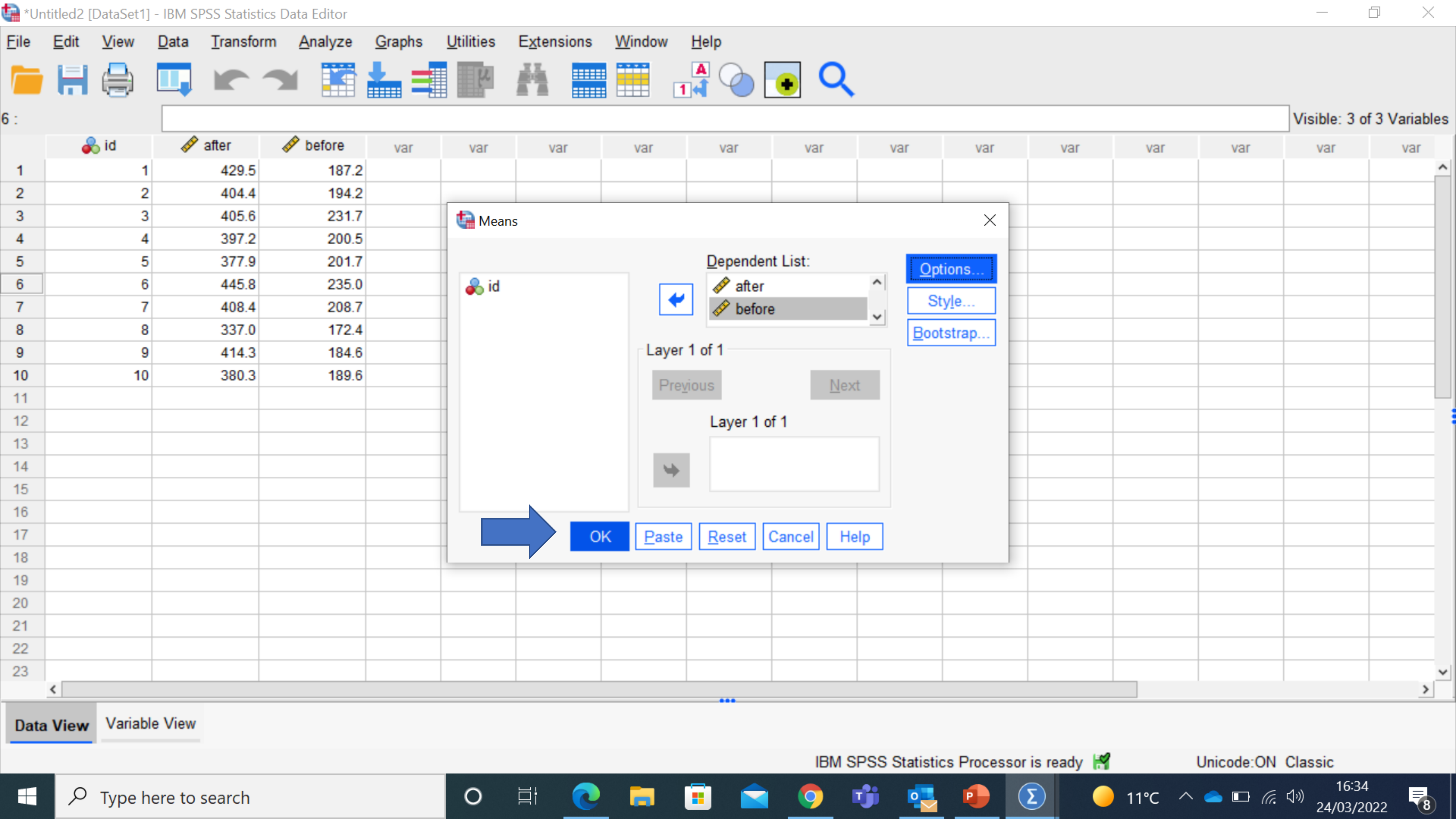
- You will notice that last time we looked at our average values this way, we only had to move one variable over… Why is there 2 this time?
- This is because our data is in a different format this time:
- Our current data is in **wide format:**
  - **Each row of data** represents a single individual
  - The treatment (the independent variable) are now colun headers!

	id	after	before
1	1	429.5	187.2
2	2	404.4	194.2
3	3	405.6	231.7
4	4	397.2	200.5
5	5	377.9	201.7
6	6	445.8	235.0
7	7	408.4	208.7
8	8	337.0	172.4
9	9	414.3	184.6
10	10	380.3	189.6
11			
12			

Data structure – Long format

- You will notice that last time we looked at our average values this way, we only had to move one variable over… Why is there 2 this time?
- This is because our data is in a different format this time:
- Our current data is in wide format: **Each row of data** represents a <u>single individual</u>
- <u>Our independent variable is listed for each row</u>

	id	Time	Weight
1	1	before	187.2
2	1	after	429.5
3	2	before	194.2
4	2	after	404.4
5	3	before	231.7
6	3	after	405.6
7	4	before	200.5
8	4	after	397.2
9	5	before	201.7
10	5	after	377.9
11	6	before	235.0
12	6	after	445.8
13	7	before	208.7
14	7	after	408.4
15	8	before	172.4
16	8	after	337.0
17	9	before	184.6
18	9	after	414.3
19	10	before	189.6
20	10	after	380.3


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
6 :
Visible: 3 of 3 Variables
id after before var var var var var var var var var var var var var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Means
id
Dependent List:
after
before
Options...
Style...
Bootstrap...
Layer 1 of 1
Previous
Next
Layer 1 of 1
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:34
24/03/2022

```
MEANS TABLES=after before
  /CELLS=MEAN COUNT STDDEV.
```

Means

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
after	10	100.0%	0	0.0%	10	100.0%
before	10	100.0%	0	0.0%	10	100.0%

Report

	after	before
Mean	400.040	200.560
N	10	10
Std. Deviation	30.0869	20.0282

Notice: This time around, our two categories (before and after) are in 2 columns… but they were in rows before, right?!

Again, this is because our data are in **wide format** as opposed to **long format**

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

9 :

Visible: 3 of 3 Variables

	id	after	before
1	1	429.5	187.2
2	2	404.4	194.2
3	3	405.6	231.7
4	4	397.2	200.5
5	5	377.9	201.7
6	6	445.8	235.0
7	7	408.4	208.7
8	8	337.0	172.4
9	9	414.3	184.6
10	10	380.3	189.6

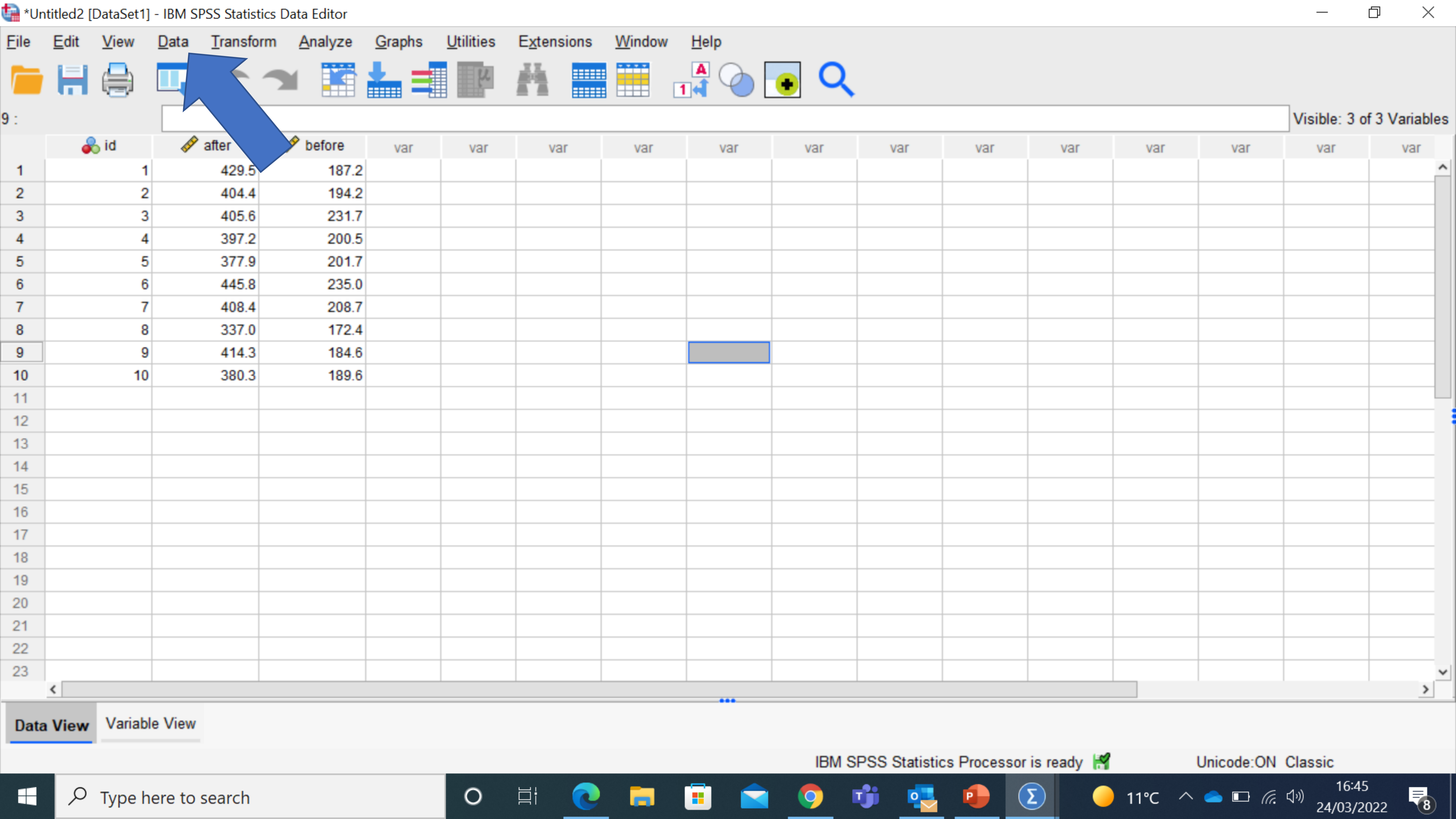
Our next course of action should be to graph our data

We could try to use the same procedure as we did before, but here is where the issue of **long** vs **wide** formatted data becomes a real problem…

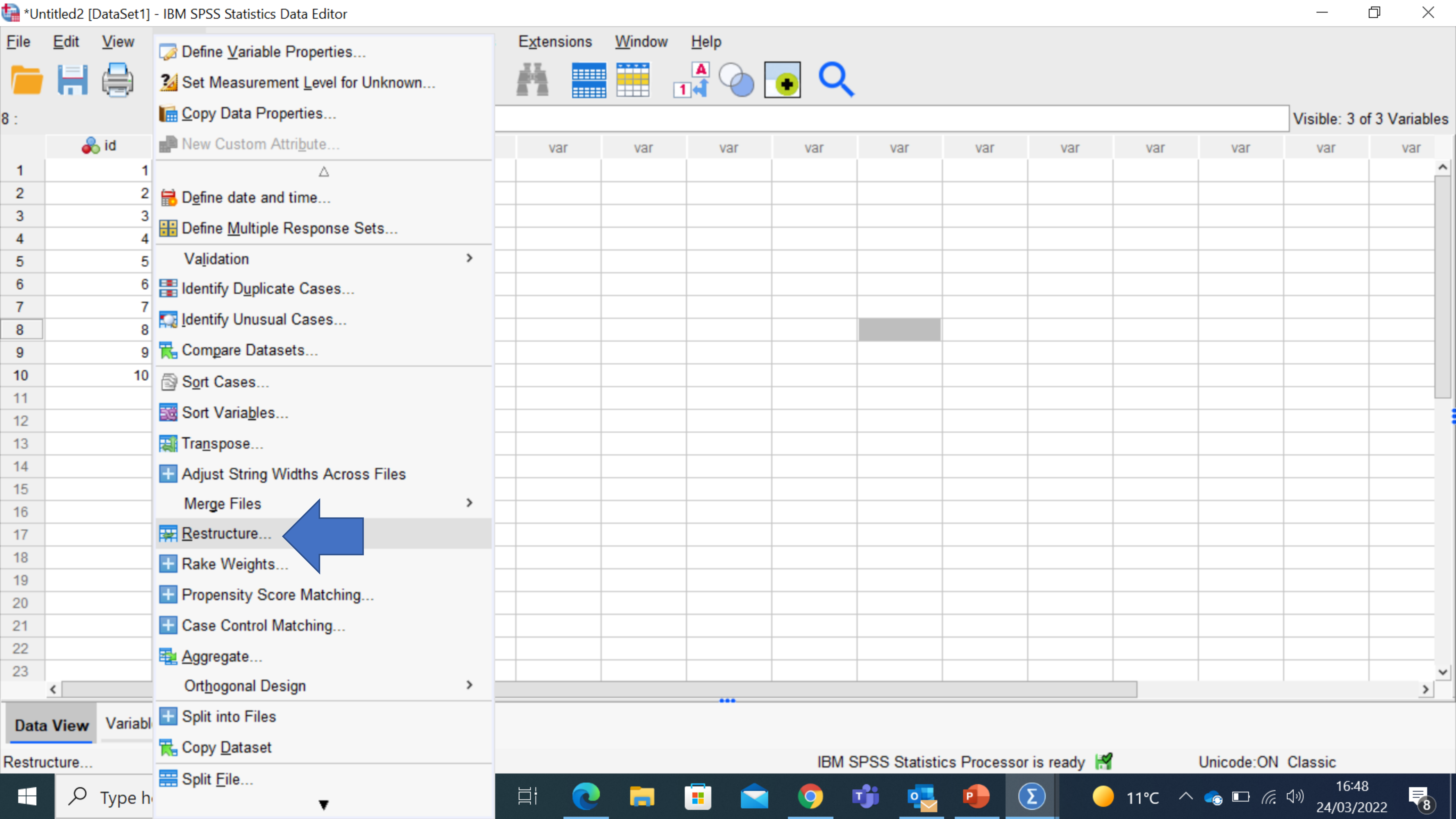
In order to make the boxplot we need, we need to covert from our current **wide format** into a **long format!**

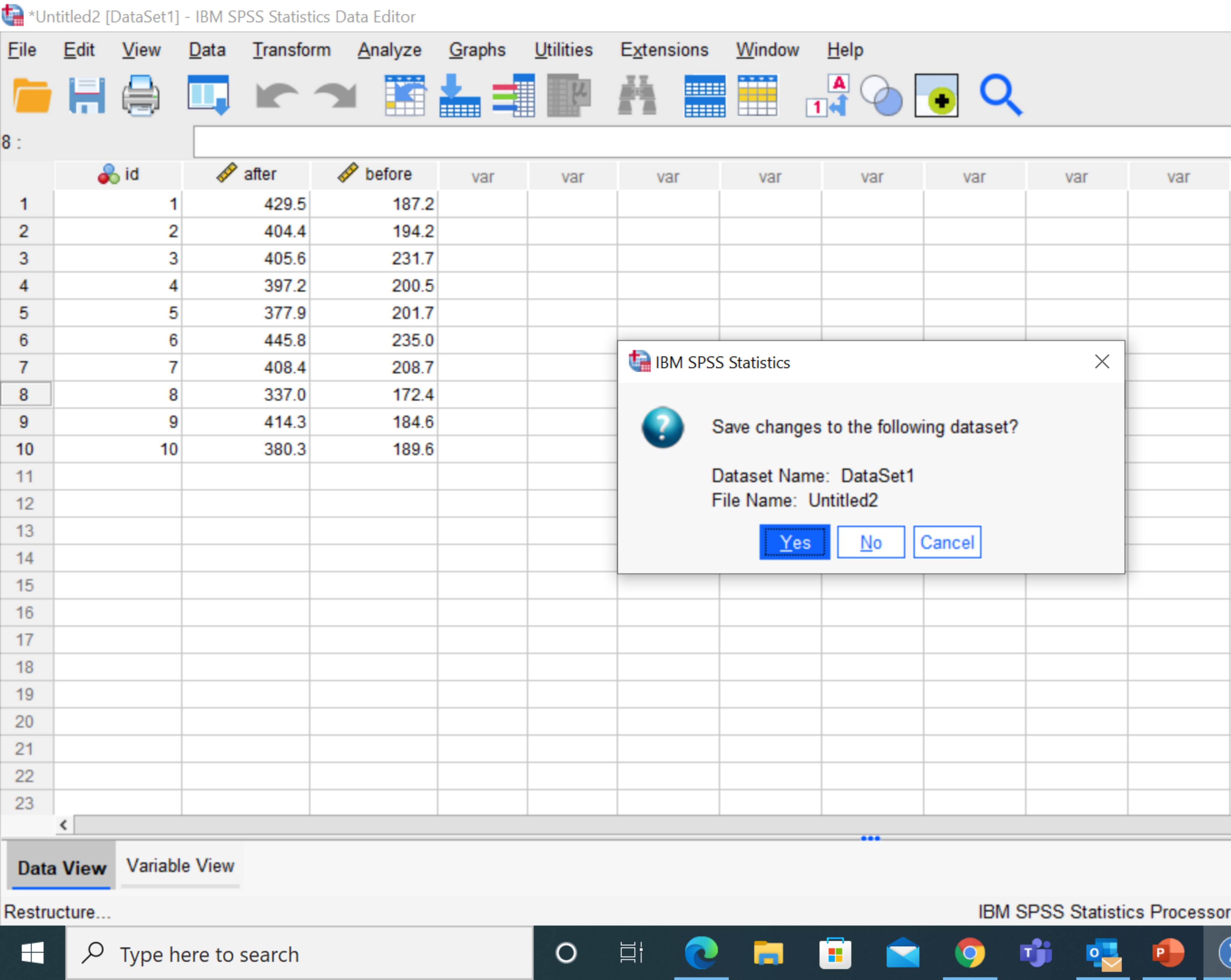
Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic



	id	after	before
1	1	429.5	187.2
2	2	404.4	194.2
3	3	405.6	231.7
4	4	397.2	200.5
5	5	377.9	201.7
6	6	445.8	235.0
7	7	408.4	208.7
8	8	337.0	172.4
9	9	414.3	184.6
10	10	380.3	189.6


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Extensions
Window
Help
Define Variable Properties...
Set Measurement Level for Unknown...
Copy Data Properties...
New Custom Attribute...
Define date and time...
Define Multiple Response Sets...
Validation
Identify Duplicate Cases...
Identify Unusual Cases...
Compare Datasets...
Sort Cases...
Sort Variables...
Transpose...
Adjust String Widths Across Files
Merge Files
Restructure...
Rake Weights...
Propensity Score Matching...
Case Control Matching...
Aggregate...
Orthogonal Design
Split into Files
Copy Dataset
Split File...
Visible: 3 of 3 Variables
id
var
Data View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
11°C
16:48
24/03/2022

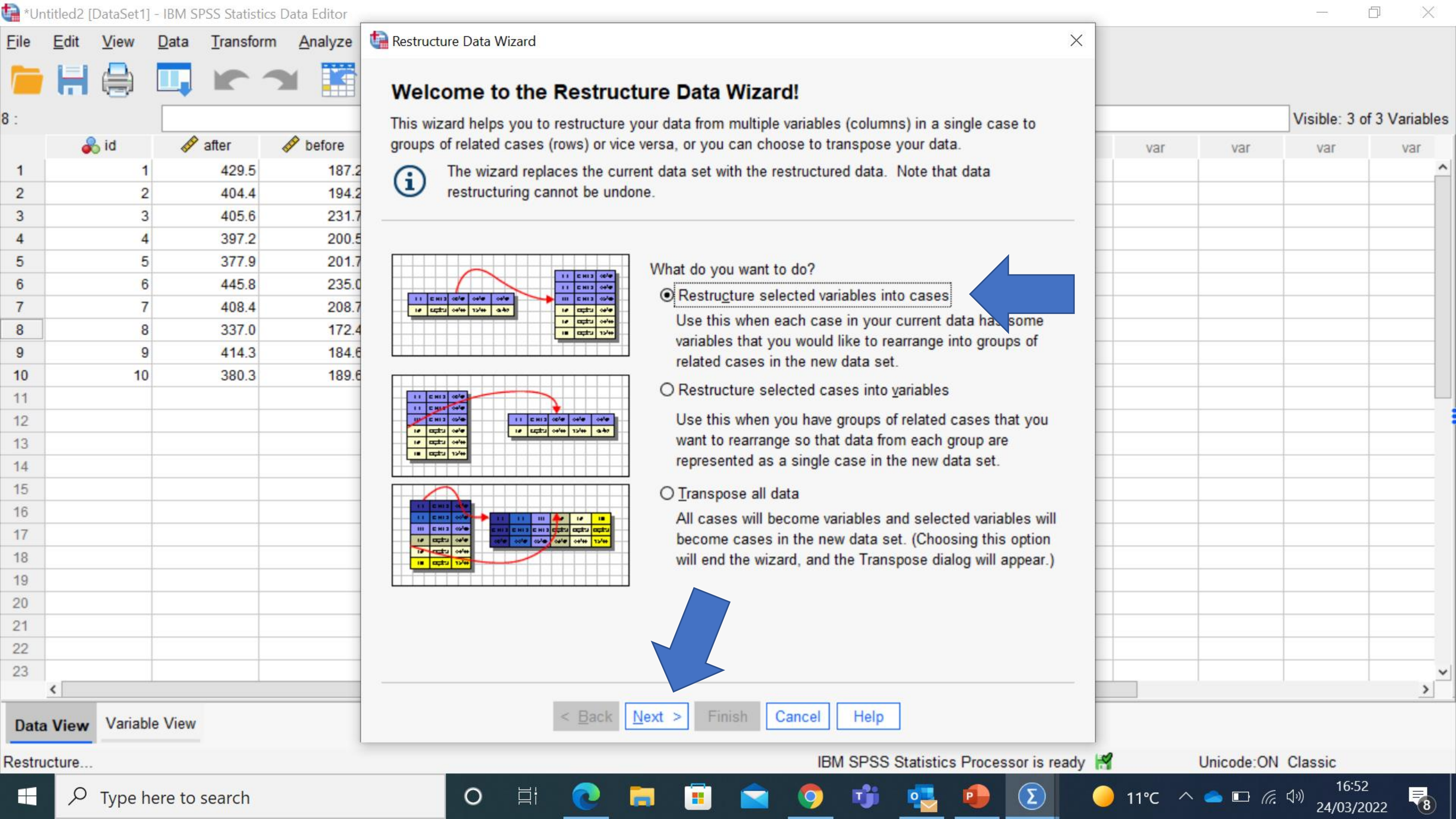


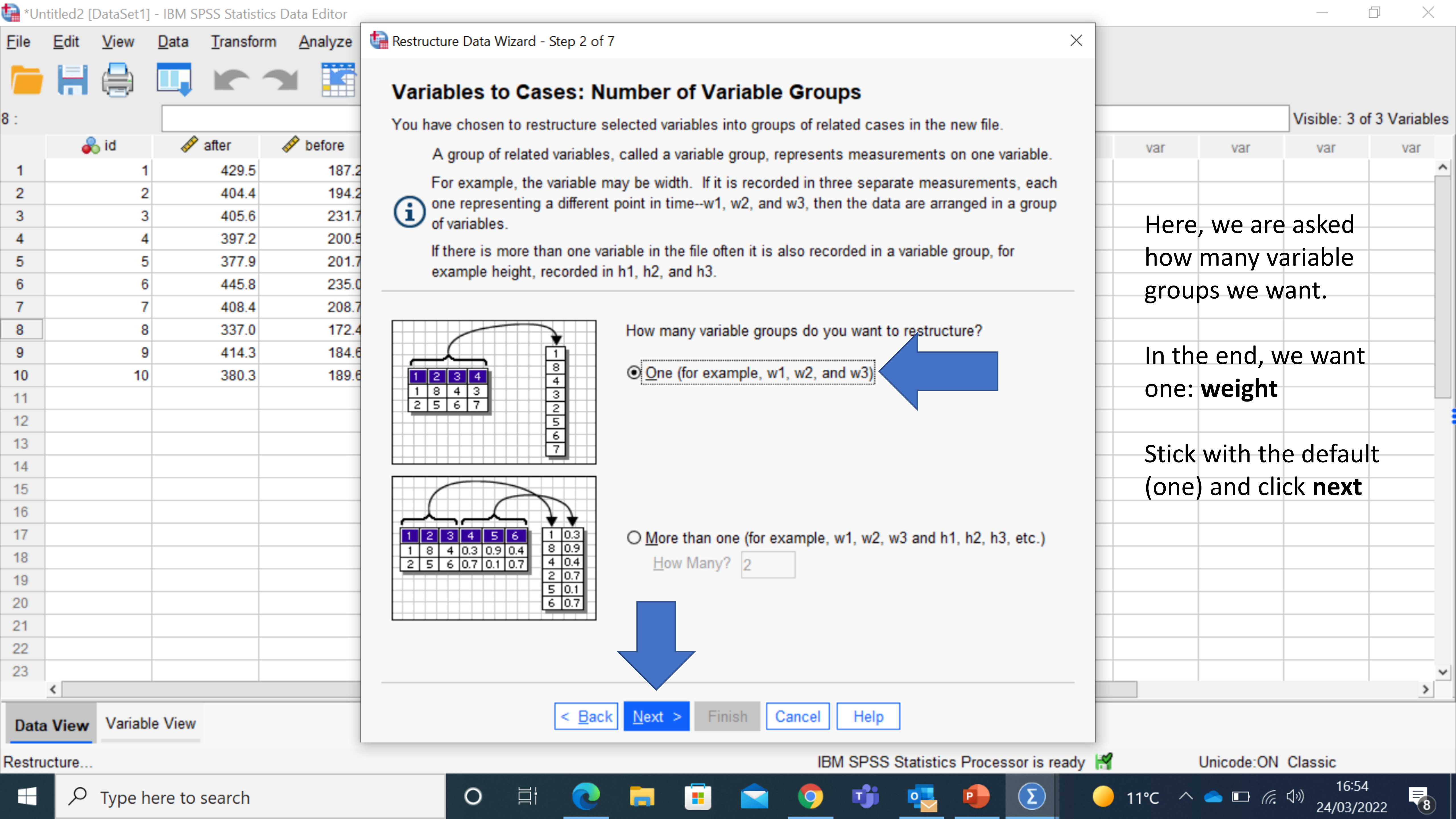
You are being asked here if you want to save your data in it's current format (wide format) before you juggle it around.

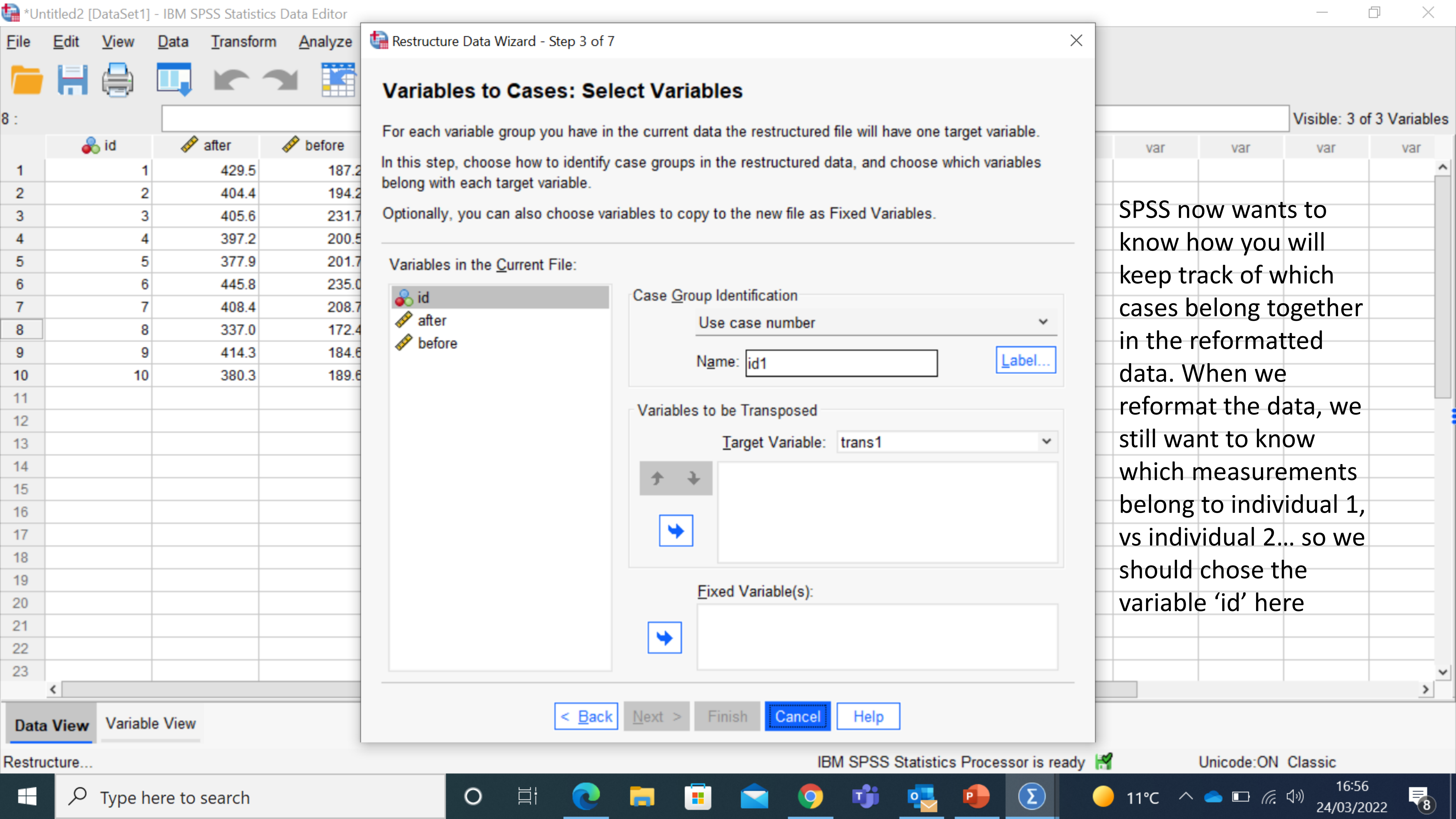
In real applications, I strongly suggest you say **YES**

Why risk it? I can be nice to have a version of your data saved in wide format, and another in long format!

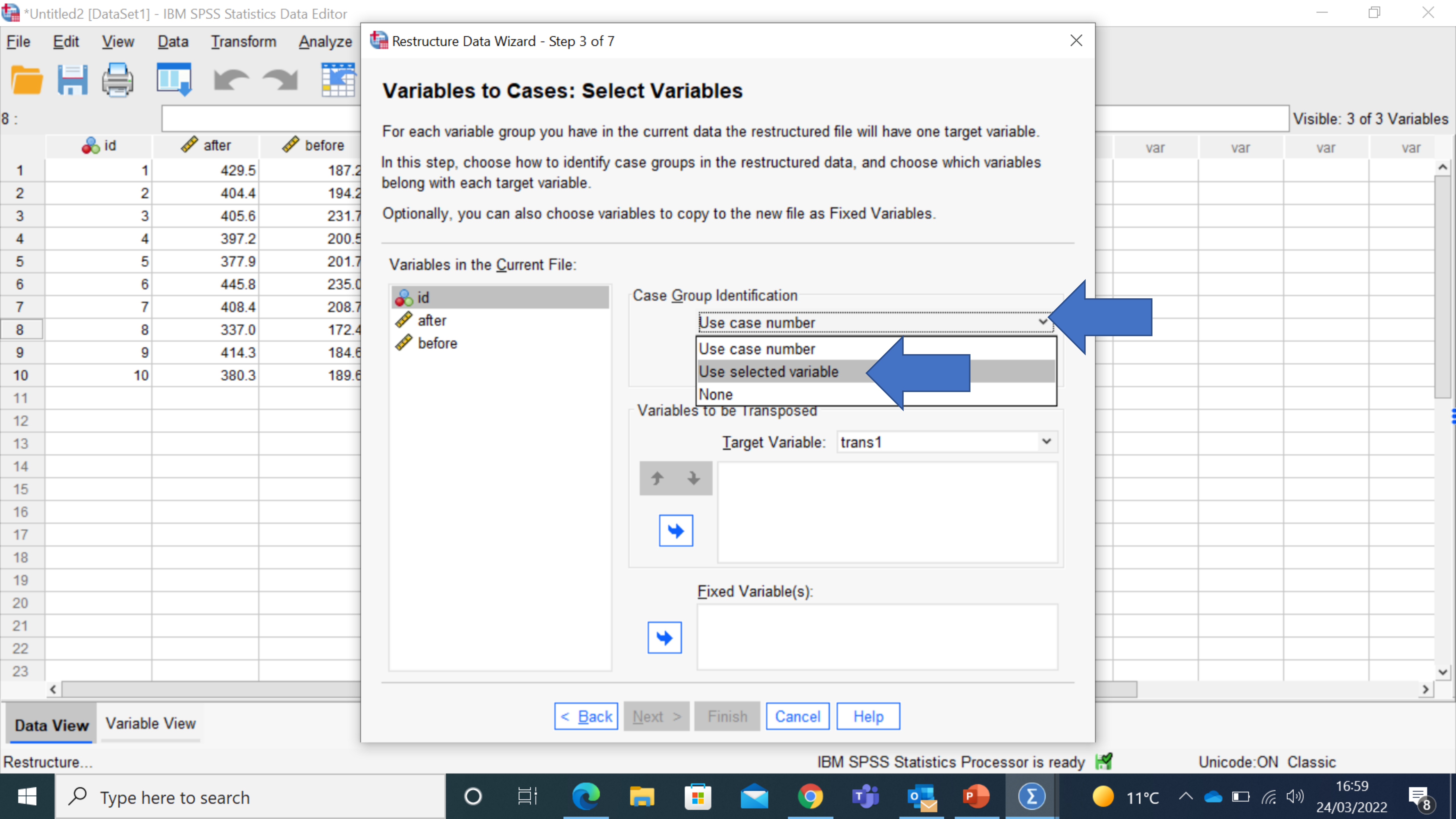
For today, I want us to experience switching back and forth, so we wont save here, but in your own practice, saving is always a clever plan!

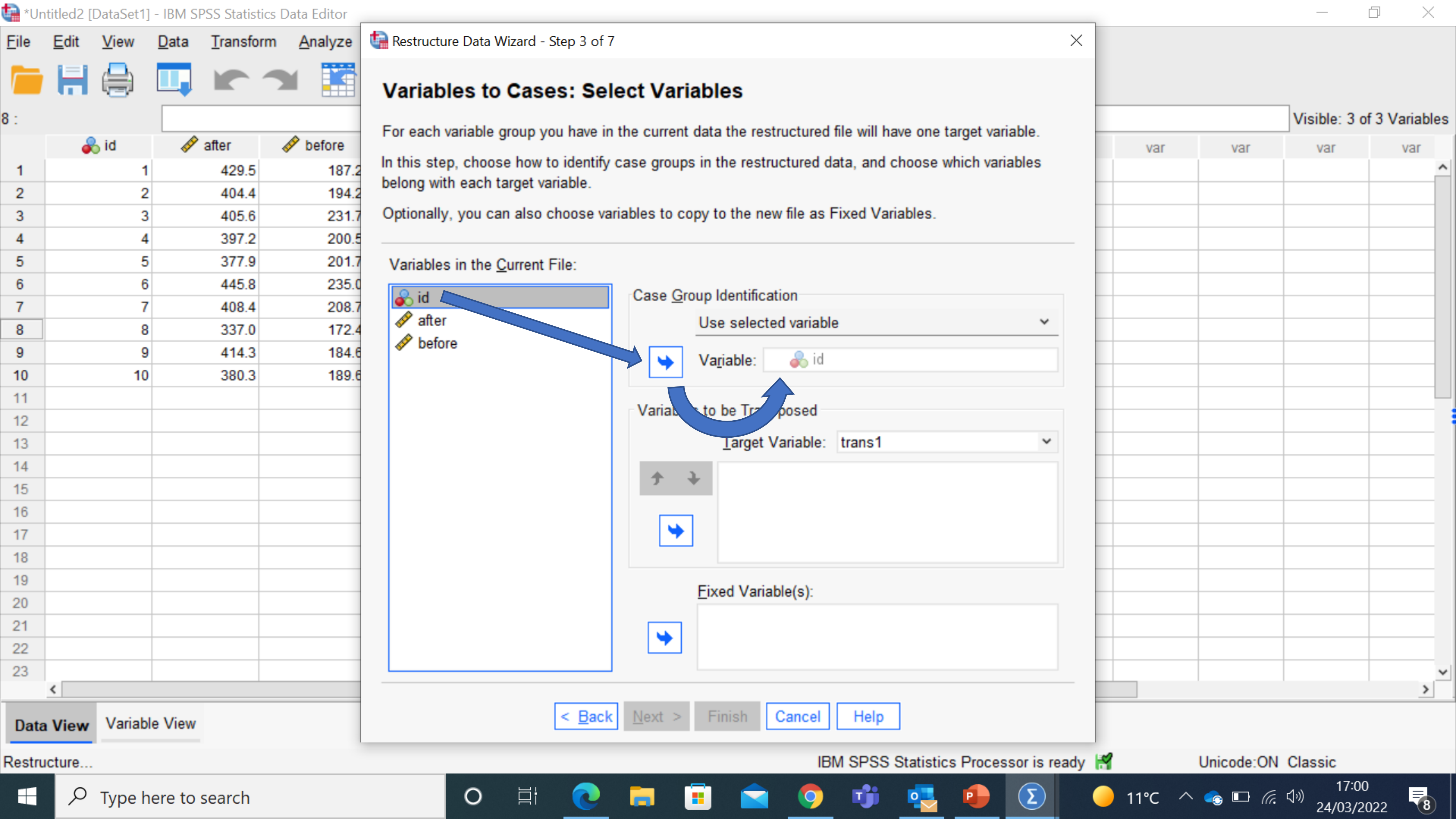

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard
Welcome to the Restructure Data Wizard!
This wizard helps you to restructure your data from multiple variables (columns) in a single case to groups of related cases (rows) or vice versa, or you can choose to transpose your data.
The wizard replaces the current data set with the restructured data. Note that data restructuring cannot be undone.
What do you want to do?
Restructure selected variables into cases
Use this when each case in your current data has some variables that you would like to rearrange into groups of related cases in the new data set.
Restructure selected cases into variables
Use this when you have groups of related cases that you want to rearrange so that data from each group are represented as a single case in the new data set.
Transpose all data
All cases will become variables and selected variables will become cases in the new data set. (Choosing this option will end the wizard, and the Transpose dialog will appear.)
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
after
before
var
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:52
24/03/2022

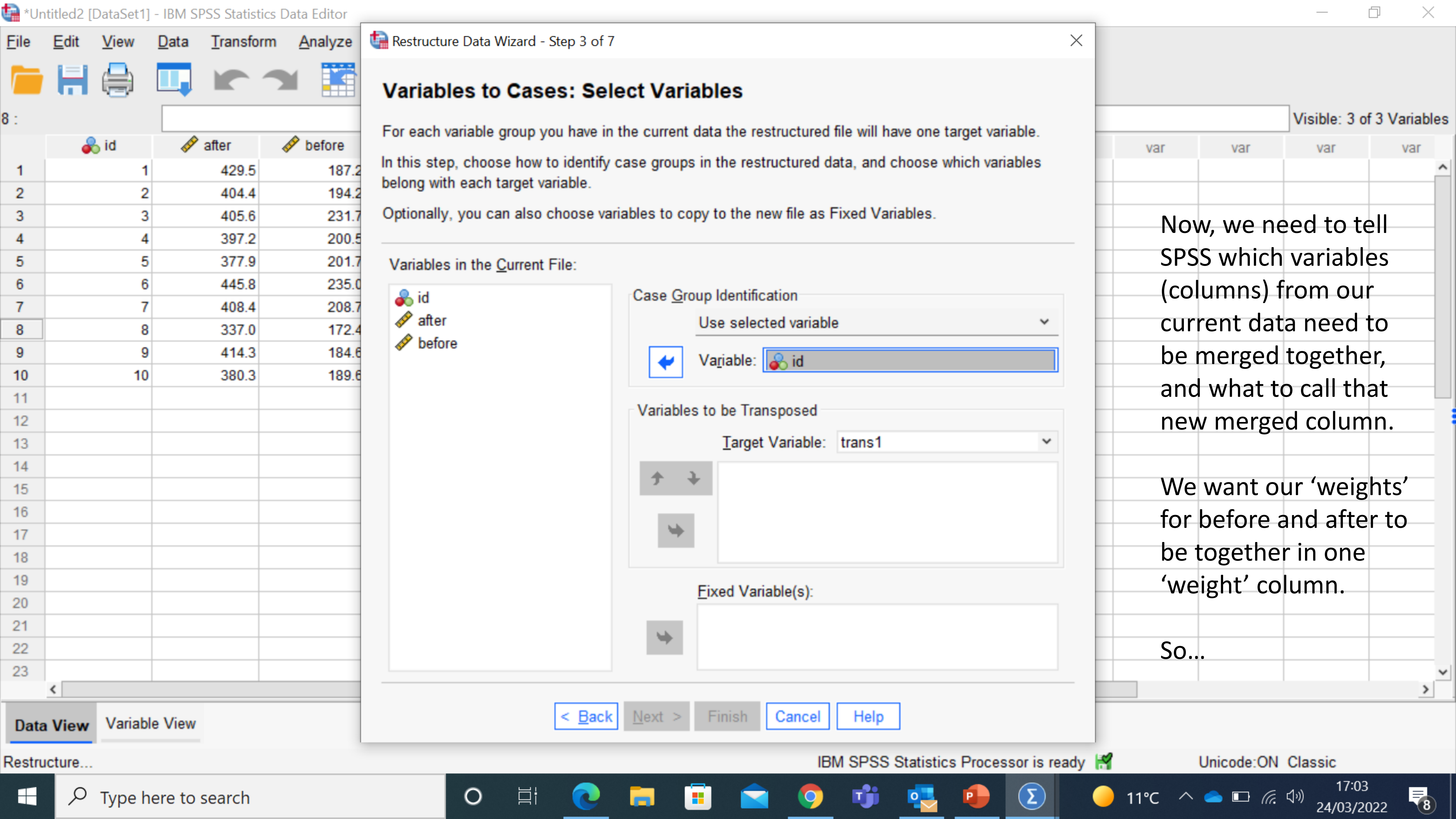

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Visible: 3 of 3 Variables
id after before
var var var var
Restructure Data Wizard - Step 2 of 7
Variables to Cases: Number of Variable Groups
You have chosen to restructure selected variables into groups of related cases in the new file.
A group of related variables, called a variable group, represents measurements on one variable.
For example, the variable may be width. If it is recorded in three separate measurements, each one representing a different point in time--w1, w2, and w3, then the data are arranged in a group of variables.
If there is more than one variable in the file often it is also recorded in a variable group, for example height, recorded in h1, h2, and h3.
How many variable groups do you want to restructure?
One (for example, w1, w2, and w3)
More than one (for example, w1, w2, w3 and h1, h2, h3, etc.)
How Many? 2
< Back
Next >
Finish
Cancel
Help
Here, we are asked how many variable groups we want.
In the end, we want one: **weight**
Stick with the default (one) and click **next**
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:54
24/03/2022



SPSS now wants to know how you will keep track of which cases belong together in the reformatted data. When we reformat the data, we still want to know which measurements belong to individual 1, vs individual 2… so we should chose the variable 'id' here


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard - Step 3 of 7
Variables to Cases: Select Variables
For each variable group you have in the current data the restructured file will have one target variable.
In this step, choose how to identify case groups in the restructured data, and choose which variables belong with each target variable.
Optionally, you can also choose variables to copy to the new file as Fixed Variables.
Variables in the Current File:
id
after
before
Case Group Identification
Use case number
Use case number
Use selected variable
None
Variables to be Transposed
Target Variable:
trans1
Fixed Variable(s):
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
after
before
var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
16:59
24/03/2022

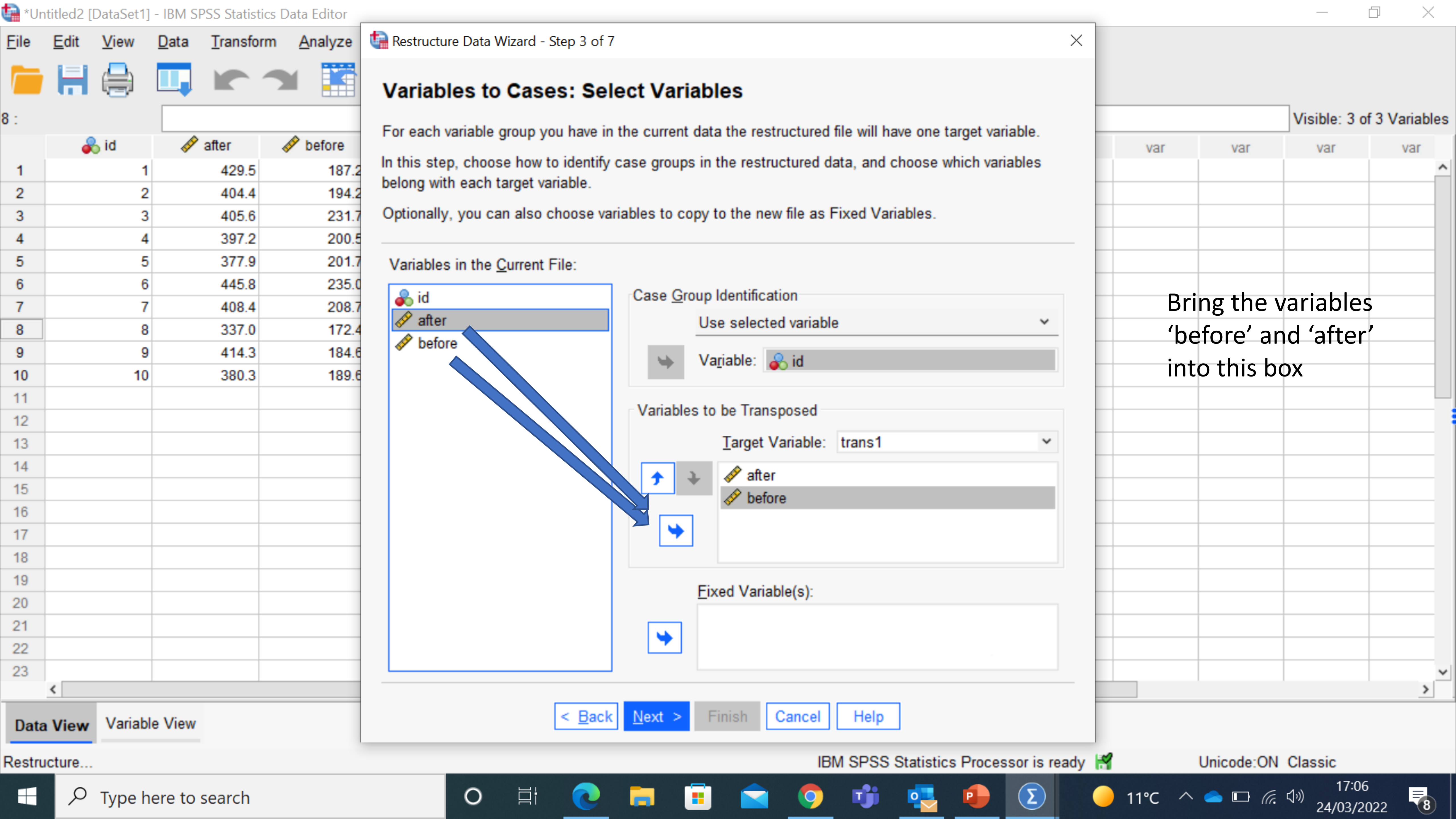

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard - Step 3 of 7
Variables to Cases: Select Variables
For each variable group you have in the current data the restructured file will have one target variable.
In this step, choose how to identify case groups in the restructured data, and choose which variables belong with each target variable.
Optionally, you can also choose variables to copy to the new file as Fixed Variables.
Variables in the Current File:
id
after
before
Case Group Identification
Use selected variable
Variable:
id
Variables to be Transposed
Target Variable:
trans1
Fixed Variable(s):
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
after
before
var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
17:00
24/03/2022



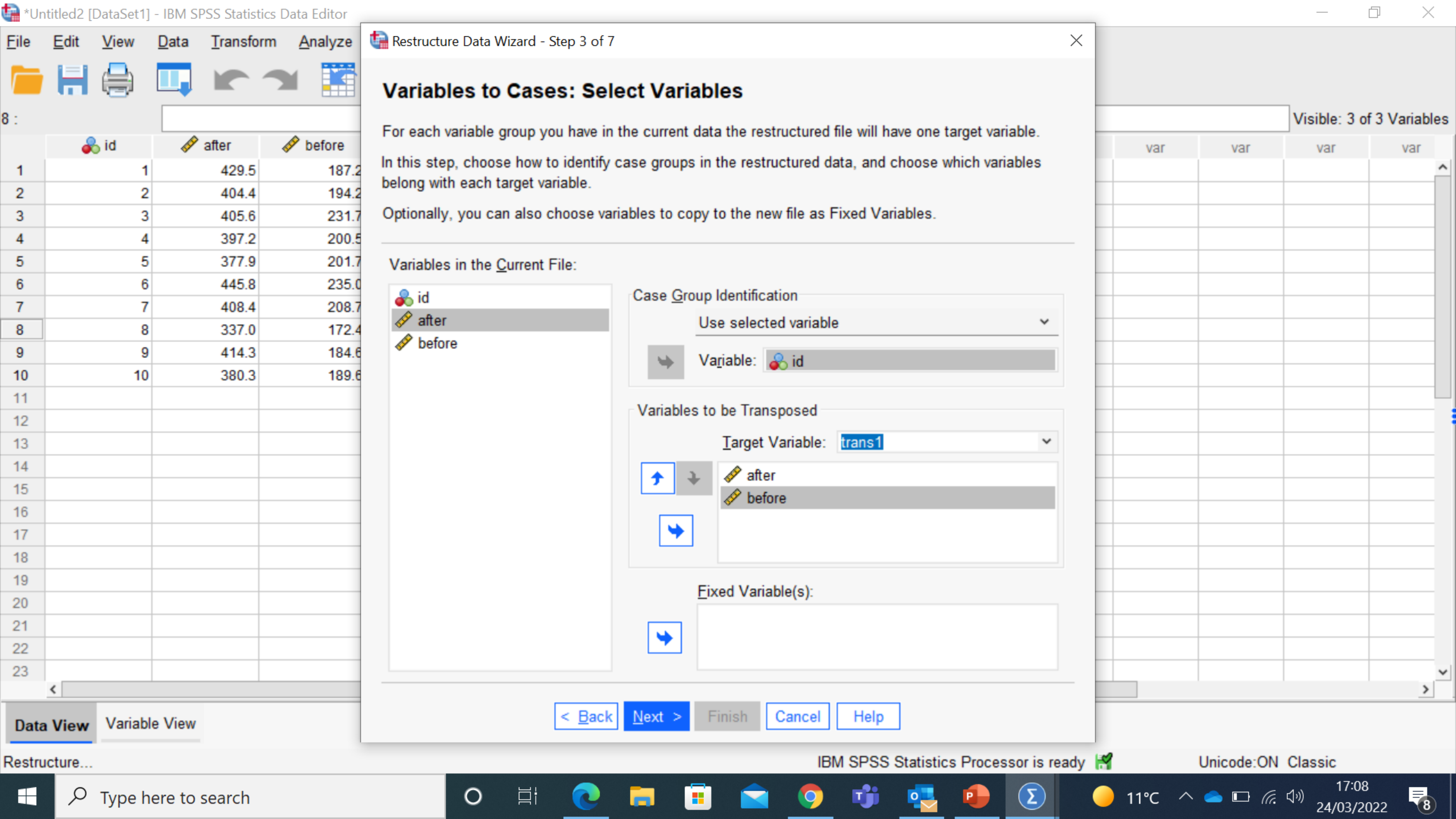
Now, we need to tell SPSS which variables (columns) from our current data need to be merged together, and what to call that new merged column.

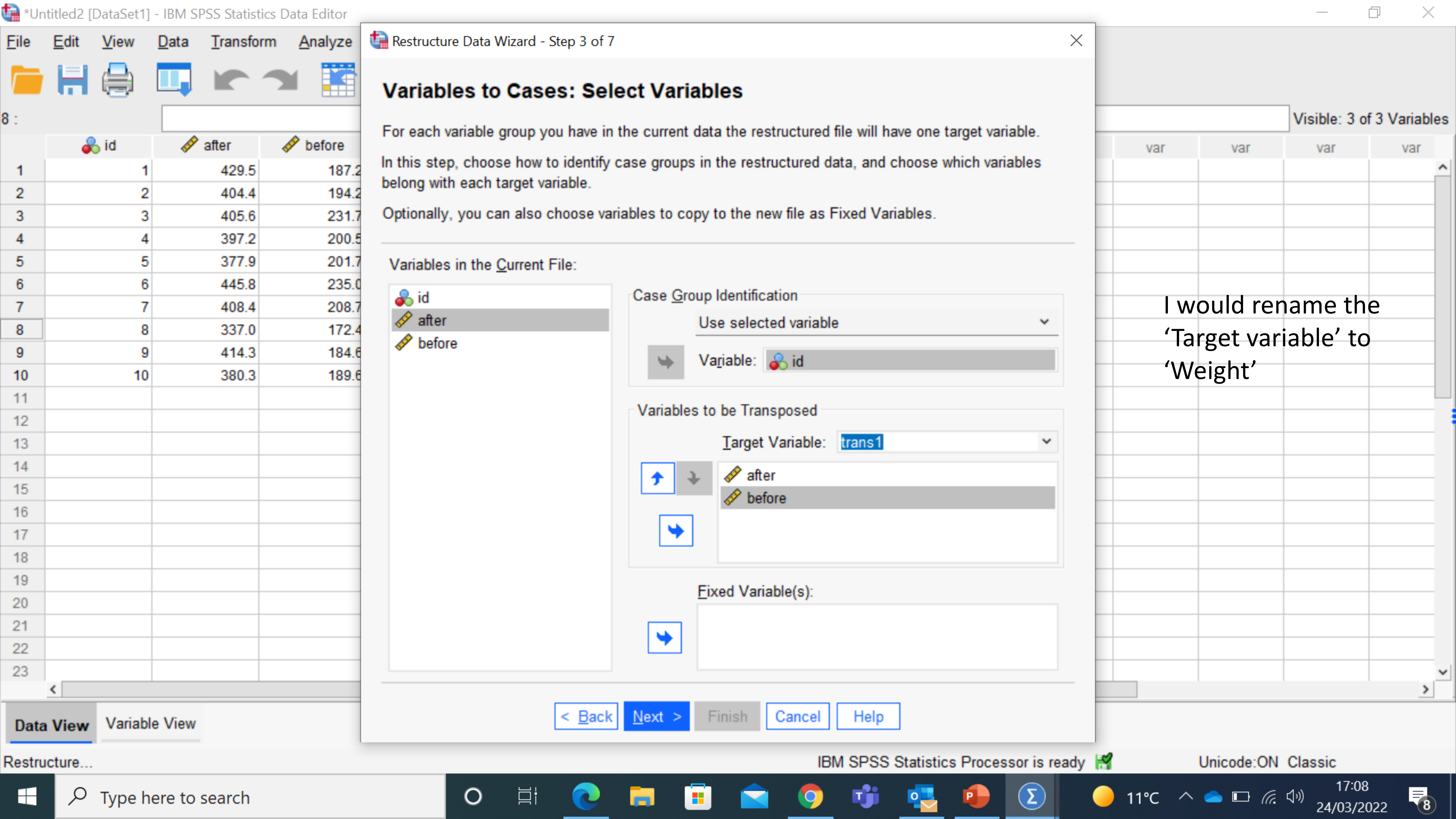
We want our ‘weights’ for before and after to be together in one ‘weight’ column.

So…

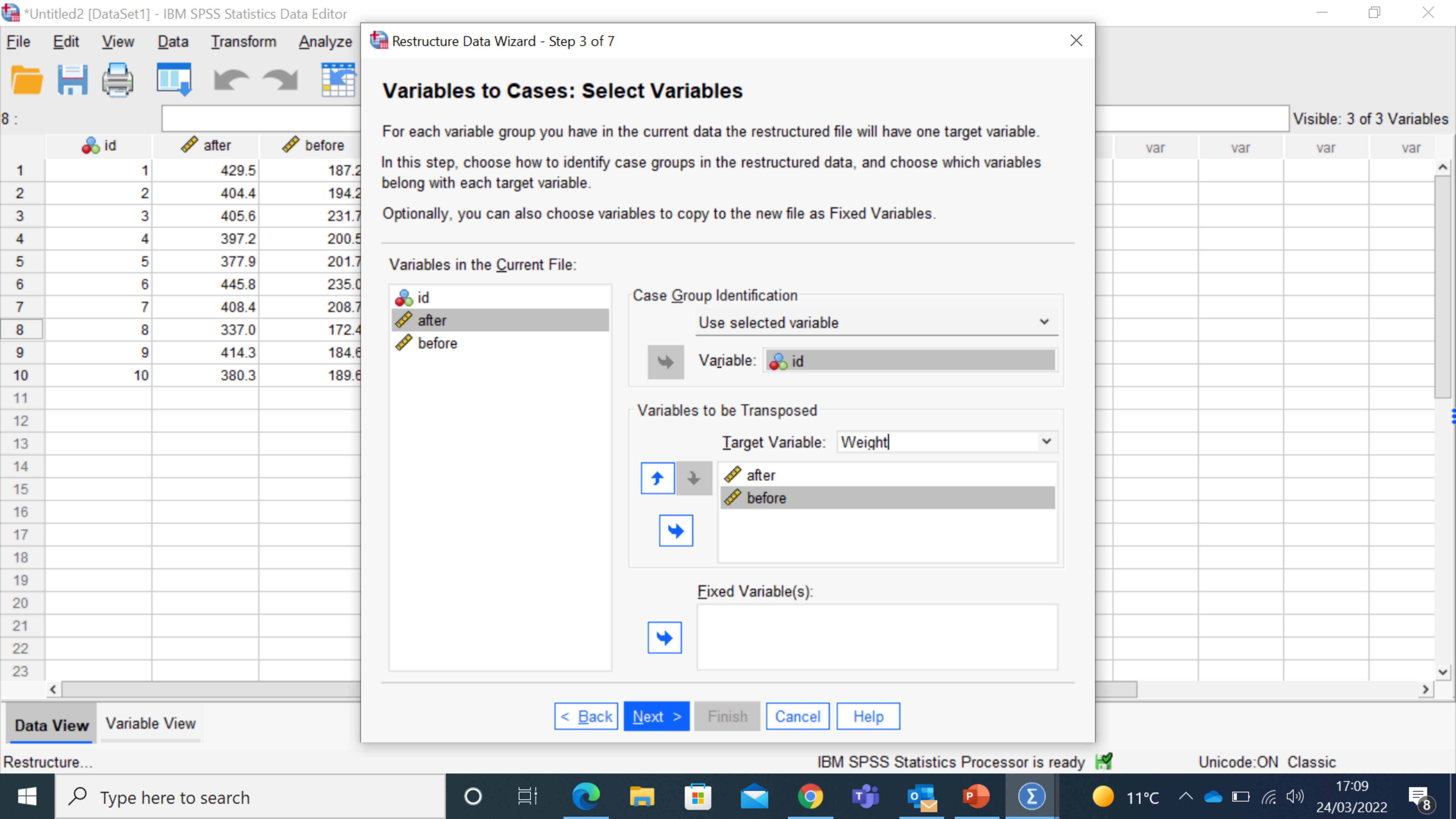


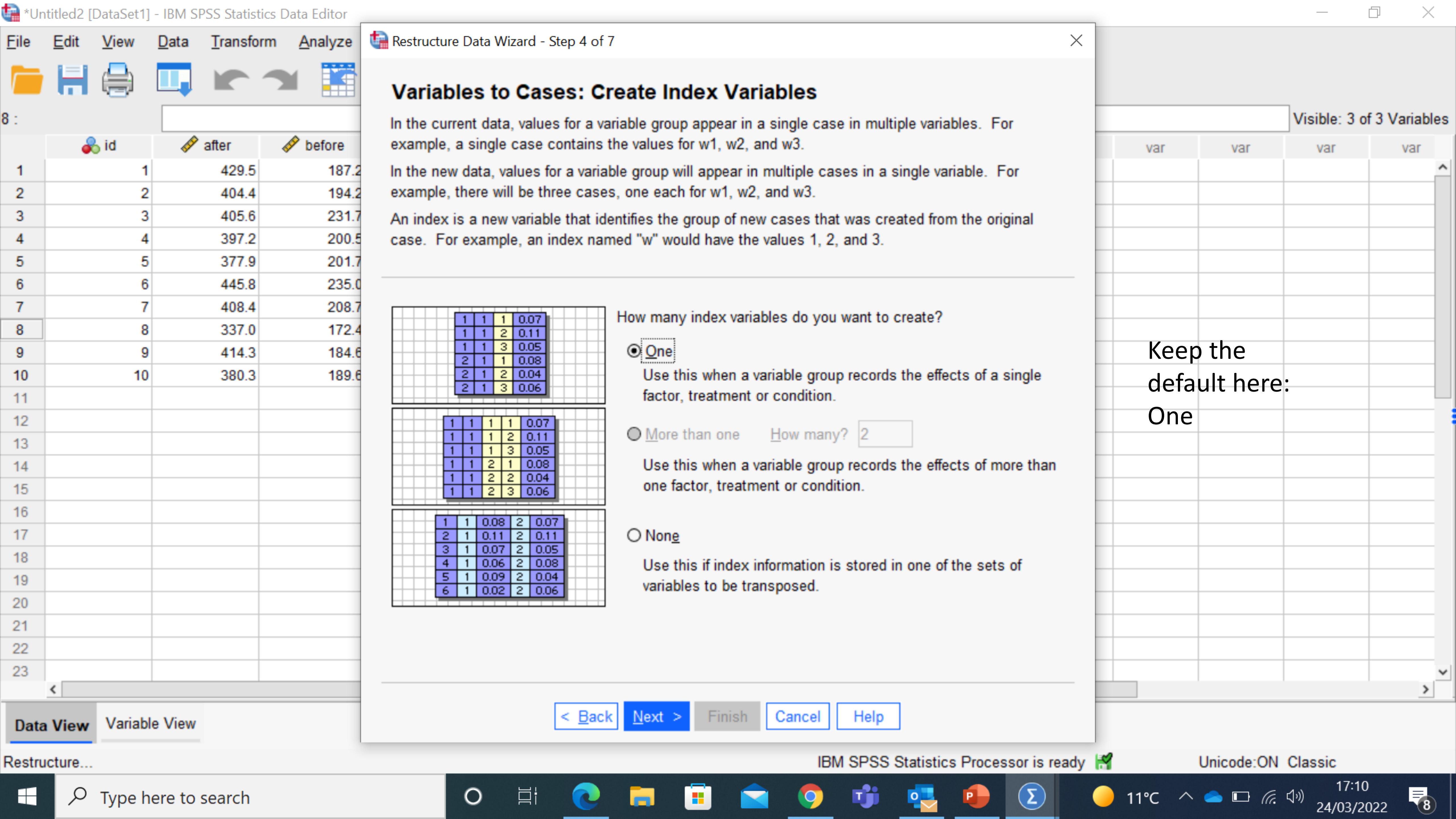
Bring the variables 'before' and 'after' into this box


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard - Step 3 of 7
Variables to Cases: Select Variables
For each variable group you have in the current data the restructured file will have one target variable.
In this step, choose how to identify case groups in the restructured data, and choose which variables belong with each target variable.
Optionally, you can also choose variables to copy to the new file as Fixed Variables.
Variables in the Current File:
id
after
before
Case Group Identification
Use selected variable
Variable: id
Variables to be Transposed
Target Variable: trans1
after
before
Fixed Variable(s):
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
after
before
var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
17:08
24/03/2022

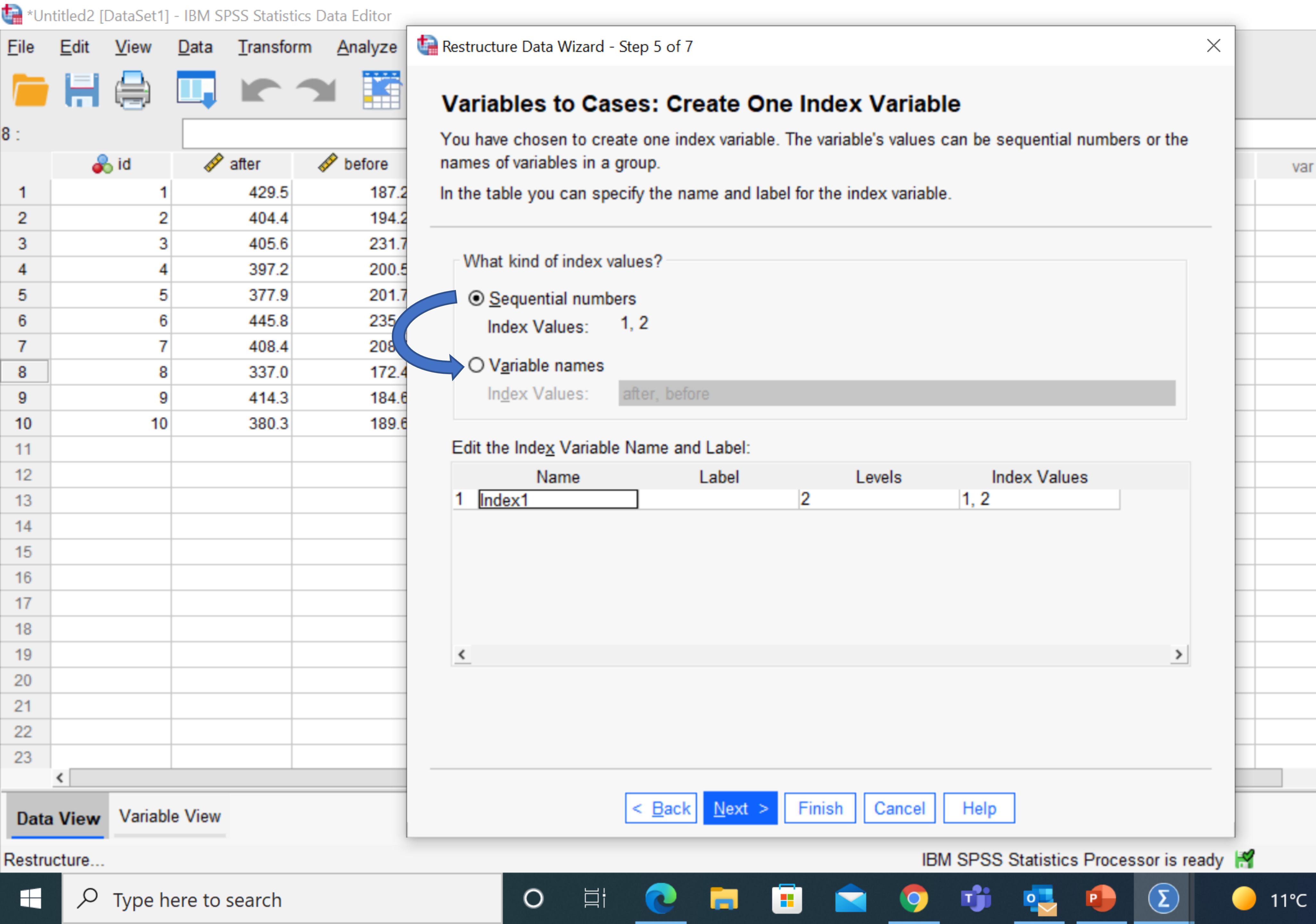


I would rename the ‘Target variable’ to ‘Weight’


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Visible: 3 of 3 Variables
id
after
before
var
1 1 429.5 187.2
2 2 404.4 194.2
3 3 405.6 231.7
4 4 397.2 200.5
5 5 377.9 201.7
6 6 445.8 235.0
7 7 408.4 208.7
8 8 337.0 172.4
9 9 414.3 184.6
10 10 380.3 189.6
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Restructure Data Wizard - Step 3 of 7
Variables to Cases: Select Variables
For each variable group you have in the current data the restructured file will have one target variable.
In this step, choose how to identify case groups in the restructured data, and choose which variables belong with each target variable.
Optionally, you can also choose variables to copy to the new file as Fixed Variables.
Variables in the Current File:
id
after
before
Case Group Identification
Use selected variable
Variable:
id
Variables to be Transposed
Target Variable:
Weight
after
before
Fixed Variable(s):
< Back
Next >
Finish
Cancel
Help
Type here to search
11°C
17:09
24/03/2022

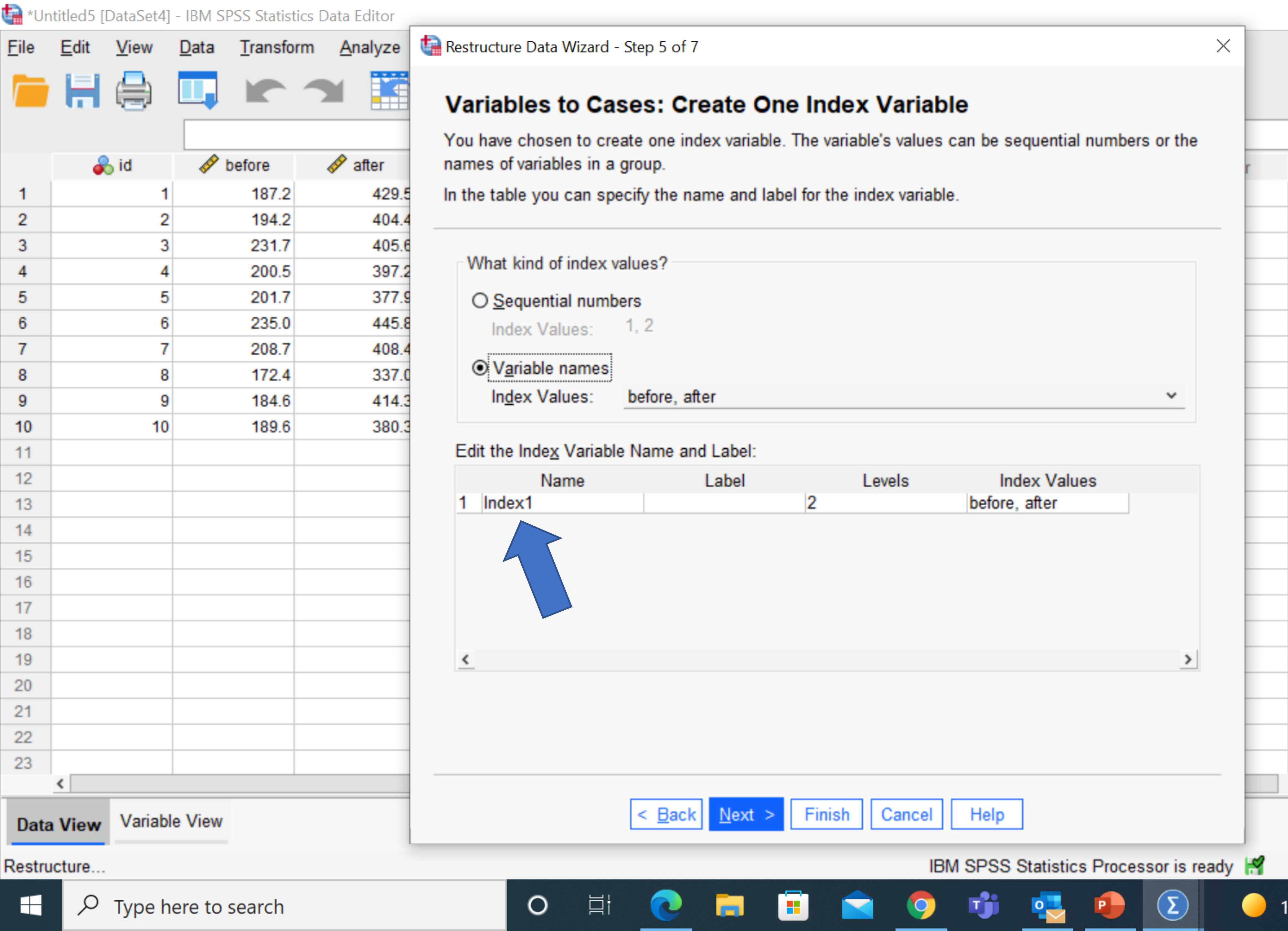


Keep the default here: One



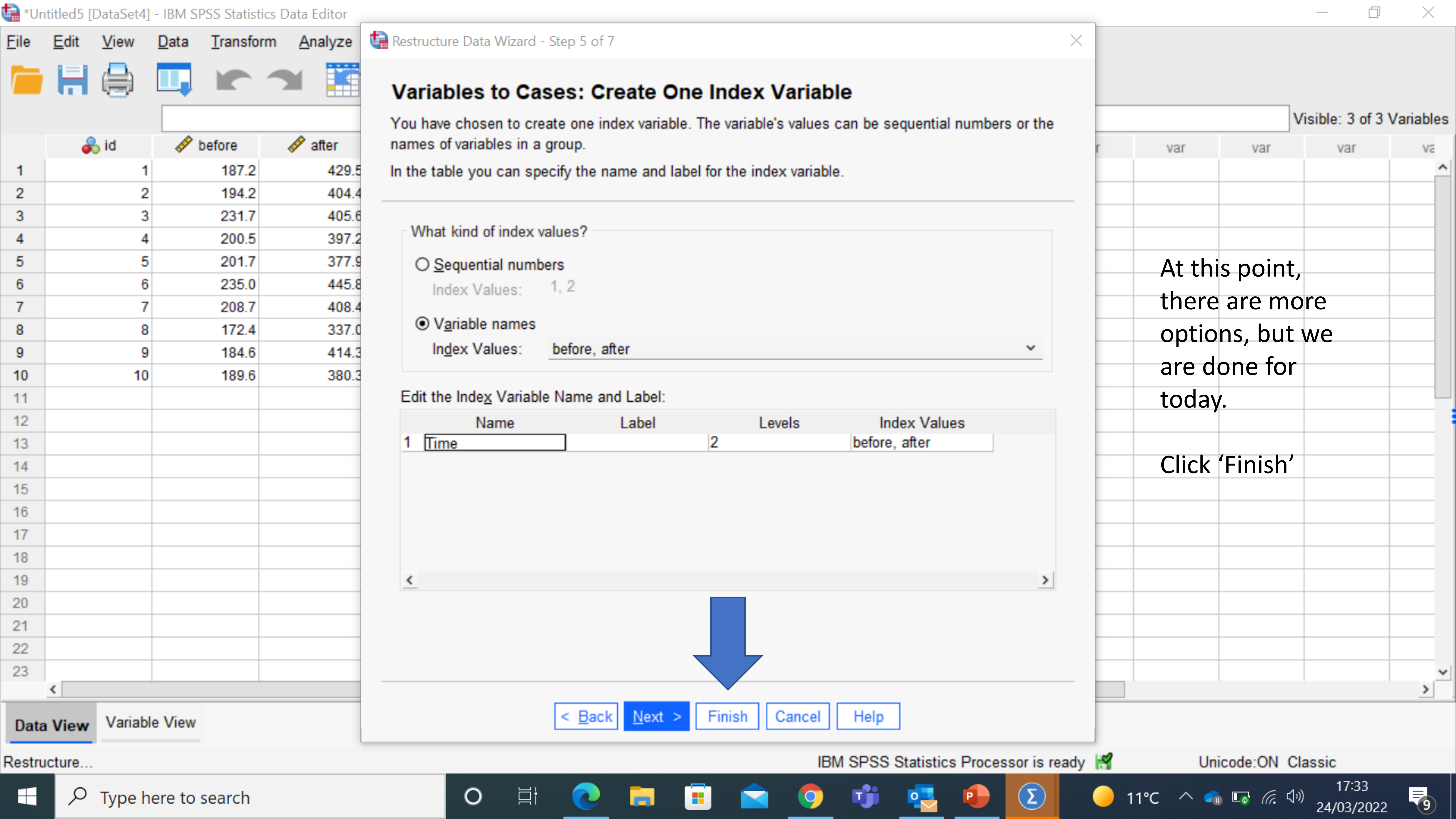
First thing here: Change the first option from 'sequential numbers' to 'variable names'

If you don't do this, your restructured data will be hard to interpret, as the terms 'after' and 'before' will be changed to '1' and '2' respectively



You can also rename the column header that will contain the values of 'before' and 'after'...

Lets call that variable 'Time'



At this point, there are more options, but we are done for today.

Click ‘Finish’

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	id	Time	Weight	var	var	var	var	var	var	var	var	var	var	var	var	var
1	1	before	187.2													
2	1	after	429.5													
3	2	before	194.2													
4	2	after	404.4													
5	3	before	231.7													
6	3	after	405.6													
7	4	before	200.5													
8	4	after	397.2													
9	5	before	201.7													
10	5	after	377.9													
11	6	before	235.0													
12	6	after	445.8													
13	7	before	208.7													
14	7	after	408.4													
15	8	before	172.4													
16	8	after	337.0													
17	9	before	184.6													
18	9	after	414.3													
19	10	before	189.6													
20	10	after	380.3													
21																
22																
23																

And out data is now in **long format**:

This looks much more like what we had for our two-sampled T Test!

Data View | Variable View

IBM SPSS Statistics Processor is ready | Unicode:ON Classic

Type here to search | 11°C | 17:40 24/03/2022

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

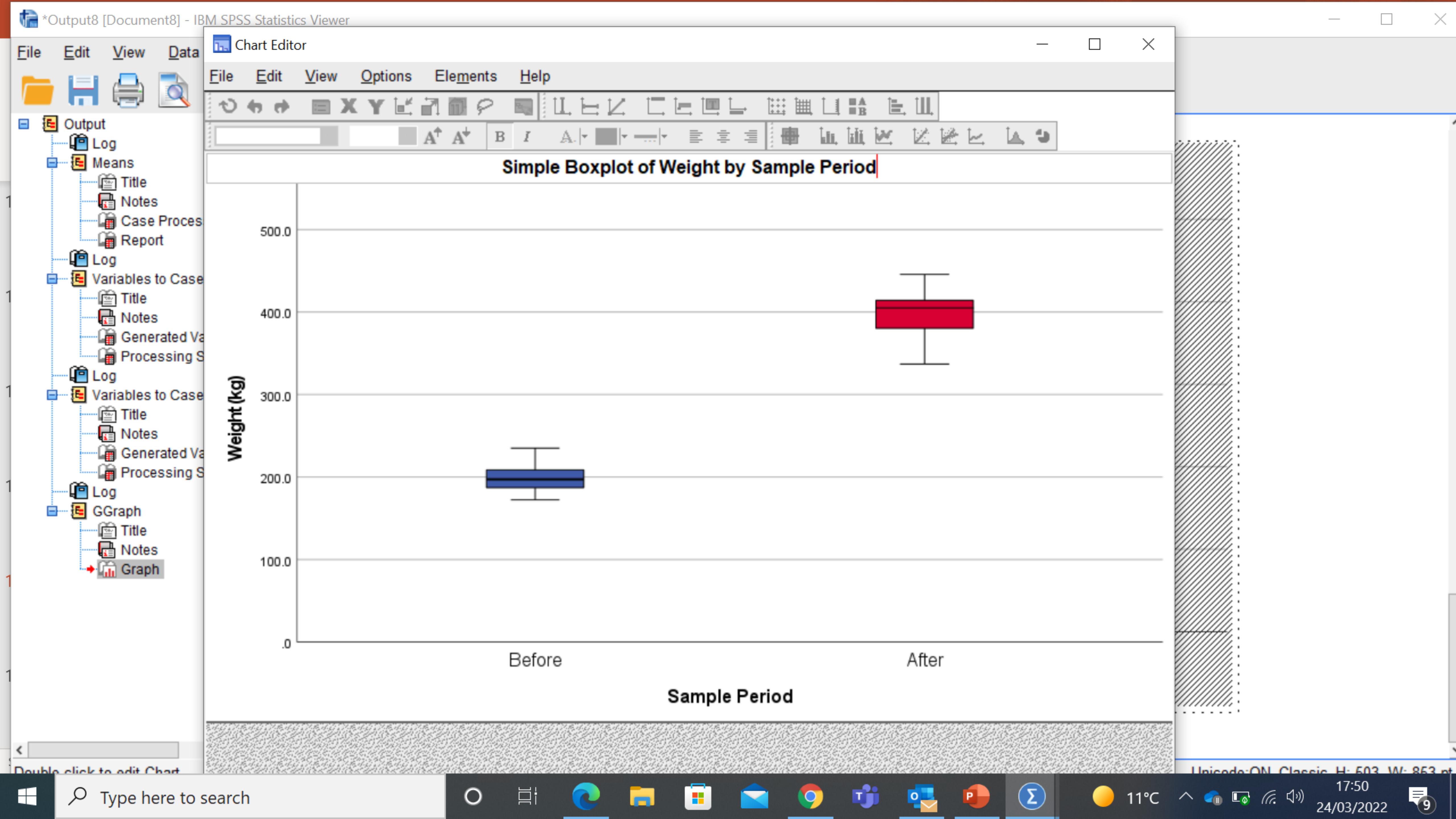
Visible: 3 of 3 Variables

	id	Time	Weight
1	1	before	187.2
2	1	after	429.5
3	2	before	194.2
4	2	after	404.4
5	3	before	231.7
6	3	after	405.6
7	4	before	200.5
8	4	after	397.2
9	5	before	201.7
10	5	after	377.9
11	6	before	235.0
12	6	after	445.8
13	7	before	208.7
14	7	after	408.4
15	8	before	172.4
16	8	after	337.0
17	9	before	184.6
18	9	after	414.3
19	10	before	189.6
20	10	after	380.3

Now you should be able to make you boxplot! Give it a try!

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic


*Output8 [Document8] - IBM SPSS Statistics Viewer
Chart Editor
File Edit View Options Elements Help
Output
Log
Means
Title
Notes
Case Proces
Report
Log
Variables to Case
Title
Notes
Generated Va
Processing S
Log
Variables to Case
Title
Notes
Generated Va
Processing S
Log
GGraph
Title
Notes
Graph
Simple Boxplot of Weight by Sample Period
Weight (kg)
500.0
400.0
300.0
200.0
100.0
.0
Before
After
Sample Period
Type here to search
11°C
17:50
24/03/2022

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	id	Time	eight
1	1	before	187.2
2	1	after	429.5
3	2	before	194.2
4	2	after	404.4
5	3	before	231.7
6	3	after	405.6
7	4	before	200.5
8	4	after	397.2
9	5	before	201.7
10	5	after	377.9
11	6	before	235.0
12	6	after	445.8
13	7	before	208.7
14	7	after	408.4
15	8	before	172.4
16	8	after	337.0
17	9	before	184.6
18	9	after	414.3
19	10	before	189.6
20	10	after	380.3

Well great job!

So, we are ready to move forward with our Paired T Test… but guess what:

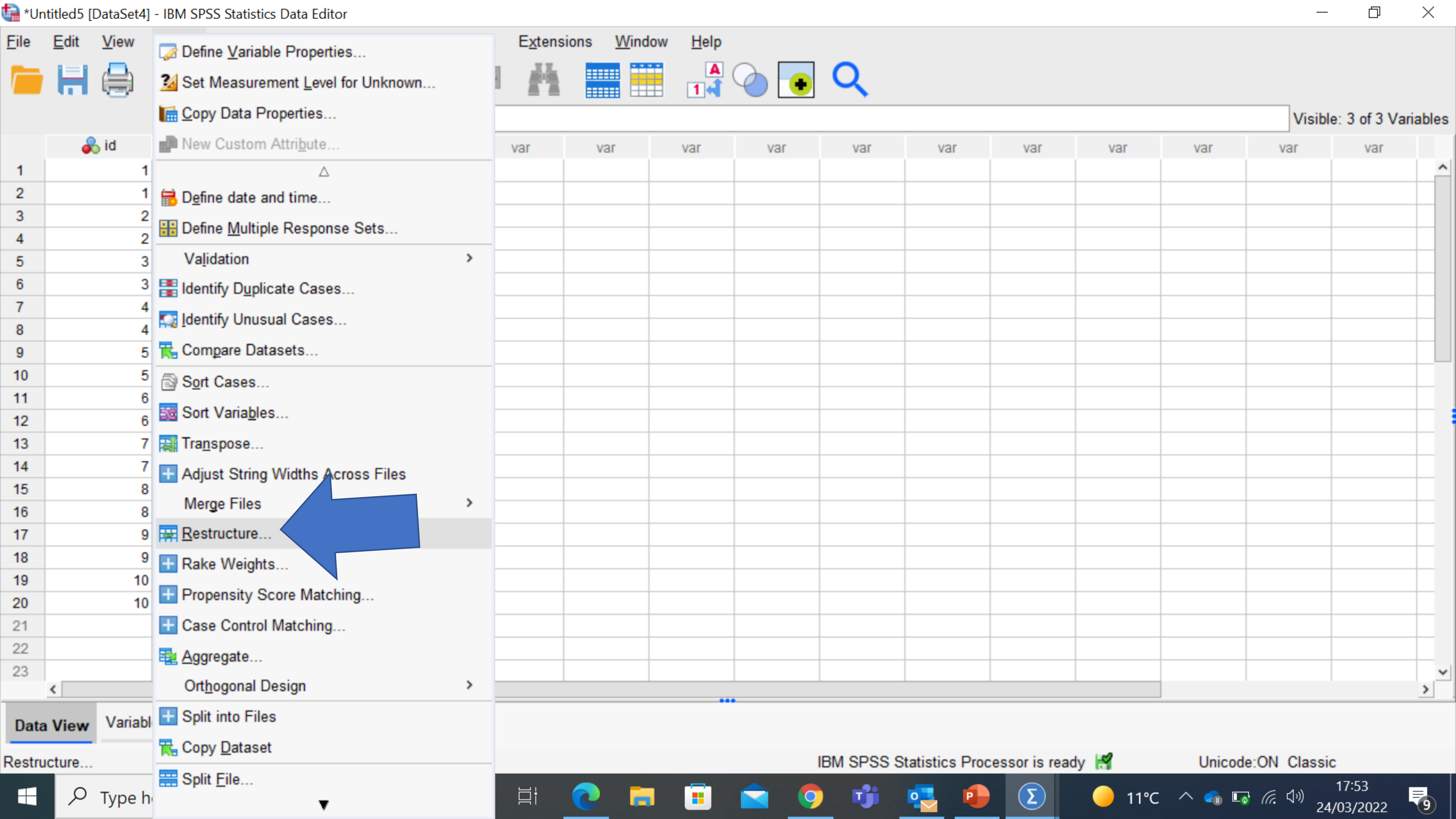
That analysis wants our data to be back in **wide format**

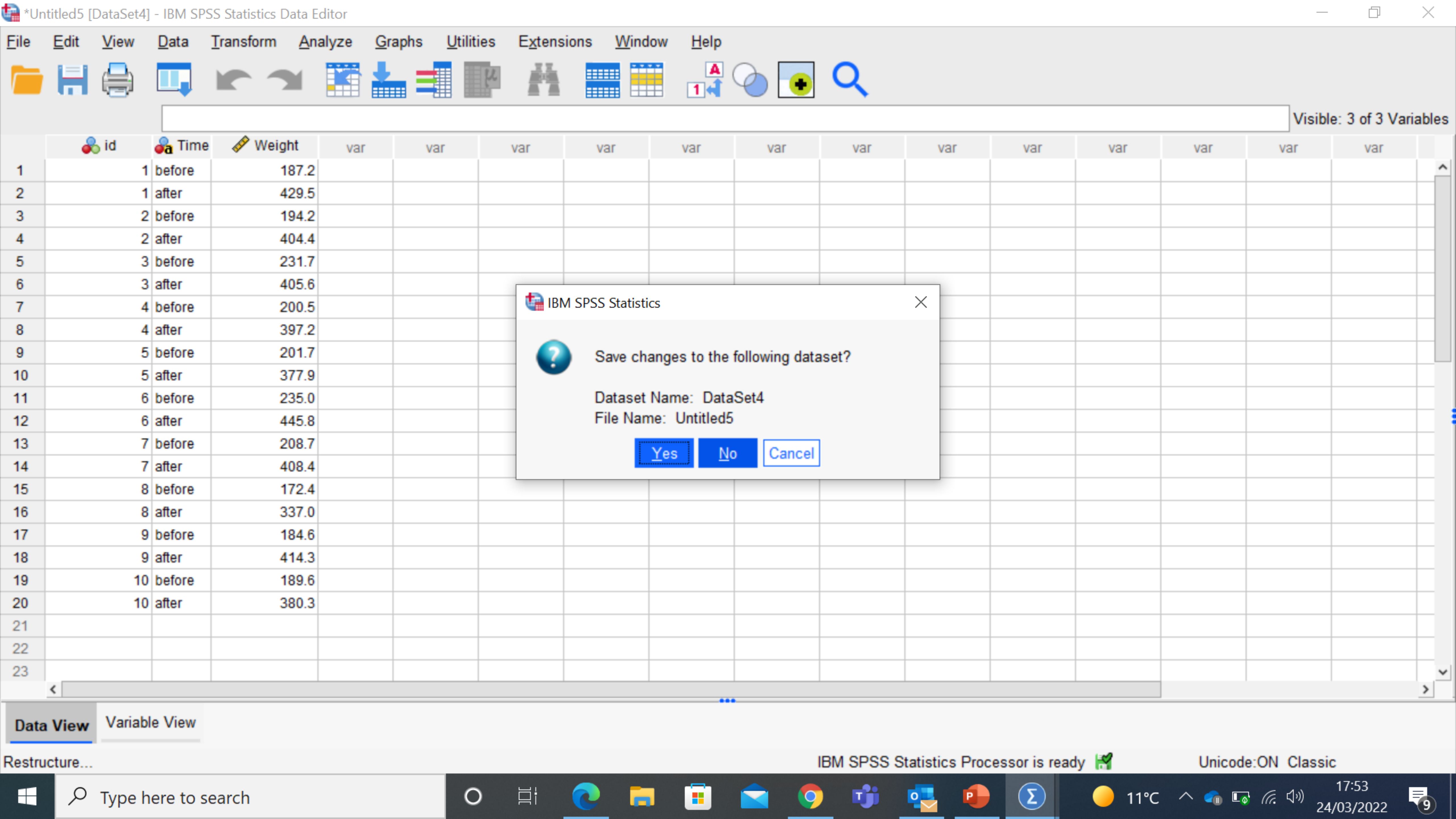
We Have to covert it back!?!?!

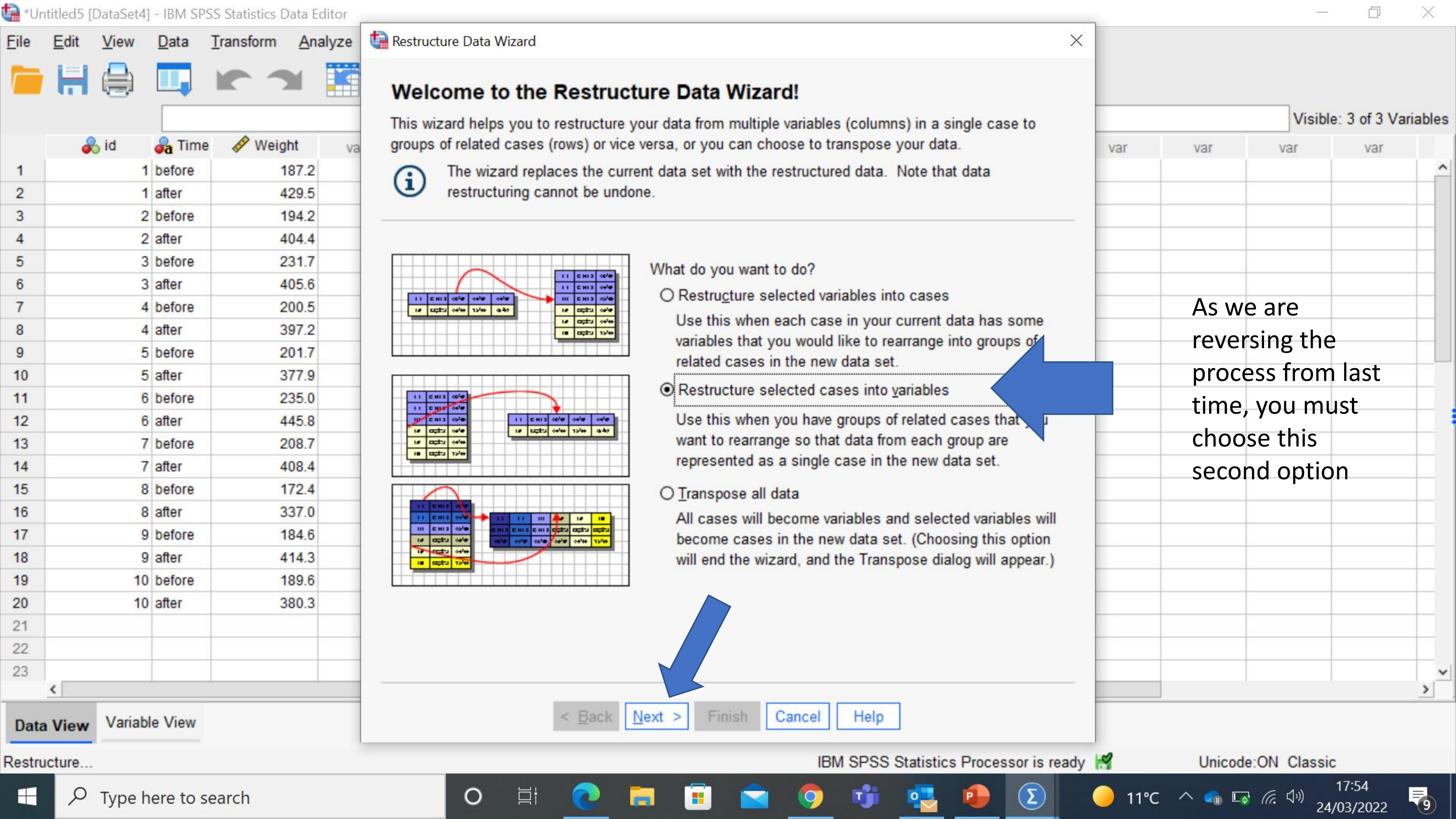
This procedure is much shorter actually

Data View Variable View

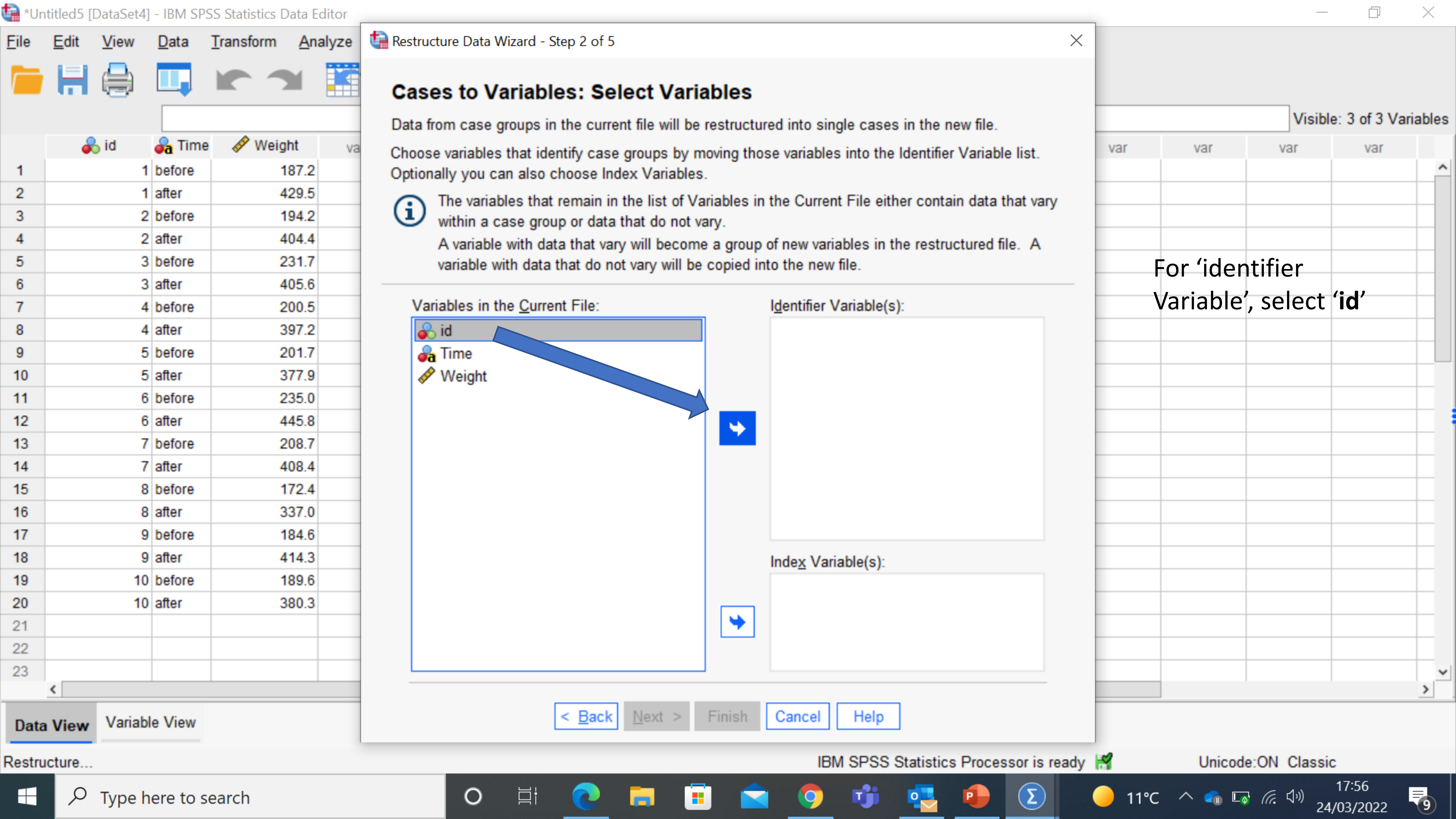
IBM SPSS Statistics Processor is ready Unicode:ON Classic


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File
Edit
View
Extensions
Window
Help
Define Variable Properties...
Set Measurement Level for Unknown...
Copy Data Properties...
New Custom Attribute...
Define date and time...
Define Multiple Response Sets...
Validation
Identify Duplicate Cases...
Identify Unusual Cases...
Compare Datasets...
Sort Cases...
Sort Variables...
Transpose...
Adjust String Widths Across Files
Merge Files
Restructure...
Rake Weights...
Propensity Score Matching...
Case Control Matching...
Aggregate...
Orthogonal Design
Split into Files
Copy Dataset
Split File...
Visible: 3 of 3 Variables
id
var
Data View
Variabl
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type h
11°C
17:53
24/03/2022

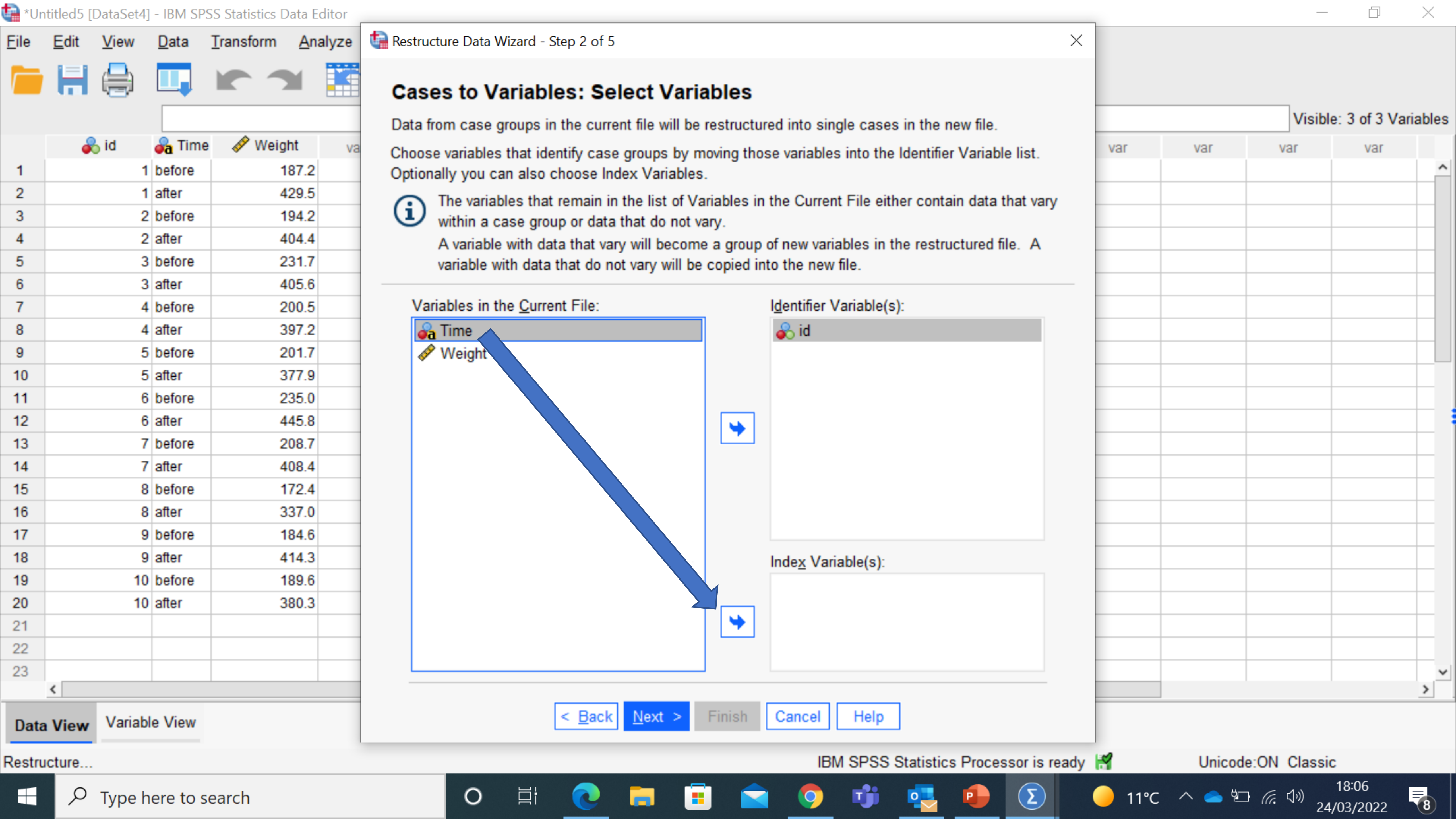

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id Time Weight var var var var var var var var var var var var var
1 1 before 187.2
2 1 after 429.5
3 2 before 194.2
4 2 after 404.4
5 3 before 231.7
6 3 after 405.6
7 4 before 200.5
8 4 after 397.2
9 5 before 201.7
10 5 after 377.9
11 6 before 235.0
12 6 after 445.8
13 7 before 208.7
14 7 after 408.4
15 8 before 172.4
16 8 after 337.0
17 9 before 184.6
18 9 after 414.3
19 10 before 189.6
20 10 after 380.3
21
22
23
IBM SPSS Statistics
Save changes to the following dataset?
Dataset Name: DataSet4
File Name: Untitled5
Yes
No
Cancel
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
17:53
24/03/2022

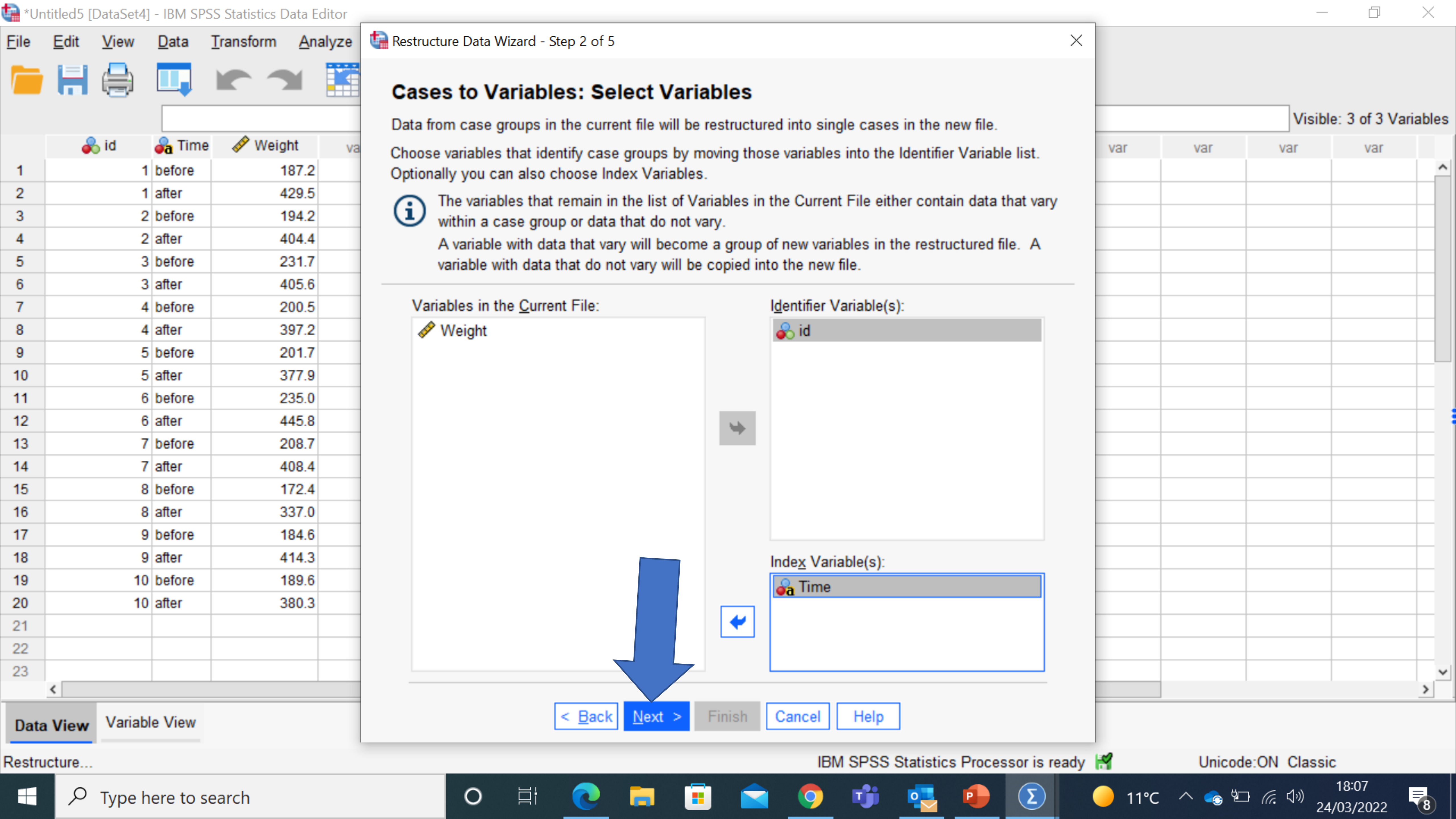


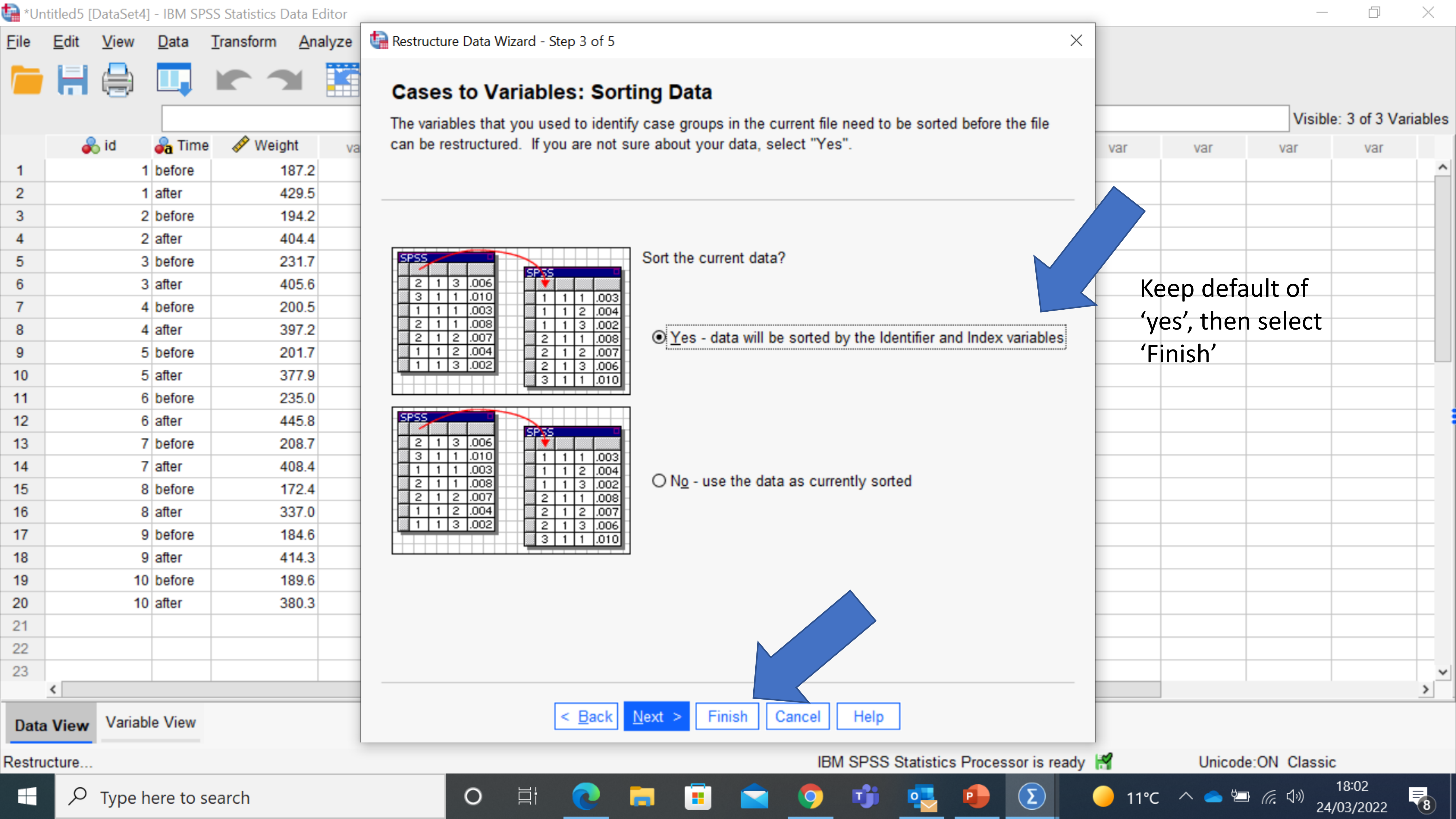
As we are reversing the process from last time, you must choose this second option



For ‘identifier Variable’, select ‘**id**’


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard - Step 2 of 5
Cases to Variables: Select Variables
Data from case groups in the current file will be restructured into single cases in the new file.
Choose variables that identify case groups by moving those variables into the Identifier Variable list. Optionally you can also choose Index Variables.
The variables that remain in the list of Variables in the Current File either contain data that vary within a case group or data that do not vary.
A variable with data that vary will become a group of new variables in the restructured file. A variable with data that do not vary will be copied into the new file.
Variables in the Current File:
Time
Weight
Identifier Variable(s):
id
Index Variable(s):
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
Time
Weight
1 before 187.2
1 after 429.5
2 before 194.2
2 after 404.4
3 before 231.7
3 after 405.6
4 before 200.5
4 after 397.2
5 before 201.7
5 after 377.9
6 before 235.0
6 after 445.8
7 before 208.7
7 after 408.4
8 before 172.4
8 after 337.0
9 before 184.6
9 after 414.3
10 before 189.6
10 after 380.3
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
18:06
24/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze
Restructure Data Wizard - Step 2 of 5
Cases to Variables: Select Variables
Data from case groups in the current file will be restructured into single cases in the new file.
Choose variables that identify case groups by moving those variables into the Identifier Variable list. Optionally you can also choose Index Variables.
The variables that remain in the list of Variables in the Current File either contain data that vary within a case group or data that do not vary.
A variable with data that vary will become a group of new variables in the restructured file. A variable with data that do not vary will be copied into the new file.
Variables in the Current File:
Weight
Identifier Variable(s):
id
Index Variable(s):
Time
< Back
Next >
Finish
Cancel
Help
Visible: 3 of 3 Variables
id
Time
Weight
var
1 before 187.2
1 after 429.5
2 before 194.2
2 after 404.4
3 before 231.7
3 after 405.6
4 before 200.5
4 after 397.2
5 before 201.7
5 after 377.9
6 before 235.0
6 after 445.8
7 before 208.7
7 after 408.4
8 before 172.4
8 after 337.0
9 before 184.6
9 after 414.3
10 before 189.6
10 after 380.3
Data View
Variable View
Restructure...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
18:07
24/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
Restructure Data Wizard - Step 3 of 5
Cases to Variables: Sorting Data
The variables that you used to identify case groups in the current file need to be sorted before the file can be restructured. If you are not sure about your data, select "Yes".
Sort the current data?
Yes - data will be sorted by the Identifier and Index variables
No - use the data as currently sorted
< Back
Next >
Finish
Cancel
Help
Keep default of 'yes', then select 'Finish'

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	id	after	before
1	1	429.5	187.2
2	2	404.4	194.2
3	3	405.6	231.7
4	4	397.2	200.5
5	5	377.9	201.7
6	6	445.8	235.0
7	7	408.4	208.7
8	8	337.0	172.4
9	9	414.3	184.6
10	10	380.3	189.6

And we are back to **wide format**!

We can start looking at our paired samples T Test…

Almost

Data View Variable View

IBM SPSS Statistics Processor is ready Unicode:ON Classic

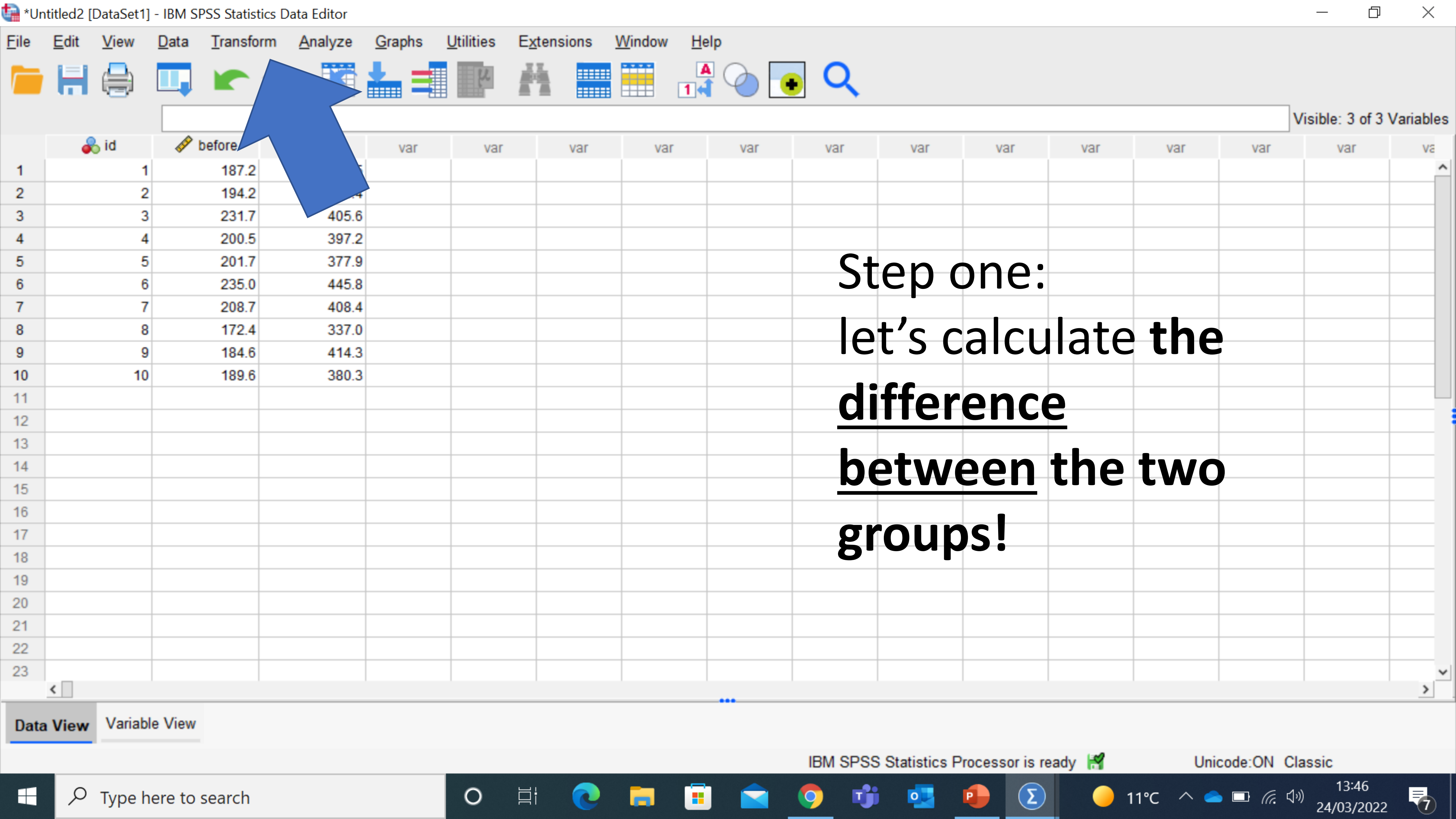
Assumptions

- Similar to the two sample T Test, we must check that our data meet some assumptions:
- The two groups are 'paired'.
  - In our example, this is the case as the data were collected measuring the **same individual rats** at **2 different time periods**.
- No significant outliers in **the <u>difference between</u> the two related groups**
- Normality: **the <u>difference between</u> the two related groups** follow a normal distribution

	id	before	after
1	1	187.2	429.5
2	2	194.2	404.4
3	3	231.7	405.6
4	4	200.5	397.2
5	5	201.7	377.9
6	6	235.0	445.8
7	7	208.7	408.4
8	8	172.4	337.0
9	9	184.6	414.3
10	10	189.6	380.3

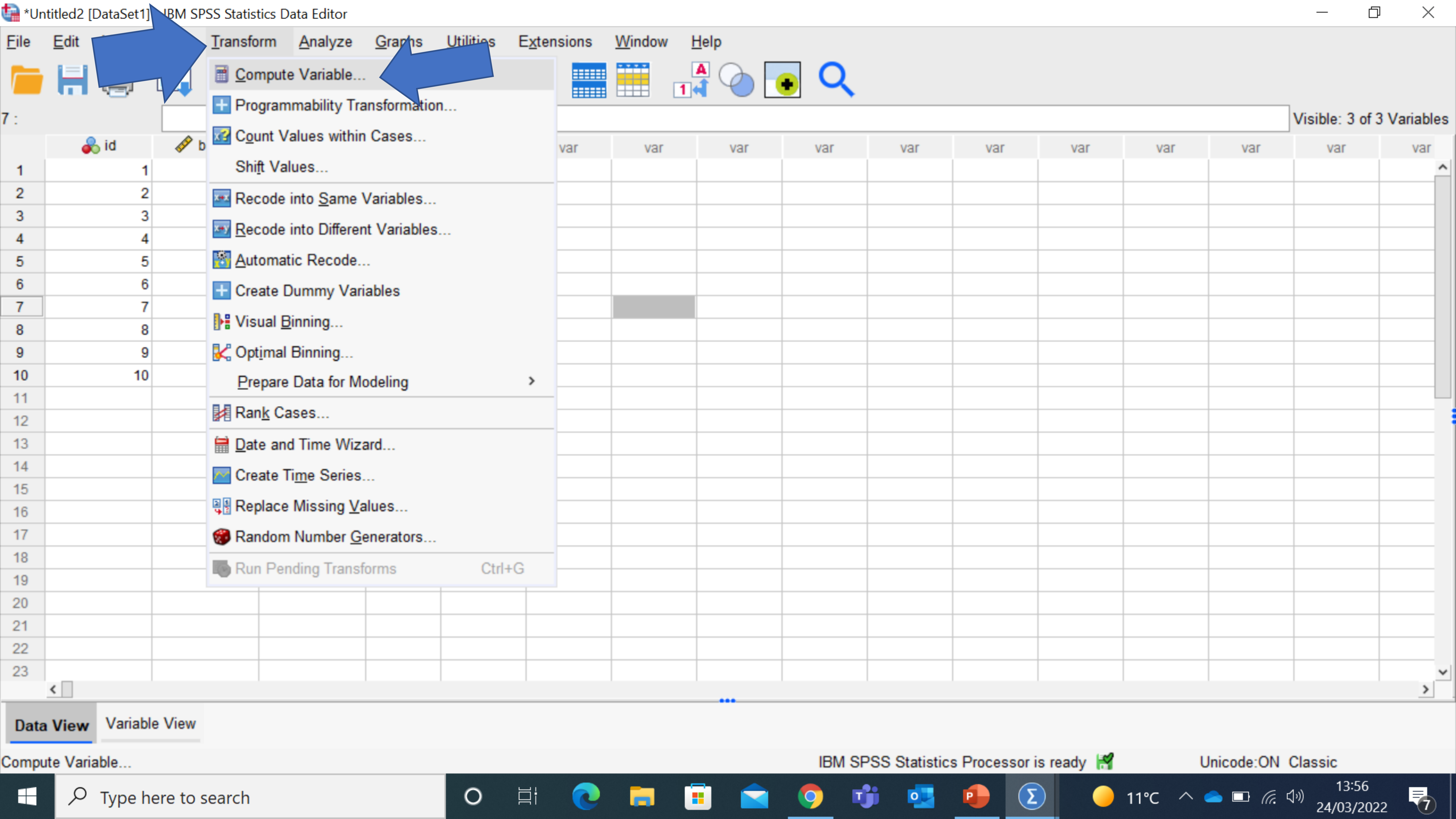
SO: We need to look at the distribution of **the difference between the two groups**. We are interested in:

1) The distribution: is it normal?
2) Are there outliers?

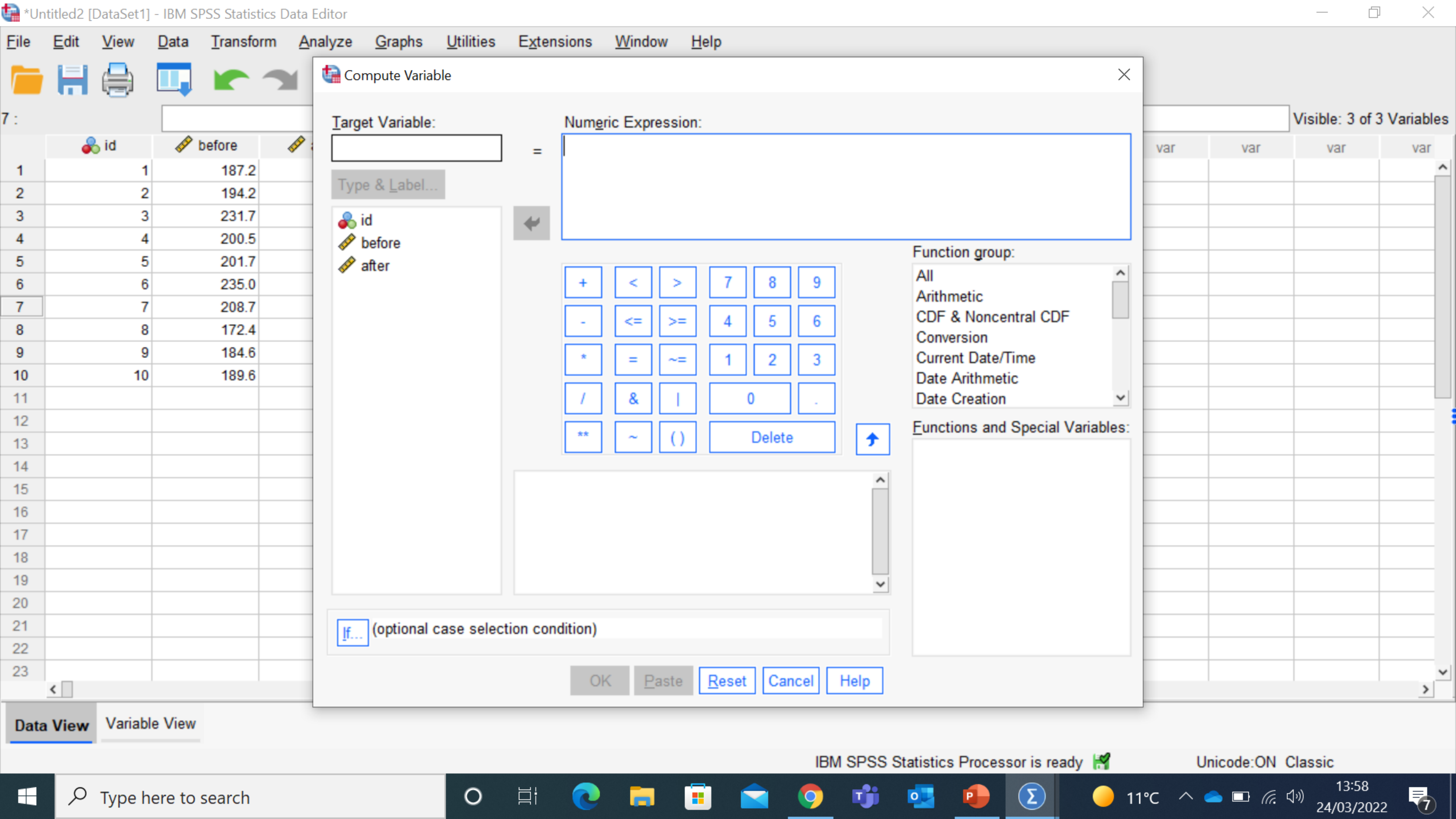


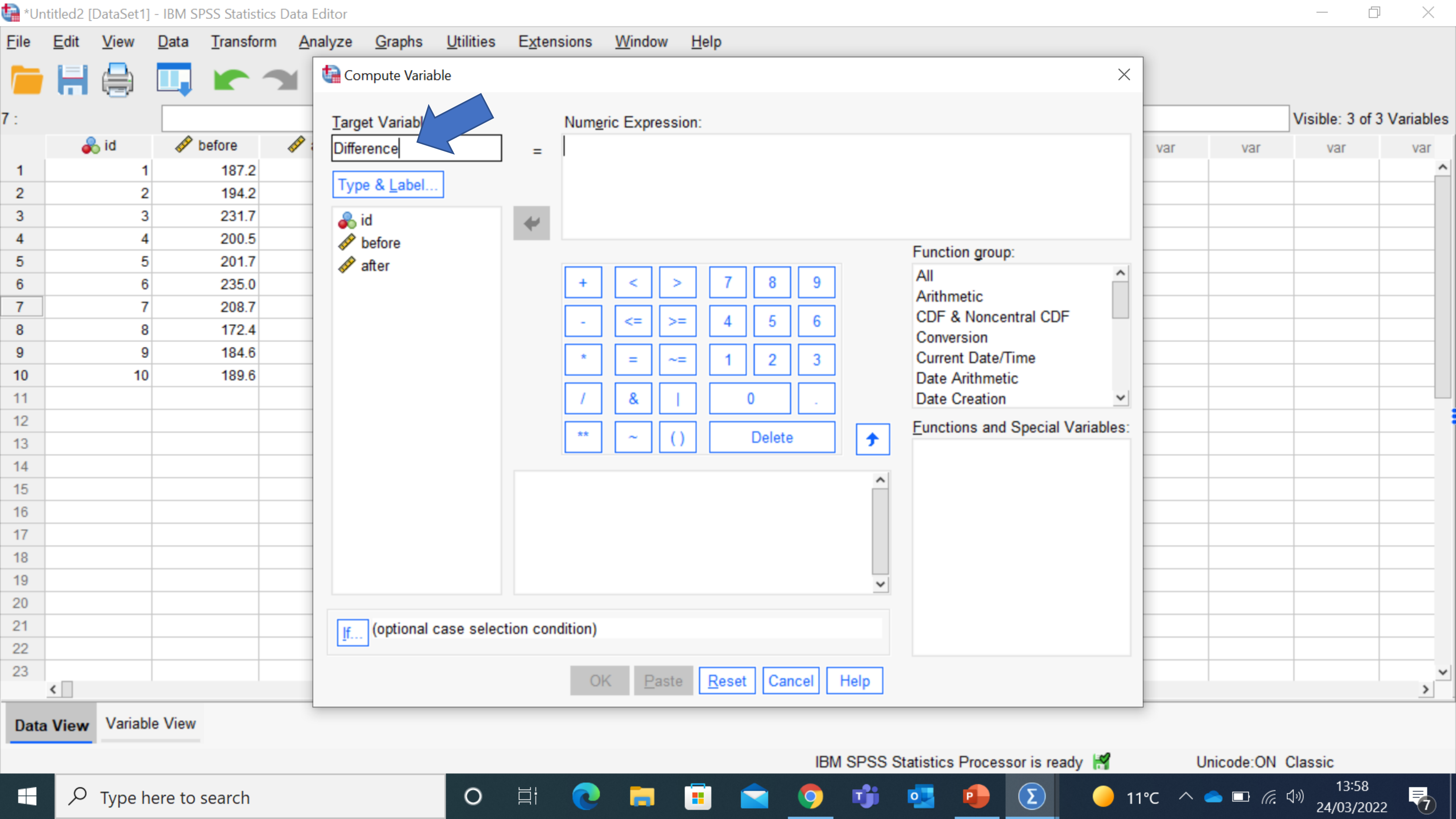
Step one:

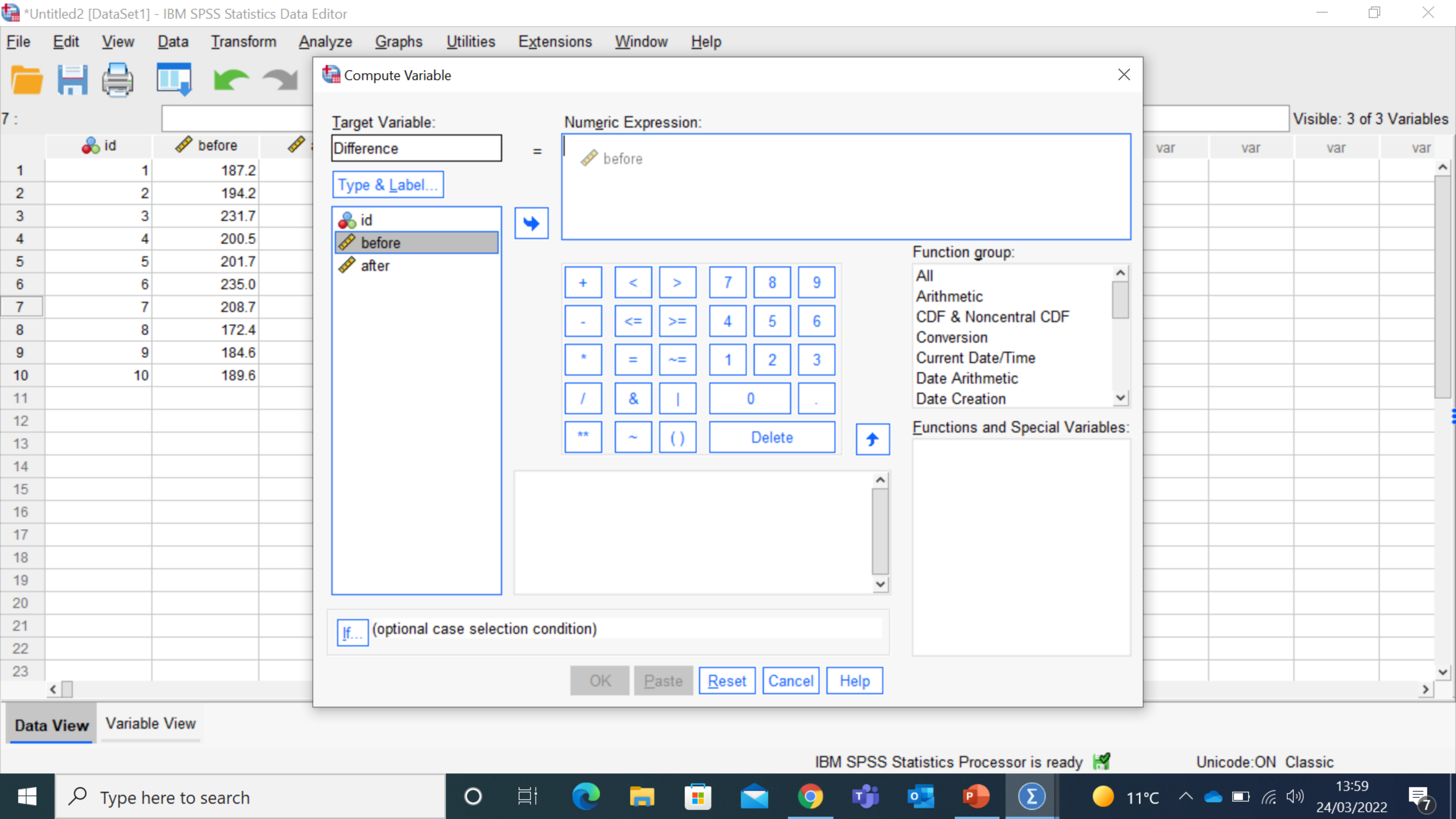
let's calculate **the <u>difference</u> <u>between</u> the two groups!**

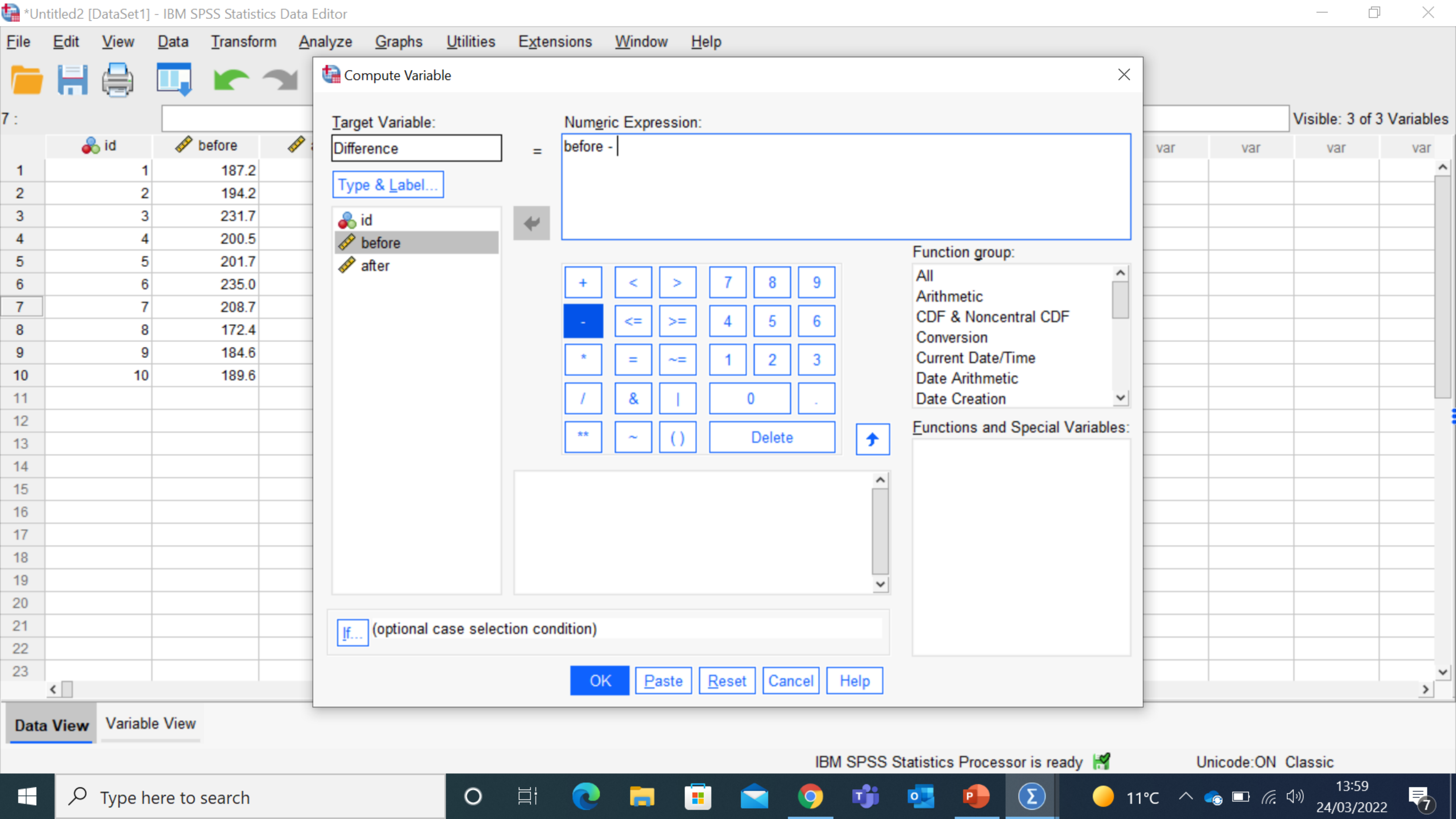


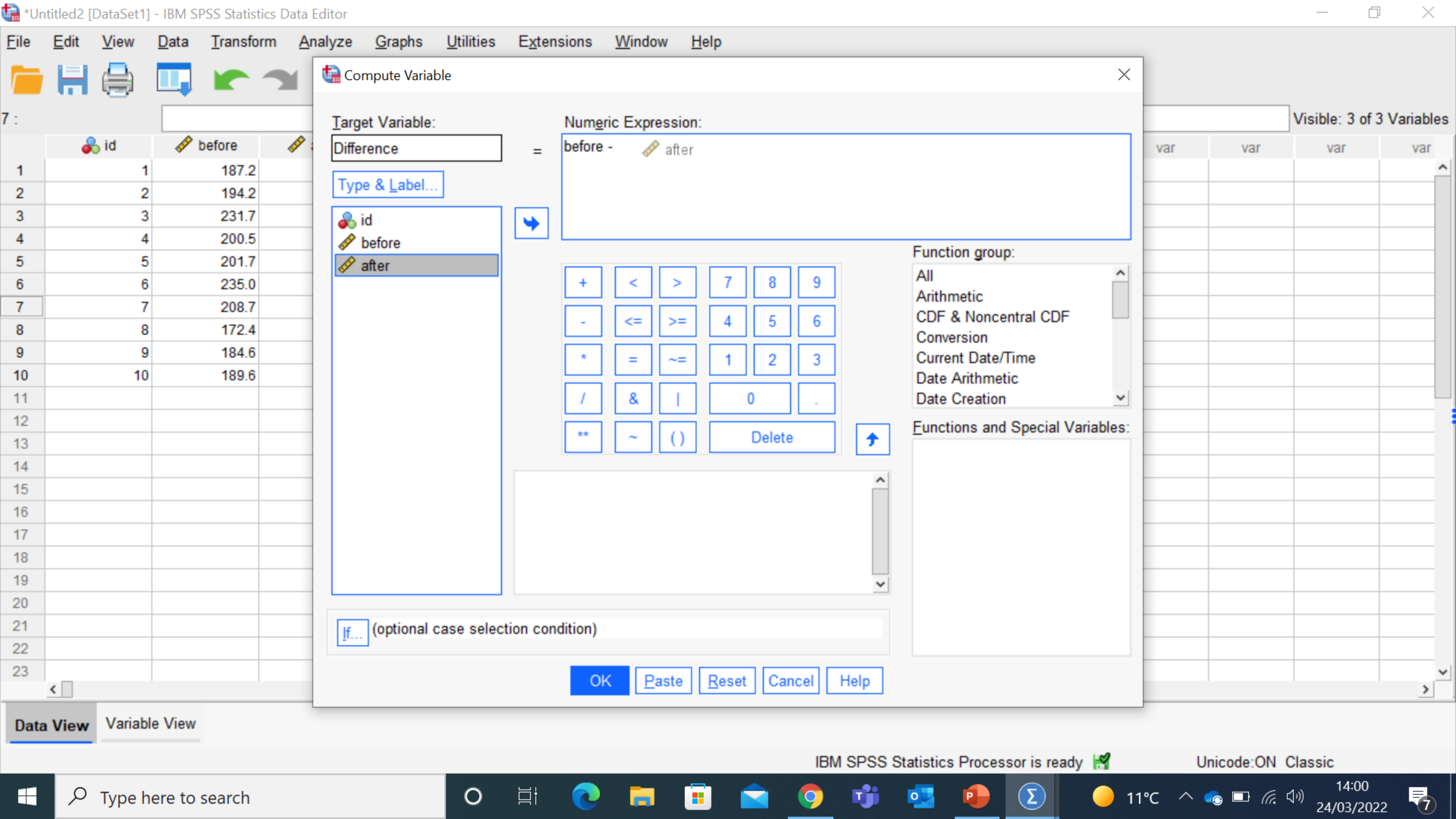
*Untitled2 [DataSet1] IBM SPSS Statistics Data Editor
File Edit Transform Analyze Graphs Utilities Extensions Window Help
Compute Variable...
Programmability Transformation...
Count Values within Cases...
Shift Values...
Recode into Same Variables...
Recode into Different Variables...
Automatic Recode...
Create Dummy Variables
Visual Binning...
Optimal Binning...
Prepare Data for Modeling
Rank Cases...
Date and Time Wizard...
Create Time Series...
Replace Missing Values...
Random Number Generators...
Run Pending Transforms Ctrl+G
Visible: 3 of 3 Variables
id
var
Data View
Variable View
Compute Variable...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
13:56
24/03/2022

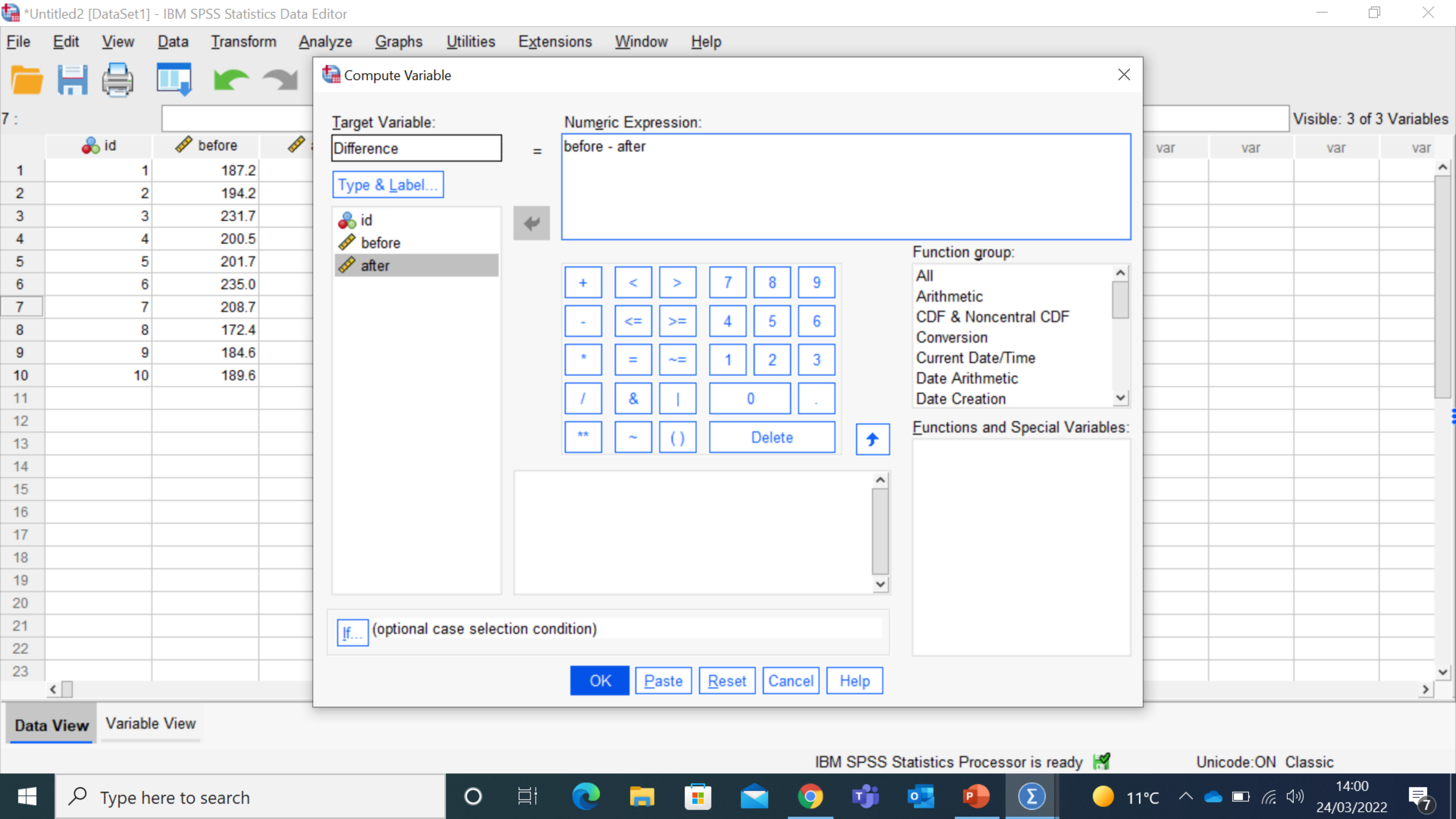

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 3 of 3 Variables
id before
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
var var var var
Compute Variable
Target Variable:
Numeric Expression:
=
Type & Label...
id
before
after
+ < > 7 8 9
- <= >= 4 5 6
* = ~= 1 2 3
/ & | 0 .
** ~ ( ) Delete
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Functions and Special Variables:
If... (optional case selection condition)
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
13:58
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Compute Variable
Target Variable:
Difference
Numeric Expression:
Type & Label...
id
before
after
+ < > 7 8 9
- <= >= 4 5 6
* = ~= 1 2 3
/ & | 0 .
** ~ ( ) Delete
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Functions and Special Variables:
If... (optional case selection condition)
OK Paste Reset Cancel Help
Visible: 3 of 3 Variables
id before
1 187.2
2 194.2
3 231.7
4 200.5
5 201.7
6 235.0
7 208.7
8 172.4
9 184.6
10 189.6
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
13:58
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id
before
var
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
Compute Variable
Target Variable:
Difference
=
Numeric Expression:
before
Type & Label...
id
before
after
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Functions and Special Variables:
Delete
If... (optional case selection condition)
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
13:59
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id before
1 187.2
2 194.2
3 231.7
4 200.5
5 201.7
6 235.0
7 208.7
8 172.4
9 184.6
10 189.6
var var var var
Compute Variable
Target Variable:
Difference
Numeric Expression:
before -
Type & Label...
id
before
after
+ < > 7 8 9
- <= >= 4 5 6
* = ~= 1 2 3
/ & | 0 .
** ~ ( ) Delete
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Functions and Special Variables:
If... (optional case selection condition)
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
13:59
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id before
1 187.2
2 194.2
3 231.7
4 200.5
5 201.7
6 235.0
7 208.7
8 172.4
9 184.6
10 189.6
Compute Variable
Target Variable:
Difference
Type & Label...
id
before
after
Numeric Expression:
before - after
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Functions and Special Variables:
Delete
If... (optional case selection condition)
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
14:00
24/03/2022



*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
id before
1 187.2
2 194.2
3 231.7
4 200.5
5 201.7
6 235.0
7 208.7
8 172.4
9 184.6
10 189.6
var var var var
Compute Variable
Target Variable:
Difference
Type & Label...
Numeric Expression:
before - after
id
before
after
Function group:
All
Arithmetic
CDF & Noncentral CDF
Conversion
Current Date/Time
Date Arithmetic
Date Creation
Delete
Functions and Special Variables:
If... (optional case selection condition)
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
14:00
24/03/2022

*Output5 [Document5] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Log

```
COMPUTE Difference=before - after.
EXECUTE.
```

IBM SPSS Statistics Processor is ready

Unicode:ON Classic

Type here to search

11°C

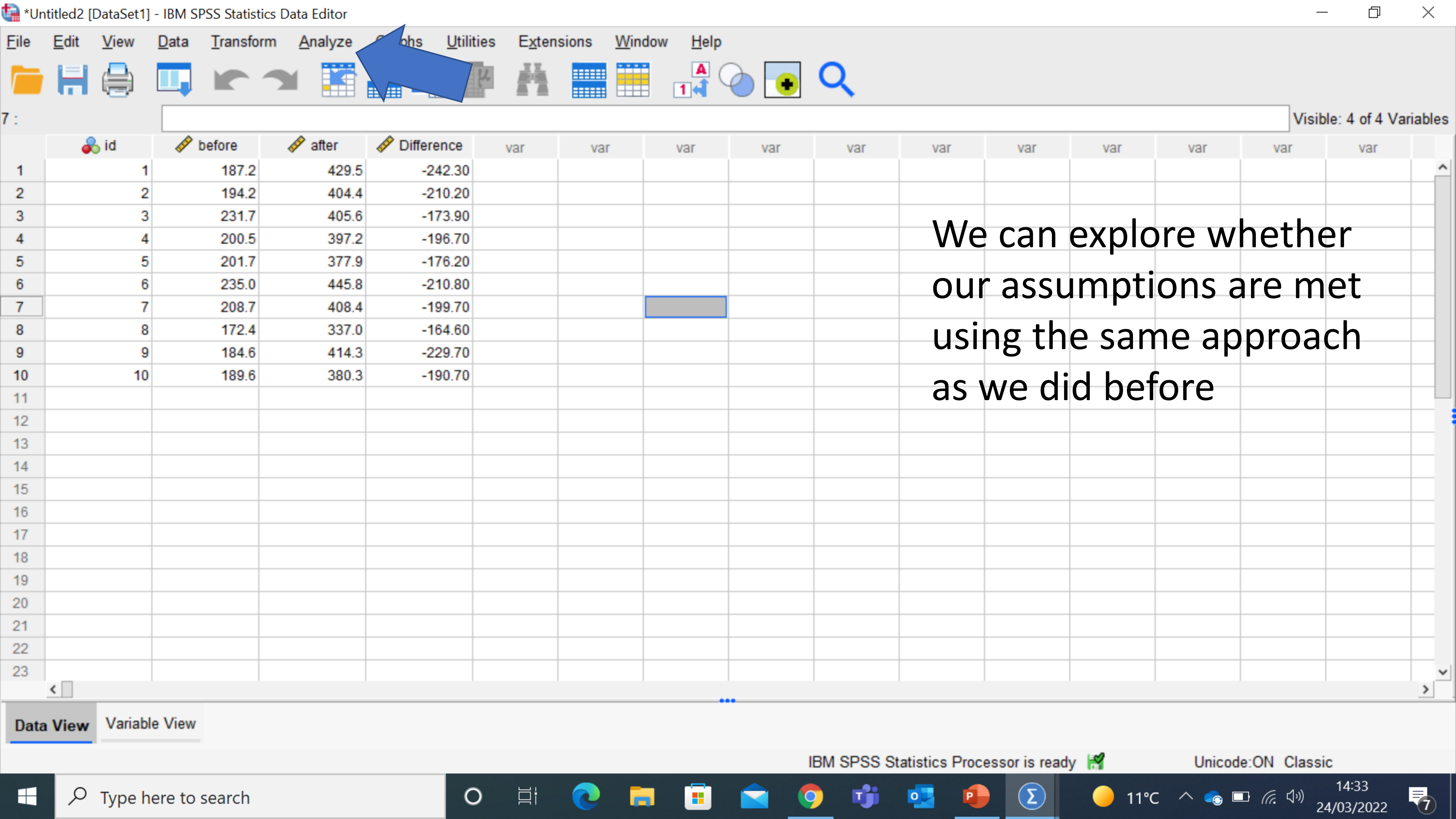
14:00
24/03/2022

id	before	after	Difference
1	187.2	429.5	-242.30
2	194.2	404.4	-210.20
3	231.7	405.6	-173.90
4	200.5	397.2	-196.70
5	201.7	377.9	-176.20
6	235.0	445.8	-210.80
7	208.7	408.4	-199.70
8	172.4	337.0	-164.60
9	184.6	414.3	-229.70
10	189.6	380.3	-190.70

Tada!

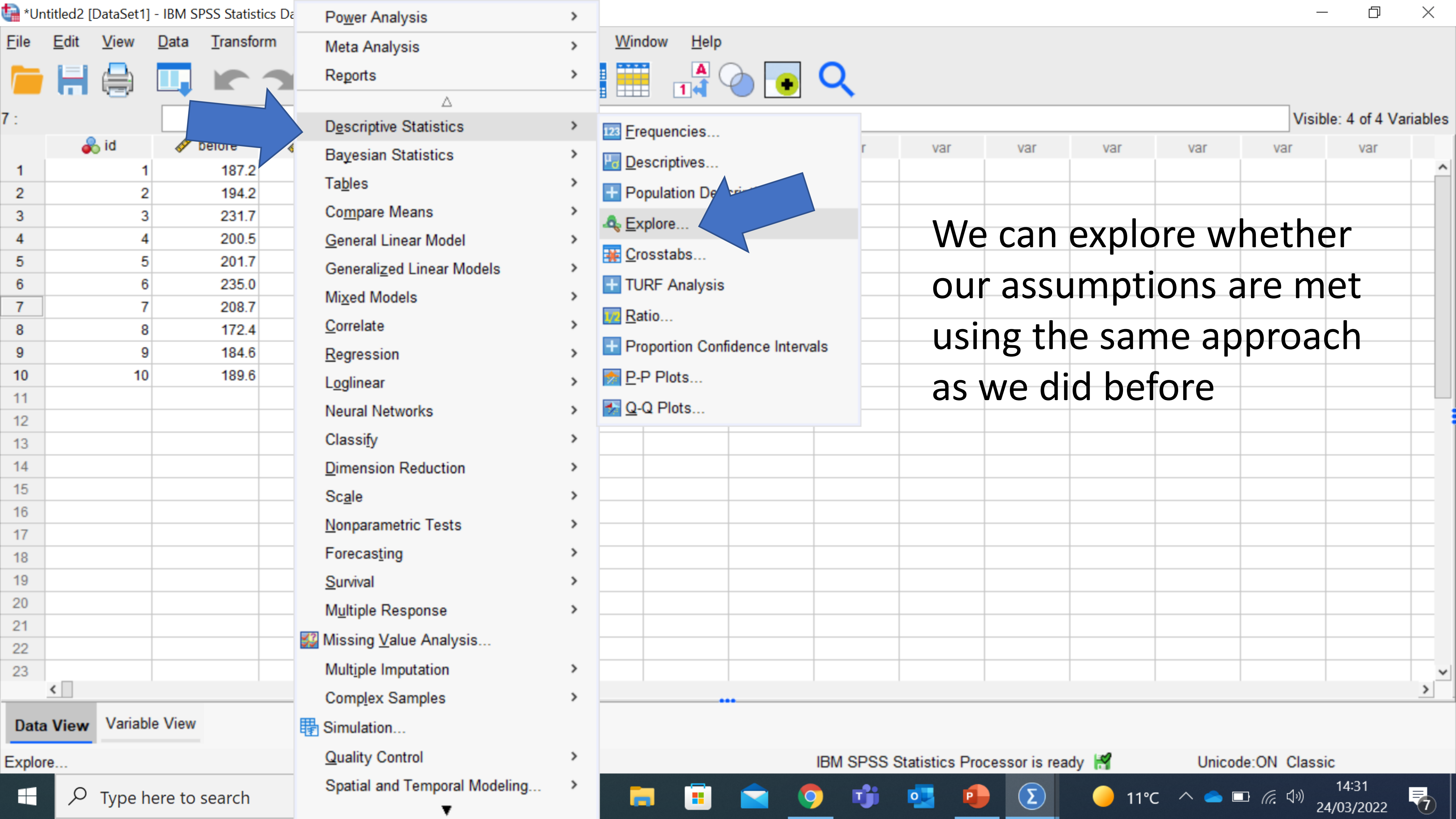
This new column represents the difference between the weights from the before and after columns!

Note the negative values: This is because I used:
before-after
I could get positive values if I used:
after-before
In truth, it makes no difference

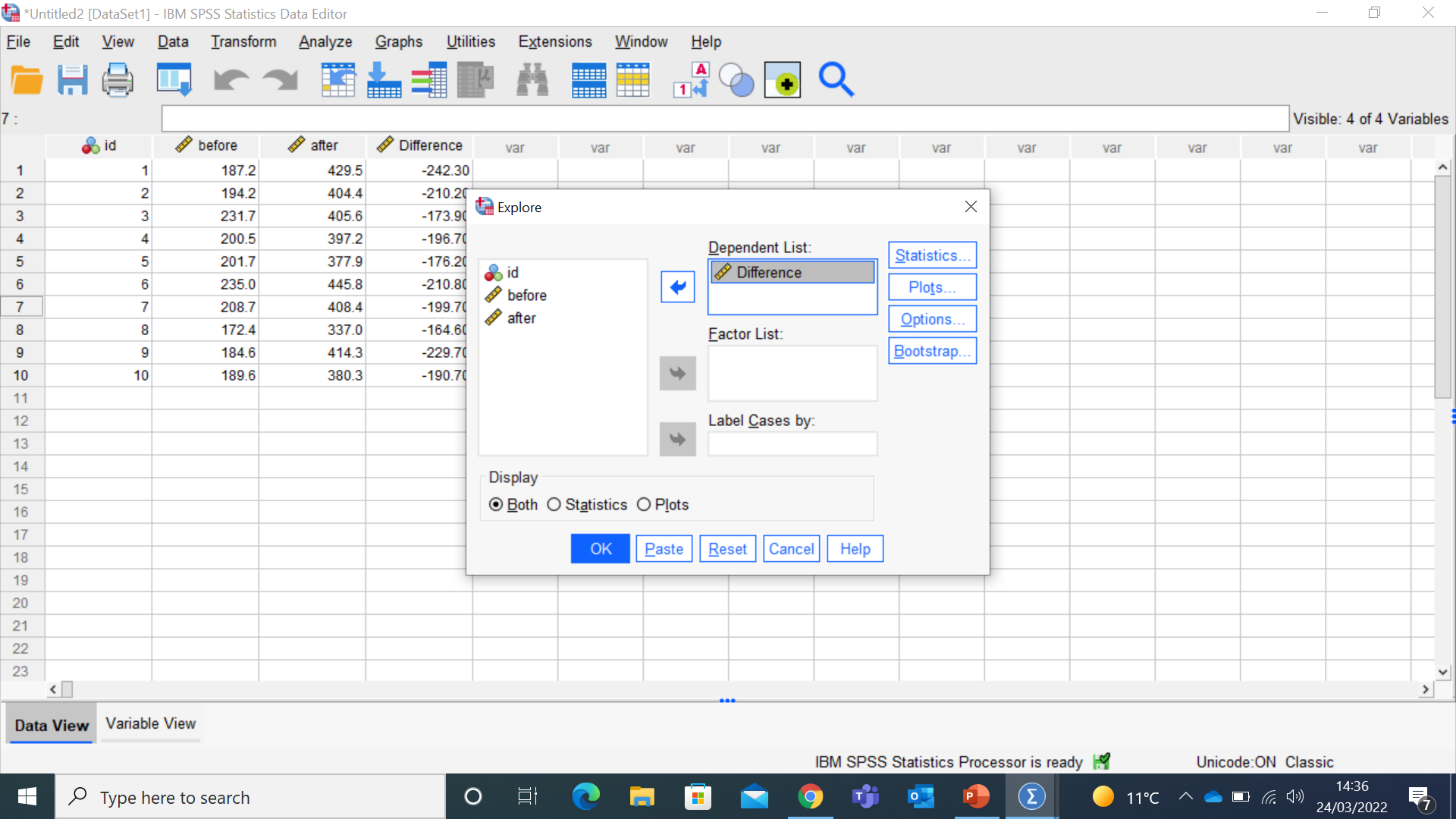


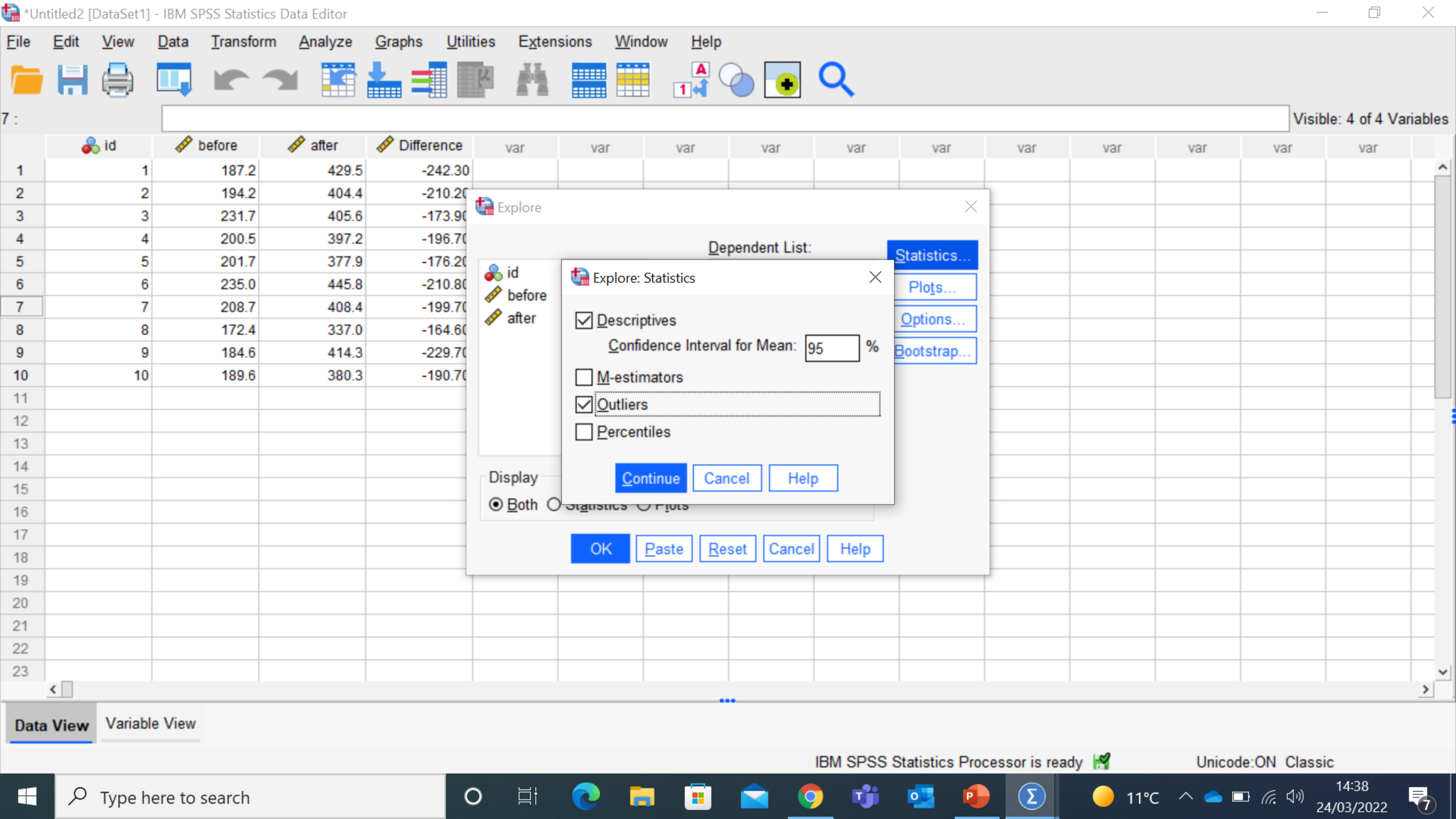
	id	before	after	Difference
1	1	187.2	429.5	-242.30
2	2	194.2	404.4	-210.20
3	3	231.7	405.6	-173.90
4	4	200.5	397.2	-196.70
5	5	201.7	377.9	-176.20
6	6	235.0	445.8	-210.80
7	7	208.7	408.4	-199.70
8	8	172.4	337.0	-164.60
9	9	184.6	414.3	-229.70
10	10	189.6	380.3	-190.70

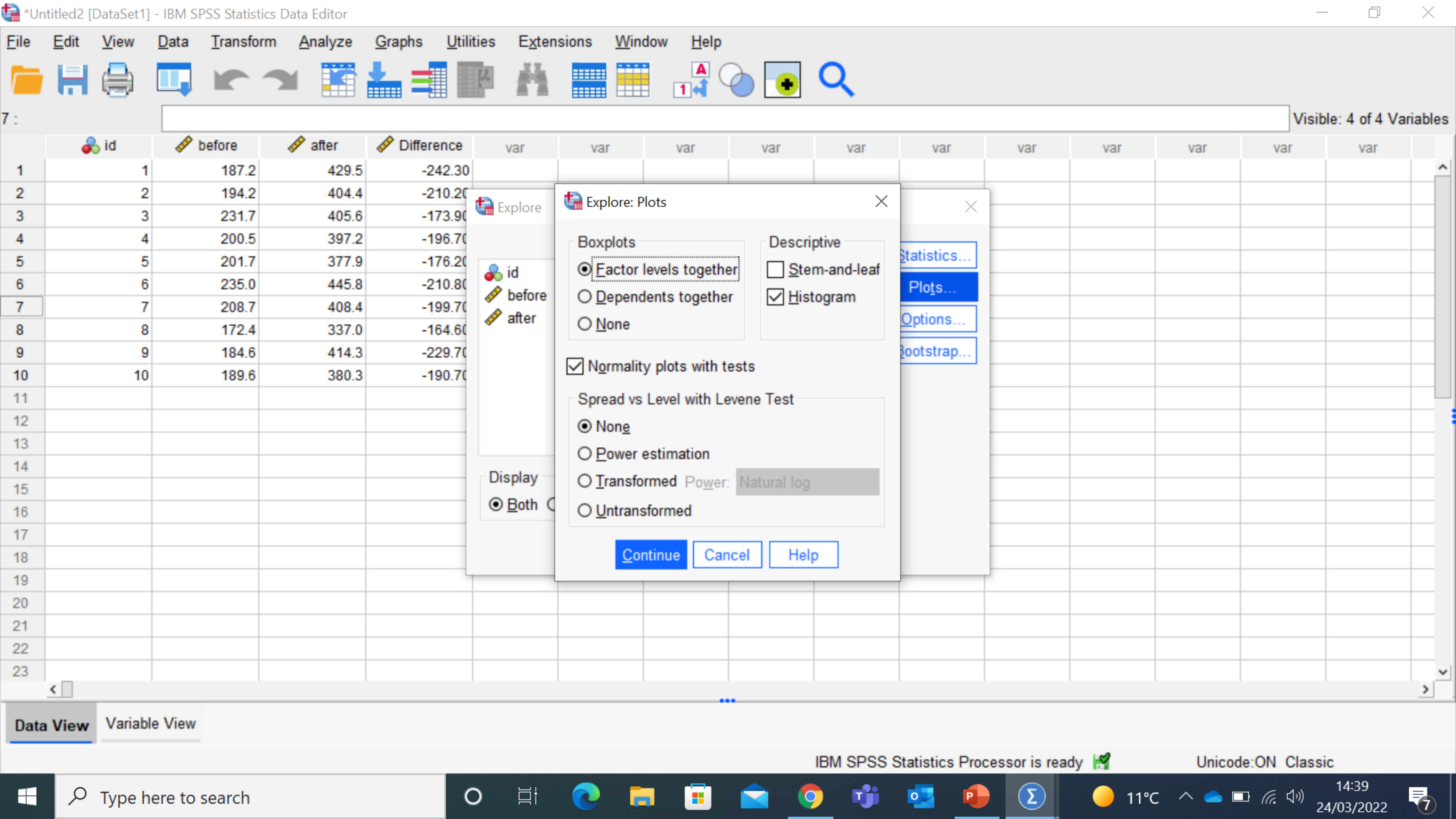
We can explore whether our assumptions are met using the same approach as we did before



We can explore whether our assumptions are met using the same approach as we did before

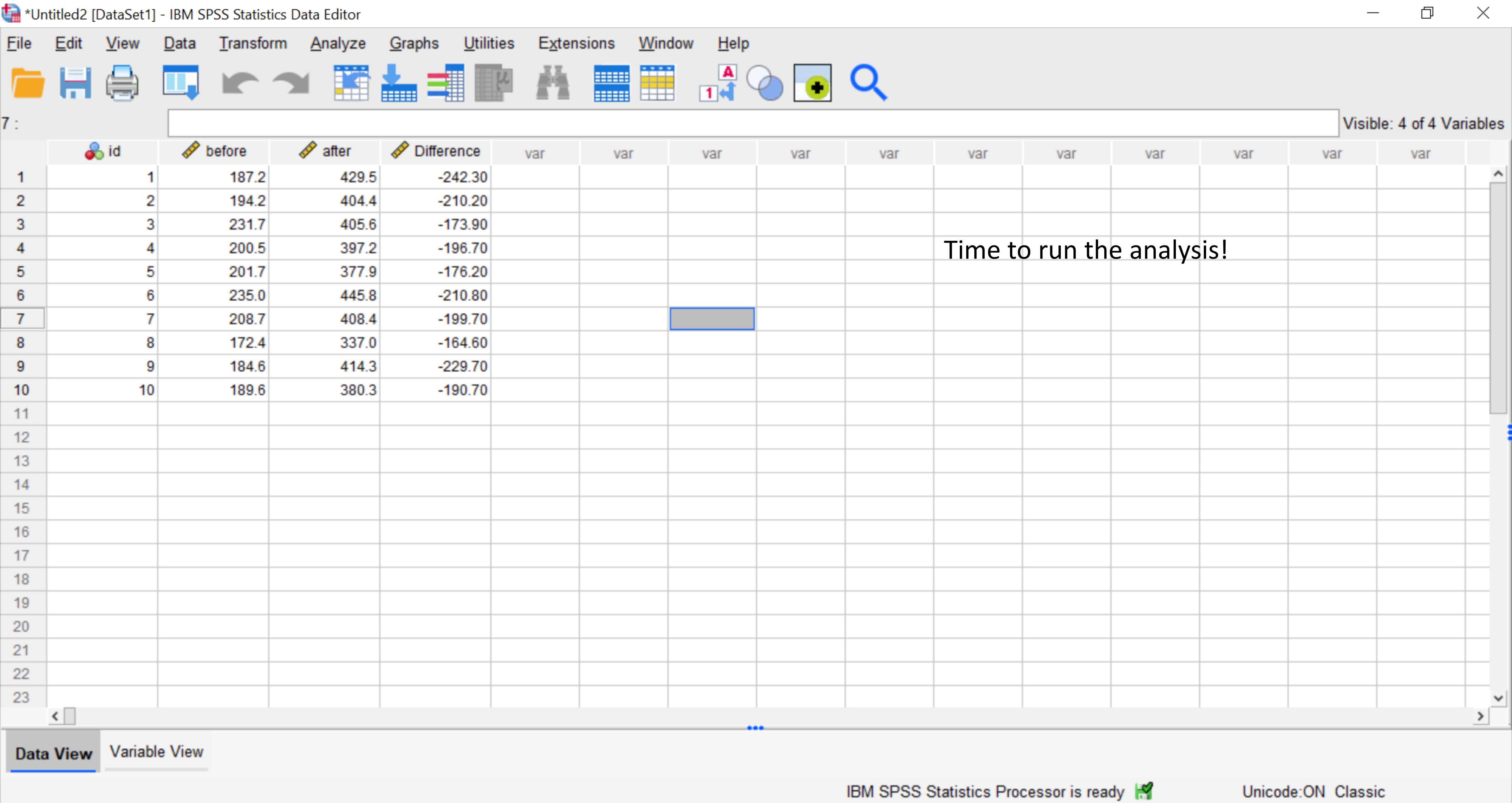

*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id before after Difference var var var var var var var var var var var
1 1 187.2 429.5 -242.30
2 2 194.2 404.4 -210.20
3 3 231.7 405.6 -173.90
4 4 200.5 397.2 -196.70
5 5 201.7 377.9 -176.20
6 6 235.0 445.8 -210.80
7 7 208.7 408.4 -199.70
8 8 172.4 337.0 -164.60
9 9 184.6 414.3 -229.70
10 10 189.6 380.3 -190.70
Explore
id
before
after
Dependent List:
Difference
Factor List:
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
14:36
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id before after Difference var var var var var var var var var var var
1 1 187.2 429.5 -242.30
2 2 194.2 404.4 -210.20
3 3 231.7 405.6 -173.90
4 4 200.5 397.2 -196.70
5 5 201.7 377.9 -176.20
6 6 235.0 445.8 -210.80
7 7 208.7 408.4 -199.70
8 8 172.4 337.0 -164.60
9 9 184.6 414.3 -229.70
10 10 189.6 380.3 -190.70
Explore
Dependent List:
Statistics...
Plots...
Options...
Bootstrap...
id
before
after
Explore: Statistics
Descriptives
Confidence Interval for Mean: 95 %
M-estimators
Outliers
Percentiles
Continue Cancel Help
Display
Both Statistics Plots
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
14:38
24/03/2022



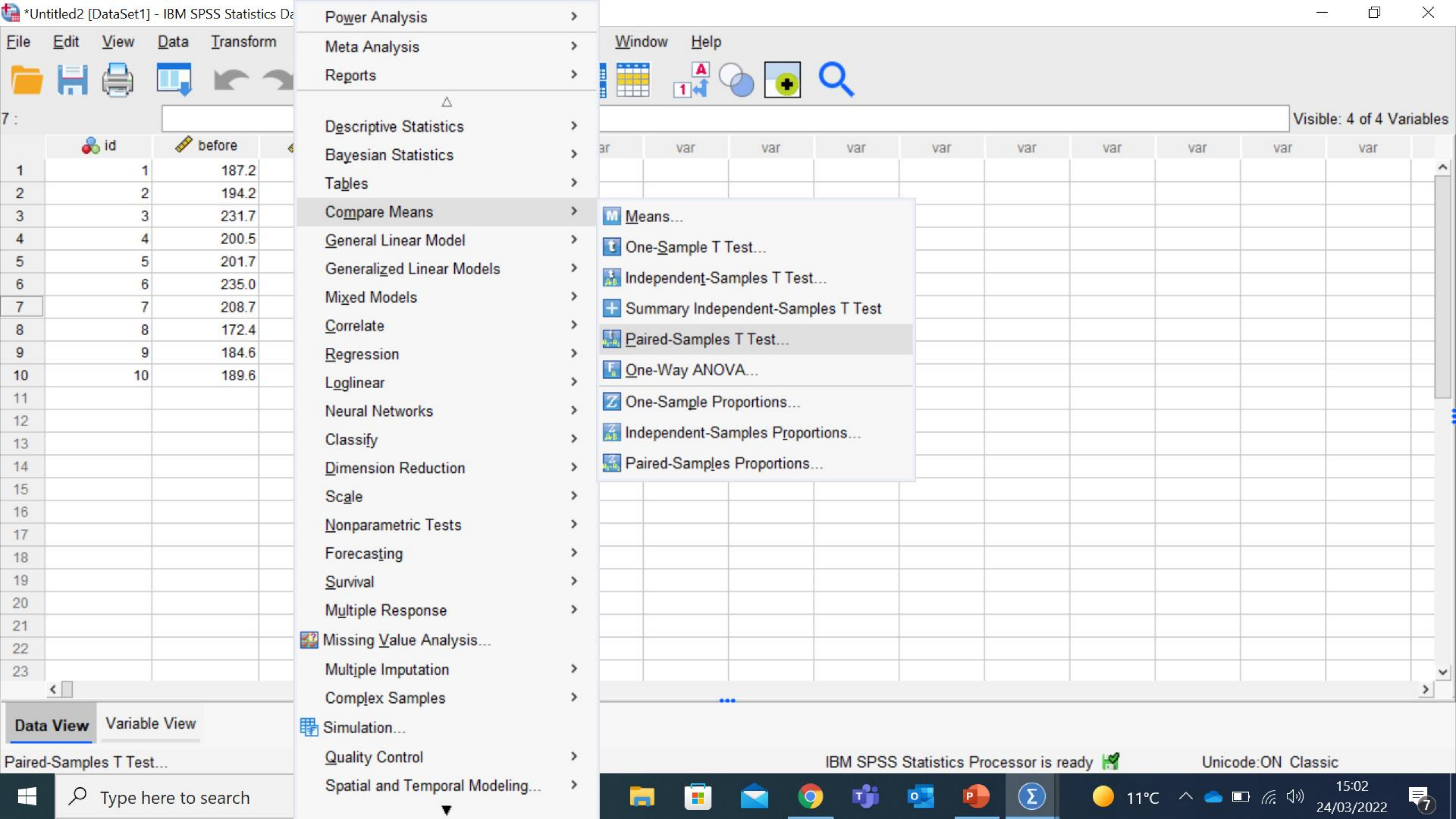
*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id before after Difference var var var var var var var var var var var
1 1 187.2 429.5 -242.30
2 2 194.2 404.4 -210.20
3 3 231.7 405.6 -173.90
4 4 200.5 397.2 -196.70
5 5 201.7 377.9 -176.20
6 6 235.0 445.8 -210.80
7 7 208.7 408.4 -199.70
8 8 172.4 337.0 -164.60
9 9 184.6 414.3 -229.70
10 10 189.6 380.3 -190.70
Explore
id
before
after
Display
Both
Statistics...
Plots...
Options...
Bootstrap...
Explore: Plots
Boxplots
Factor levels together
Dependents together
None
Descriptive
Stem-and-leaf
Histogram
Normality plots with tests
Spread vs Level with Levene Test
None
Power estimation
Transformed Power: Natural log
Untransformed
Continue Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
14:39
24/03/2022

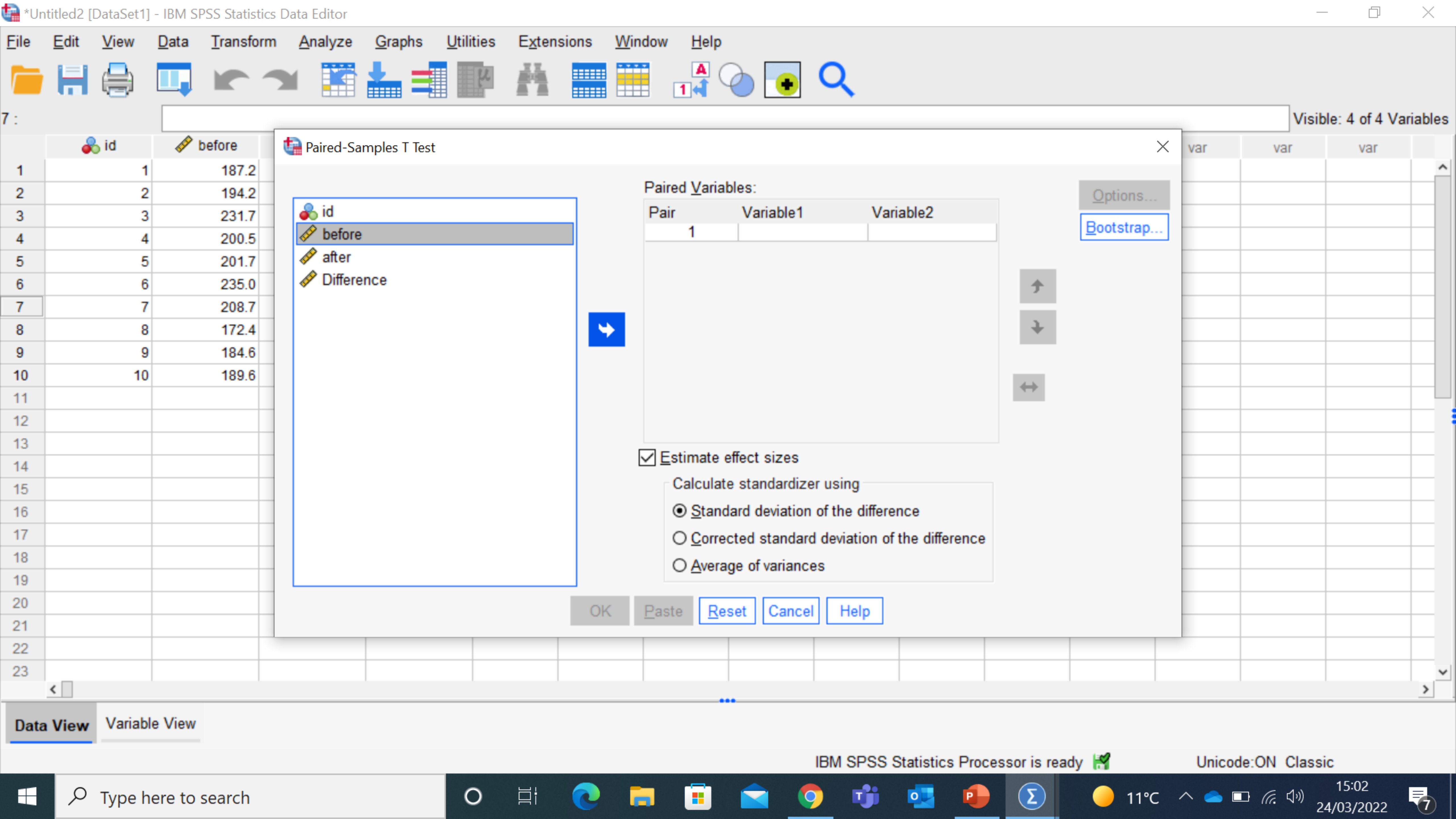
Let’s check our data together!

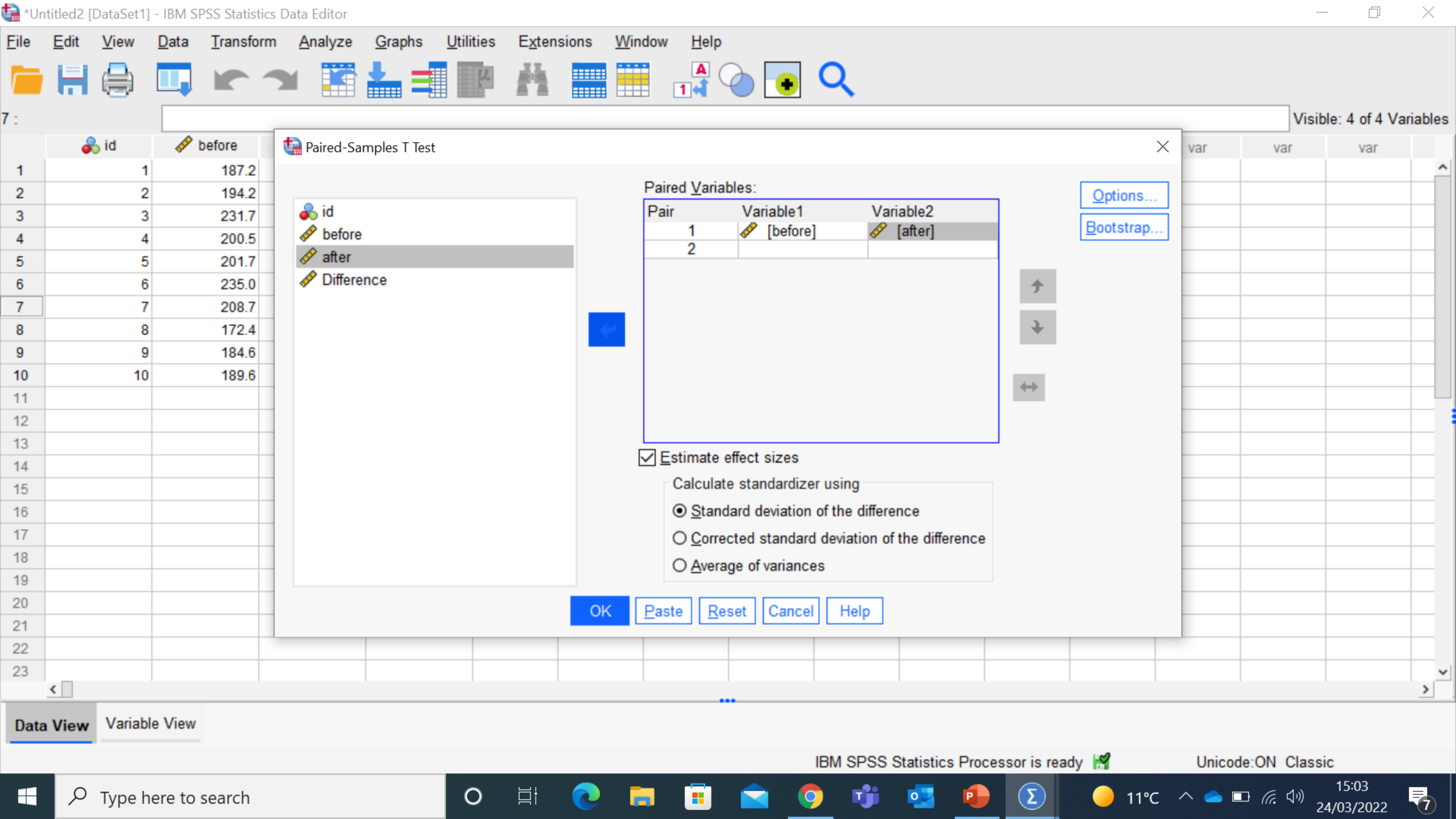


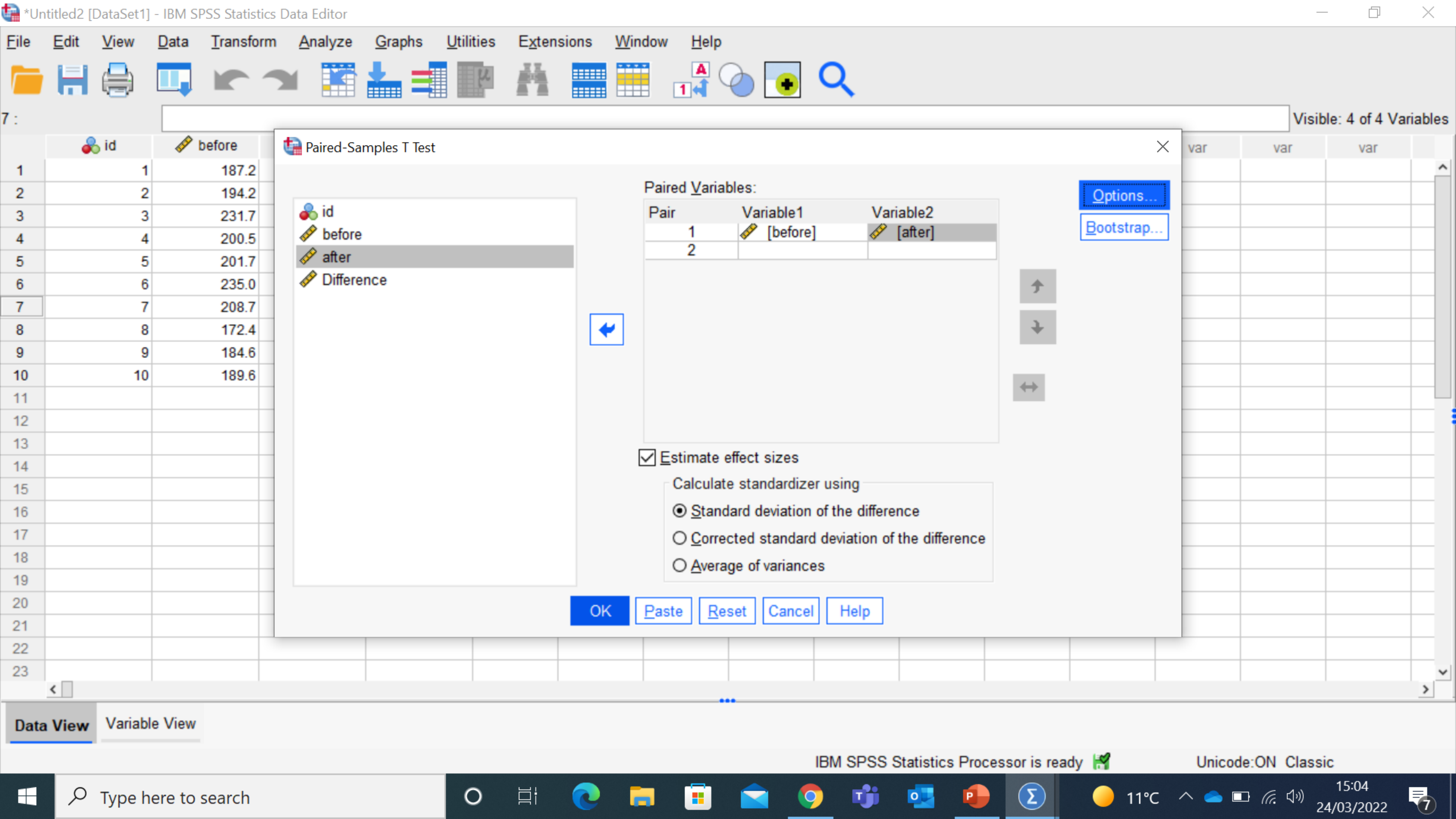
	id	before	after	Difference
1	1	187.2	429.5	-242.30
2	2	194.2	404.4	-210.20
3	3	231.7	405.6	-173.90
4	4	200.5	397.2	-196.70
5	5	201.7	377.9	-176.20
6	6	235.0	445.8	-210.80
7	7	208.7	408.4	-199.70
8	8	172.4	337.0	-164.60
9	9	184.6	414.3	-229.70
10	10	189.6	380.3	-190.70

Time to run the analysis!


*Untitled2 [DataSet1] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Means...
One-Sample T Test...
Independent-Samples T Test...
Summary Independent-Samples T Test
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
Visible: 4 of 4 Variables
id
before
var
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
Data View
Variable View
Paired-Samples T Test...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
15:02
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id
before
var var var
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
Paired-Samples T Test
id
before
after
Difference
Paired Variables:
Pair Variable1 Variable2
1
Options...
Bootstrap...
Estimate effect sizes
Calculate standardizer using
Standard deviation of the difference
Corrected standard deviation of the difference
Average of variances
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
15:02
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id
before
var
var
var
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
Paired-Samples T Test
id
before
after
Difference
Paired Variables:
Pair Variable1 Variable2
1 [before] [after]
2
Options...
Bootstrap...
Estimate effect sizes
Calculate standardizer using
Standard deviation of the difference
Corrected standard deviation of the difference
Average of variances
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
15:03
24/03/2022


*Untitled2 [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
7 :
Visible: 4 of 4 Variables
id
before
var
var
var
1 1 187.2
2 2 194.2
3 3 231.7
4 4 200.5
5 5 201.7
6 6 235.0
7 7 208.7
8 8 172.4
9 9 184.6
10 10 189.6
Paired-Samples T Test
id
before
after
Difference
Paired Variables:
Pair Variable1 Variable2
1 [before] [after]
2
Options...
Bootstrap...
Estimate effect sizes
Calculate standardizer using
Standard deviation of the difference
Corrected standard deviation of the difference
Average of variances
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
11°C
15:04
24/03/2022

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
T-Test
Title
Notes
Paired Samples St
Paired Samples C
Paired Samples Te
Paired Samples Ef

T-Test

The means for each group…

Worth noting as we should report this!

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	200.560	10	20.0282	6.3335
	after	400.040	10	30.0869	9.5143

Paired Samples Correlations

				Significance	
		N	Correlation	One-Sided p	Two-Sided p
Pair 1	before & after	10	.578	.040	.080

Paired Samples Test

		Paired Differences							Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	One-Sided p	Two-Sided p
Pair 1	before - after	-199.4800	24.6929	7.8086	-217.1442	-181.8158	-25.546	9	<.001	<.001

Paired Samples Effect Sizes

					95% Confidence Interval	
			Standardizer[a]	Point Estimate	Lower	Upper
Pair 1	before - after	Cohen's d	24.6929	-8.078	-11.791	-4.366
		Hedges' correction	27.0200	-7.383	-10.776	-3.990

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 9°C 19:50 22/03/2022

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
T-Test
Title
Notes
Paired Samples St
Paired Samples C
Paired Samples Te
Paired Samples Ef

T-Test

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	200.560	10	20.0282	6.3335
	after	400.040	10	30.0869	9.5143

Paired Samples Correlations

		N	Correlation	Significance	
				One-Sided p	Two-Sided p
Pair 1	before & after	10	.578	.040	.080

- First column: The pair of variables you are testing. The order the subtraction is indicated.
- **Mean:** The average difference between the two treatments → also a good thing to report
- **Standard deviation:** The standard deviation of the difference scores.
- **Standard error mean:** The standard error (standard deviation divided by the square root of the sample size). Used in computing the test statistic

Paired Samples Test

		Paired Differences					t	df	Significance	
					95% Confidence Interval of the Difference					
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			One-Sided p	Two-Sided p
Pair 1	before - after	-199.4800	24.6929	7.8086	-217.1442	-181.8158	-25.546	9	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval	
					Lower	Upper
Pair 1	before - after	Cohen's d	24.6929	-8.078	-11.791	-4.366
		Hedges' correction	27.0200	-7.383	-10.776	-3.990

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 9°C 19:50 22/03/2022

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
T-Test
Title
Notes
Paired Samples St
Paired Samples C
Paired Samples Te
Paired Samples Ef

T-Test

- **t:** The test statistic
- **df:** The degrees of freedom
- **Sig. (2-tailed)**: The *p*-value corresponding to the given test statistic *t* with degrees of freedom *df*.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	200.560	10	20.0282	6.3335
	after	400.040	10	30.0869	9.5143

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	before & after	10	.578	.040	.080

Paired Samples Test

		Paired Differences Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Significance One-Sided p	Two-Sided p
Pair 1	before - after	-199.4800	24.6929	7.8086	-217.1442	-181.8158	-25.546	9	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	Upper
Pair 1	before - after	Cohen's d	24.6929	-8.078	-11.791	-4.366
		Hedges' correction	27.0200	-7.383	-10.776	-3.990

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.

IBM SPSS Statistics Processor is ready Unicode:ON Classic

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
T-Test
Title
Notes
Paired Samples St
Paired Samples C
Paired Samples Te
Paired Samples Ef

T-Test

• **Cohen's d: our measurement of effect size**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	200.560	10	20.0282	6.3335
	after	400.040	10	30.0869	9.5143

Paired Samples Correlations

		N	Correlation	Significance One-Sided p	Significance Two-Sided p
Pair 1	before & after	10	.578	.040	.080

Paired Samples Test

		Paired Differences Mean	Paired Differences Std. Deviation	Paired Differences Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Significance One-Sided p	Significance Two-Sided p
Pair 1	before - after	-199.4800	24.6929	7.8086	-217.1442	-181.8158	-25.546	9	<.001	<.001

Paired Samples Effect Sizes

			Standardizer[a]	Point Estimate	95% Confidence Interval Lower	95% Confidence Interval Upper
Pair 1	before - after	Cohen's d	24.6929	-8.078	-11.791	-4.366
		Hedges' correction	27.0200	-7.383	-10.776	-3.990

a. The denominator used in estimating the effect sizes.
Cohen's d uses the sample standard deviation of the mean difference.

IBM SPSS Statistics Processor is ready Unicode:ON Classic

Type here to search 9°C 19:50 22/03/2022

We have a significant result! So what do we report?

- The average weight of the 10 rats in our sample before the food court opened up was 200.56g (SD=20.03), and increased to 400.04g (SD=30.09) after the food court opened. A paired samples T Test revealed this difference to be significant, $t(9) = 25.55$, $p < 0.001$, $d = 8.078$.

- Notice I removed the negative from cohen’s d. This is fine

If normality isn't met

- Just like with our 2 samples T Test, there is a non-parametric equivalent for when our data do not meet the required assumption of normality

- Lets open the data 'rats3.xlsx'

*Untitled8 [DataSet7] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	id	Before	After
1	1	200.1	200.5
2	2	201.9	200.8
3	3	192.7	198.7
4	4	216.1	227.7
5	5	170.0	271.8
6	6	196.9	265.5
7	7	172.2	259.6
8	8	185.5	288.5
9	9	160.3	249.0
10	10	193.7	189.8

Let's say you run a similar experiment in a different town with a different food court (rats were different to start off with, and the choice of garbage they can eat after the mall opens up is different)

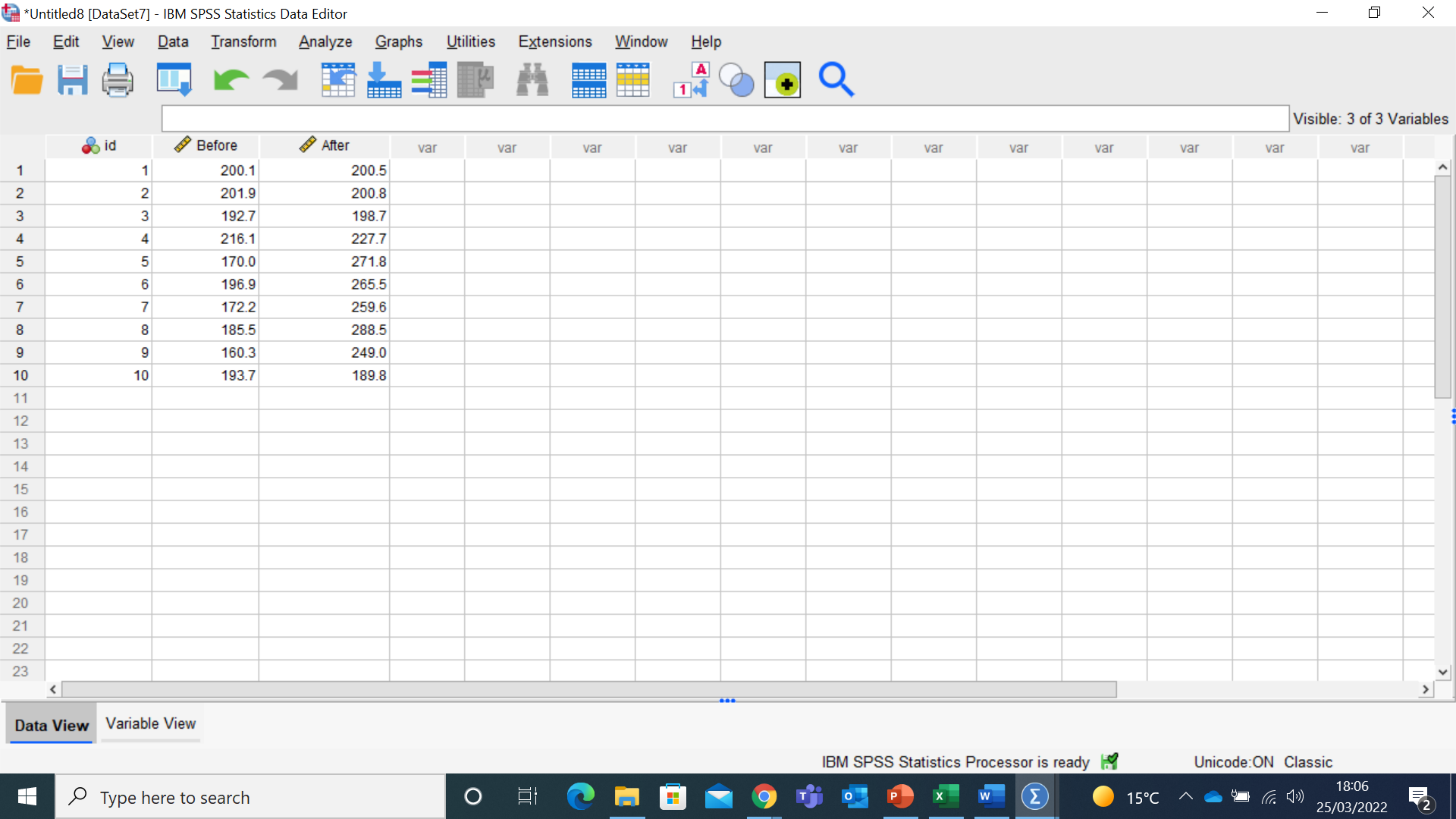
Let's run the same diagnostics to check for normality as we have been in the past…

Data View Variable View

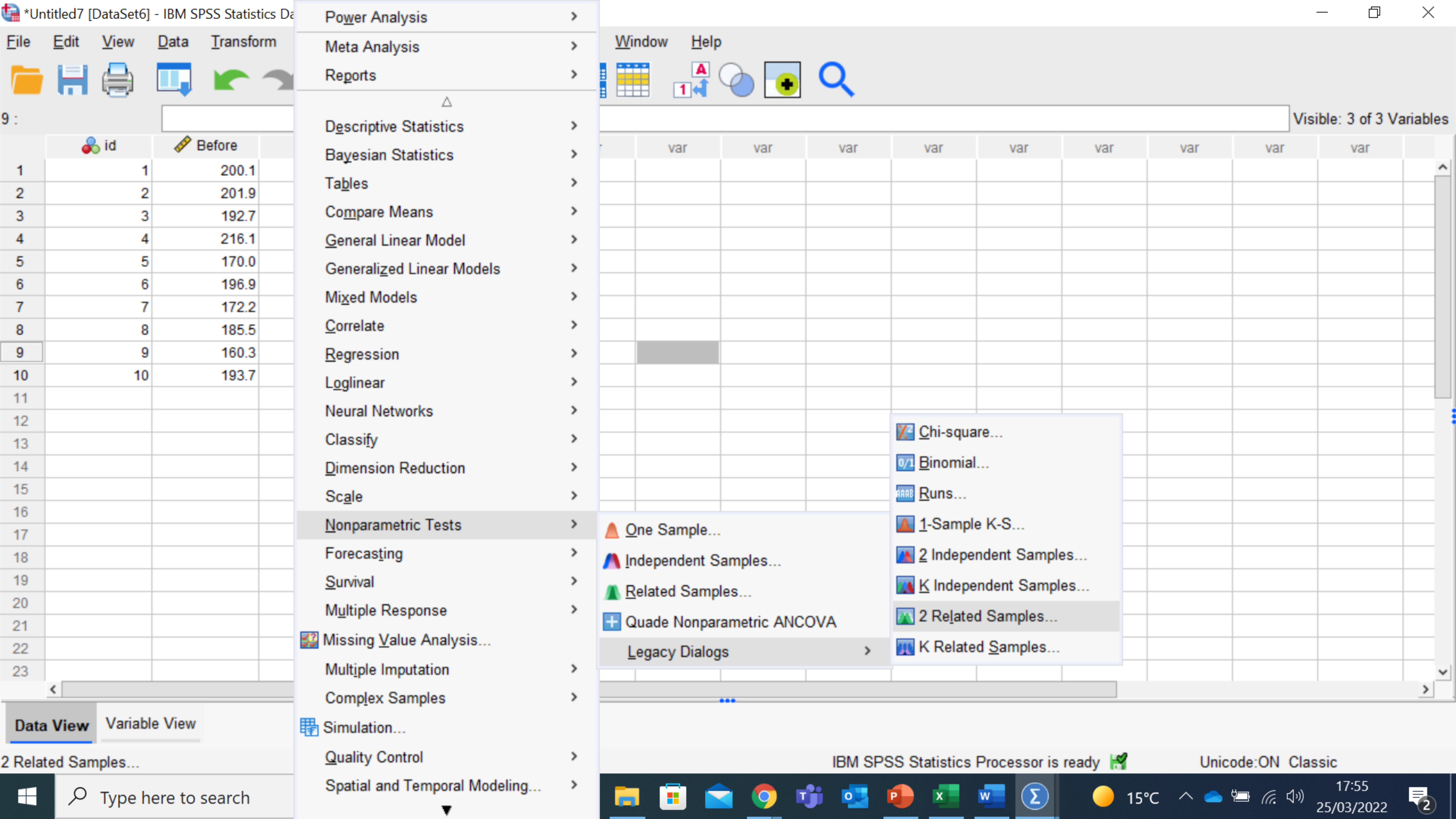
IBM SPSS Statistics Processor is ready Unicode:ON Classic

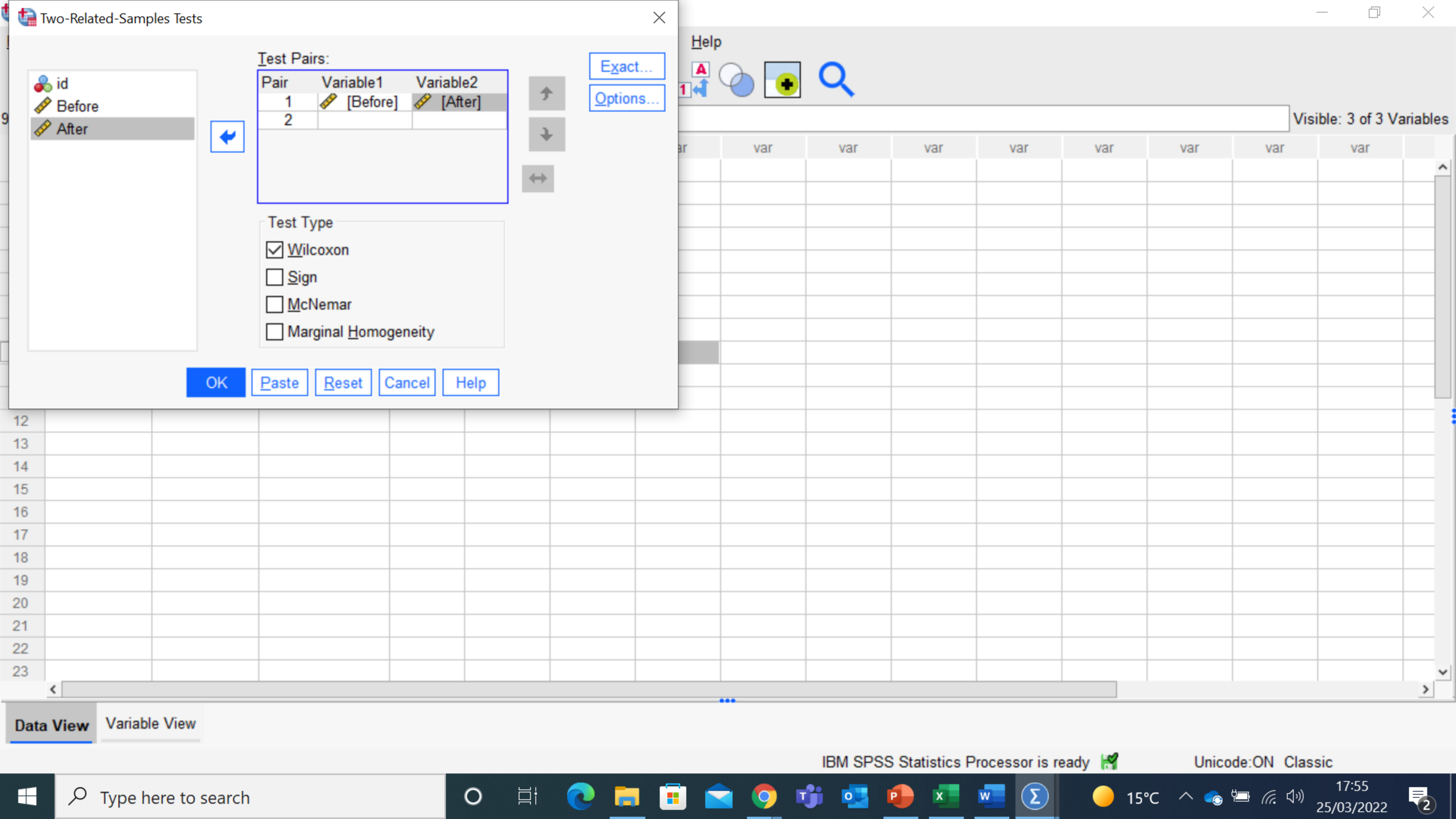
- **the difference between the two groups** are not normally distributed
- This means we cannot use a paired samples T Test!
- We do have a non-parametric equivalent though!

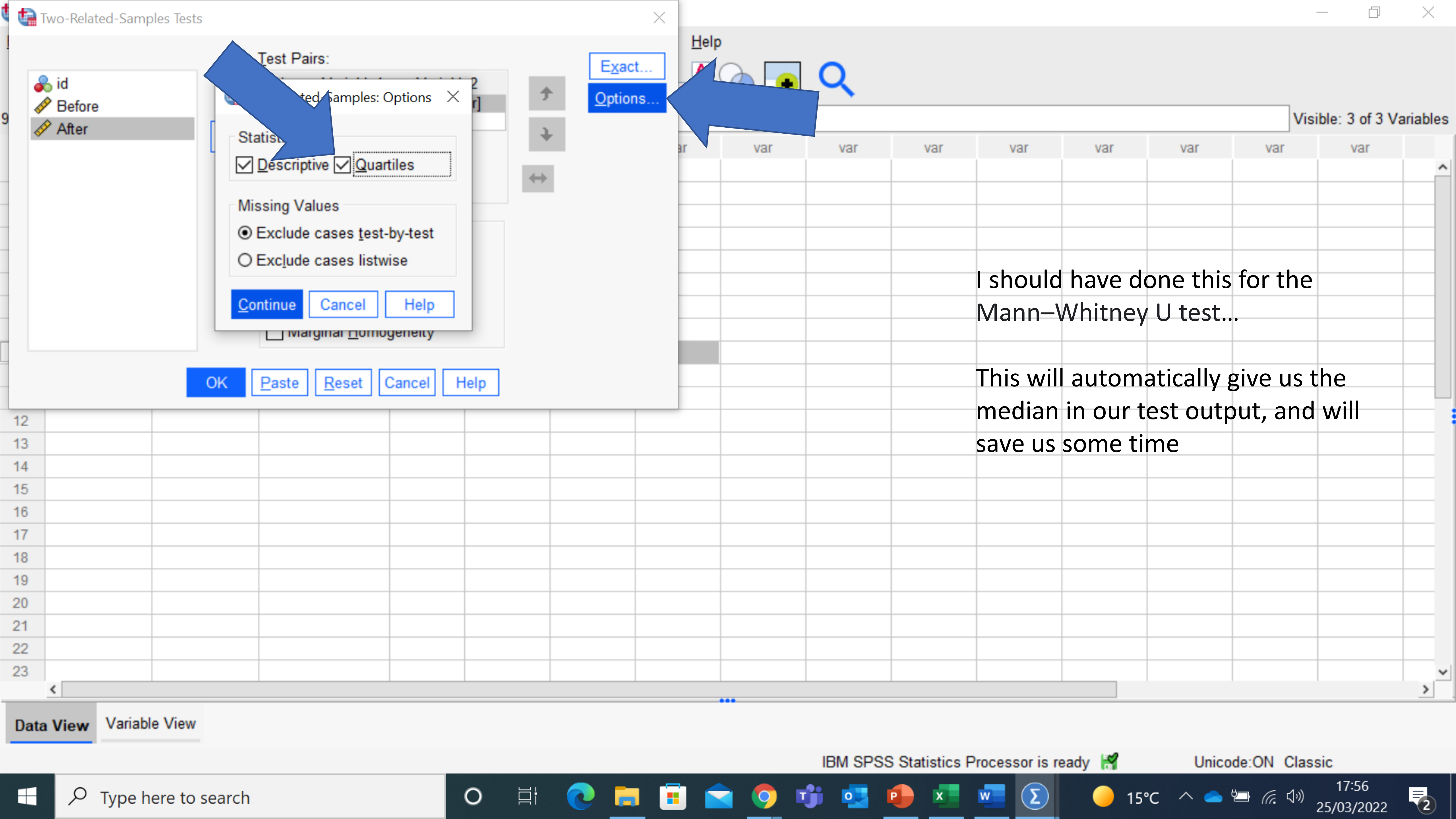
The **Wilcoxon signed rank test for paired samples**



	id	Before	After
1	1	200.1	200.5
2	2	201.9	200.8
3	3	192.7	198.7
4	4	216.1	227.7
5	5	170.0	271.8
6	6	196.9	265.5
7	7	172.2	259.6
8	8	185.5	288.5
9	9	160.3	249.0
10	10	193.7	189.8

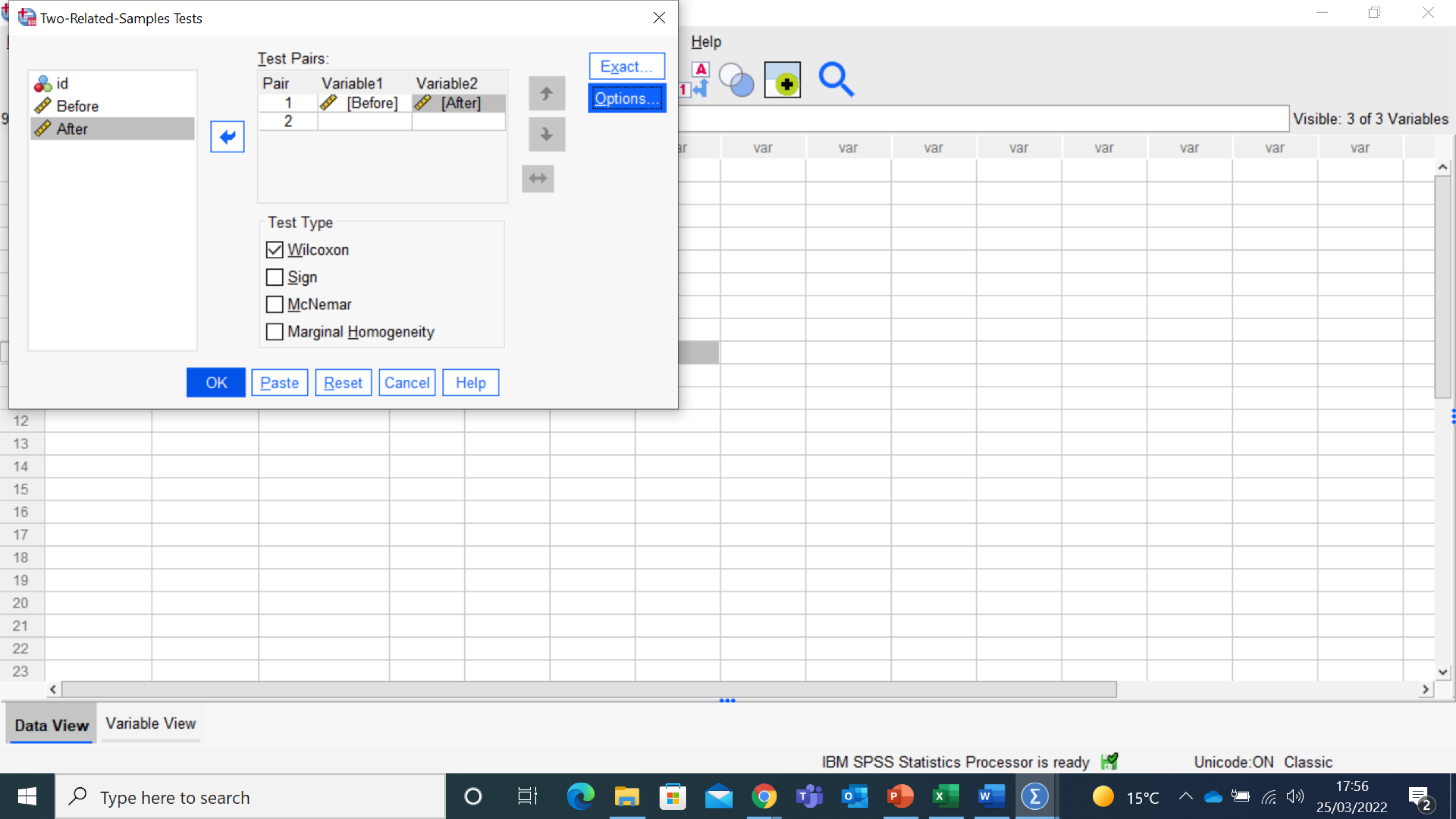

*Untitled7 [DataSet6] - IBM SPSS Statistics Da
File Edit View Data Transform
Window Help
Visible: 3 of 3 Variables
id Before
1 200.1
2 201.9
3 192.7
4 216.1
5 170.0
6 196.9
7 172.2
8 185.5
9 160.3
10 193.7
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
One Sample...
Independent Samples...
Related Samples...
Quade Nonparametric ANCOVA
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...
Data View Variable View
2 Related Samples...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
15°C
17:55
25/03/2022


Two-Related-Samples Tests
Test Pairs:
Pair
Variable1
Variable2
1
[Before]
[After]
2
id
Before
After
Exact...
Options...
Test Type
Wilcoxon
Sign
McNemar
Marginal Homogeneity
OK
Paste
Reset
Cancel
Help
Visible: 3 of 3 Variables
var
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
15°C
17:55
25/03/2022



I should have done this for the Mann–Whitney U test…

This will automatically give us the median in our test output, and will save us some time


Two-Related-Samples Tests
id
Before
After
Test Pairs:
Pair
Variable1
Variable2
1
[Before]
[After]
2
Exact...
Options...
Test Type
Wilcoxon
Sign
McNemar
Marginal Homogeneity
OK
Paste
Reset
Cancel
Help
Visible: 3 of 3 Variables
var
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
15°C
17:56
25/03/2022

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - NPar Tests
    - Title
    - Notes
    - Descriptive Statisti
    - Wilcoxon Signed R
      - Title
      - Ranks
      - Test Statistics

NPar Tests

Descriptive Statistics

	N	Mean	Deviation	Minimum	Maximum	Percentiles 25th	50th (Median)	75th
Before	10	188.94	17.0012	160.3	216.1	171.650	193.200	200.550
After	10	235.190	36.0885	189.8	288.5	200.050	238.350	267.075

Wilcoxon Signed Ranks T

Ranks

		N	Mean Rank	Sum of Ranks
After - Before	Negative Ranks	2[a]	2.50	5.00
	Positive Ranks	8[b]	6.25	50.00
	Ties	0[c]		
	Total	10		

a. After < Before

b. After > Before

c. After = Before

Test Statistics[a]

	After - Before
Z	-2.293[b]
Asymp. Sig. (2-tailed)	.022

a. Wilcoxon Signed Ranks Test

The **Ranks** table provides data on the comparison of participants' Before and After weight.

Notice legend at the bottom of table:

2 rats had lower weights after the treatment vs before

8 rats had higher weights after the treatment vs before

0 rats maintained the same weight

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

Type here to search   14°C   18:08 25/03/2022

*Output9 [Document9] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - NPar Tests
    - Title
    - Notes
    - Descriptive Statisti
    - Wilcoxon Signed R
      - Title
      - Ranks
      - Test Statistics

NPar Tests

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Before	10	188.940	17.0012	160.3	216.1	171.650	193.200	200.550
After	10	235.190	36.0885	189.8	288.5	200.050	238.350	267.075

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
After - Before	Negative Ranks	2[a]	2.50	5.00
	Positive Ranks	8[b]	6.25	50.00
	Ties	0[c]		
	Total	10		

a. After < Before

b. After > Before

c. After = Before

Test Statistics[a]

	After - Before
Z	-2.293[b]
Asymp. Sig. (2-tailed)	.022

a. Wilcoxon Signed Ranks Test

It seems we have a significant p value, suggesting to us there was a significant change in weight between our two sampling periods!

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

Type here to search   14°C   18:08 25/03/2022

What do we report?

- Our data indicate that rats in the neighbourhood were heavier after the vegan restaurant opened (MDN = 238.35) compared to baseline measurements beforehand (MDN = 193.20). A Wilcoxon Signed-ranks test indicated that there was a significant difference in rat weight across the two sample periods, $z = -2.293$, $p = 0.022$.

What if we don’t have a continuous response variable?

- In all four of the statistical analyses we’ve looked at so far, we have been looking at the effects of a dichotomous independent variable (2 classes) on a continuous dependant variable
- Effect of sex (2 levels) on body weight (continuous)
- Effect of sex (2 levels) on number of aggressive interactions over a study period (continuous)
- Effect of sampling period (**before** vs **after** restaurant opening) on body weight (continuous)

What if we don’t have a continuous response variable?

- What if we are interested in how a categorical variable effects another categorical variable?

Example

- You are running a natural experiment:
- You have 2 age classes of arboreal monkeys (Capuchin monkeys in Costa Rica - immature and mature)
- You are interested whether immature and mature individuals bias their foraging to different parts of a tree crown
  - Maybe smaller immatures want to avoid being on the edge of the tree crown, as birds of prey are adept at catching small, inexperienced youngsters
  - Alternatively, the it may be better for smaller animals to be at the edge of the tree crown, because there is less feeding competition at the edge from larger monkeys that don’t allow them to eat enough!

- As a wildlife biologist, you are extremely interested in knowing how age affects where these monkeys are foraging.
- You can easily identify whether a monkey is an adult versus as immature based on the characteristics of their genitalia
- You also come up with a rudimentary measure of whether the monkey is at the ‘centre’ of a tree crown, or at the ‘edge’

- Let’s open the data ‘Monkey_tree1.xlsx’

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

Visible: 3 of 3 Variables

	V1	Age	Position	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	51	Imm	Edge														
2	52	Imm	Edge														
3	53	Imm	Edge														
4	54	Imm	Center														
5	55	Imm	Edge														
6	56	Imm	Center														
7	57	Imm	Edge														
8	58	Imm	Center														
9	59	Imm	Edge														
10	60	Imm	Center														
11	61	Imm	Center														
12	62	Imm	Edge														
13	63	Imm	Edge														
14	64	Imm	Edge														
15	65	Imm	Center														
16	66	Imm	Edge														
17	67	Imm	Center														
18	68	Imm	Edge														
19	69	Imm	Edge														
20	70	Imm	Center														
21	71	Imm	Edge														
22	72	Imm	Edge														

Data View | Variable View

IBM SPSS Statistics Processor is ready   Unicode:ON Classic

Type here to search   6°C   12:59 31/03/2022

Assumptions

- **Assumption 1: Both variables are categorical**
  - **They sure are in this case!**
- **Assumption 2: All observations are independent**
  - **The value of one measurement does not depend on the other**
- **Assumption 3: Expected value of cells should be 5 or greater in at least 80% of cells.**
  - **We will check this on our output**

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help

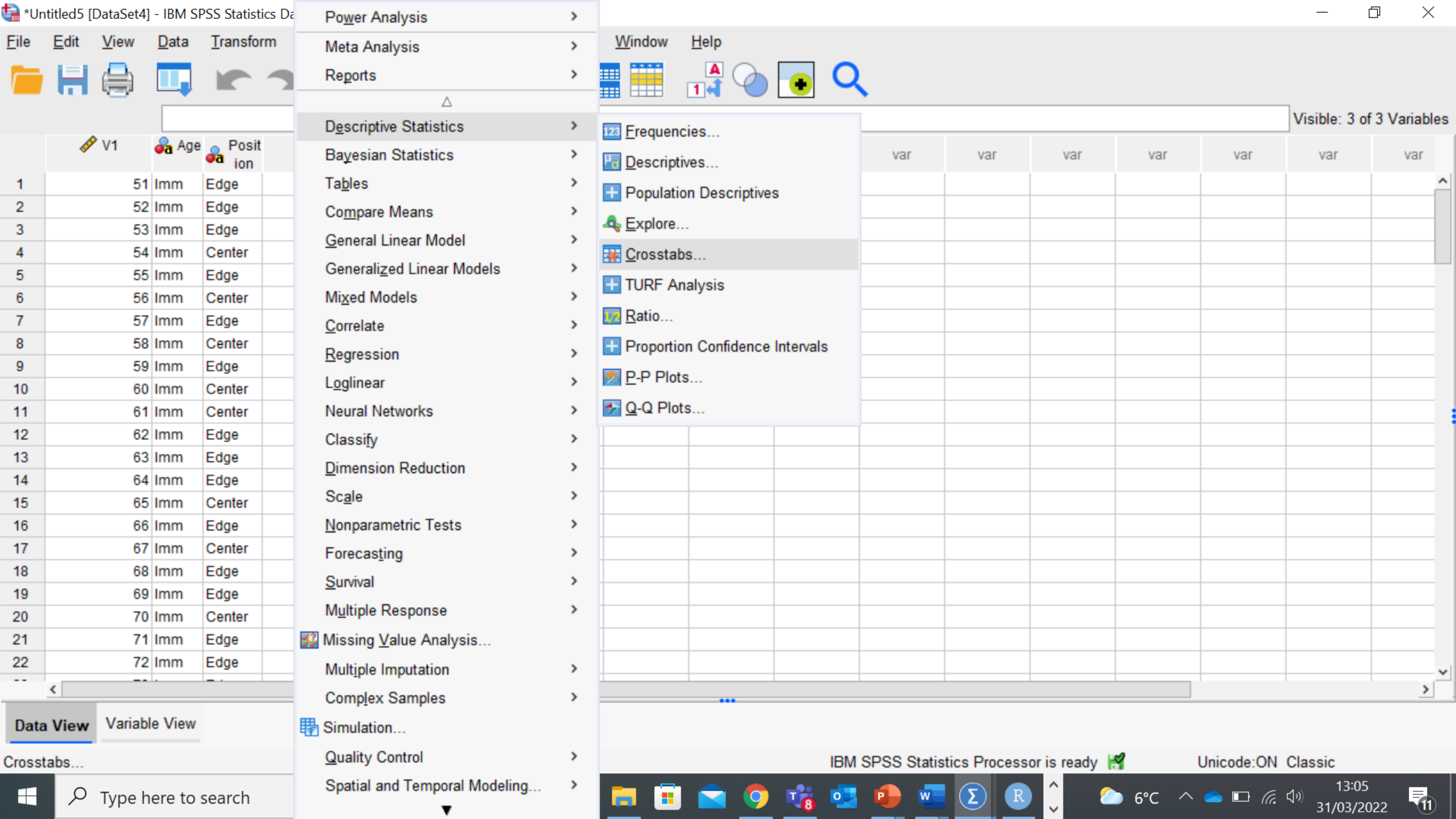
Visible: 3 of 3 Variables

	V1	Age	Position	var	var	var	var	var	var	var	var	var	var	var	var	var	var
1	51	Imm	Edge														
2	52	Imm	Edge														
3	53	Imm	Edge														
4	54	Imm	Center														
5	55	Imm	Edge														
6	56	Imm	Center														
7	57	Imm	Edge														
8	58	Imm	Center														
9	59	Imm	Edge														
10	60	Imm	Center														
11	61	Imm	Center														
12	62	Imm	Edge														
13	63	Imm	Edge														
14	64	Imm	Edge														
15	65	Imm	Center														
16	66	Imm	Edge														
17	67	Imm	Center														
18	68	Imm	Edge														
19	69	Imm	Edge														
20	70	Imm	Center														
21	71	Imm	Edge														
22	72	Imm	Edge														

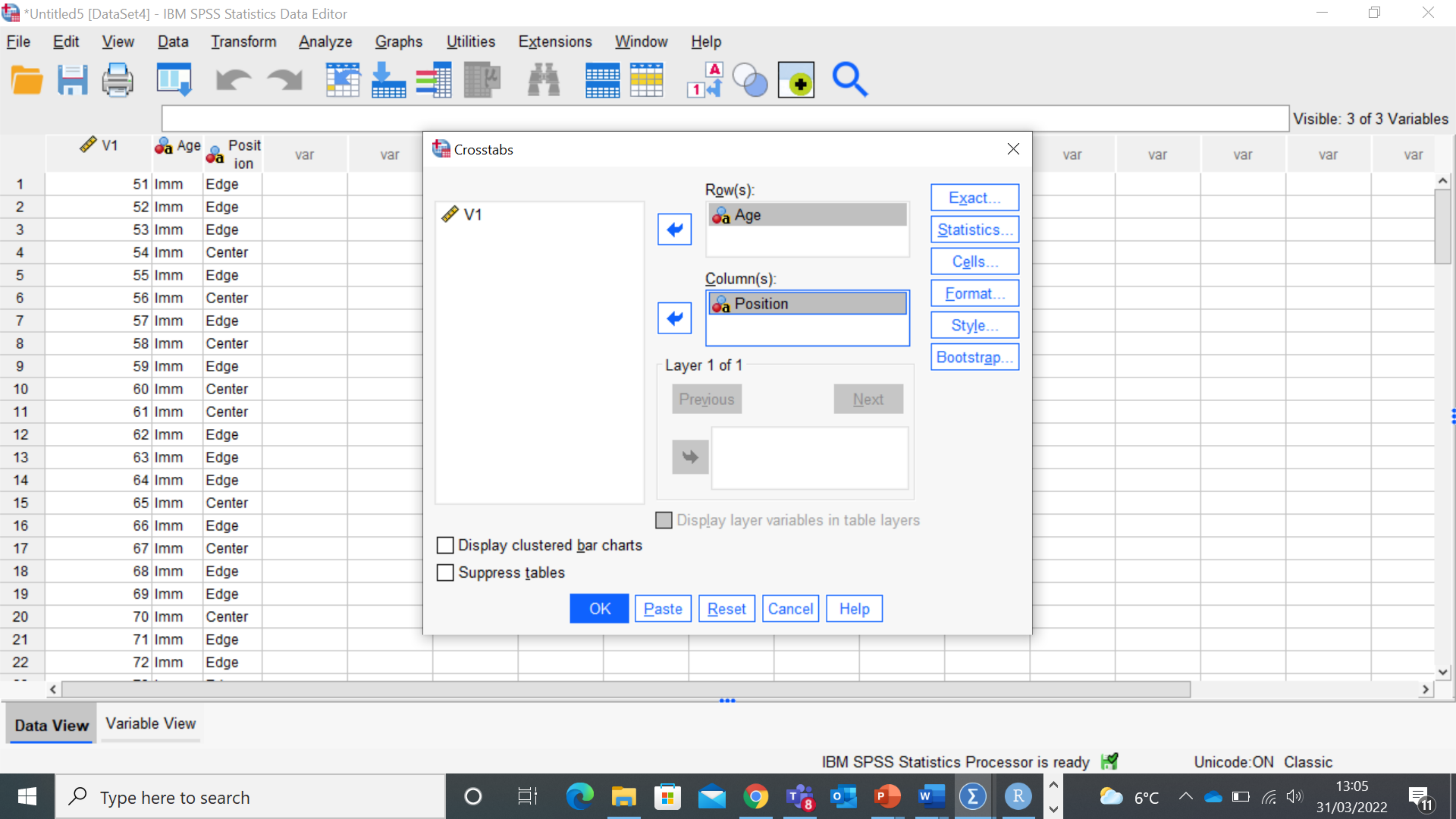
Data View Variable View

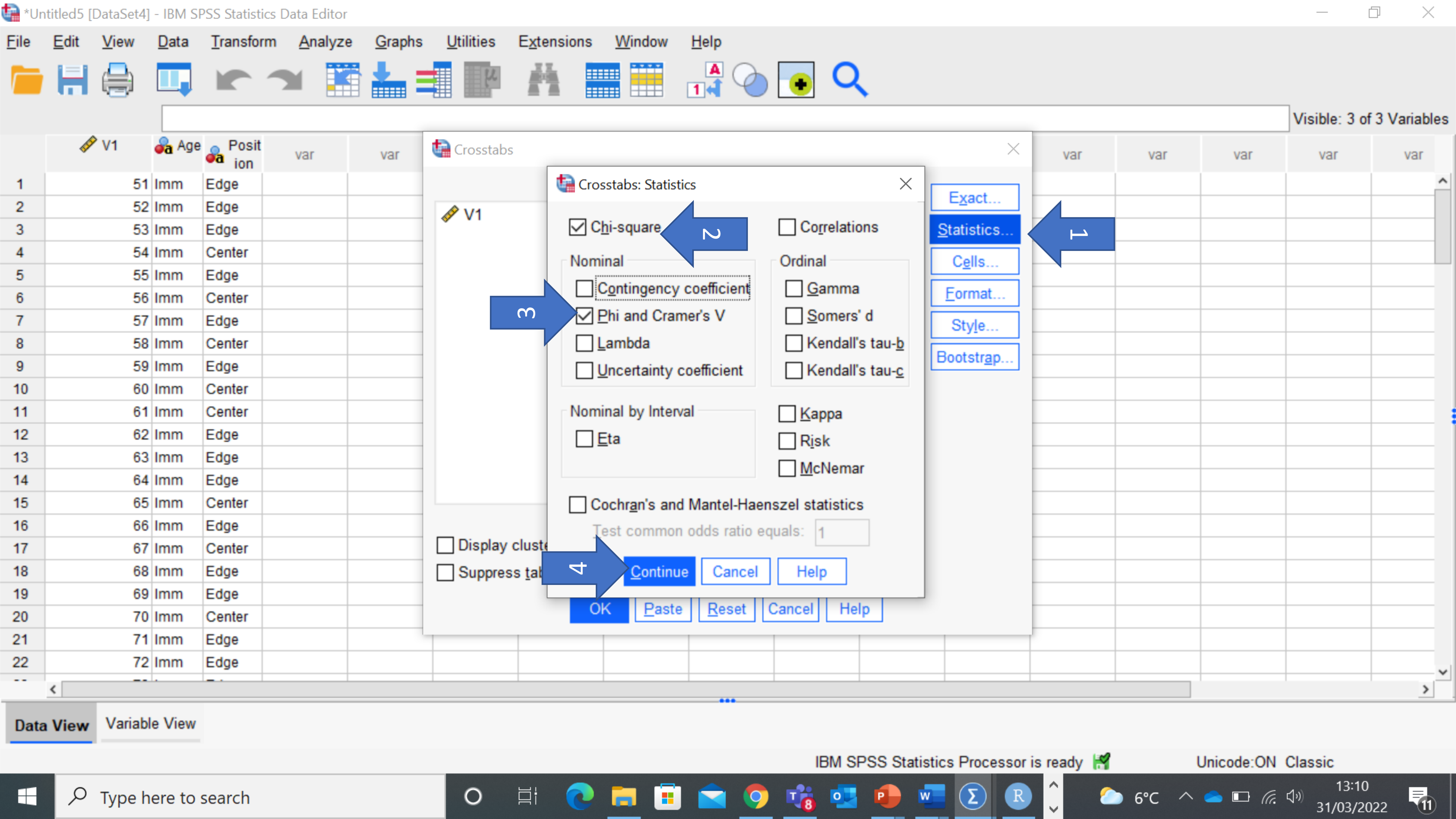
IBM SPSS Statistics Processor is ready Unicode:ON Classic

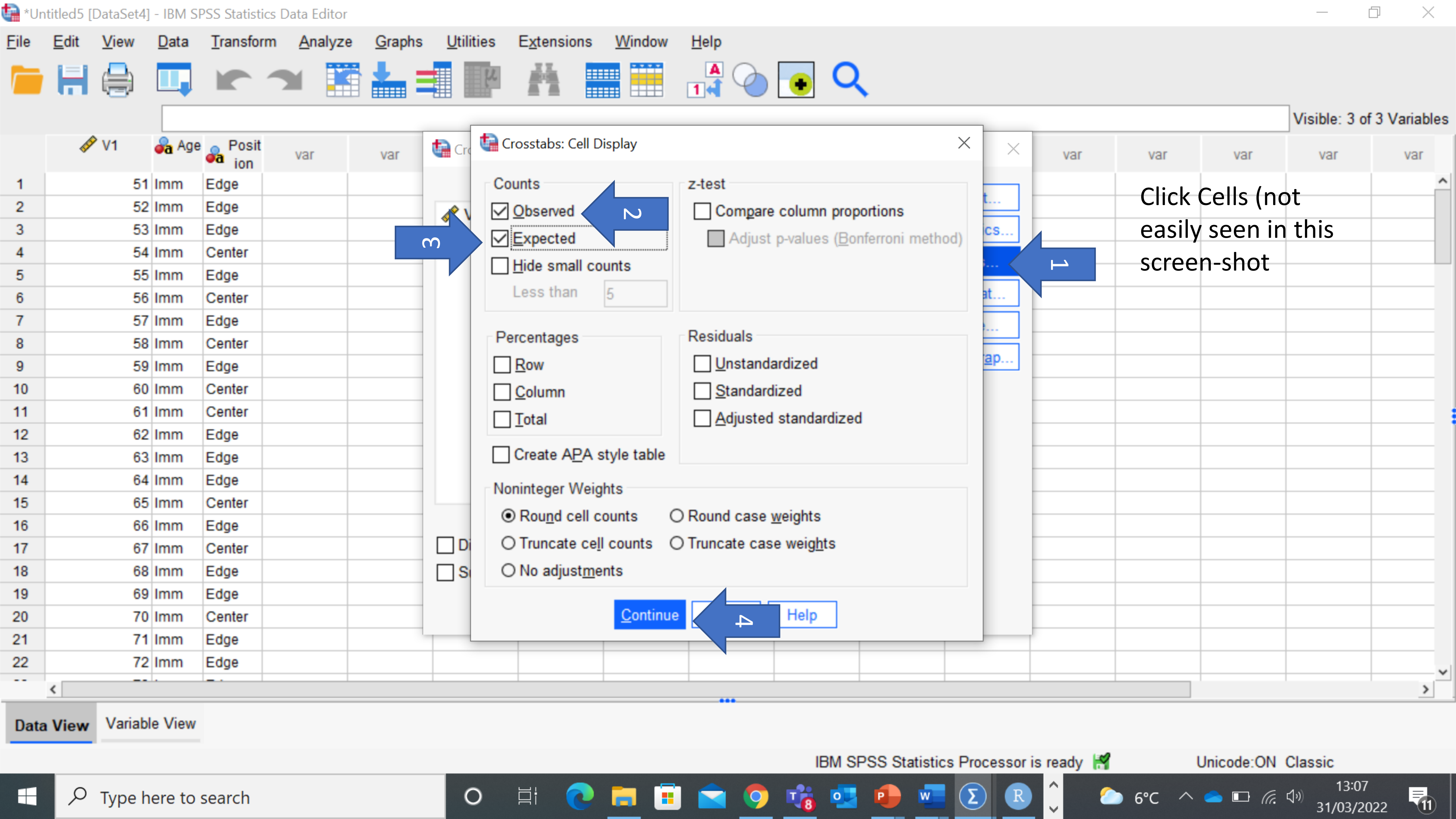
Type here to search 6°C 12:59 31/03/2022

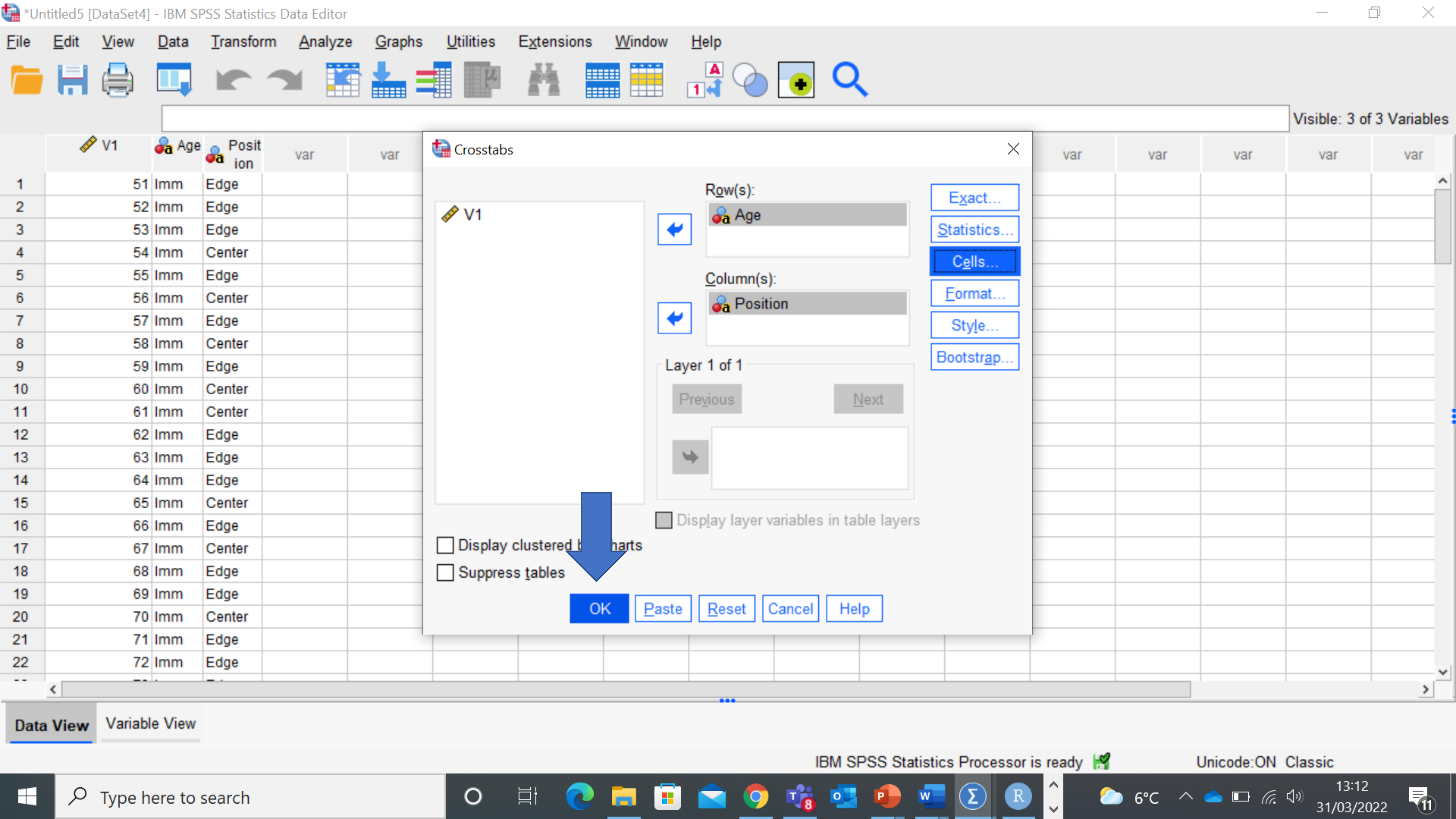


*Untitled5 [DataSet4] - IBM SPSS Statistics Da
File
Edit
View
Data
Transform
Window
Help
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Frequencies...
Descriptives...
Population Descriptives
Explore...
Crosstabs...
TURF Analysis
Ratio...
Proportion Confidence Intervals
P-P Plots...
Q-Q Plots...
Visible: 3 of 3 Variables
V1
Age
Posit ion
var
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
Data View
Variable View
Crosstabs...
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
6°C
13:05
31/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 3 of 3 Variables
V1
Age
Posit ion
var
Crosstabs
V1
Row(s):
Age
Column(s):
Position
Layer 1 of 1
Previous
Next
Display layer variables in table layers
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Type here to search
6°C
13:05
31/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Visible: 3 of 3 Variables
V1
Age
Position
var
Crosstabs
Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
1
2
3
4
Display cluste
Suppress tab
OK
Paste
Reset
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
6°C
13:10
31/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
Crosstabs: Cell Display
Counts
Observed
Expected
Hide small counts
Less than 5
z-test
Compare column proportions
Adjust p-values (Bonferroni method)
Percentages
Row
Column
Total
Residuals
Unstandardized
Standardized
Adjusted standardized
Create APA style table
Noninteger Weights
Round cell counts
Round case weights
Truncate cell counts
Truncate case weights
No adjustments
Continue
Help
1
2
3
4
Click Cells (not easily seen in this screen-shot
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
6°C
13:07
31/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Visible: 3 of 3 Variables
V1 Age Position var var var var var var var
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
Crosstabs
V1
Row(s):
Age
Column(s):
Position
Layer 1 of 1
Previous
Next
Display layer variables in table layers
Exact...
Statistics...
Cells...
Format...
Style...
Bootstrap...
Display clustered bar charts
Suppress tables
OK Paste Reset Cancel Help
Data View Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic
Type here to search
6°C
13:12
31/03/2022

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - GGraph
    - Title
    - Notes
    - Active Dataset
    - Graph
  - Crosstabs
    - Title
    - Notes
    - Case Processing S
    - Age * Position Cro:
    - Chi-Square Tests
    - Symmetric Measur

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Age * Position	100	100.0%	0	0.0%	100	100.0%

Age * Position Crosstabulation

			Position		
			Center	Edge	Total
Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

Crosstabulation table:
Shows us the actual count for each category…
So there were 6 instances where we found Adults at the 'centre' of the tree crown

There were 44 instances where we saw an adult at the edge!

We also have an **expected value**, assuming no association between age and foraging position
: Based on the numbers of adults we observed, and the number of 'edge' observations we observed, we would expect 36.5 instances of adults at the edge!

IBM SPSS Statistics Processor is ready   Unicode:ON Classic H: 503, W: 853 pt

Type here to search   6°C   13:14 31/03/2022

Are the differences between expected and actual values different enough to say something is going on?

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Age * Position	100	100.0%	0	0.0%	100	100.0%

Age * Position Crosstabulation

			Position		
			Center	Edge	Total
Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

*Output1 [Document1] - IBM SPSS Statistics Viewer

Age * Position Crosstabulation

			Position: Center	Position: Edge	Total
Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

b. Comp[illegible]nly for a 2x2 table

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-.338	<.001
	Cramer's V	.338	<.001

Note: Chisquared needs enough data so that the ‘expected’ value is above **5 in at least 80% of cells**

So here, we need this % of values with expeceted counts less than 5 to be less than 20%

We are A-okay here!

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - GGraph
    - Title
    - Notes
    - Active Dataset
    - Graph
  - Crosstabs
    - Title
    - Notes
    - Case Processing S
    - Age * Position Cros
    - Chi-Square Tests
    - Symmetric Measur

Age * Position Crosstabulation

			Position		Total
			Center	Edge	
Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-.338	<.001
	Cramer's V	.338	<.001

Our chi squared value is 11.416

Our degrees of freedom is 1

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - GGraph
    - Title
    - Notes
    - Active Dataset
    - Graph
  - Crosstabs
    - Title
    - Notes
    - Case Processing S
    - Age * Position Cro
    - Chi-Square Tests
    - Symmetric Measur

Age * Position Crosstabulation

			Position		
			Center	Edge	Total
Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-.338	<.001
	Cramer's V	.338	<.001

Our p value is well below 0.05

We can accept our alternate hypothesis that there IS a significant relationship between age category and 'foraging position' within the tree

IBM SPSS Statistics Processor is ready Unicode:ON Classic H: 503, W: 853 pt

Type here to search 6°C 13:19 31/03/2022

*Output1 [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

- Output
  - GGraph
    - Title
    - Notes
    - Active Dataset
    - Graph
  - Crosstabs
    - Title
    - Notes
    - Case Processing S
    - Age * Position Cro:
    - Chi-Square Tests
    - Symmetric Measur

Age	Adult	Count	6	44	50
		Expected Count	13.5	36.5	50.0
	Imm	Count	21	29	50
		Expected Count	13.5	36.5	50.0
Total		Count	27	73	100
		Expected Count	27.0	73.0	100.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.416[a]	1	<.001		
Continuity Correction[b]	9.944	1	.002		
Likelihood Ratio	11.930	1	<.001		
Fisher's Exact Test				.001	<.001
N of Valid Cases	100				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.50.

b. Computed only for a 2x2 table

Symmetric Measure

		Value	Approximate Significance
Nominal by Nominal	Phi	-.338	<.001
	Cramer's V	.338	<.001
N of Valid Cases		100	

For effect size, we can look at our value for 'Phi'

As was the case with cohen's d, there are some general rules of thumb available for this. These can vary from textbook to textbook. My current favourite text uses:

0.1 is considered a small effect, 0.3 a medium effect and 0.5 a large effect.

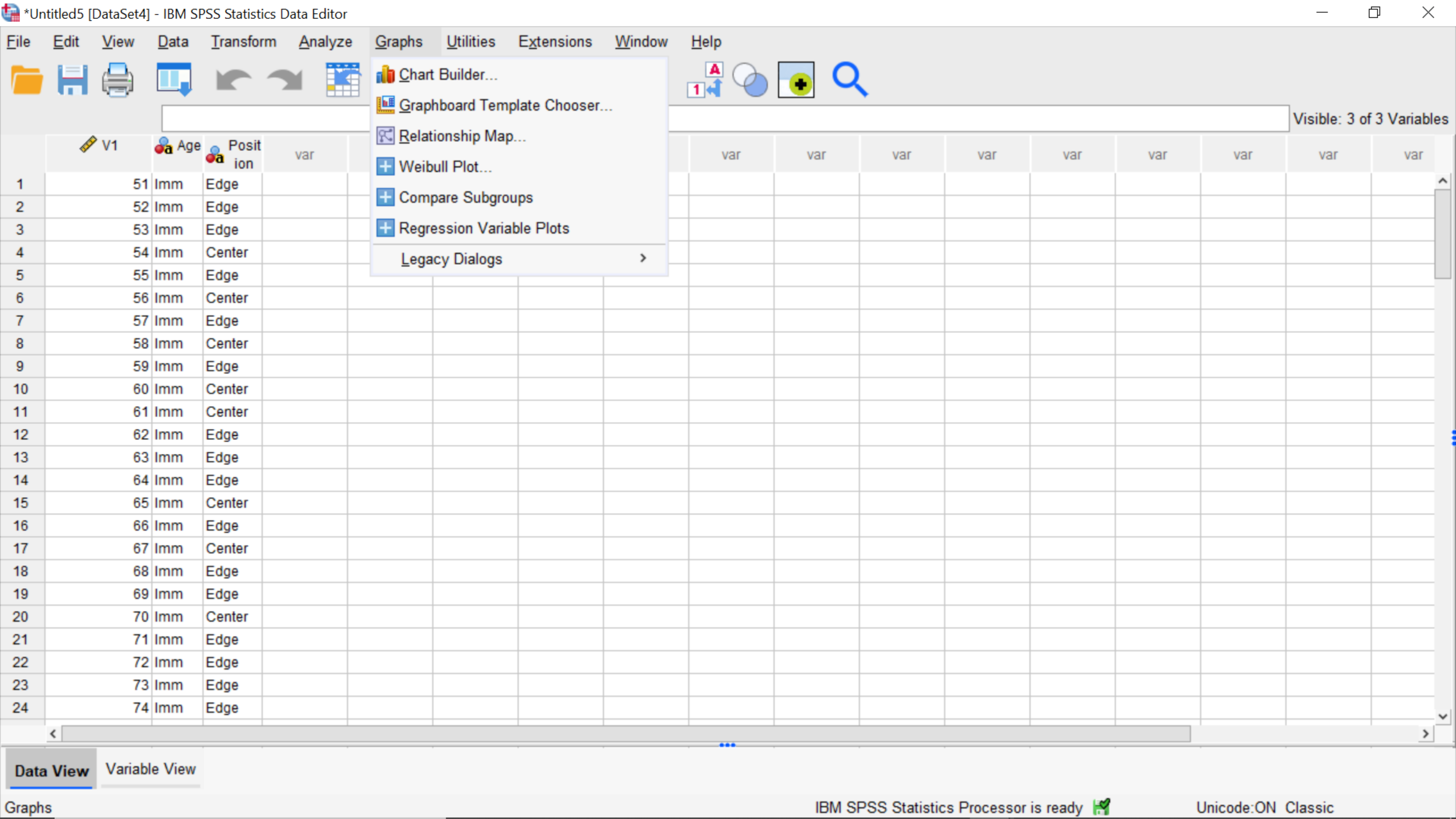
IBM SPSS Statistics Processor is ready Unicode:ON Classic H: 503, W: 853 pt

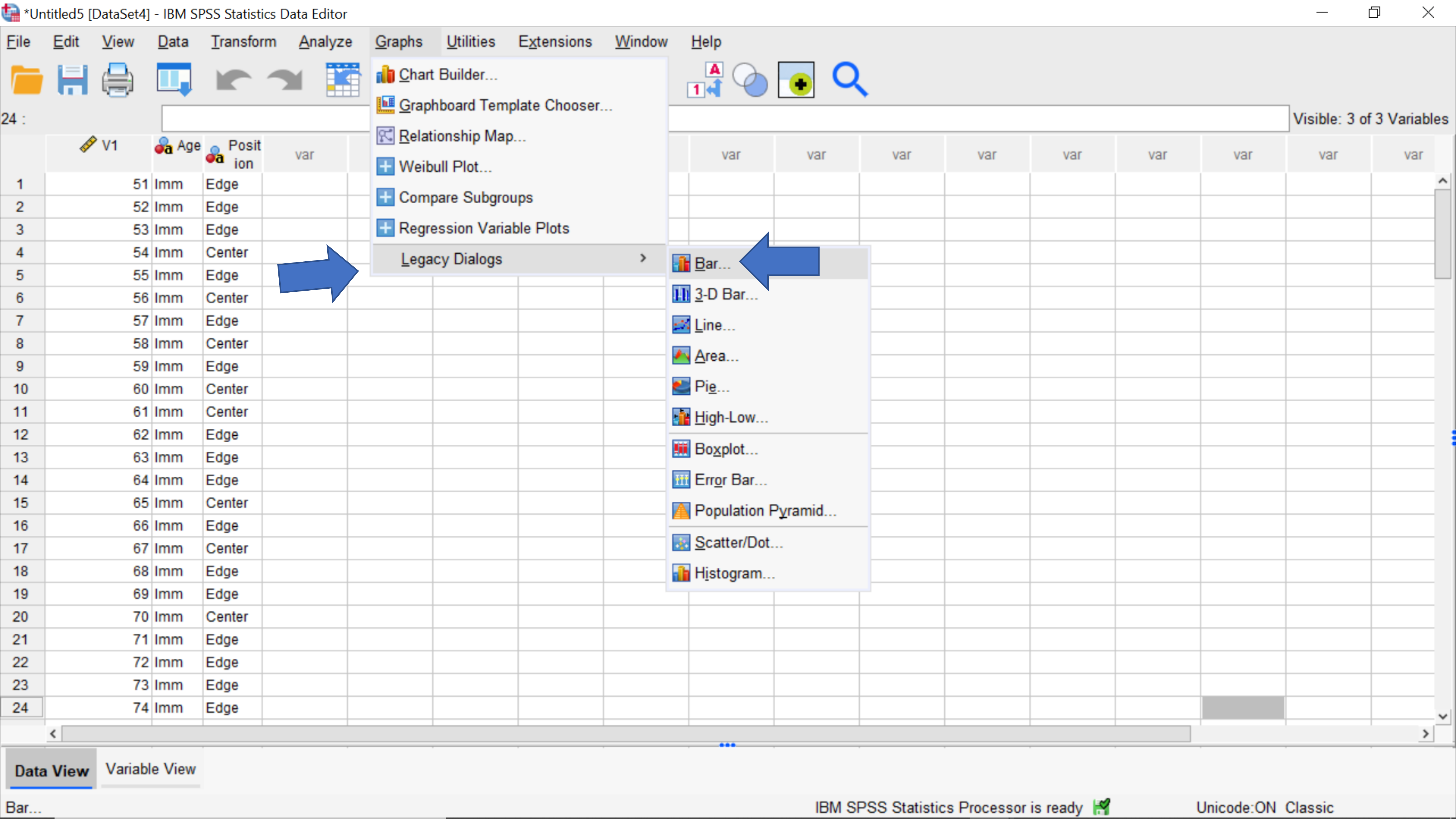
Type here to search 6°C 13:25 31/03/2022

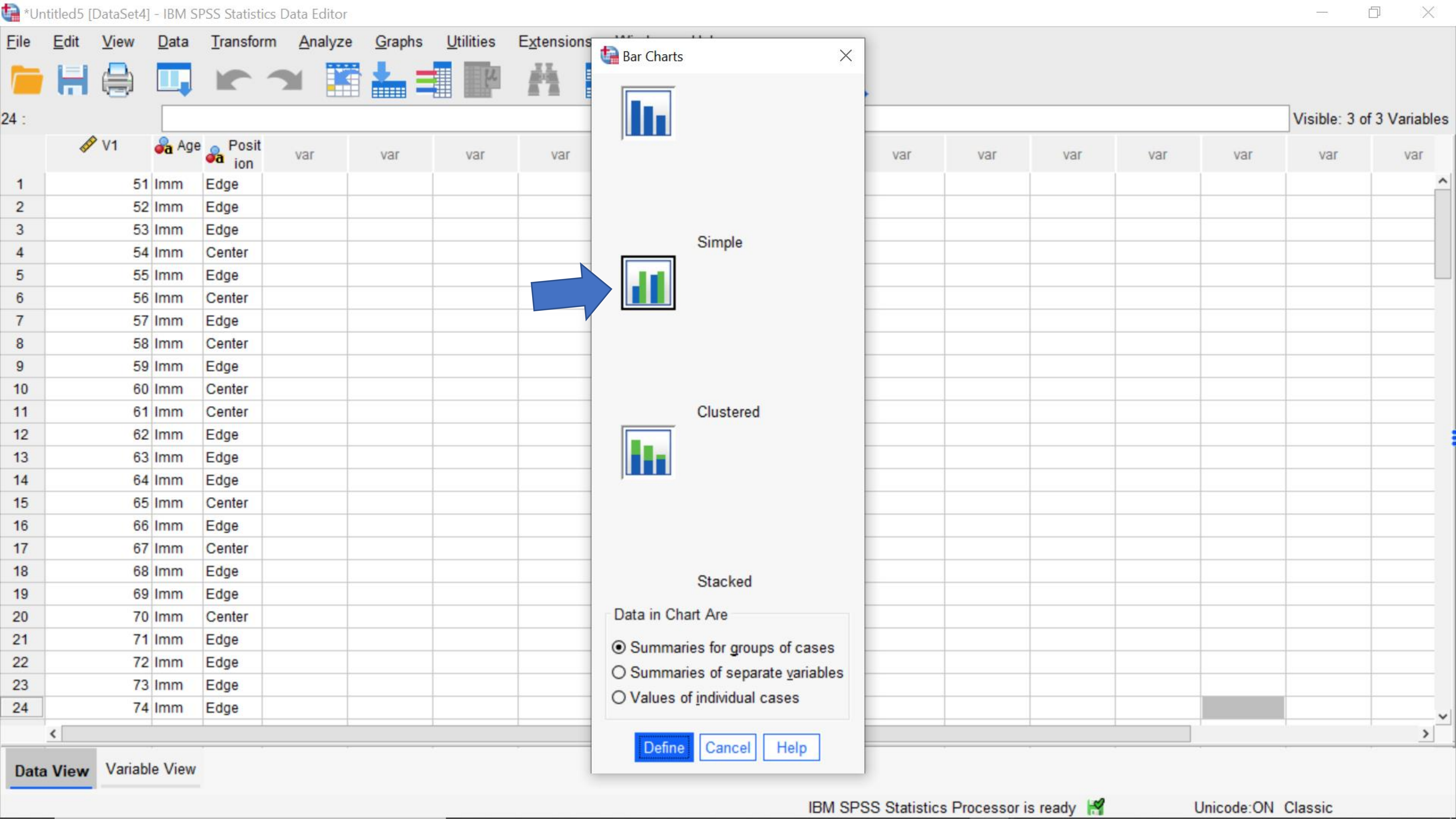
- X2 (degress of freedom, N = sample size) = chi-square statistic value, p = p value.

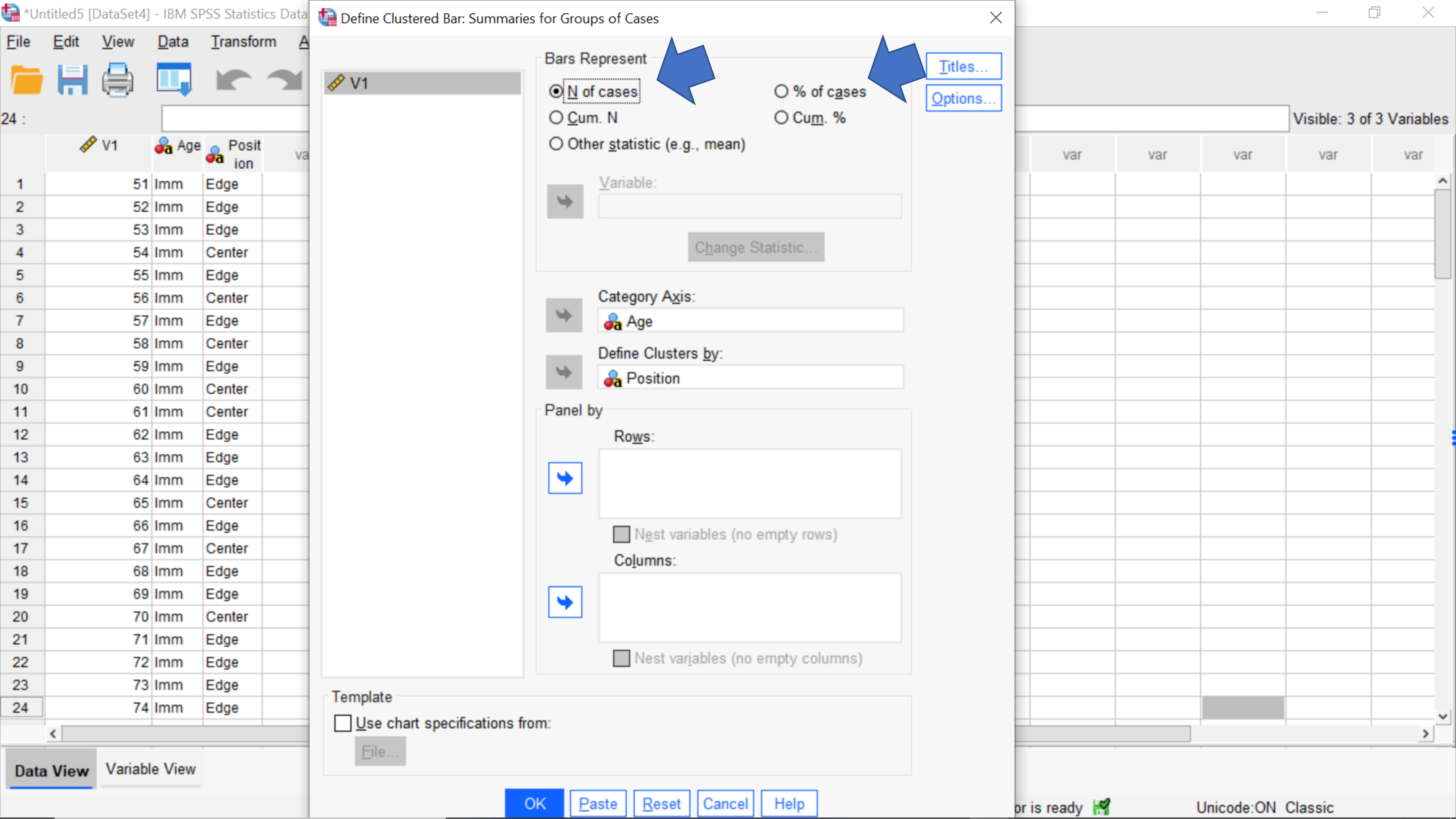
- Example
- A chi-square test of independence was performed to examine the relation between age group and foraging position within tree crowns. The relation between these variables was significant, X2 (1, N = 100) = 11.416, $p < .001$.

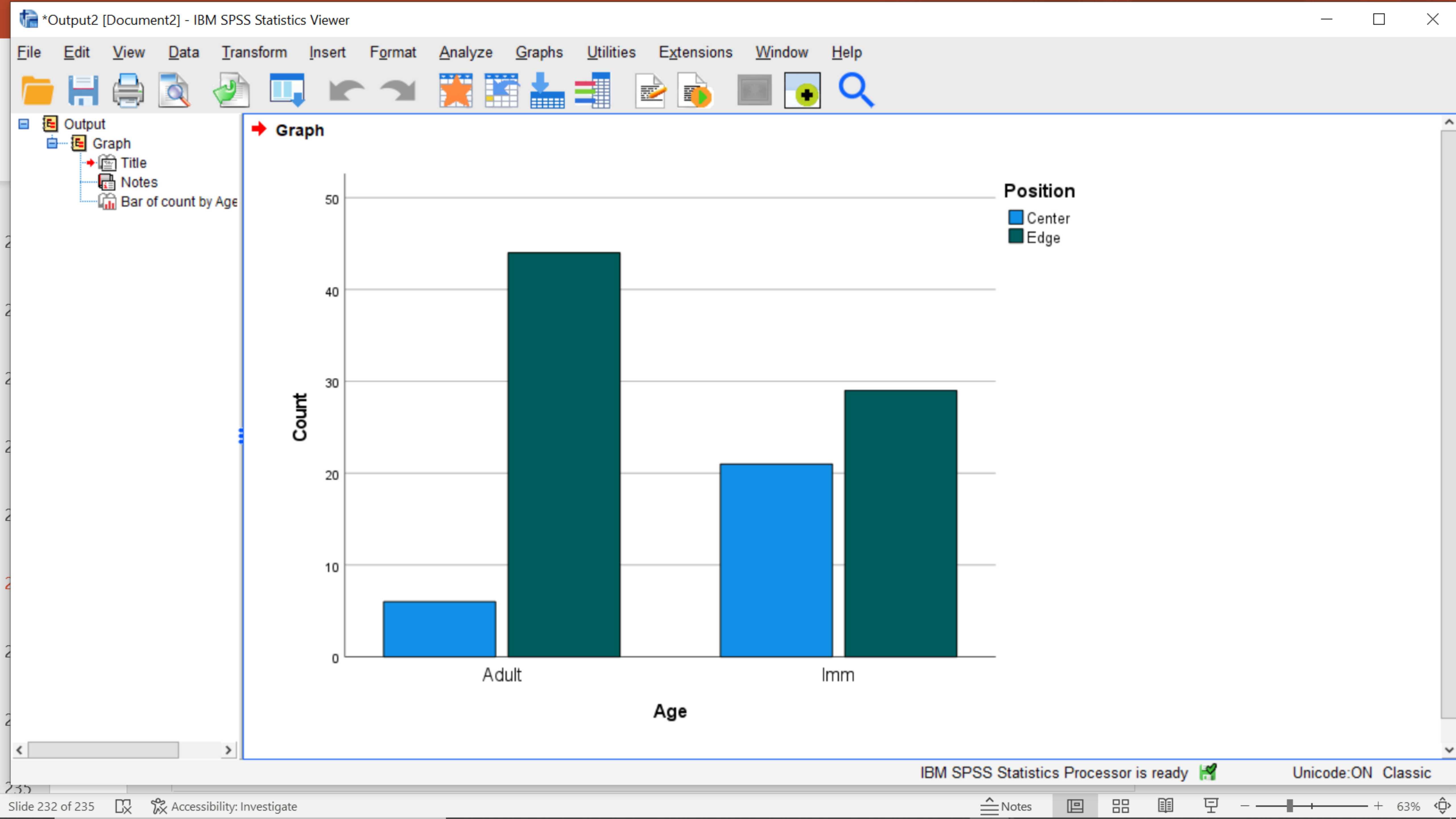
Let’s make a figure!

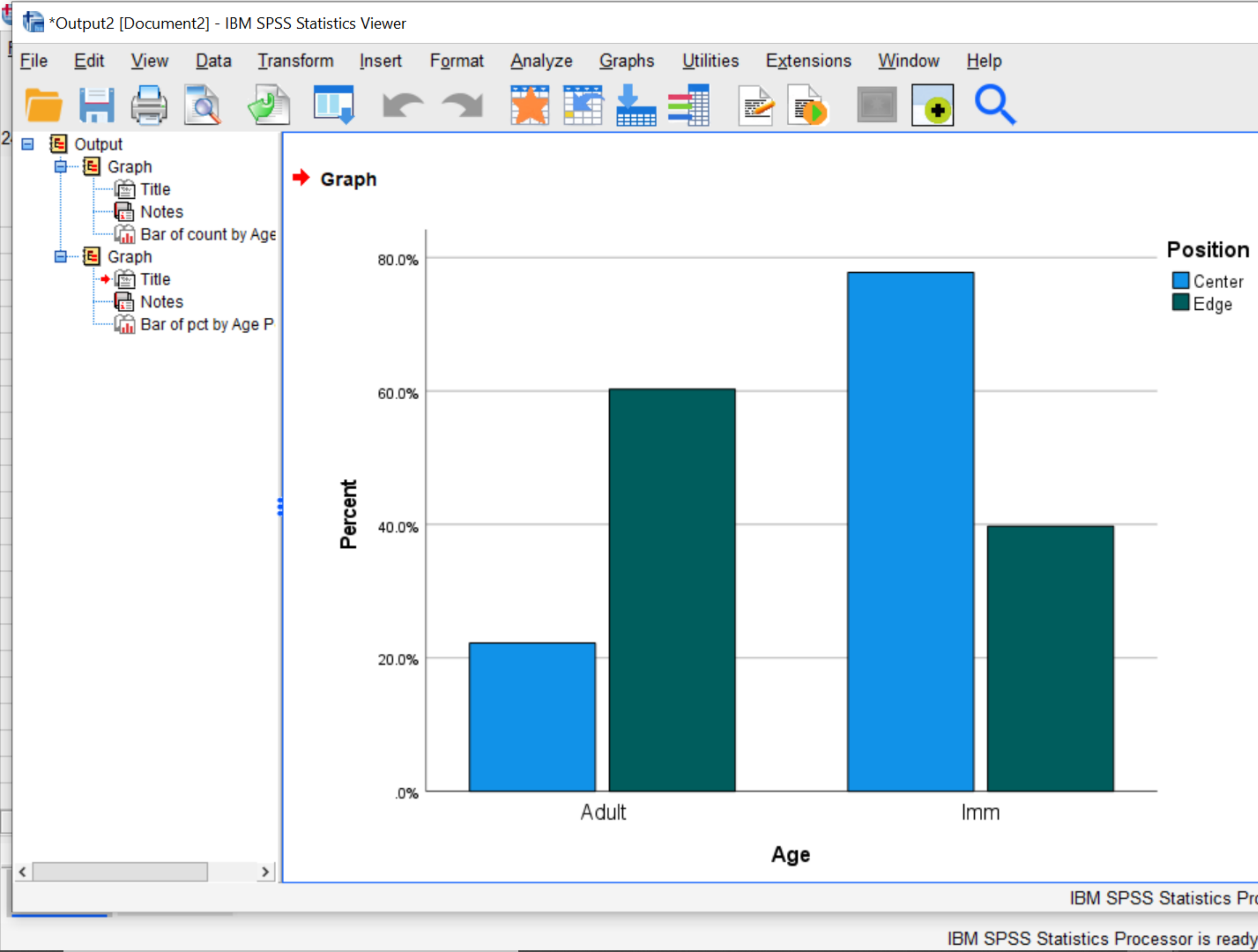

*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Chart Builder...
Graphboard Template Chooser...
Relationship Map...
Weibull Plot...
Compare Subgroups
Regression Variable Plots
Legacy Dialogs
Visible: 3 of 3 Variables
V1 Age Position var var var var var var var var var var
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
23 73 Imm Edge
24 74 Imm Edge
Data View Variable View
Graphs
IBM SPSS Statistics Processor is ready
Unicode:ON Classic


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Extensions
Window
Help
Chart Builder...
Graphboard Template Chooser...
Relationship Map...
Weibull Plot...
Compare Subgroups
Regression Variable Plots
Legacy Dialogs
Bar...
3-D Bar...
Line...
Area...
Pie...
High-Low...
Boxplot...
Error Bar...
Population Pyramid...
Scatter/Dot...
Histogram...
24 :
Visible: 3 of 3 Variables
V1
Age
Posit ion
var
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
23 73 Imm Edge
24 74 Imm Edge
Data View
Variable View
Bar...
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions
24 :
Visible: 3 of 3 Variables
V1 Age Position
var
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
23 73 Imm Edge
24 74 Imm Edge
Bar Charts
Simple
Clustered
Stacked
Data in Chart Are
Summaries for groups of cases
Summaries of separate variables
Values of individual cases
Define
Cancel
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready
Unicode:ON Classic

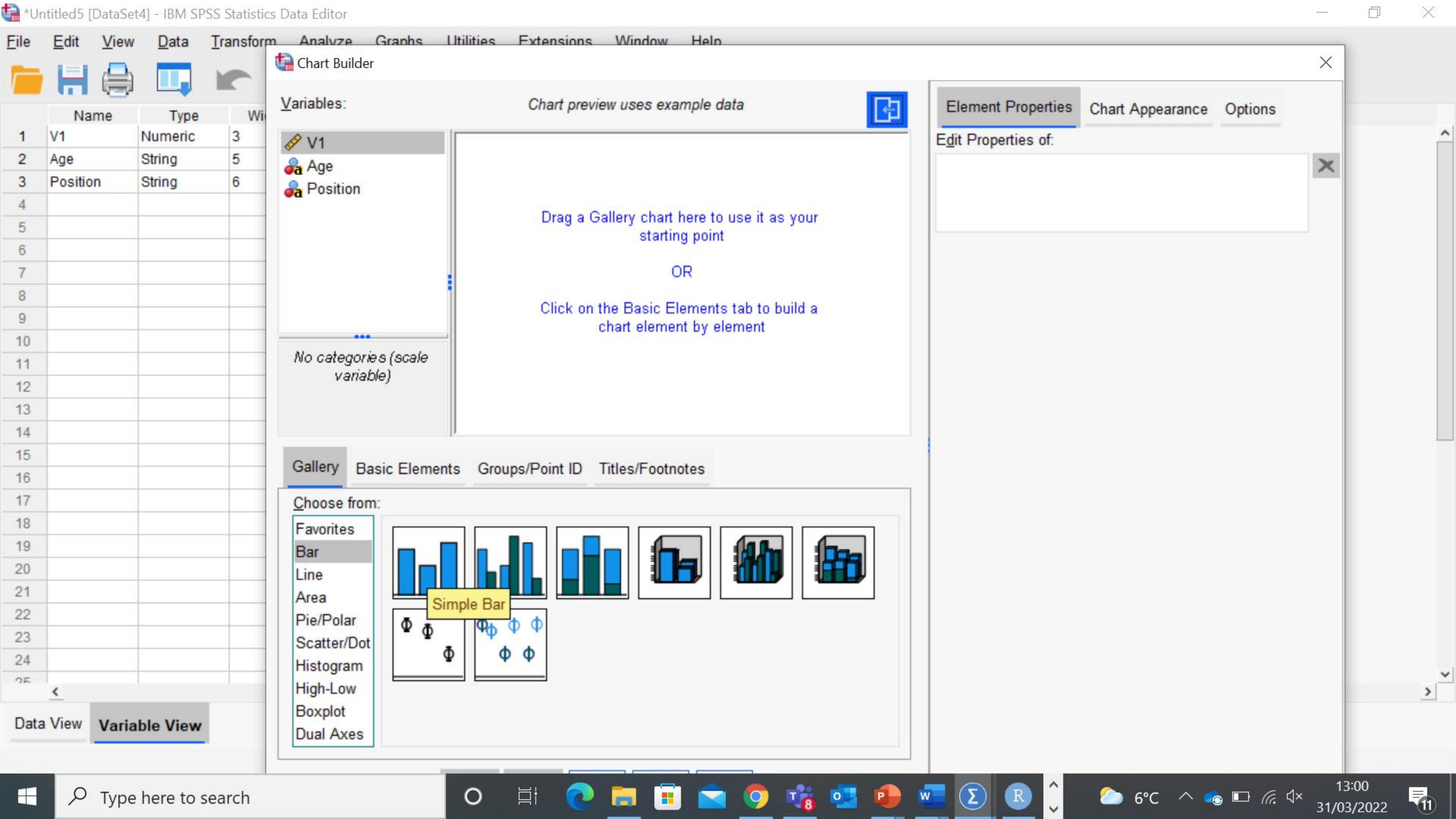

*Untitled5 [DataSet4] - IBM SPSS Statistics Data
File Edit View Data Transform A
24 :
V1 Age Position
1 51 Imm Edge
2 52 Imm Edge
3 53 Imm Edge
4 54 Imm Center
5 55 Imm Edge
6 56 Imm Center
7 57 Imm Edge
8 58 Imm Center
9 59 Imm Edge
10 60 Imm Center
11 61 Imm Center
12 62 Imm Edge
13 63 Imm Edge
14 64 Imm Edge
15 65 Imm Center
16 66 Imm Edge
17 67 Imm Center
18 68 Imm Edge
19 69 Imm Edge
20 70 Imm Center
21 71 Imm Edge
22 72 Imm Edge
23 73 Imm Edge
24 74 Imm Edge
Data View
Variable View
Define Clustered Bar: Summaries for Groups of Cases
V1
Bars Represent
N of cases
% of cases
Cum. N
Cum. %
Other statistic (e.g., mean)
Variable:
Change Statistic...
Titles...
Options...
Category Axis:
Age
Define Clusters by:
Position
Panel by
Rows:
Nest variables (no empty rows)
Columns:
Nest variables (no empty columns)
Template
Use chart specifications from:
File...
OK
Paste
Reset
Cancel
Help
Visible: 3 of 3 Variables
var
var
var
var
var
or is ready
Unicode:ON Classic

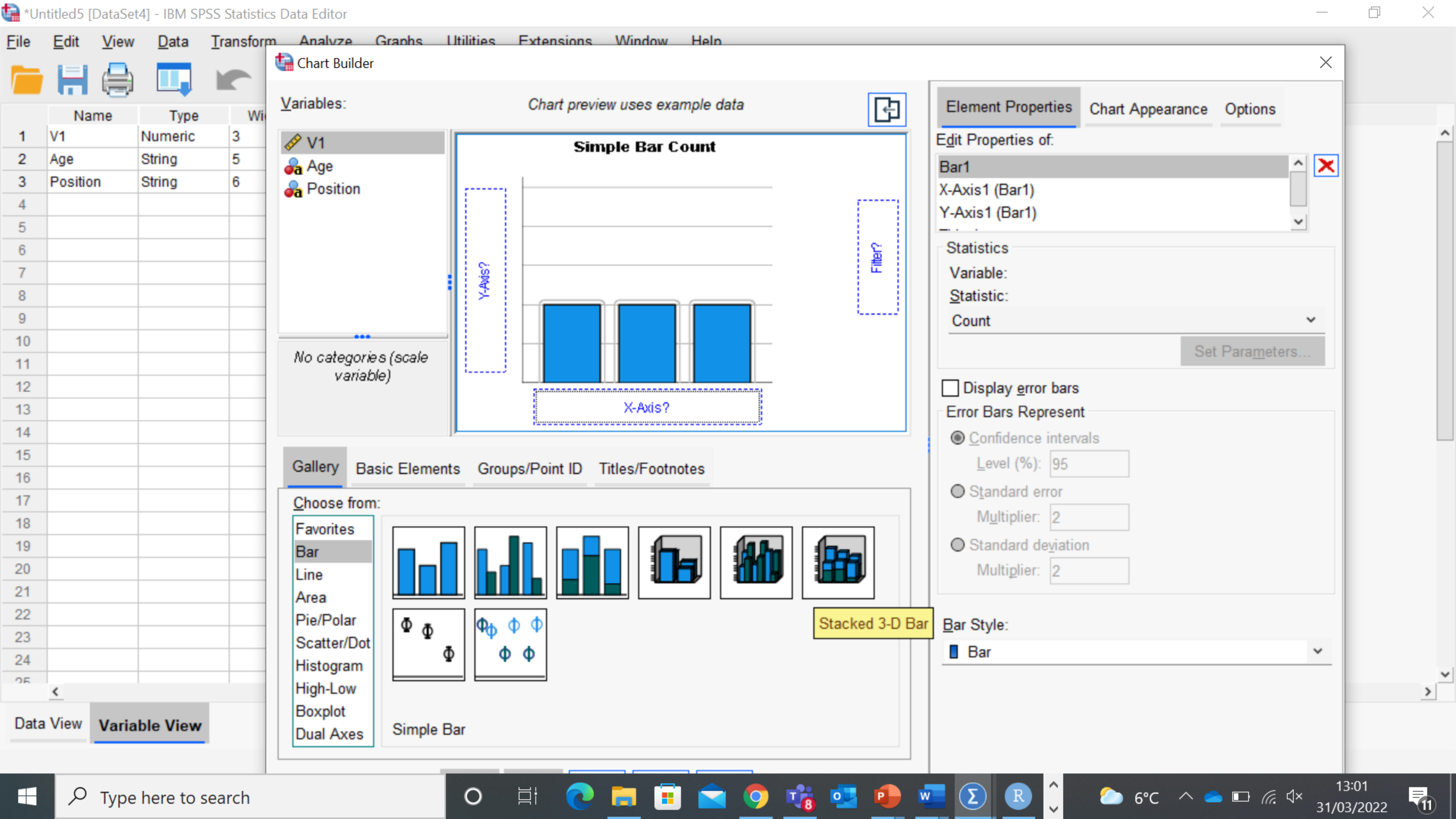

*Output2 [Document2] - IBM SPSS Statistics Viewer
File
Edit
View
Data
Transform
Insert
Format
Analyze
Graphs
Utilities
Extensions
Window
Help
Output
Graph
Title
Notes
Bar of count by Age
Graph
Position
Center
Edge
Count
50
40
30
20
10
0
Adult
Imm
Age
IBM SPSS Statistics Processor is ready
Unicode:ON
Classic
Slide 232 of 235
Accessibility: Investigate
Notes
63%



I prefer to use % of cases here, because we see that when an animal was at the centre of the tree, it was more often than not an immature individual

Meanwhile, it tended to be an adult at the centre of the group


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Name Type
1 V1 Numeric 3
2 Age String 5
3 Position String 6
Data View
Variable View
Chart Builder
Variables:
Chart preview uses example data
V1
Age
Position
No categories (scale variable)
Drag a Gallery chart here to use it as your starting point
OR
Click on the Basic Elements tab to build a chart element by element
Gallery
Basic Elements
Groups/Point ID
Titles/Footnotes
Choose from:
Favorites
Bar
Line
Area
Pie/Polar
Scatter/Dot
Histogram
High-Low
Boxplot
Dual Axes
Simple Bar
Element Properties
Chart Appearance
Options
Edit Properties of:
Type here to search
6°C
13:00
31/03/2022


*Untitled5 [DataSet4] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Name Type
1 V1 Numeric 3
2 Age String 5
3 Position String 6
Data View
Variable View
Chart Builder
Variables:
Chart preview uses example data
V1
Age
Position
No categories (scale variable)
Simple Bar Count
Y-Axis?
Filter?
X-Axis?
Gallery
Basic Elements
Groups/Point ID
Titles/Footnotes
Choose from:
Favorites
Bar
Line
Area
Pie/Polar
Scatter/Dot
Histogram
High-Low
Boxplot
Dual Axes
Stacked 3-D Bar
Simple Bar
Element Properties
Chart Appearance
Options
Edit Properties of:
Bar1
X-Axis1 (Bar1)
Y-Axis1 (Bar1)
Statistics
Variable:
Statistic:
Count
Set Parameters...
Display error bars
Error Bars Represent
Confidence intervals
Level (%): 95
Standard error
Multiplier: 2
Standard deviation
Multiplier: 2
Bar Style:
Bar
Type here to search
6°C
13:01
31/03/2022

- You sit down with Dr. Lynn Besenyei, who teaches you how to bait and use traps for small mammals
- For your masters project, you spend your time traveling to your field sites, capturing mice, weighing them, marking them, and releasing them