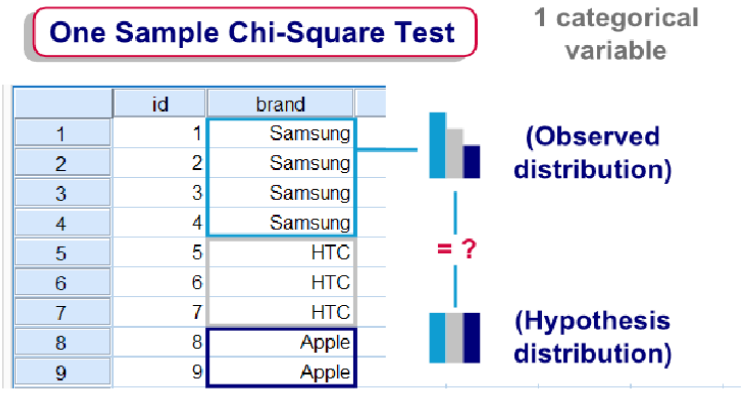
Statistical Tests
(Comparison Tests - One Variable)
One Sample Chi-Square Test

Dony Waluya Firdaus

SPSS One Sample Chi-Square Test

SPSS one-sample chi-square test is used to test whether a single categorical variable follows a hypothesized population distribution. The figure below illustrates the basic idea.



SPSS One-Sample Chi-Square Test Example

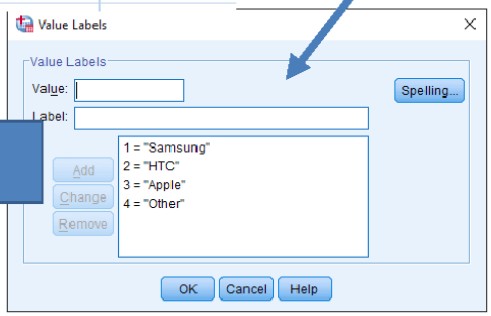
A marketeer believes that 4 smartphone brands are equally attractive. He asks 43 people which brand they prefer, resulting in brands.sav. If the brands are really equally attractive, each brand should be chosen by roughly the same number of respondents. In other words, the **expected frequencies** under the null hypothesis are (43 cases / 4 brands =) 10.75 cases for each brand. The more the observed frequencies differ from these expected frequencies, the less likely it is that the brands really are equally attractive.

SPSS One-Sample Chi-Square Test Example
➔ Variabel View

Name	Type	Width	Decimals	Label	Values
id	Numeric	2	0		None
brand	Numeric	2	0	Which brand of smartphone do you prefer?	{1, Samsun...

Missing	Columns	Align	Measure	Role
None	10	Right	Scale	Input
None	10	Right	Nominal	Input

1 = "Samsung"
2 = "HTC"
3 = "Apple"
4 = "Other"



SPSS One-Sample Chi-Square Test Example
➔ Data View

	id	brand
1	3	1
2	4	1
3	7	1
4	9	4
5	11	4
6	13	1
7	15	2
8	17	1
9	18	2
10	21	3
11	22	2
12	24	3
13	27	4
14	28	3
15	30	3
16	31	1
17	33	3
18	36	4
19	37	3
20	39	3
21	40	1
22	41	1
23	42	1
24	43	3
25	44	3
26	45	3
27	48	3
28	50	3
29	51	4
30	52	2
31	53	2

SPSS One-Sample Chi-Square Test Example
➔ Data View

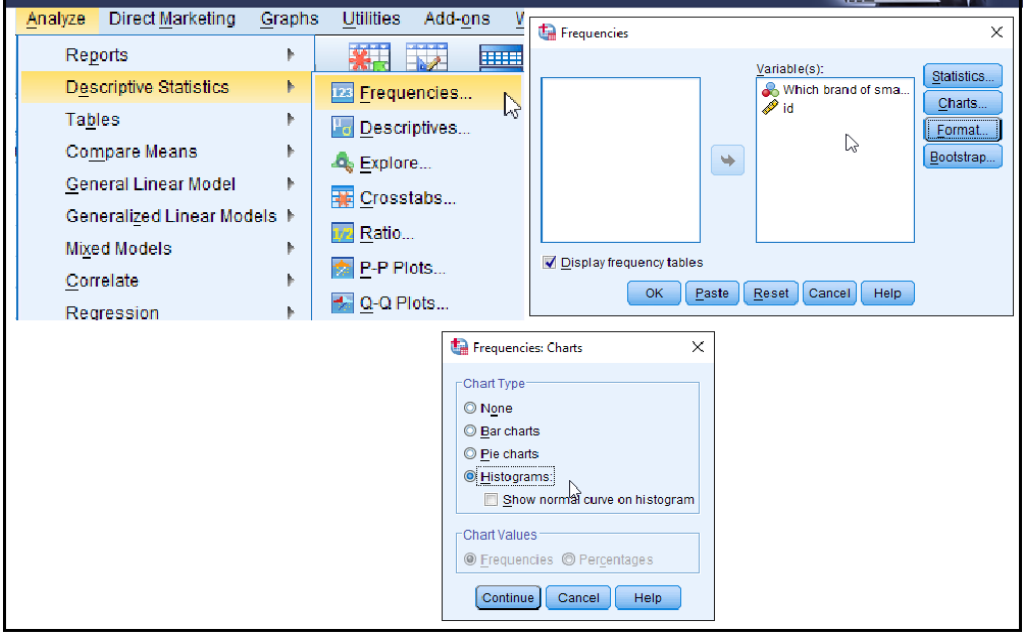
32	56	3
33	58	1
34	61	3
35	63	1
36	64	4
37	66	1
38	69	3
39	70	3
40	71	2
41	72	1
42	73	1
43	76	2

1. Quick Data Check

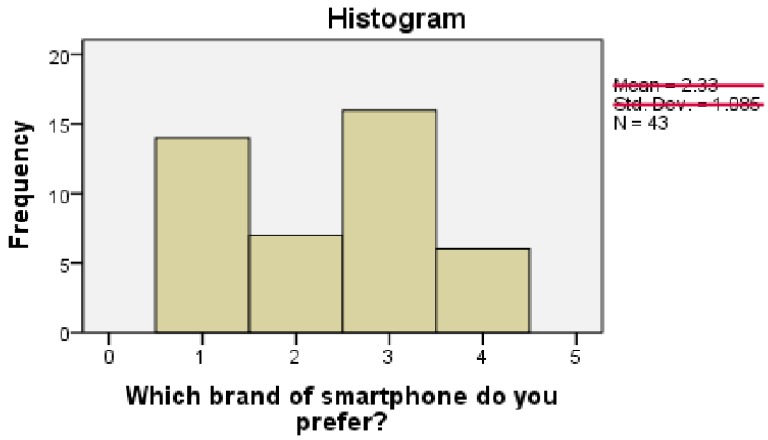
Before running any statistical tests, we always want to have an idea what our data basically look like. In this case we'll inspect a histogram of the preferred brand by running FREQUENCIES.

We'll open the data file and create our histogram by running the syntax below. Since it's very simple, we won't bother about clicking through the menu here.

1. Quick Data Check



1. Quick Data Check… Cont'



1. Quick Data Check… Cont'

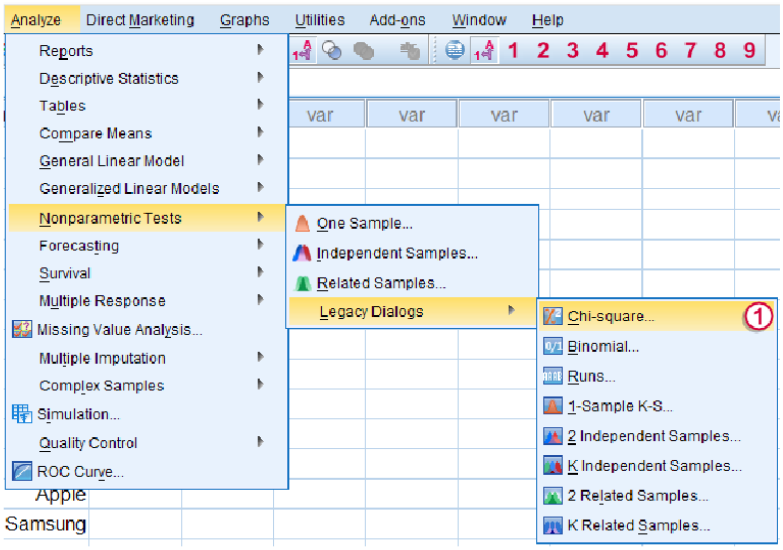
First, N = 43 means that the histogram is based on 43 cases. Since this is our sample size, we conclude that no missing values ar e present. SPSS also calculates a mean and standard deviation but these are not mean ingful for nominal variables so we'll just ig nore them. Second, the preferred brands h ave rather unequal frequencies, casting so me doubt upon the null hypothesis of thos e being equal in the population.

2. Assumptions One-Sample Chi-Square Test

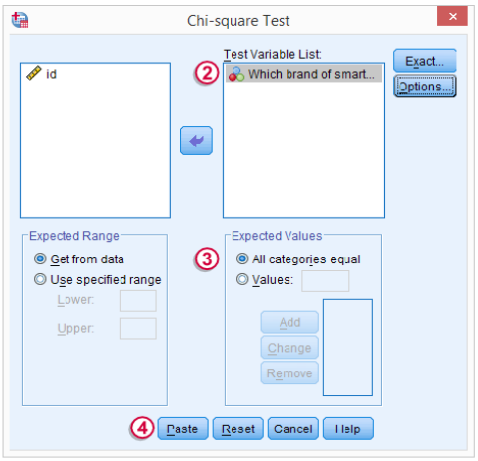
1. the sample size is much smaller than the population size;
2. the sample is representative for the target population;
3. Assumption of Independent and Identically Distributed Variables (part of which is "independent observations");
4. none of the expected frequencies are < 5;

The first three assumptions are beyond the scope of this tutorial. We'll presume they've been met by our data. Whether the fourth assumption holds is reported by SPSS whenever we run a one-sample chi-square test. However, we already saw that all expected frequencies are 10.75 for our data.

3. Run SPSS One Sample Chi-Square Test



3. Run SPSS One Sample Chi-Square Test… Cont'



3 **Expected Values** refers to the expected frequencies, the aforementioned 10.75 cases for each brand. We could enter these values but selecting All categor*i*es equal is a faster option and yields identical results.

4 Clicking *P*aste results in the syntax below.

3. Run SPSS One Sample Chi-Square Test… Cont'

brand

	1 Observed N	2 Expected N	3 Residual
1 Samsung	14	10.8	3.3
2 HTC	7	10.8	-3.8
3 Apple	16	10.8	5.3
4 Other	6	10.8	-4.8
Total	43		

3. Run SPSS One Sample Chi-Square Test… Cont'

1. Under **Observed N** we find the observed frequencies that we saw previously;
2. under **Expected N** we find the theoretically expected frequencies;*
3. for each frequency the **Residual** is the difference between the observed and the expected frequency and thus expresses a deviation from the null hypothesis;

3. Run SPSS One Sample Chi-Square Test… Cont'

Test Statistics

	brand
Chi-Square	6.953[a] (4)
df	3 (5)
Asymp. Sig.	.073 (6)

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 10.8.

3. Run SPSS One Sample Chi-Square Test… Cont'

④ the **Chi-Square** test statistic sort of summarizes the residuals and hence indicates the overall difference between the data and the hypothesis. The larger the chi-square value, the less the data "fit" the null hypothesis;

⑤ degrees of freedom (**df**) specifies which chi-square distribution applies;

⑥ **Asymp. Sig.** refers to the p value and is .073 in this case. If the brands are exactly equally attractive in the population, there's a 7.3% chance of finding our observed frequencies or a larger deviation from the null hypothesis. We usually reject the null hypothesis if $p < .05$. Since this is not the case, we conclude that the brands are equally attractive in the population.

Reporting a One-Sample Chi-Square Test

When reporting a one-sample chi-square test, we always report the observed frequencies. The expected frequencies usually follow readily from the null hypothesis so reporting them is optional. Regarding the significance test, we usually write something like ***"we could not demonstrate that the four brands are not equally attractive; $\chi^2(3) = 6.95$, $p = .073$."***

Refensi : Ruben Geert van den Berg on September 16, 2014