

STATISTICS (SPSS FOR BEGINNERS)

Chapter 6

Descriptive Statistics (Part 2)

Dimaz Ramananda, S.E., M.Ak

- In Statistical activities are not only making tables and graphs, other activities are presenting a summary of data which can be interpreted as an effort to describe data through various statistical measures.
- Descriptive statistics are more related to collecting and summarizing data, as well as presenting the results of the summary.
- SPSS provides two menus related to descriptive statistics, namely reports and descriptive statistics.
- Reports menu are more related to report from the variables that are categorical types and have relation with variable scale.
- Descriptive statistics menu has more deeper use of descriptive data (eg. Skewness etc). Submenu on the descriptive statististics are:
 - Frequencies: gives simple image from group of data
 - Descriptives: shows many statistics size on a table for one or more quantitative variables
 - Explore: checking more details of group of data.
 - Crosstabs: uses for present descriptive data on the form of crosstab
 - Ratio: gives statistics summary of comparation from two specific data
 - P-P Plots & Q-Q Plots: shows data distribution test chart
 - Codebook: gives information about the variable properties
 - Case Summaries: uses to see the content of statistical descriptive
 - Olap Cubes: uses to make a table that connect some variables

- Normality tests is a test with the aim of assessing the distribution of data in a group of data or variables, whether the distribution of the data is normally distributed or not.
- Using Descriptive Statistics

A. Normality Test

How to:

1. Open height.sav
2. Choose Analyze
3. Choose Descriptive statistics
4. Choose explore
5. Put height on Dependent list
6. Put gender on Factor list
7. Click Statistics
8. Follow the tutorial
9. Click Plots
10. Follow the tutorial
11. Choose Both on Displays
12. Click OK

*if the Significance number of Kolmogorov-Smirnov and Shapiro-Wilk are below than 0.05 (<0.05) then the distribution are not normal. But, if the number are greater than 0.05 (>0.05) then the distribution are normal.

Tests of Normality

		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
respondent_gender		Statistic	df	Sig.	Statistic	df	Sig.
respondet_height	female	.160	10	.200*	.932	10	.469
	male	.161	10	.200*	.932	10	.465

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

a. Lilliefors Significance Correction

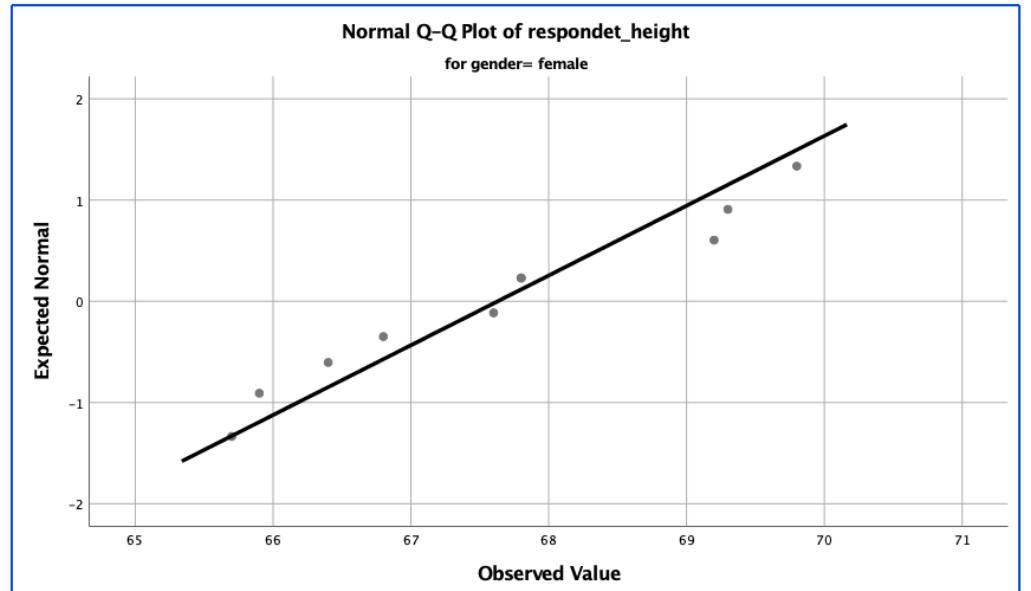
*if the Significance number from the test of homogeneity of variance are below than 0.05 (<0.05) then the data are coming from the population that don't have same variance. But, if the number are greater than 0.05 (>0.05) then the data are coming from the population that have same variance.

Test of Homogeneity of Variance







		Levene Statistic	df1	df2	Sig.
respondet_height	Based on Mean	2.217	1	18	.154
	Based on Median	2.185	1	18	.157
	Based on Median and with adjusted df	2.185	1	16.298	.158
	Based on trimmed mean	2.227	1	18	.153

*if the dot are located around (near) the line then the data distribution is normally distributed, but if all the dot are located on the line then the data is not normally distributed.

Normal Q-Q Plots



*ratio_data.sav dataset

	 name	 gender	 age	 number_of_child	 salary	 working_hour
1	BEN	male	34	3	3500	40
2	JOSH	male	29	0	2400	25
3	EVA	female	30	1	5000	35
4	TOM	male	28	1	2800	25
5	KIM	female	29	2	3400	40
6	JAYDEN	male	40	3	5700	40
7	JESSICA	female	33	2	4400	40
8	CHRISTINE	female	38	0	4000	40
9	MIKE	male	27	0	3000	25
10	ASHLEY	female	30	1	3200	40
11						

B. Using Ratio submenu (compare salary/each kids for every gender group)

1. Open ratio_data.sav
2. Choose Analyze menu
3. Choose descriptive statistics
4. Click ratio
5. Put salary on the numerator box
6. Put number of child on the denominator box
7. Put gender on the group variable box
8. Click statistics
9. Follow the tutorial
10. Click OK

*employee_2.sav dataset

	number	level	education	batch	employe_status	age	first_sal_ary	current_salary
1	1	manager	bachelor	2015	married	48	15000	18000
2	2	supervisor	bachelor	2015	married	37	8000	11000
3	3	staff	vocational	2015	single	30	5000	7500
4	4	staff	bachelor	2015	married	28	5000	8000
5	5	assistant manager	vocational	2015	single	42	11000	14000
6	6	manager	bachelor	2016	married	51	15250	17500
7	7	staff	high school	2016	single	27	4000	7000
8	8	assistant manager	bachelor	2016	married	40	11500	13500
9	9	supervisor	bachelor	2016	married	35	8500	10500
10	10	staff	high school	2017	married	31	5500	6500
11	11	staff	vocational	2017	single	28	5500	6500
12	12	staff	high school	2017	married	29	4500	6500
13	13	staff	bachelor	2017	married	27	5500	6500
14	14	supervisor	bachelor	2017	single	29	8750	10000
15	15	manager	bachelor	2017	married	46	16000	17000
16	16	assistant manager	bachelor	2018	married	38	12000	13000
17	17	supervisor	high school	2018	single	30	8750	9500
18	18	staff	bachelor	2018	married	25	6000	9500
19	19	staff	bachelor	2018	single	24	5750	6000
20	20	supervisor	bachelor	2019	single	26	9000	9250

C. Using Codebook submenu

1. Open employee_2.sav
2. Choose Analyze menu
3. Choose descriptive statistics
4. Click Codebook
5. Put all of the variables on the Variable(s)
6. Click Statistics
7. Follow the tutorial
8. Click OK

D. Using Case Summaries submenu

1. Open employee_2.sav
2. Choose Analyze menu
3. Choose descriptive statistics
4. Click Case Summaries
5. Put age on the Variable(s)
6. Put education and status on the Grouping Variable(s)
7. Click Statistics
8. Follow the tutorial
9. Click OK