

# Running a Factor analysis with XLSTAT

[demoFA.xls](#)

## Dataset for running a Factor Analysis

An Excel sheet containing both the data and the results for use in this tutorial can be downloaded by clicking [here](#).

The data are from [Kendall M. (1975). Multivariate analysis. Griffin, London] and correspond to 48 applicants for a position in firm who have been judged on 15 variables:

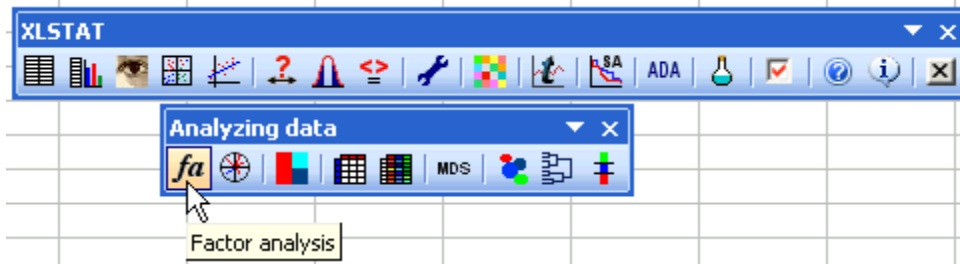
- Form of letter of application
- Appearance
- Academic ability
- Likeability
- Self-confidence
- Lucidity
- Honesty
- Salesmanship
- Experience Drive
- Ambition Grasp
- Potential Keeness to join
- Suitability

Because many correlations between the variables are high, it was felt that the judge might be confusing some of the variables, or that some variables might be redundant. Therefore a factor analysis was conducted to determine the fewer underlying factors.

Several methods are available for computing factor analysis. XLSTAT default method is the Principal factor method applied iteratively. We will apply it here to generate seven factors, and we will do a varimax rotation to facilitate the interpretation of the results.

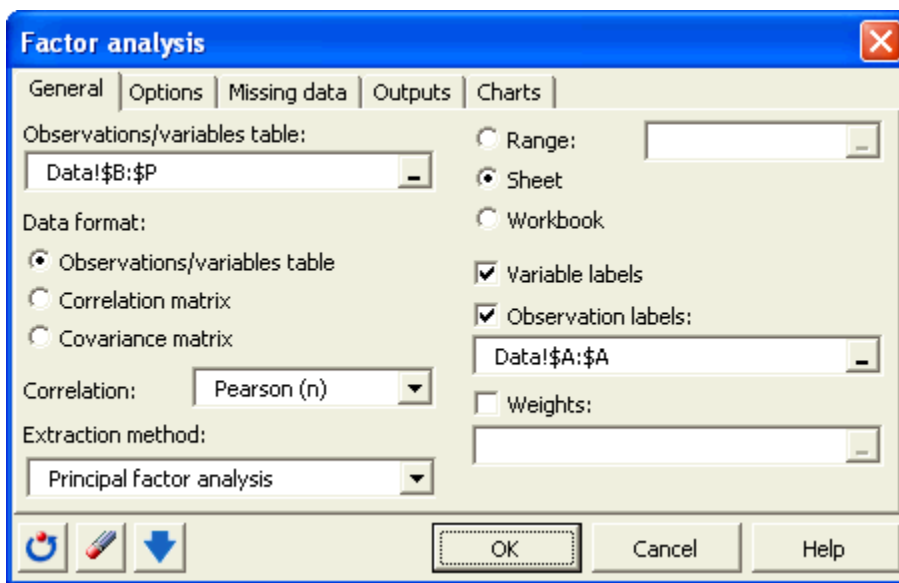
## Setting up a Factor Analysis

After opening XLSTAT, select the **XLSTAT / Analyzing data / Factor analysis** command, or click on the corresponding button of the **Analyzing data** toolbar (see below).

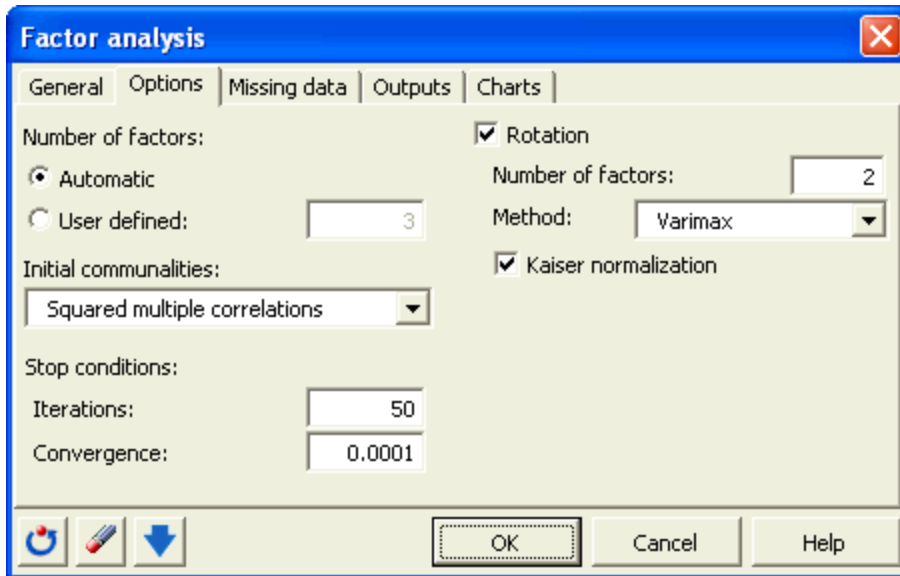


Once you've clicked on the button, the Factor analysis dialog box appears. Select the data on the Excel sheet.

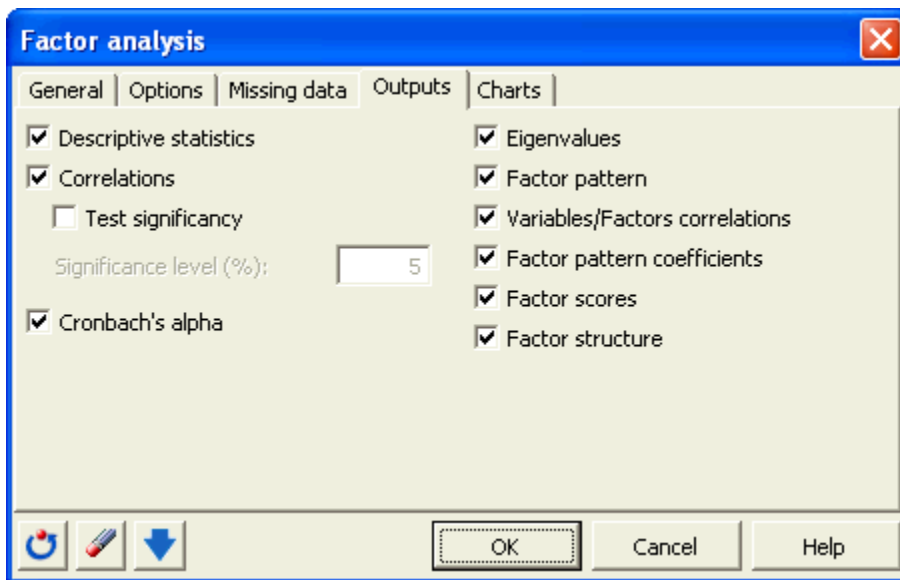
The **Observations labels** are also selected in the corresponding field.

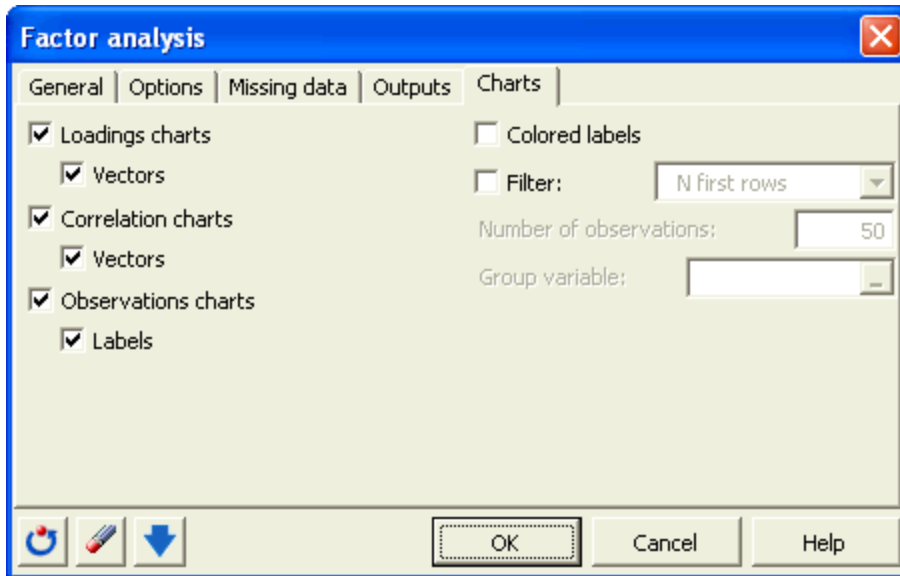


In the **Options** tab we select the **varimax** option for the rotation that will be applied to the first two factors.



The following options have been activated for the outputs and the charts.





The computations begin once you have clicked on **OK**. The results will then be displayed.

## Interpreting the results of a Factor Analysis

The first results that are displayed are the summary statistics of the selected variables, and the correlation matrix between the variables. We can see that some of the correlations are quite high (0.883 for Grasp and Lucidity).

The standardized Cronbach's alpha is computed for the whole input table. An alpha of 0.914 means that there is some redundancy among the selected variables.

The reproduced and residual correlation matrices allow you to verify if the factor analysis model is fine or not, and where it fails to reproduce correlations.

The next table shows the eigenvalues resulting from the factor analysis. We can see that with 4 factors we keep 75.5 % of the variability of the initial data.

Eigenvalues:										
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Eigenvalue	7,325	1,738	1,253	0,859	0,216	0,184	0,119	0,078	0,011	0,003
Variability	48,835	11,587	8,352	5,726	1,440	1,228	0,792	0,520	0,073	0,021
Cumulative	48,835	60,422	68,774	74,500	75,940	77,168	77,960	78,480	78,552	78,574

*Note: the eigenvalues displayed above are those obtained with the principal factors extraction method.*

With the principal components analysis we would have obtained the following results:

Eigenvalues:										
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
Eigenvalue	7,514	2,056	1,456	1,198	0,739	0,495	0,351	0,310	0,257	0,185
Variability	50,092	13,709	9,705	7,986	4,928	3,297	2,342	2,066	1,713	1,233
Cumulative	50,092	63,801	73,506	81,492	86,420	89,717	92,059	94,125	95,838	97,071

Next, we can see that the varimax rotation has changed the way each factor explains part of the variance.

The varimax rotation makes the interpretation easier by maximizing the variance of the squared factors loadings by column. For a given factor, high loadings become higher, low loadings become lower, and intermediate loadings become either lower or higher.

Percentage of variance after Varimax rotation:		
	D1	D2
Variability (%)	44,389	16,032
Cumulative %	44,389	60,421

The next results we want to look at, are the factor loadings after the varimax rotation. These results are used to interpret the meaning of the (rotated) factors.

Factor pattern after Varimax rotation:		
	D1	D2
Form of letter	0,215	0,640
Appearance	0,515	0,157
Academic abil	0,001	0,292
Likeability	0,636	0,039
Self-confidenc	0,866	-0,055
Lucidity	0,857	0,133
Honesty	0,582	-0,358
Salesmanship	0,840	0,281
Experience	0,080	0,805
Drive	0,767	0,364
Ambition	0,844	0,220
Grasp	0,854	0,291
Potential	0,843	0,356
Keeness to jo	0,719	0,134
Suitability	0,389	0,772

From this table we can see that the first factor is highly positively related to Ambition, Self-confidence, Salesmanship and Lucidity.

The second factor is loaded on Form of application, Experience, and Suitability.

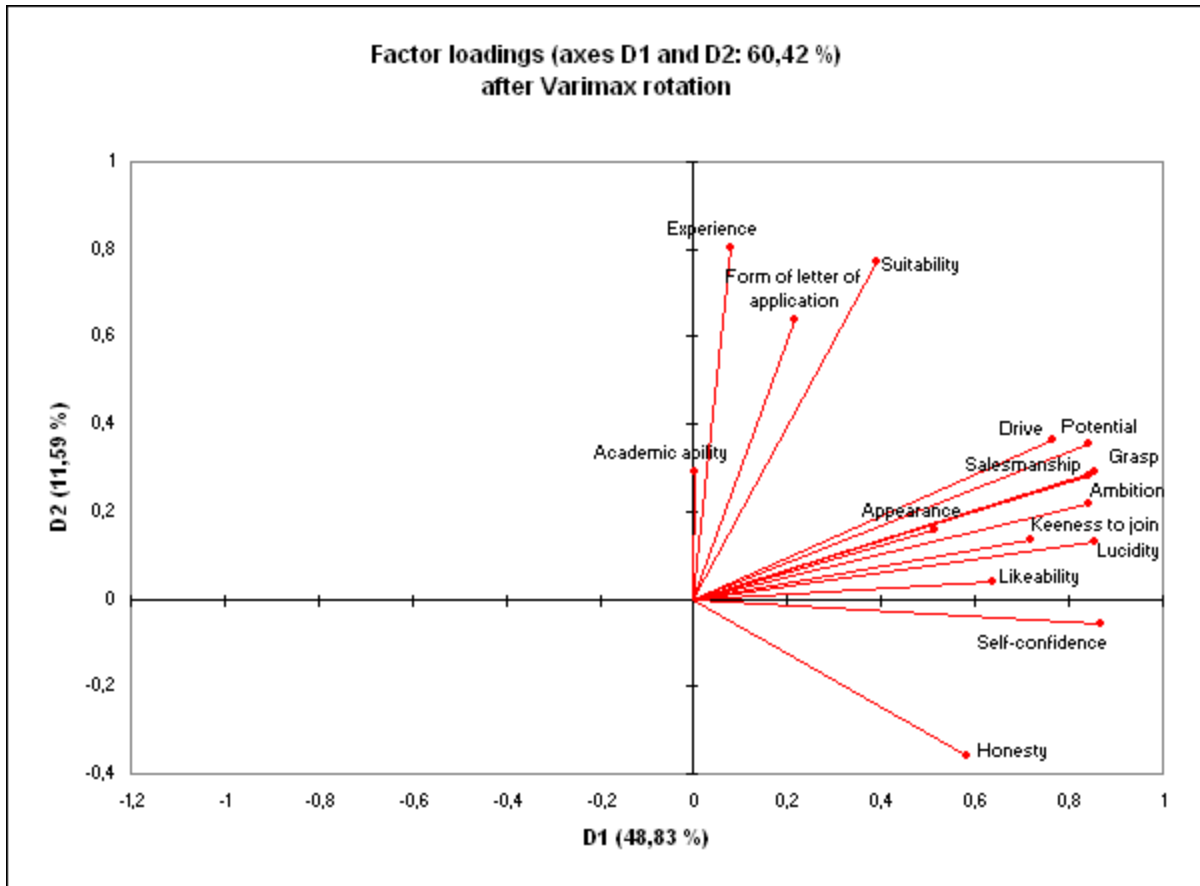
The third factor is heavily loaded on Honesty.

The fourth factor is positively loaded on Academic and negatively on Keeness to join.

The Appearance appears to be significant only on the fifth factor.

From these results, we can understand that the individuals that have high scores on the first factor are promising salesmen, while for other jobs such as management, individuals with high coordinates on the second and third factors might be more appropriate.

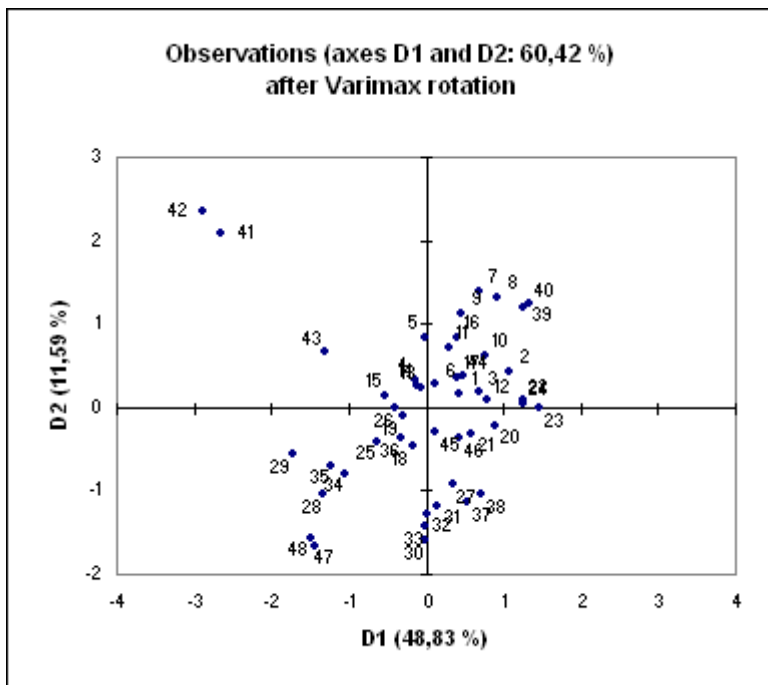
The following chart gives the position of the variables on axes F1 and F2. Other charts mixing other factors can be displayed.



The following table gives the factor scores after the varimax rotation, which are the estimated coordinates of the observations on the factor axes.

Factor scores after Varimax rotation:		
	D1	D2
1	0,440	0,164
2	1,068	0,427
3	0,691	0,177
4	-0,147	0,343
5	-0,025	0,829
6	0,110	0,277
7	0,688	1,388
8	0,911	1,310
9	0,464	1,131
10	0,760	0,631
11	0,288	0,723
12	0,779	0,084
13	-0,066	0,246
14	-0,124	0,265
15	-0,528	0,137
16	0,390	0,841
17	0,398	0,362
18	-0,172	-0,458
19	0,798	0,101

XLSTAT displays the 2D maps on the selected factors. The chart below represents the map for F1 and F2.



Watch this video to see the demonstration of this tutorial.

[http://www.youtube.com/watch?feature=player\\_embedded&v=-52vvNceI90](http://www.youtube.com/watch?feature=player_embedded&v=-52vvNceI90)