Lab# 7.

Starting a new dataset

Once you open SPSS, in the **New Files** field choose **New Dataset** and click **OK**

You will now have two windows: output window and dataset window.

A large company buys thousands of lightbulbs every year. The company is currently considering four brands of lightbulbs to choose from. Before the company decides which lightbulb to buy, it wants to investigate if the mean lifetimes of the four types of lightbulbs are the same. The company's research department randomly selected a few bulbs of each type and tested them. The following table lists the number of hours (in thousands) that each of the bulbs in each brand lasted before being burned out:

Brand I	Brand II	Brand III	Brand IV
23	19	23	26
24	23	27	24
19	18	25	21
26	24	26	29
22	20	23	28
23	22	21	24
25	19	27	28

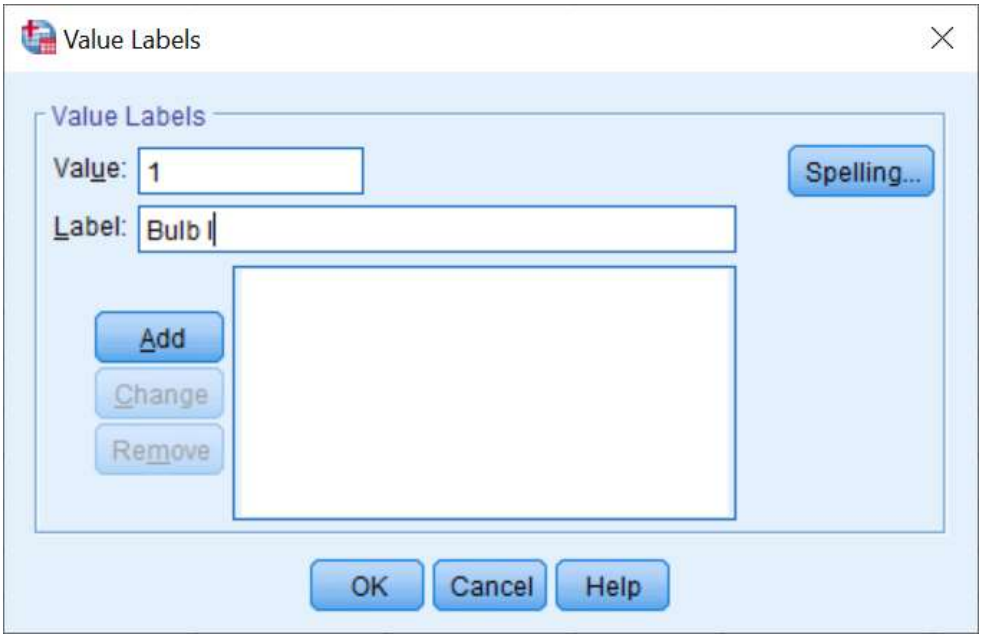
Creating variables in the dataset

The dataset has two tabs (Data View and Variable View). To create the variables you will need to click on **Variable View** tab.

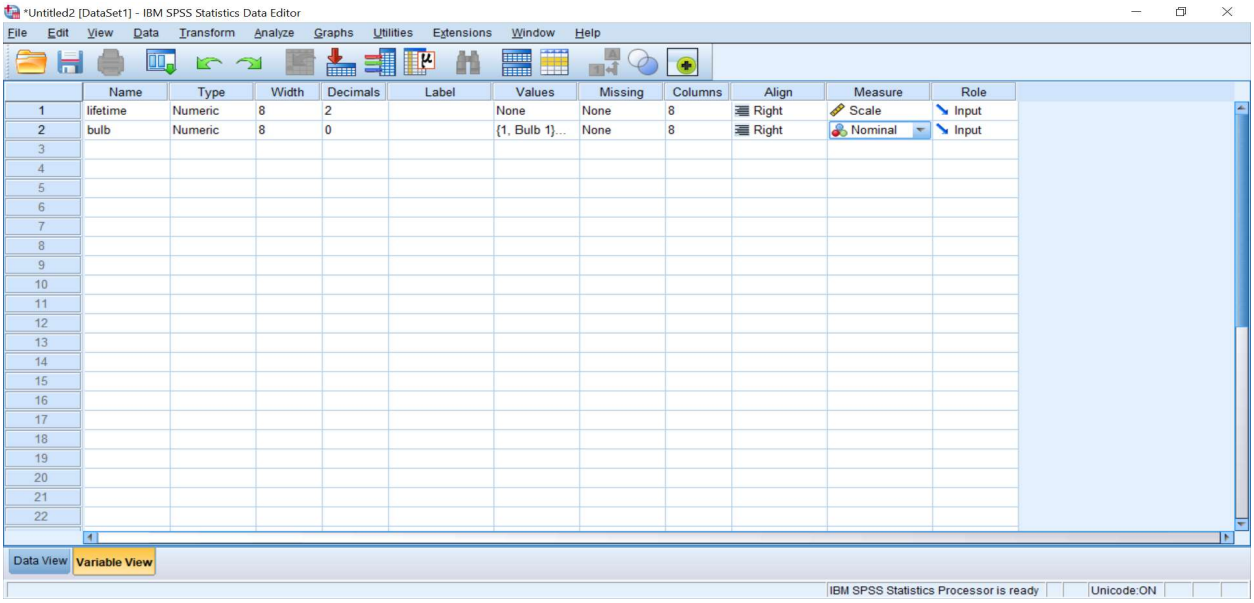
For the Lightbulb example, we need two variables: lifetime (the response variable) and bulb (the treatment variable).

Enter the name of the variable in the field labelled "Name", in the "Type" keep the default option 'numeric'. Next you can choose the "Width" of your variable and how many decimals it should have under "Decimals" (it is set by default to width=8 and decimals=2). The "Label" and "Values" fields are filled for treatment variable, so skip them for the response variable. The last field that should be filled is "Measure". Since the lifetime is the response variable of quantitative type, set it to 'scale'.

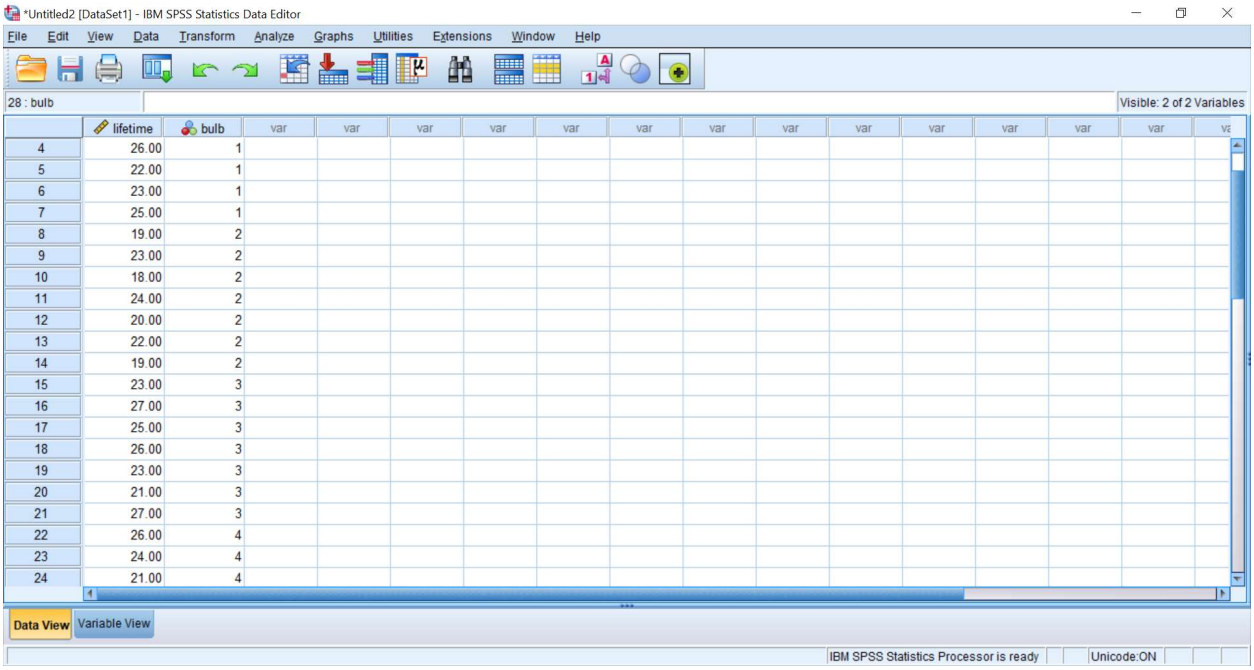
For the factor (i.e. treatment) variable, we repeat: "Name" = bulb, we have no need for decimals here since the bulbs are numbered 1,2,3,4. Next we click on "Values" and here we can label/code each bulb as we wish.



Choose the first numerical value, e.g. 1 and then enter its label, e.g. Bulb 1. Click **'add'** or "+" and continue with rest of the bulbs. At the end click **OK**. After we labeled the levels of our treatment variable, we need to change the "Measure" to 'nominal', since the bulbs have no specific order and they are not continuous rather discrete/categorical.

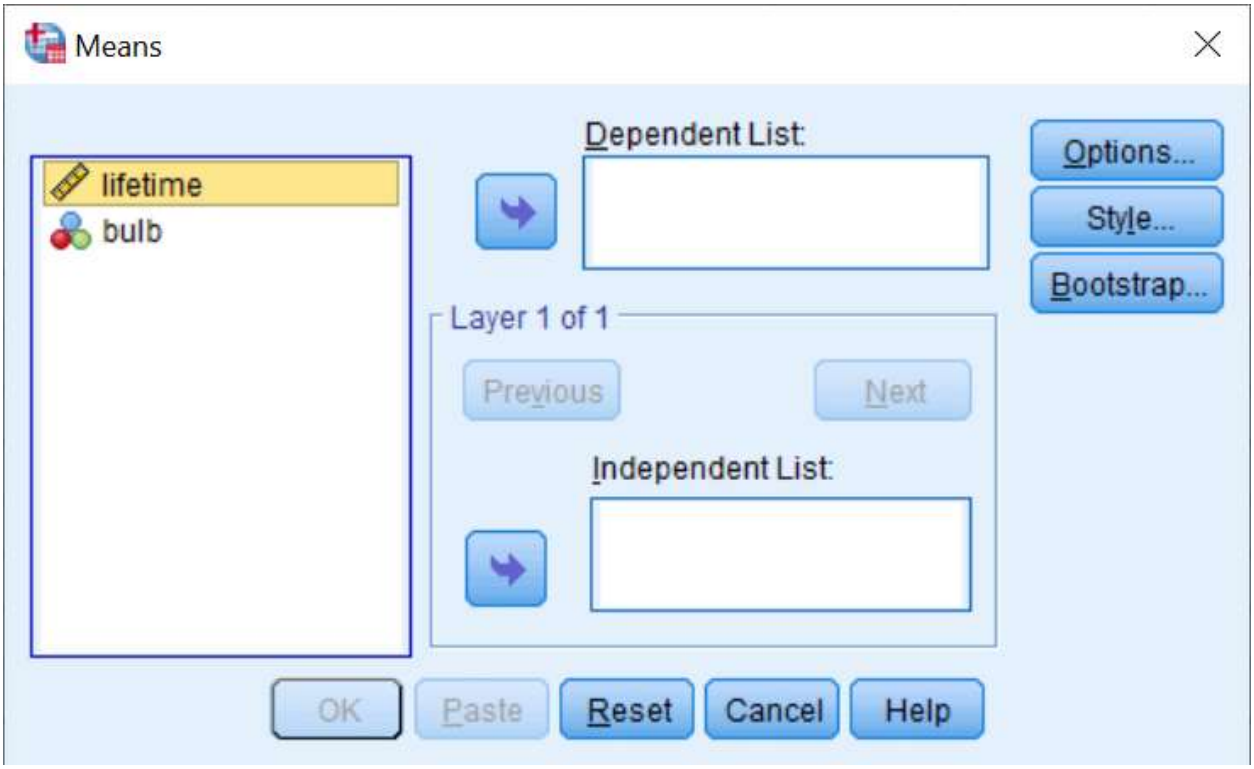


To enter the data, we switch to **Data View** tab and start entering the values.



Obtaining the descriptive statistics for treatments

From the main menu on top, choose '**Analyze**' → '**Compare Means**' → '**Means**'. You will get the following window:



Choose lifetime as the 'Dependent List' and bulb as the 'Independent List', then choose "**Options**" to select statistics you wish to calculate. For our example we can get means, std.

deviations, number of cases and the medians. Once you selected what you want, click **Continue** and you will come back to the previous window. Then click **OK**. In the output window you will get the selected statistics.

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
lifetime * bulb	28	100.0%	0	0.0%	28	100.0%

Report

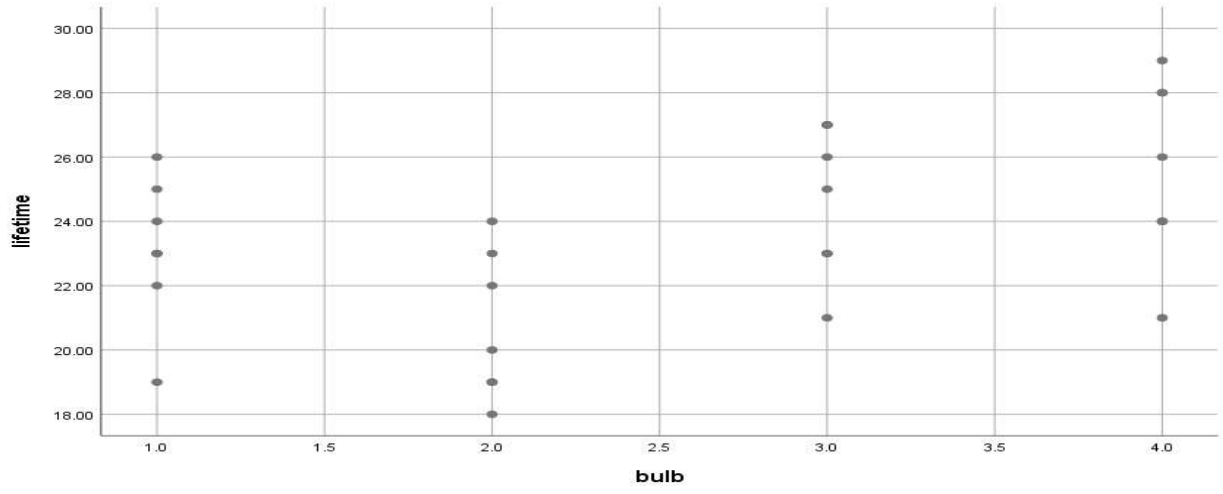
lifetime

bulb	Mean	N	Std. Deviation	Median
Bulb 1	23.1429	7	2.26779	23.0000
Bulb 2	20.7143	7	2.28869	20.0000
Bulb 3	24.5714	7	2.29907	25.0000
Bulb 4	25.7143	7	2.87021	26.0000
Total	23.5357	28	2.98741	23.5000

OPTIONAL

Obtaining a scatter plot of the factor variable: to see if we have same variance/distribution across the treatments

Choose '**Graphs**' → '**Legacy Dialogs**' → '**Scatter/Dot**'. Select '**Simple Scatter**' and then click "**Define**". Select lifetime for the y-axis and bulb for the x-axis and then click **OK**.



Obtaining ANOVA table for One-Factor Study (i.e. for CRD):

Choose '**Analyze**' → '**Compare Means**' → '**One-Way ANOVA**'. Choose lifetime for the 'Dependent List' and bulb for the 'Factor/Independent List', then click **OK**.

lifetime

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	97.536	3	32.512	5.440	.005
Within Groups	143.429	24	5.976		
Total	240.964	27			

Obtaining Simultaneous C.I.'s (i.e. doing Multiple Comparisons) **to compare treatment means**:

Choose '**Analyze**' → '**Compare Means**' → '**One-Way ANOVA**'. As before, select lifetime as your dependent variable and bulb as factor/treatment/independent variable. Then click on "**Post Hoc**"

Here, you can choose the methods for Multiple Comparisons that you want to do and change the significance level if it is different than 5% (set by default). We did Fisher's lsd, Tukey's hsd, Bonferroni.

Multiple Comparisons

Dependent Variable: lifetime

	(I) bulb	(J) bulb	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
LSD	Bulb 1	Bulb 2	2.42857*	1.30671	.075	.1929	4.6642
		Bulb 3	-1.42857	1.30671	.285	-3.6642	.8071
		Bulb 4	-2.57143*	1.30671	.061	-4.8071	-.3358
	Bulb 2	Bulb 1	-2.42857*	1.30671	.075	-4.6642	-.1929
		Bulb 3	-3.85714*	1.30671	.007	-6.0928	-1.6215
		Bulb 4	-5.00000*	1.30671	.001	-7.2356	-2.7644
	Bulb 3	Bulb 1	1.42857	1.30671	.285	-.8071	3.6642
		Bulb 2	3.85714*	1.30671	.007	1.6215	6.0928
		Bulb 4	-1.14286	1.30671	.390	-3.3785	1.0928
	Bulb 4	Bulb 1	2.57143*	1.30671	.061	.3358	4.8071
		Bulb 2	5.00000*	1.30671	.001	2.7644	7.2356
		Bulb 3	1.14286	1.30671	.390	-1.0928	3.3785
Bonferroni	Bulb 1	Bulb 2	2.42857	1.30671	.452	-.9344	5.7916
		Bulb 3	-1.42857	1.30671	1.000	-4.7916	1.9344

		Bulb 4	-2.57143	1.30671	.364	-5.9344	.7916
	Bulb 2	Bulb 1	-2.42857	1.30671	.452	-5.7916	.9344
		Bulb 3	-3.85714*	1.30671	.042	-7.2201	-.4941
		Bulb 4	-5.00000*	1.30671	.005	-8.3630	-1.6370
	Bulb 3	Bulb 1	1.42857	1.30671	1.000	-1.9344	4.7916
		Bulb 2	3.85714*	1.30671	.042	.4941	7.2201
		Bulb 4	-1.14286	1.30671	1.000	-4.5059	2.2201
	Bulb 4	Bulb 1	2.57143	1.30671	.364	-.7916	5.9344
		Bulb 2	5.00000*	1.30671	.005	1.6370	8.3630
		Bulb 3	1.14286	1.30671	1.000	-2.2201	4.5059

*. The mean difference is significant at the 0.10 level.

			Subset	
	bulb	N	1	2

lifetime

Tukey B[a]	Bulb 2	7	20.7143	
	Bulb 1	7	23.1429	23.1429
	Bulb 3	7		24.5714
	Bulb 4	7		25.7143

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 7.000.

Here: means for 1 and 2 are the same and means for 1, 3 and 4 are the same

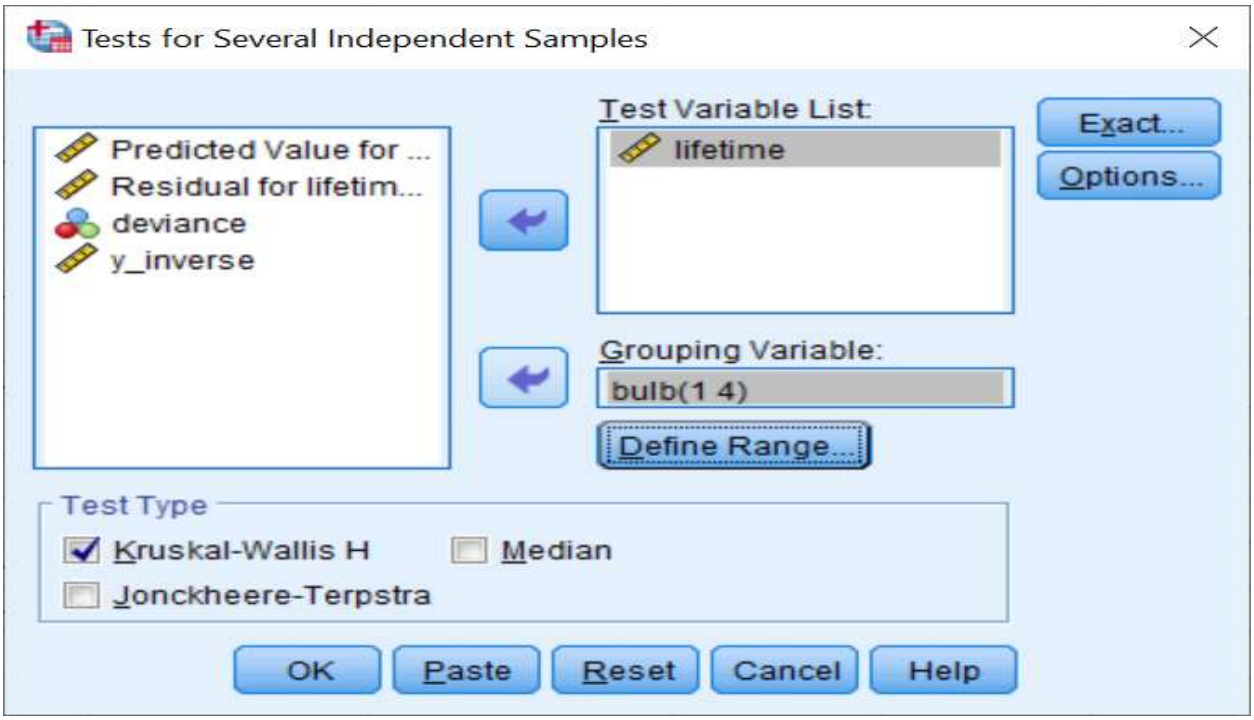
NOTE: Harmonic mean = $\left(\frac{\sum_{i=1}^{n} \frac{1}{x_i}}{n} \right)^{-1}$, Geometric mean = $\sqrt[n]{x_1 x_2 \cdots x_n}$, Mean = $\frac{\sum_{i=1}^{n} x_i}{n}$

Relationship between these 3 means: $H \le G \le \bar{X}$

Kruskal-Wallis Test:

Select '**Analyze**' → '**Nonparametric Test**' → '**Legacy Dialogs**' → '**K Independent Samples**'

1. Select one or more numeric variables.
2. Select a grouping variable and click **Define Range** to specify minimum and maximum integer values for the grouping variable.



(To define the range, enter integer values for **Minimum** and **Maximum** that correspond to the lowest and highest categories of the grouping variable. Cases with values outside of the bounds are excluded. For example, if you specify a minimum value of 1 and a maximum value of 3, only the integer values of 1 through 3 are used. The minimum value must be less than the maximum value, and both values must be specified).

We will use min=1 and max=4

Then click "**Continue**" and then "**OK**"

NPar Tests

Kruskal-Wallis Test

Ranks

	bulb	N	Mean Rank
lifetime	Bulb 1	7	13.36
	Bulb 2	7	7.00
	Bulb 3	7	17.29
	Bulb 4	7	20.36
	Total	28	

$TR_1 = 13.36 \times 7$

$TR_2 = 7 \times 7$

$TR_3 = 17.29 \times 7$

$TR_4 = 20.36 \times 7$

– to get treat't rank totals, we multiply each "Mean Rank" by N

Test Statistics[a,b]

	lifetime
Kruskal-Wallis H	10.428 ← H
df	3
Asymp. Sig.	.015 p-value

a. Kruskal Wallis Test

b. Grouping Variable: bulb