The background features a dark blue gradient with a subtle pattern of white dots. On the left side, there are several circular elements: a large scale with numerical markings from 140 to 260, and several smaller circles with dashed lines and arrows, suggesting a technical or scientific theme.

CORRESPONDENCE ANALYSIS WITH SPSS

ZAINEB MEZDOUD

INTRODUCTION

The method of Correspondence Analysis is considered one of the most important factorial methods for data exploration. It is among the most prominent in multidimensional analysis and is widely used in scientific research and studies, especially in the field of humanities, due to its strong association with qualitative data analysis. This method resembles Principal Component Analysis (PCA), which is used for quantitative data analysis, but it specifically examines the relationships between two categorical variables, whether nominal or ordinal.

Example: The table below shows a group of companies and the evaluation of the characteristics of Visa cards provided by each company.

	Amazon	Flipcart	Myntra	Naaptol	Japong	Snapdeal	Loacalbania	Total
Ease of Use	56	65	15	11	12	8	8	175
Secure Payment	48	53	8	3	3	34	25	174
Special Offers	64	51	36	28	16	17	12	224
After-Sales Service	41	35	3	6	8	46	26	165
Delivery Time	32	19	5	31	25	13	14	139
Competitive Price	51	37	13	5	9	15	14	144
Total	292	260	80	84	69	95	123	1062



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	companies	Numeric	8	2	name of compa...	{1.00, Amaz...	None	8	Center	Nominal	Input
2	evaluation	Numeric	8	2		None	None	8	Center	Nominal	Input
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											
28											
29											

Value Labels

Value Labels

Value:

Label:

Spelling...

Add

Change

Remove

- 1.00 = "Amazon"
- 2.00 = "Flipcart"
- 3.00 = "Myntra"
- 4.00 = "Naaptol"
- 5.00 = "Japong"
- 6.00 = "Snapdeal"

OK Cancel Help

AFC.sav [DataSet1] - IBM SPSS Statistics Data Editor

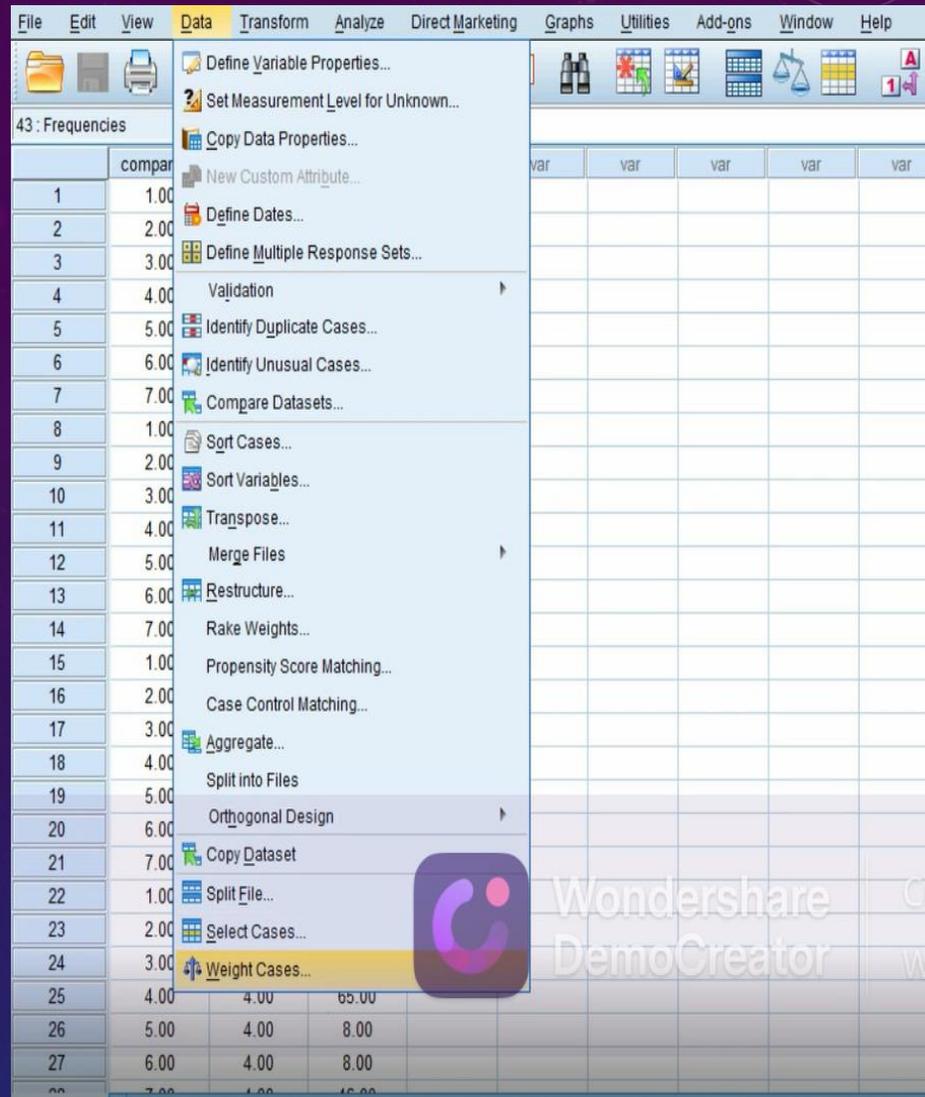
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

43: Frequencies

	companies	evaluation	Frequencies	var											
1	1.00	1.00	56.00												
2	2.00	1.00	65.00												
3	3.00	1.00	15.00												
4	4.00	1.00	11.00												
5	5.00	1.00	12.00												
6	6.00	1.00	8.00												
7	7.00	1.00	8.00												
8	1.00	2.00	48.00												
9	2.00	2.00	53.00												
10	3.00	2.00	8.00												
11	4.00	2.00	3.00												
12	5.00	2.00	3.00												
13	6.00	2.00	34.00												
14	7.00	2.00	25.00												
15	1.00	3.00	64.00												
16	2.00	3.00	51.00												
17	3.00	3.00	36.00												
18	4.00	3.00	28.00												
19	5.00	3.00	12.00												
20	6.00	3.00	16.00												
21	7.00	3.00	17.00												
22	1.00	4.00	41.00												
23	2.00	4.00	35.00												
24	3.00	4.00	3.00												
25	4.00	4.00	65.00												
26	5.00	4.00	8.00												
27	6.00	4.00	8.00												
28	7.00	4.00	16.00												

Data View Variable View

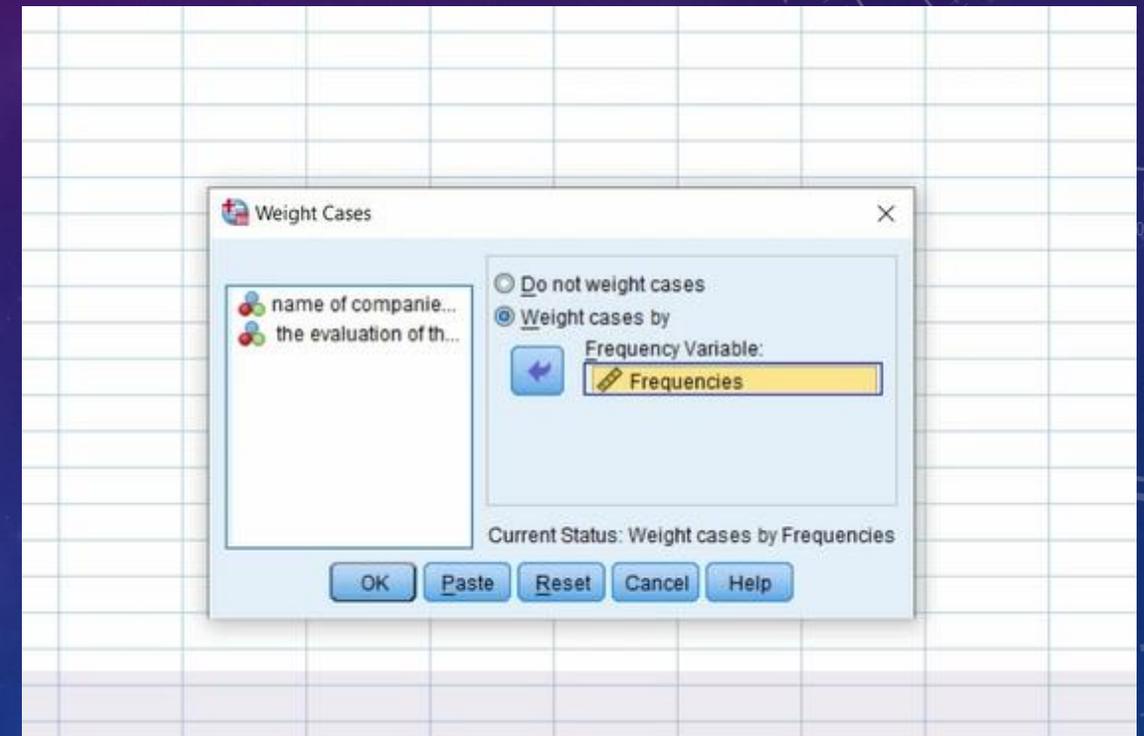




In the first step, we need to weight the cases using the 'Weight Cases' option from the 'Data' menu."

Select the option "Weight cases by", then move the frequency variable to the box and click OK.

The cases will be weighted, but no visible change will appear in the data because the process happens in the background.



*AFC.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Reports
Descriptive 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
ROC Curve...

Factor...
Correspondence Analysis...
Optimal Scaling...

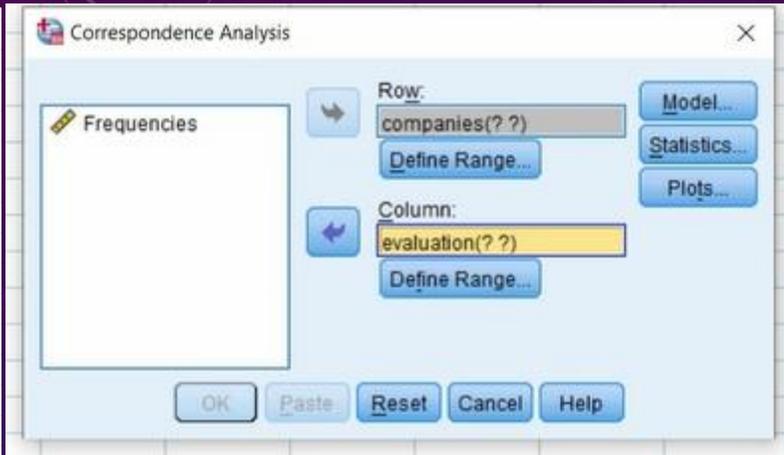
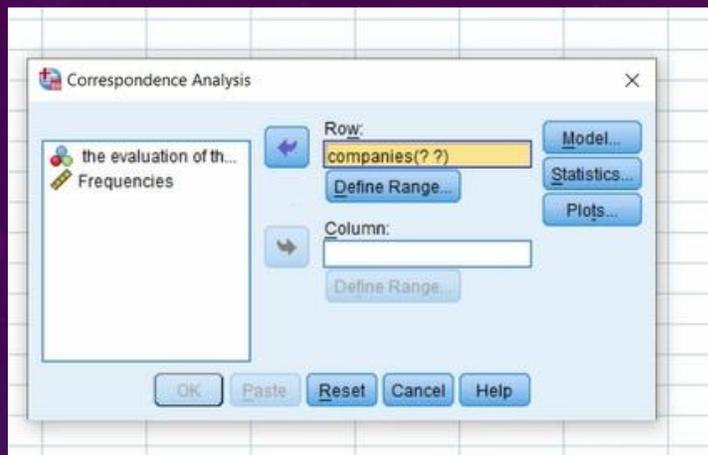
	companies	evaluation		
1	1.00	1.00		
2	2.00	1.00		
3	3.00	1.00		
4	4.00	1.00		
5	5.00	1.00		
6	6.00	1.00		
7	7.00	1.00		
8	1.00	2.00		
9	2.00	2.00		
10	3.00	2.00		
11	4.00	2.00		
12	5.00	2.00		
13	6.00	2.00		
14	7.00	2.00		
15	1.00	3.00		
16	2.00	3.00		
17	3.00	3.00		
18	4.00	3.00		
19	5.00	3.00		
20	6.00	3.00		
21	7.00	3.00		
22	1.00	4.00	41.00	
23	2.00	4.00	35.00	
24	3.00	4.00	3.00	
25	4.00	4.00	65.00	
26	5.00	4.00	8.00	
27	6.00	4.00	8.00	
28	7.00	4.00	16.00	

Data View Variable View

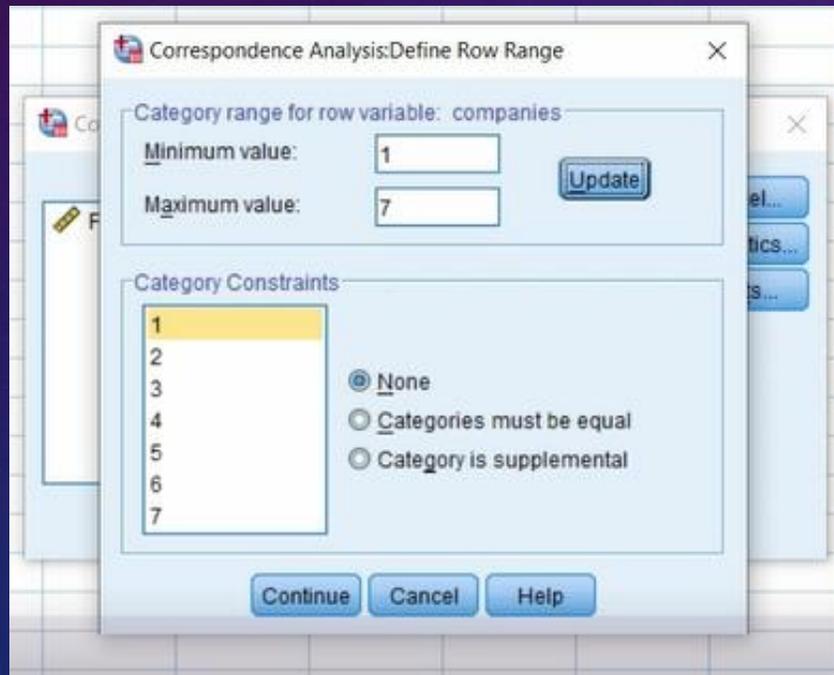
Dimension Reduction

Wondershare DemoCreator

Then, from the main menu, go to **Analyze > Dimension Reduction > Correspondence Analysis**.

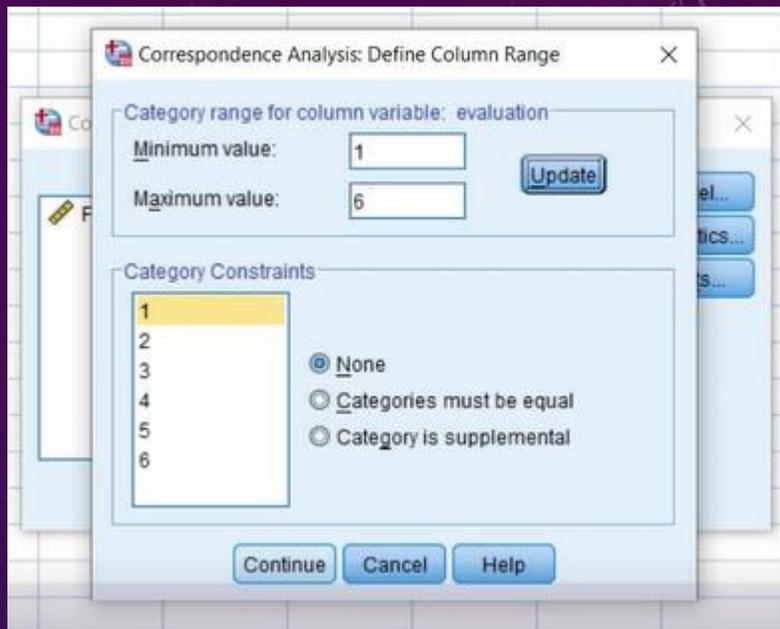


1- Move the nominal variable (Company) to the Row field, and the nominal variable (Attributes) to the Column field."

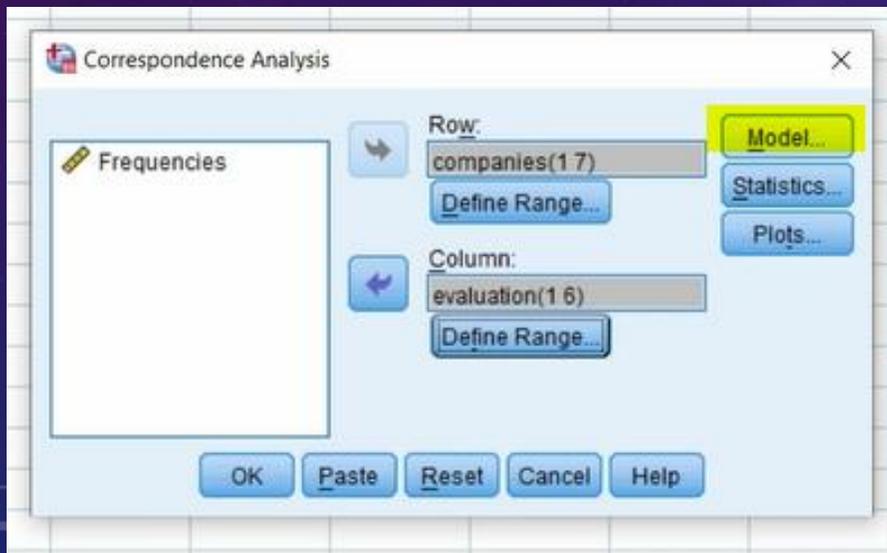


2- Select 'Define Range' and set the appropriate range. In this case, since there are 7 companies, the range will be from 1 to 7.

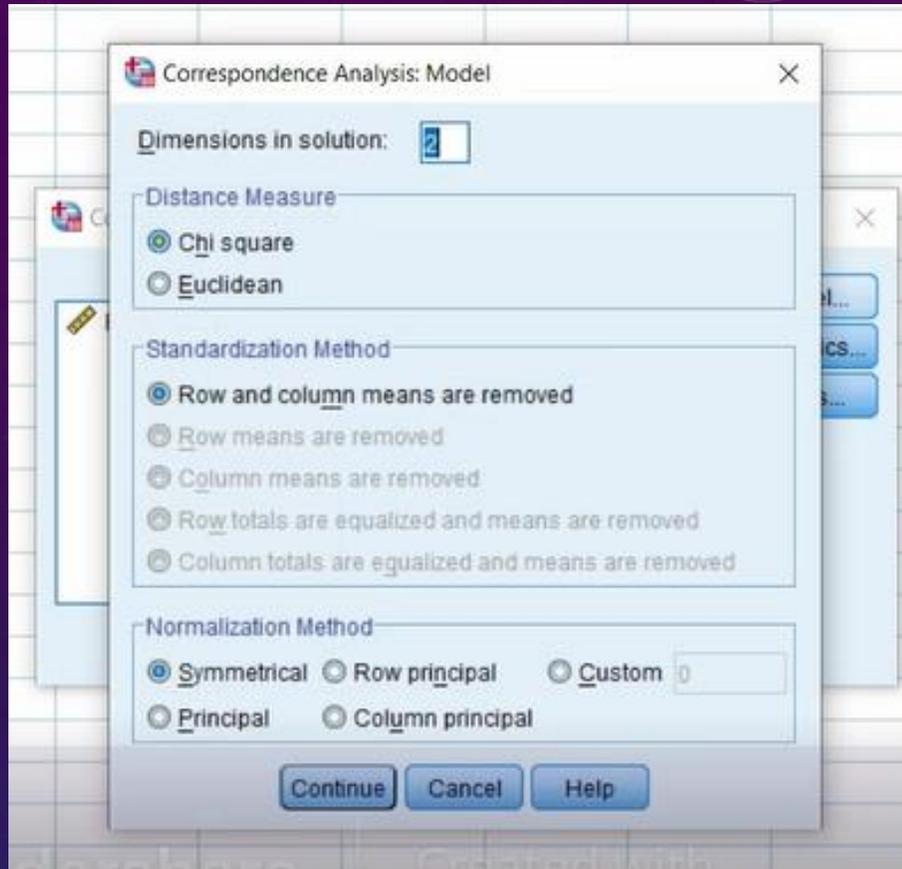
"Enter 1 as the minimum value and 7 as the maximum value, then click 'Update', and then 'Continue'."



3- Select 'Define Range' and set the range accordingly. Since there are 6 attributes in our case, the range should be from 1 to 6."

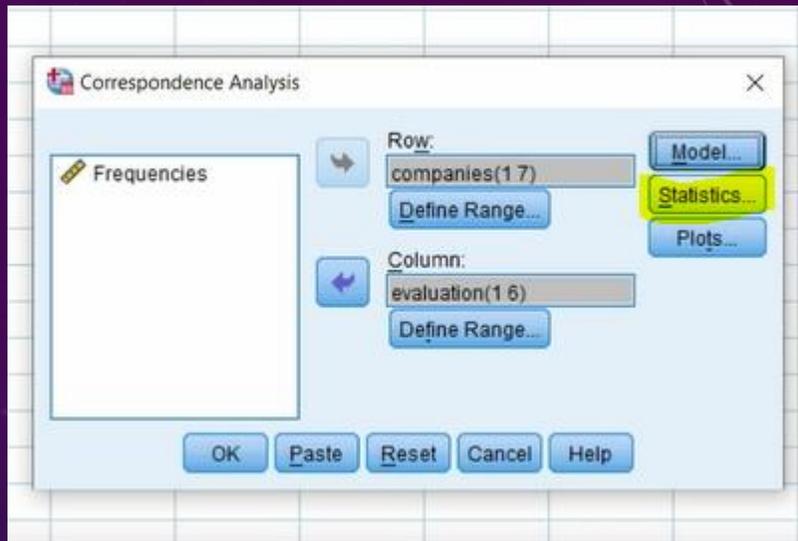


4- Click on 'Model'

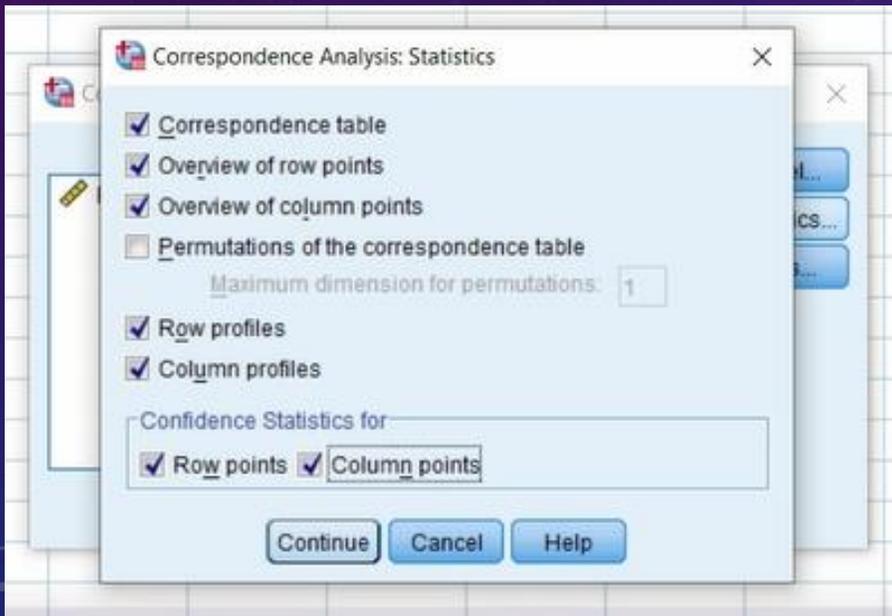


1. You can define the number of dimensions in the solution, with 2 as the default.
2. Select the distance measure, either using the **Chi-square** method or the **Euclidean distance** method.
3. Specify the normalization method: **Row and column means are removed**: This centers both rows and columns. This method is used in **standard correspondence analysis**.
4. Choose the **Symmetrical** method (or any of the other five available methods). In the **Symmetrical method**: For each dimension, the row scores are the weighted averages of the column scores divided by the corresponding cell value, and the column scores are the weighted averages of the row scores divided by the corresponding cell value. Use this method if you want to examine the **differences or similarities** between the categories of the two variables.

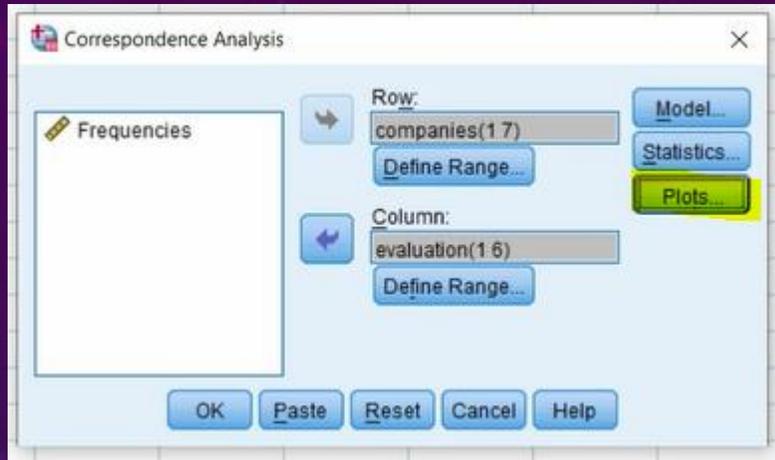
5. Click **Continue**.



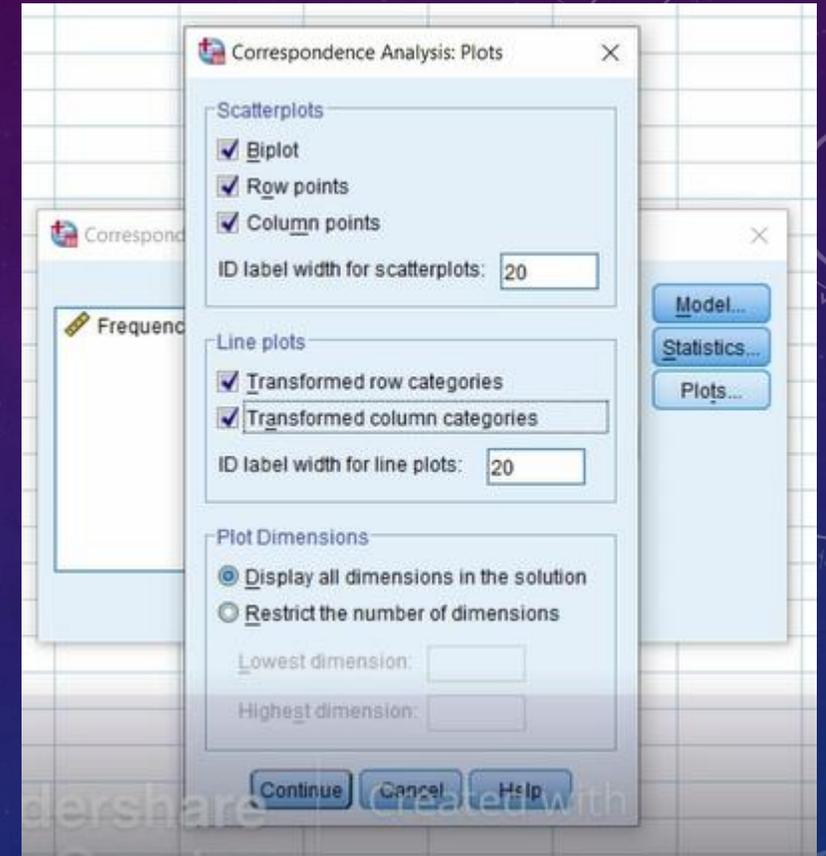
6- Click **Statistics**.



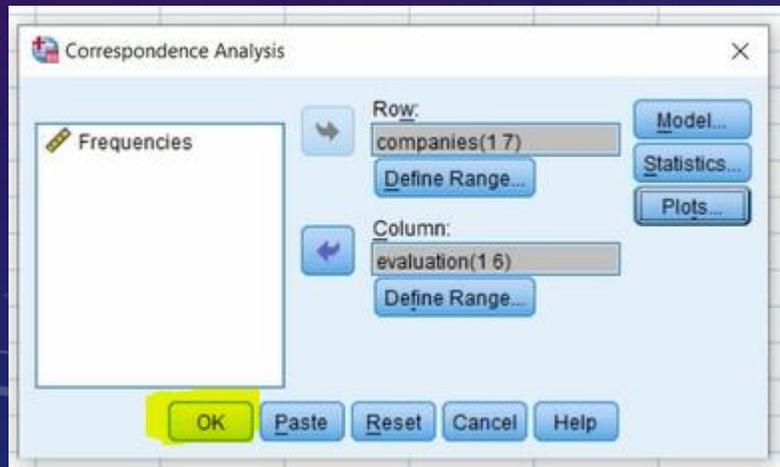
This option allows us to display the **contingency table**, as well as the **relative frequency tables** for both rows and columns. It also provides a **general overview** of the row and column points, such as **inertia**, and the **contribution** of each row or column to the total inertia, among other details.



"7. Click on 'Plots', select the desired charts, then click 'Continue'."



"Return to the main dialog box and click 'OK'"



Correspondence Table

name of companies	the evaluation of the characteristics of Visa cards						
	Ease of Use	Secure Payment	Special Offers	After-Sales Service	Delivery Time	Competitive Price	Active Margin
Amazon	56	48	64	41	32	51	292
Flipcart	65	53	51	35	19	37	260
Myntra	15	8	36	3	5	13	80
Naaptol	11	3	28	65	31	5	143
Japong	12	3	12	8	25	9	69
Snapdeal	8	34	16	8	14	15	95
Loacalbania	8	25	17	46	13	14	123
Active Margin	175	174	224	206	139	144	1062

Table 1: Contingency Table (Cross-tabulation).

Row Profiles							
name of companies	the evaluation of the characteristics of Visa cards						
	Ease of Use	Secure Payment	Special Offers	After-Sales Service	Delivery Time	Competitive Price	Active Margin
Amazon	.192	.164	.219	.140	.110	.175	1.000
Flipcart	.250	.204	.196	.135	.073	.142	1.000
Myntra	.188	.100	.450	.038	.063	.163	1.000
Naaptol	.077	.021	.196	.455	.217	.035	1.000
Japong	.174	.043	.174	.116	.362	.130	1.000
Snapdeal	.084	.358	.168	.084	.147	.158	1.000
Loacalbania	.065	.203	.138	.374	.106	.114	1.000
Mass	.165	.164	.211	.194	.131	.136	1.000

Table 2: Relative Frequency Table for Rows

Column Profiles

name of companies	the evaluation of the characteristics of Visa cards						
	Ease of Use	Secure Payment	Special Offers	After-Sales Service	Delivery Time	Competitive Price	Mass
Amazon	.320	.276	.286	.199	.230	.354	.275
Flipcart	.371	.305	.228	.170	.137	.257	.245
Myntra	.086	.046	.161	.015	.036	.090	.075
Naaptol	.063	.017	.125	.316	.223	.035	.135
Japong	.069	.017	.054	.039	.180	.063	.065
Snapdeal	.046	.195	.071	.039	.101	.104	.089
Loacalbania	.046	.144	.076	.223	.094	.097	.116
Active Margin	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: Relative Frequency Table for Columns.

significance level, and we note that the Chi-square value is 269.57 with a significance level of 0.0001, which is less than the threshold significance level of 0.05. Therefore, we can conclude that the row and column variables are not independent. In other words, there is a relationship between the companies and the measured attributes. We also observe in the third column the **inertia** explained by each principal component, and it is evident that the first two components together explain 92.3% of the total variance, indicating that the representation is very good.

Summary								
Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation
								2
1	.382	.146			.575	.575	.029	-.031-
2	.231	.053			.210	.784	.028	
3	.188	.035			.139	.923		
4	.129	.017			.065	.988		
5	.054	.003			.012	1.000		
Total		.254	269.570	.000 ^a	1.000	1.000		

a. 30 degrees of freedom

Table 04: The summary table.

From column six, we observe that Amazon contributed 0.052 in forming the first axis, which is 5.2%. Its contribution to forming the second axis was 0.7%, meaning that its contribution to the first axis was better. The highest contribution was from Naaptol in the first axis, with a contribution of 58%. Then, Jabong contributed 30% to the second axis, followed by Localbania with a 27% contribution to the second axis.

Overview Row Points ^a									
name of companies	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
Amazon	.275	-.268	.077	.009	.052	.007	.798	.040	.838
Flipcart	.245	-.412	-.109	.024	.109	.013	.654	.028	.682
Myntra	.075	-.640	.773	.035	.081	.195	.335	.296	.631
Naaptol	.135	1.290	.198	.088	.586	.023	.972	.014	.985
Japong	.065	.305	1.042	.035	.016	.305	.066	.467	.533
Snapdeal	.089	-.451	-.694	.031	.048	.187	.225	.322	.547
Loacalbania	.116	.600	-.734	.031	.109	.270	.509	.460	.968
Active Total	1.000			.254	1.000	1.000			

a. Symmetrical normalization

Table 5: Contribution of Columns in Forming the Axes.

The table shows the percentage contribution of each attribute to the first and second axes.

We observe that **After-Sales Service** has the highest contribution to the first axis with **56%**, while **Secure Payment** contributes **51%** to the second axis.

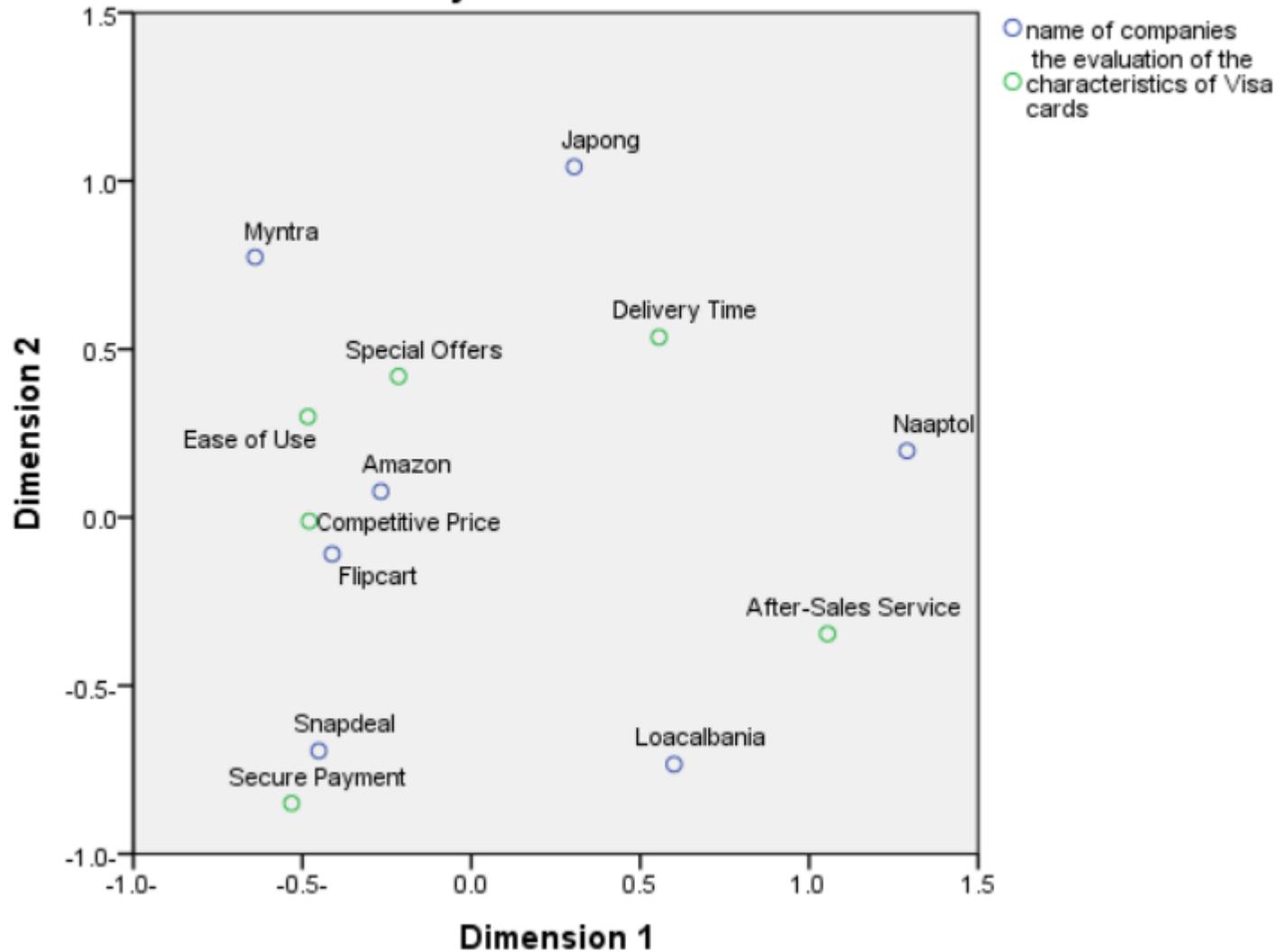
It is clear that the attribute *After-Sales Service* is associated with the company **Naaptol**, and the attribute *Secure Payment* is more prominently linked to the company **Snapdeal**.

Overview Column Points ^a									
the evaluation of the characteristics of Visa cards	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	Total
Ease of Use	.165	-.483-	.299	.029	.101	.064	.505	.117	.622
Secure Payment	.164	-.532-	-.849-	.048	.121	.512	.365	.563	.929
Special Offers	.211	-.215-	.419	.025	.026	.161	.149	.342	.492
After-Sales Service	.194	1.054	-.346-	.092	.565	.101	.894	.058	.953
Delivery Time	.131	.556	.535	.045	.106	.163	.344	.193	.537
Competitive Price	.136	-.480-	-.011-	.014	.082	.000	.828	.000	.829
Active Total	1.000			.254	1.000	1.000			

a. Symmetrical normalization

Table 6: Contribution of Columns in Forming the Axes.

Row and Column Points Symmetrical Normalization



It is evident that the attribute *After-Sales Service* is associated with the company **Naaptol**, while the attribute *Secure Payment* is more strongly present in the company **Snapdeal**

THANK YOU