FACTOR ANALYSIS

By:
Dr. Akash Asthana
Assistant Professor,
Dept. of Statistics,
University of Lucknow, Lucknow

FACTOR ANALYSIS

- It is a dimension reduction technique.
- It is used when in analysis a large number of variables and it is not possible to deal with all the variables simultaneously.
- The factor analysis is of two types:

1. Exploratory Factor Analysis (EFA)
2. Confirmatory Factor Analysis (CFA)

- The EFA is used when the structure of underlying factors is unknown and is to be determine.
- The CFA is used when the structure of underlying factors is already known and it is required to check whether the data collected confirm that structure or not.

EXPLORATORY FACTOR ANALYSIS

- The main objectives of the EFA are:

1. To identifying the underlying dimensions or factors that explain the variation (or correlations) among the set of variables.
2. To obtain a new smaller set of uncorrelated variables to replace the original set of correlated variables in subsequent analysis.
3. To obtain a smaller set of salient variables from a large set for use in subsequent analysis.

EFA VERSUS PRINCIPAL COMPONENT ANALYSIS

- Both the techniques the Exploratory Factor Analysis (EFA) and Principal Component Analysis (PCA) are termed as data reduction techniques. But EFA and PCA can't be seprated from each other. PCA can be termed as a method of performing the EFA. The PCA is a technique in which we obtain the uncorrelated linear combinations of the variables under study which are able to explain the variation (or correlation) in the dataset, but is unable to answer the 2nd and 3rd objectives of the EFA i.e. How many factors should be retained in the data and which variable should be considered within which factor.

MODEL FOR FACTOR ANALYSIS

- Factor Analysis is based on a model in which the observed vector is partitioned into an unobserved systematic part and an unobserved error part.
- The components of error part are considered as independent whereas systematic part is taken as a linear combination of relatively small number of unobserved factor variables.
- This model separates the effect of factors from the error.
- The model for Factor Analysis is defined as:

$$X = \mu + \Lambda f + U$$

- X is the px1 vector of observed variables. It may be considered as the score on a battery of test.
- μ is (px1) vector of the average score of this test in the population.

MODEL FOR FACTOR ANALYSIS

- f is (mx1) vector of unobserved variables called as common factors. These are the scores on hidden (underlying) ability whose linear combinations enters to the test scores.
- Λ is (pxm) matrix of the component loadings or factor loadings. It consists of the coefficients of the linear combinations of factors.
- U is (px1) vector of random error terms.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ . \\ . \\ . \\ X_i \\ . \\ . \\ . \\ X_p \end{bmatrix}, \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ . \\ . \\ . \\ \mu_i \\ . \\ . \\ . \\ \mu_p \end{bmatrix}, f = \begin{bmatrix} f_1 \\ f_2 \\ . \\ . \\ . \\ f_j \\ . \\ . \\ . \\ f_m \end{bmatrix}, U = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_i \\ . \\ . \\ . \\ u_p \end{bmatrix}, \Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1j} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2j} & \dots & \lambda_{2m} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ \lambda_{i1} & \lambda_{i2} & \dots & \lambda_{ij} & \dots & \lambda_{im} \\ . & . & & . & & . \\ . & . & & . & & . \\ . & . & & . & & . \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pj} & \dots & \lambda_{pm} \end{bmatrix}$$

ASSUMPTIONS FOR FACTOR ANALYSIS

1. The mean of random error term is 0, i.e. $E[U] = 0$
2. The mean of common factor is 0, i.e. $E[f] = 0$
3. The variance of error term is ψ_i, i.e. $V(u_i) = \psi_i; i=1, 2, \ldots, p$.
4. The error terms are independent of each other, i.e.,

   $Cov(u_i, u_j) = 0; i \neq j=1, 2, \ldots, p$

   The assumption 3 and 4 can be collectively written as: $V(U) = \Psi = diag[\psi_1 \psi_2 \ldots \psi_p]$.
5. The variance of the common factor is given by: $V(f) = \Phi$.

   If the factors are considered to be orthogonal then $V(f) = I$.
6. The common factors and error terms are independent of each other, i.e. $Cov(u_i, f_j) = 0; i = 1, 2, \ldots, p \ \& \ j = 1, 2, \ldots, m$.

ESTIMATION OF PARAMETERS

- Now consider the variance of X vector:

$$V(X) = V(\mu+\Lambda f+U)$$
$$or, \quad \Sigma = \Lambda\Phi\Lambda' + \Psi$$

  If factors are considered to be orthogonal then: $\Sigma = \Lambda\Lambda' + \Psi$
- Therefore, in factor analysis there are basically two type of parameters involved:

1. pm parameters in matrix Λ.
2. m parameters in diagonal matrix Ψ.

- Therefore there are a total of p(m+1) parameters which are required to be estimated.
- There are several methods for obtaining the estimates of these parameters among which two most commonly used methods are:

1. Principal component method
2. Method of maximum likelihood.

USING PRINCIPAL COMPONENT METHOD

- It is discussed in detail in the lecture of Principal component Analysis.
- It has following steps:

1. First transform the matrix of all variables under consideration to a matrix X such that mean of X will be 0.
2. Obtain the Variance-covariance matrix of X, Σ (or its MLE) under the assumption that X is Normally Distributed.
3. Obtain the Characteristic roots of Σ and arrange them in descending order ($\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$).
4. For each distinct Eigen root obtain Eigen vector.

USING PRINCIPAL COMPONENT METHOD

5. Normalize these Eigen vectors dividing these by their norms ($\beta^{(1)}$, $\beta^{(2)}$, …,$\beta^{(p)}$).
6. Then obtain the principal components by multiplying these β_i's with X (i.e. $\beta^{(1)}X$, $\beta^{(2)}X$, …, $\beta^{(p)}X$)
7. In the situation if the unit of measurements for variables are not same it is better to use correlation matrix in place of variance-covariance matrix.

- Using these steps the estimates of elements of Λ can be obtained.
- Now for obtaining the estimate of elements of Ψ, we can use:

$$\Psi = \Sigma - \Lambda\Lambda'$$

USING METHOD OF MAXIMUM LIKELIHOOD

- In this method it is assumed that the vector X have Multivariate Normal distribution with mean μ and variance-covariance matrix Σ, i.e. $X \sim N_p(\mu, \Sigma)$.
- Let X1, X2, …, Xn be the random sample from above distribution. Then the log-likelihood function can be written as:

$$\ln L(\mu, \Sigma) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln|\Sigma| - \frac{1}{2}\sum_{i=1}^{n}(X_i - \mu)'\Sigma^{-1}(X_i - \mu)$$

- Putting $\Sigma = \Lambda\Lambda' + \Psi$ in log-likelihood we get:

$$\ln L(\mu, \Sigma) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln|\Lambda\Lambda' + \Psi| - \frac{1}{2}\sum_{i=1}^{n}(X_i - \mu)'(\Lambda\Lambda' + \Psi)^{-1}(X_i - \mu)$$

- However, it is not quite easy to obtain the estimates of Λ and Ψ.
- A lot of methods can be used to maximize it among which main methods are steepest descent method, Newton-Raphson iterative procedure and scoring method.

COMPUTATION OF FACTOR SCORES

- For obtaining the estimate of factor scores factor (*f*) analysis model is reconsidered:

$$X = \mu + \Lambda f + U$$

- It is fitted in same manner as Linear Regression model. Instead of Λ its estimate obtained by above stated method is used and model become: $X = \mu + \hat{\Lambda} f + U$
- For estimating f following methods are used:
- The estimate of f can be obtained by using:

1. Ordinary Least Square (OLS Method)
2. Weighted Least Square (Bartlett's Method)
3. Regression Method

COMPUTATION OF FACTOR SCORES

1. **OLS Method:**

- In this method the estimates are obtained by minimizing the error sum of square (U'U). The estimate of factor score is given by:

$$\hat{f} = \left(\hat{\Lambda}'\hat{\Lambda}\right)^{-1}\hat{\Lambda}'(X - \bar{X})$$

2. **Bartlett's Method:**

- In OLS method V(U) is considered as identity matrix but in factor analysis it is considered as Ψ matrix. Identity matrix will be its one special case therefore Bartlett had suggested to use the weighted least square method. Using this method the estimate of factor score is obtained as:

$$\hat{f} = \left(\hat{\Lambda}'\hat{\Psi}^{-1}\hat{\Lambda}\right)^{-1}\hat{\Lambda}'\hat{\Psi}^{-1}(X - \bar{X})$$

COMPUTATION OF FACTOR SCORES

3. Regression Method:

- In this method the factor scores are obtained by using maximum likelihood method.
- Here the joint distribution of X and *f* is taken as:

$$\begin{pmatrix} X \\ f \end{pmatrix} \sim N\left[\begin{pmatrix} \mu \\ 0 \end{pmatrix}, \begin{pmatrix} \Lambda\Lambda' + \Psi & \Lambda \\ \Lambda' & I \end{pmatrix}\right]$$

- The by using conditional expectation it is obtained that:

$$E(f \mid X) = L'(LL' + \Psi)^{-1}(X - \mu)$$

- Using the estimates of L and Ψ the estimate of factor scores will be:

$$\hat{f} = \hat{\Lambda}'\left(\hat{\Lambda}\hat{\Lambda}' + \Psi\right)^{-1}(X - \bar{X})$$

- Here $\bar{X}$ is the estimate of μ.

ROTATION OF FACTORS

The unrotated output maximizes variance accounted for by the first and subsequent factors, and forces the factors to be orthogonal. This data-compression comes at the cost of having most items load on the early factors, and usually, of having many items load substantially on more than one factor. Rotation serves to make the output more understandable, by seeking so-called “Simple Structure” which is a pattern of loadings where each item loads strongly on only one of the factors, and much more weakly on the other factors. It is of two types:

1. Orthogonal rotation
2. Oblique rotation

ORTHOGONAL ROTATION

- It is a transformational system used in factor analysis in which the different underlying or latent variables are required to remain separated from or uncorrelated with one another. There are three different methods that can be used for Orthogonal rotation:

1. **Varimax rotation:** It is an orthogonal rotation of the factor axes to maximize the variance of the squared loadings of a factor (column) on all the variables (rows) in a factor matrix, which has the effect of differentiating the original variables by extracted factor. A varimax solution yields results which make it as easy as possible to identify each variable with a single factor. This is the most common and most frequently used rotation method.
2. **Quartimax rotation:** It is an orthogonal alternative which minimizes the number of factors needed to explain each variable. This type of rotation often generates a general factor on which most variables are loaded to a high or medium degree.
3. **Equimax rotation:** It is a compromise between Varimax and Quartimax criteria.

OBLIQUE ROTATION

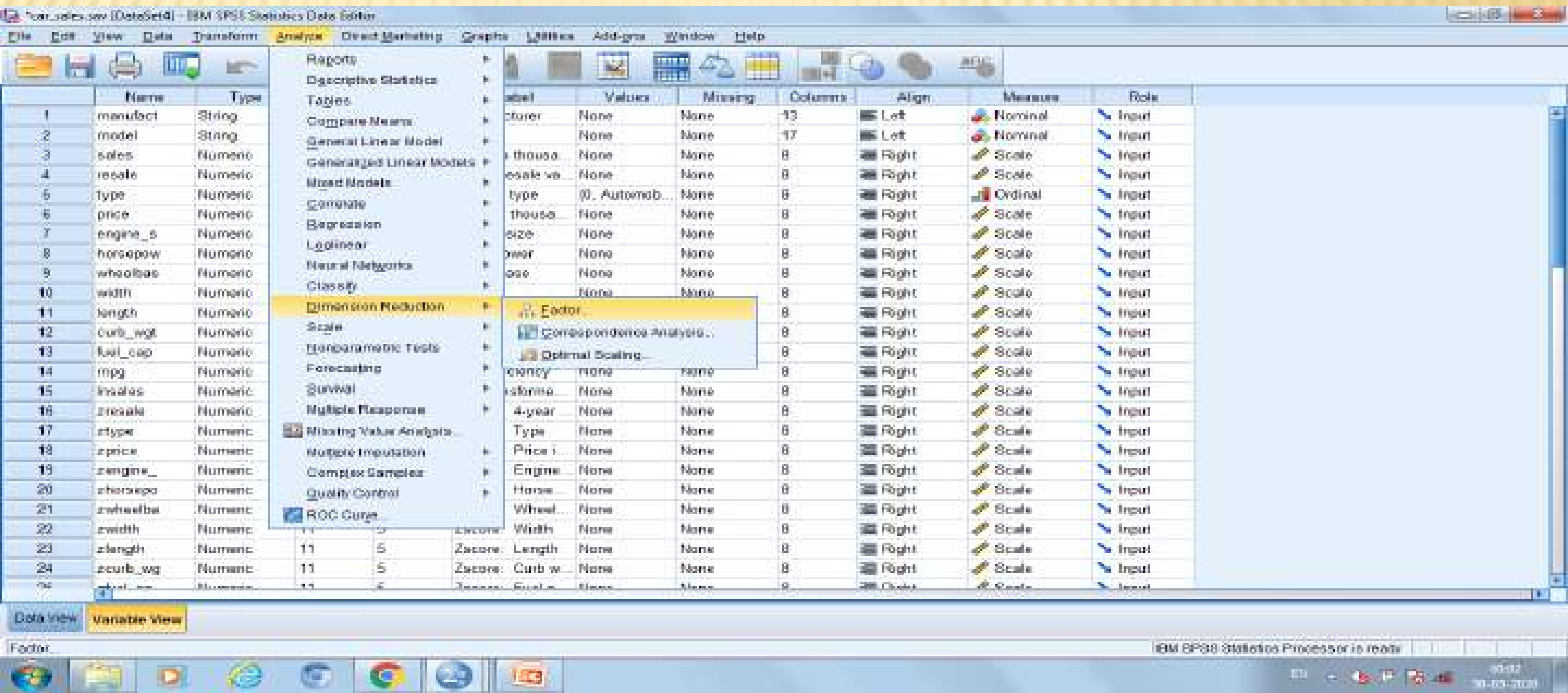
- It is a transformational system used in factor analysis when two or more factors (i.e., latent variables) are correlated. Oblique rotation reorients the factors so that they fall closer to clusters of vectors representing manifest variables, thereby simplifying the mathematical description of the manifest variables. There are two methods used for the oblique rotation:

1. **Direct oblimin rotation:**
2. **Promax Rotation**

- Promax method is similar to Direct oblimin method but is computationally faster than it.

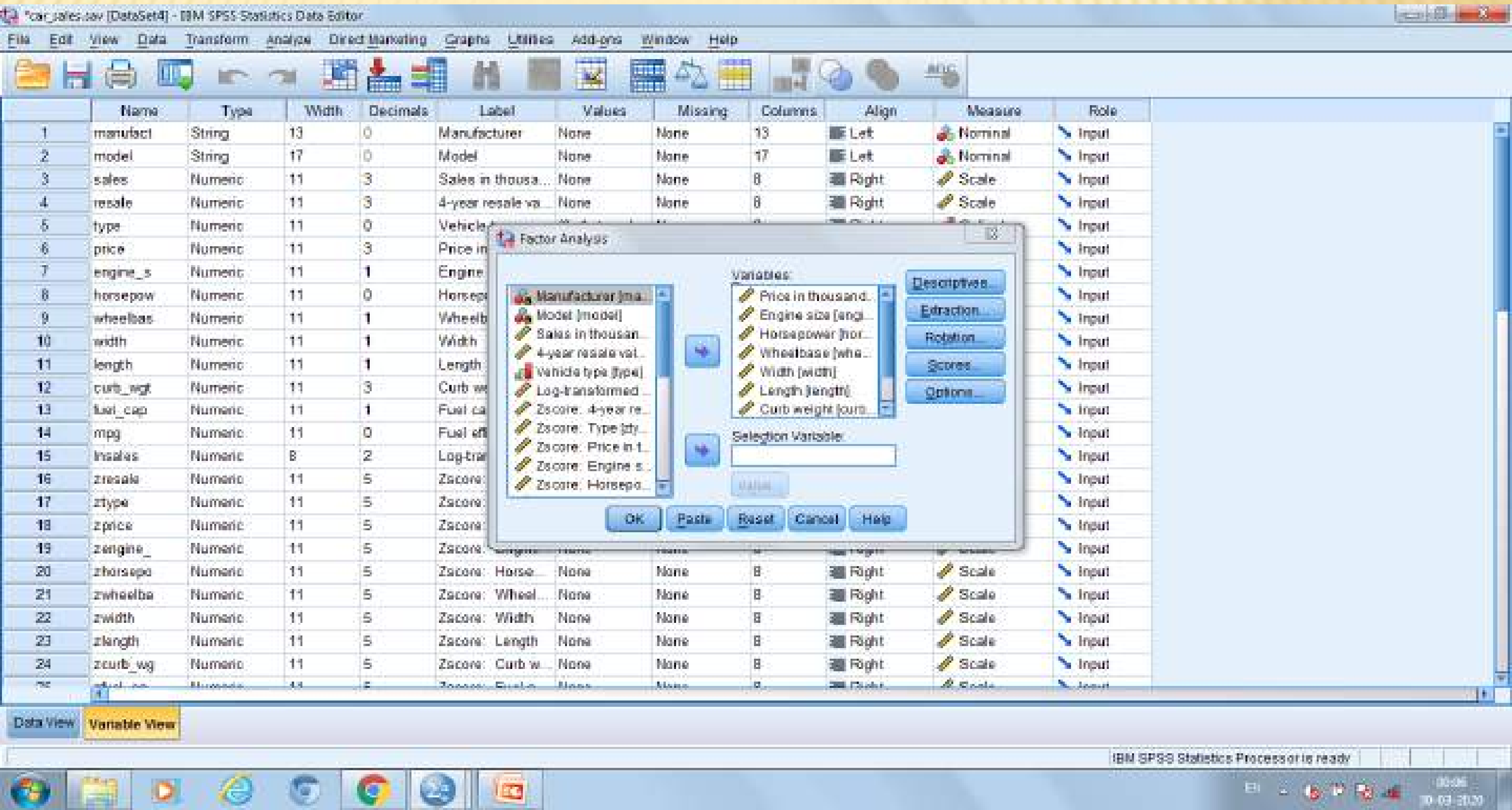
FACTOR ANALYSIS: AN EXAMPLE USING SPSS

- For performing Exploratory Factor Analysis (EFA) using SPSS Following steps are used.
- Click on Analyze → Dimension Reduction → Factor



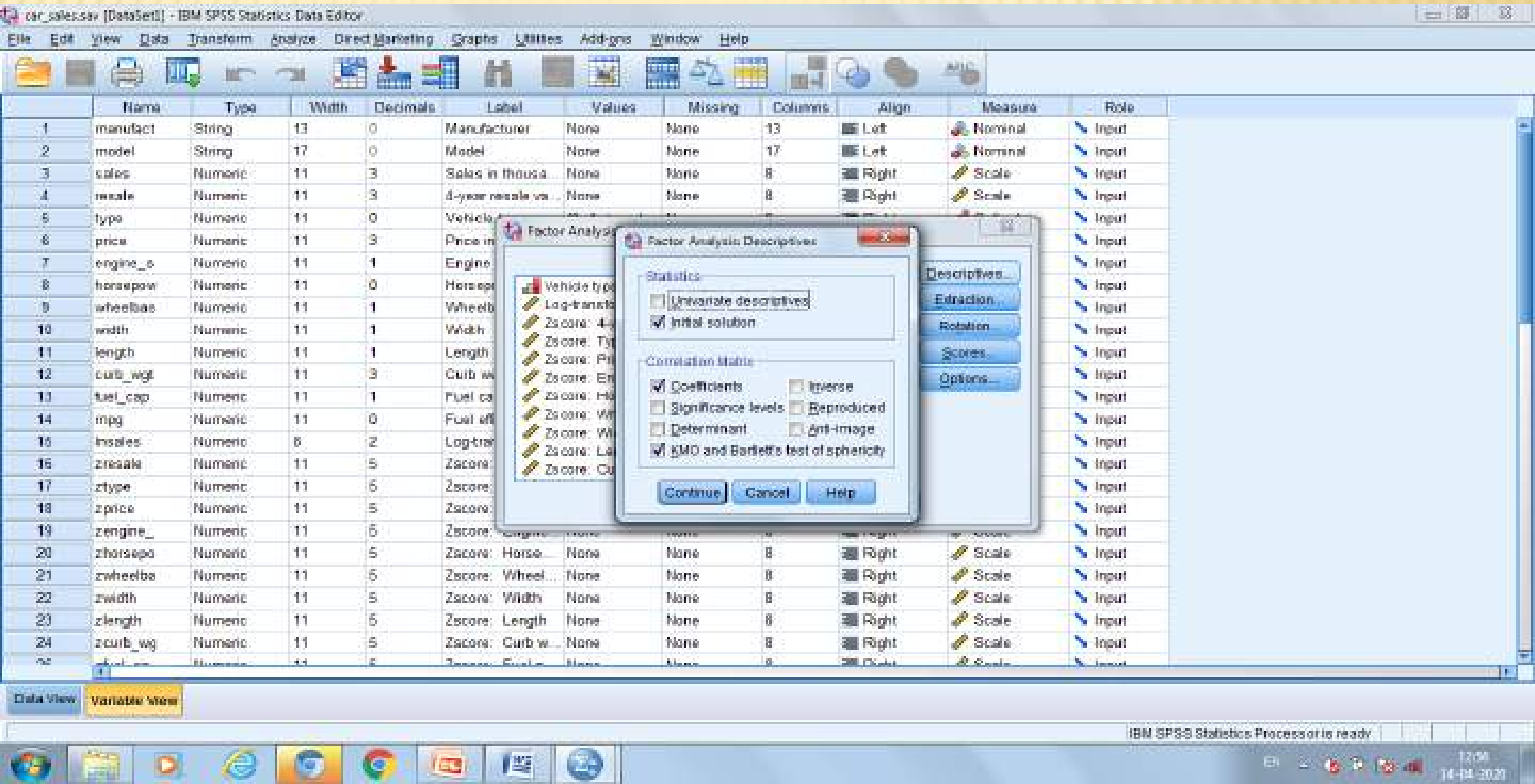
EXAMPLE (CONTD.)

- It will open the factor analysis window put all the variables required for EFA in variable box. Then click on Extraction.



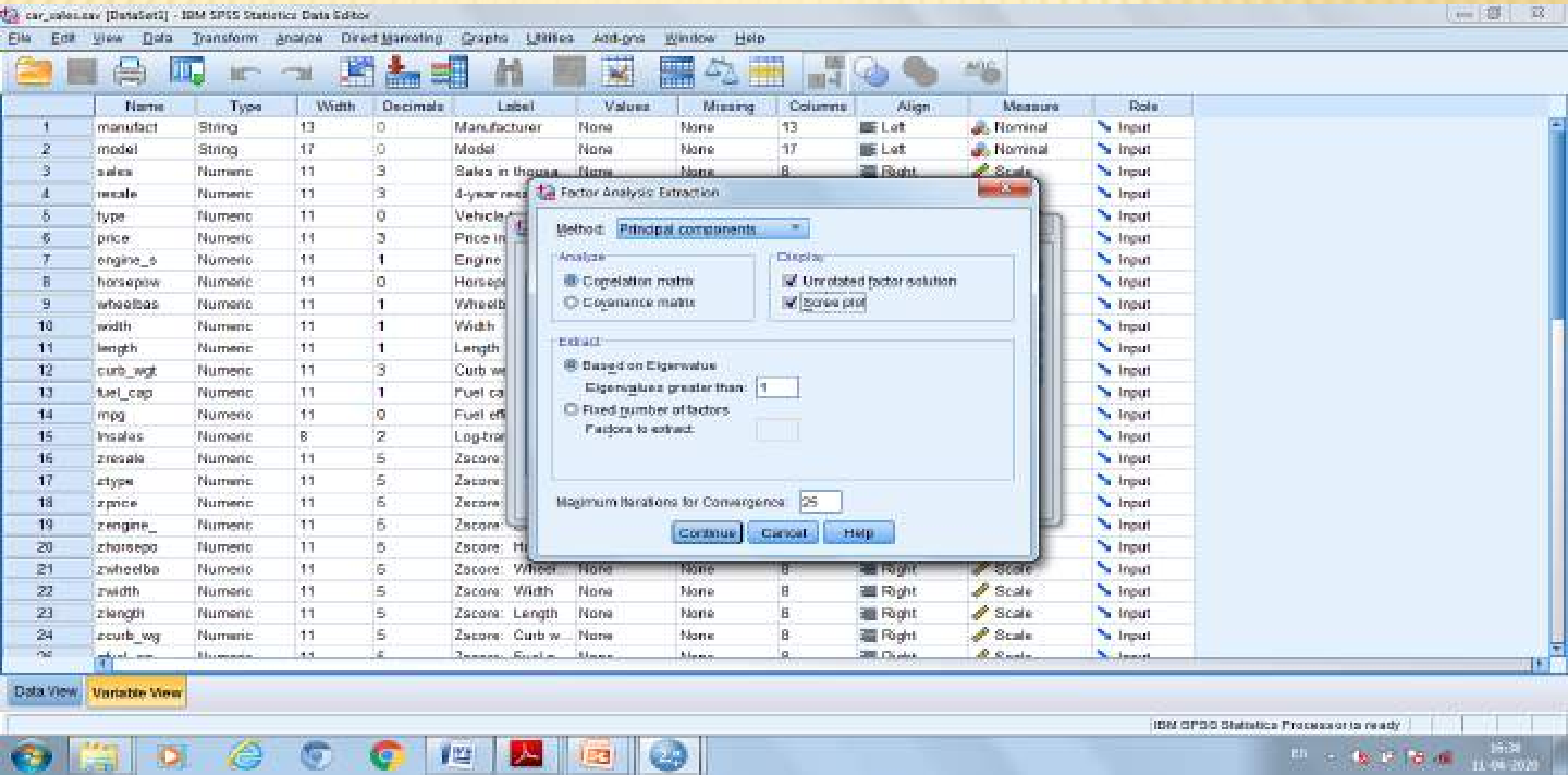
EXAMPLE (CONTD.)

- Click on Descriptive button it will open a new window. In this window select coefficients in correlation matrix and KMO and Bartlett's test for sphericity. Click on continue.



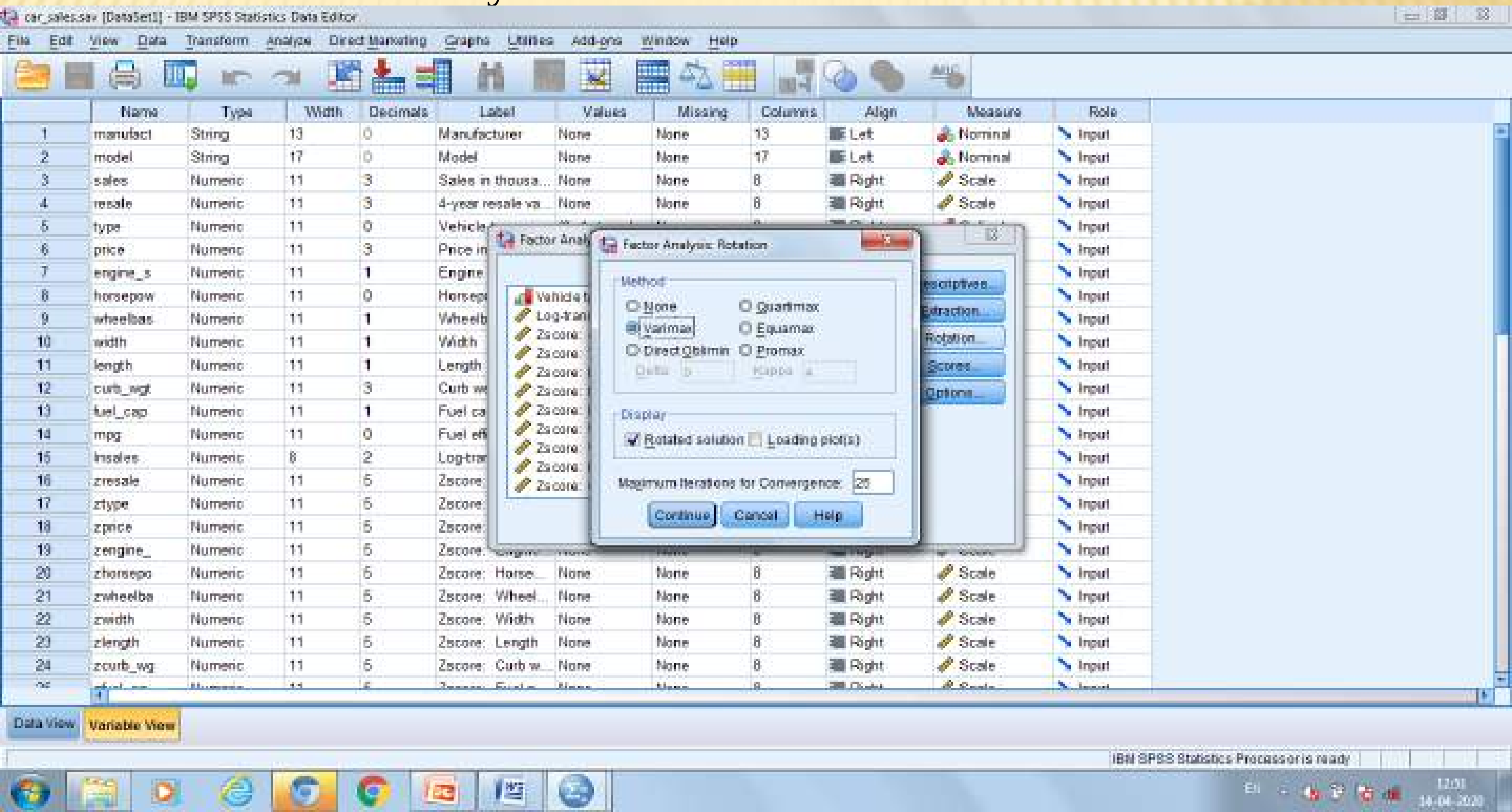
EXAMPLE (CONTD.)

- On clicking Extraction window will be open. Click on Correlation matrix and Scree plot. For number of factors to extracted you can choose any option. In this based on Eigen values is selected. By default it take Eigen Value > 1 which can be changed. Click on continue.



EXAMPLE (CONTD.)

- On clicking Rotation a window will be open. Click on Varimax rotation as it is most commonly used method (As per requirement one can choose any other rotation method. Click on continue.

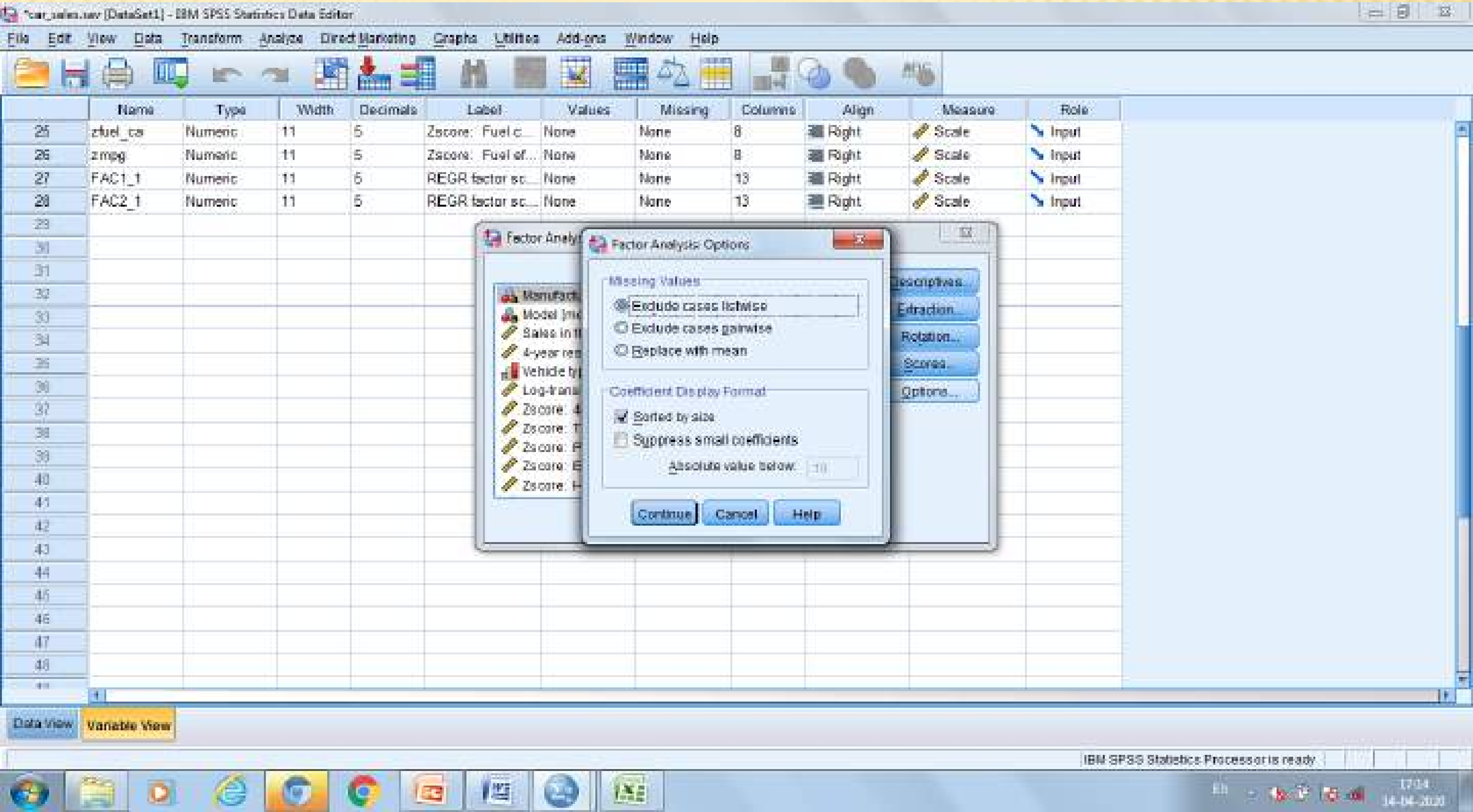


EXAMPLE (CONTD.)

- Click on continue, then click on Scores → Save as variable → Display factor score coefficient matrix. Click on continue.

EXAMPLE (CONTD.)

- Click on Options. It will open a new window. Click on Sorted by Size and then on continue. Then click on OK.



EXAMPLE (CONTD.)

- The output of SPSS shows a no. of tables. The interpretation of these tables is as follows:
- Table1: Correlation Matrix

  As most of the variables are highly correlated it can be said that Factor Analysis is suitable for the data and will give very good results.

Correlation Matrix									
	Price in thousands	Engine size	Horse power	Wheelbase	Width	Length	Curb weight	Fuel capacity	Fuel efficiency
Price in thousands	1.000	0.624	0.841	0.108	0.328	0.155	0.527	0.424	-0.492
Engine size	0.624	1.000	0.837	0.473	0.692	0.542	0.761	0.667	-0.737
Horse power	0.841	0.837	1.000	0.282	0.535	0.385	0.611	0.505	-0.616
Wheelbase	0.108	0.473	0.282	1.000	0.681	0.840	0.651	0.657	-0.497
Width	0.328	0.692	0.535	0.681	1.000	0.706	0.723	0.663	-0.602
Length	0.155	0.542	0.385	0.840	0.706	1.000	0.629	0.571	-0.448
Curb weight	0.527	0.761	0.611	0.651	0.723	0.629	1.000	0.865	-0.820
Fuel capacity	0.424	0.667	0.505	0.657	0.663	0.571	0.865	1.000	-0.802
Fuel efficiency	-0.492	-0.737	-0.616	-0.497	-0.602	-0.448	-0.820	-0.802	1.000

EXAMPLE (CONTD.)

- Table2: It shows the result of KMO and Bartlett's test. It shows the results of two results:

1. **Kaiser-Meyer-Olkin Measure of Sampling Adequacy:** It shows the proportion of variance in your variables that might be caused by underlying factors. Higher value of it indicates the usefulness of the analysis.
2. **Bartlett's test of sphericity**: It is used to test the null hypothesis that the correlation matrix is identity. P-value smaller than 0.05 shows that correlation matrix is not Identity and Factor Analysis may be useful.

- Here the value of KMO measure is 0.843 which shows that FA is useful in this case and Bartlett's test shows that the correlation matrix is not identity.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.843
Bartlett's Test of Sphericity	Approx. Chi-Square	1407.020
	df	36.000
	Sig.	<0.001

EXAMPLE (CONTD.)

- **Table3: Communalities:** It shows two values Initial and Extraction. Initial communalities shows how much percentage of the variation in the variable is caused by the other variables. The Extraction communalities shows how much percentage of the variation in the variable is caused by the factors.

Communalities

	Initial	Extraction
Price in thousands	1.000	0.853
Engine size	1.000	0.838
Horsepower	1.000	0.878
Wheelbase	1.000	0.868
Width	1.000	0.745
Length	1.000	0.797
Curb weight	1.000	0.854
Fuel capacity	1.000	0.762
Fuel efficiency	1.000	0.726

EXAMPLE (CONTD.)

- **Table3: Total Variance Explained:** Table is divide into three parts. First part shows initial Eigen Values, which indicates how much percent of variance can be explained by a particular factor (% of variance) and the factor along with previous factors how much percent of variance can be explained (cumulative %). Second part shows how many factors are extracted from the data or in other words how many factor are sufficient to explain the variation in the data. As per rule of thumb the factors having Eigen value >1.0 or cumulative % extraction more the 70 % are sufficient to explain the data. Third part shows the rotated sum of square loadings, which is the result obtained by the rotation of the factor. It distribute the % of variance explained by the factors approximately equal to each factor.
- In our results the Eigen values for first two factors are more than 1.0 and it can explain 81% of total variation in the data. Therefore these two factor can be considered sufficient for the data. In initial solution first factor explain 64% whereas second factor explain 17% of the total variation, however in rotated solution first factor explains the 43% and second factor 38% of the total variation.

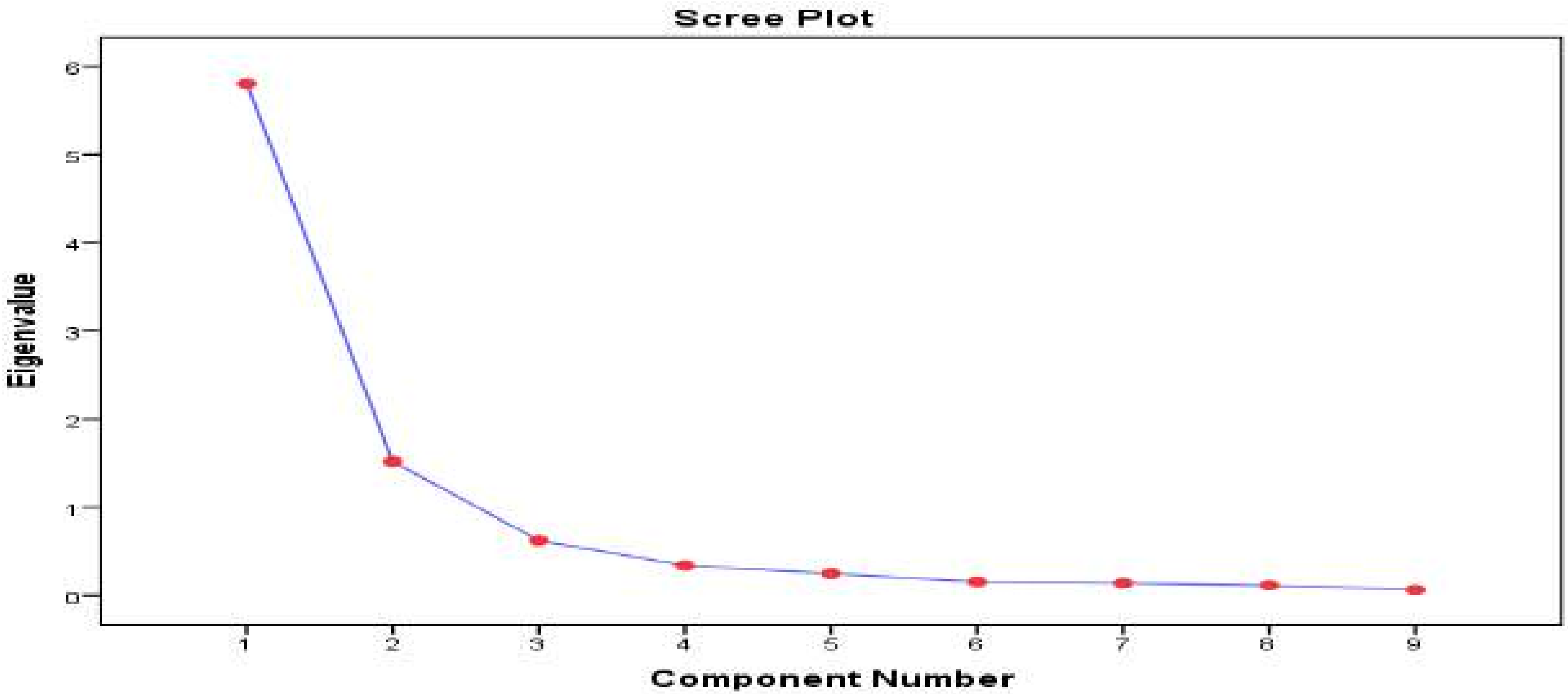
EXAMPLE (CONTD.)

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.804	64.490	64.490	5.804	64.490	64.490	3.911	43.457	43.457
2	1.517	16.860	81.349	1.517	16.860	81.349	3.410	37.892	81.349
3	0.623	6.918	88.267						
4	0.338	3.757	92.025						
5	0.247	2.747	94.772						
6	0.155	1.719	96.491						
7	0.139	1.547	98.038						
8	0.114	1.266	99.305						
9	0.063	0.695	100.000						

EXAMPLE (CONTD.)

- **Scree Plot:** It is another method to obtain the required number of factors. In this the Eigen value is plotted against the number of factor. The point after which the curve become parallel to the horizontal axis will be the last factor selected. In the given example after 2nd factor curve become parallel to the horizontal axis therefore only two factors are retained.



EXAMPLE (CONTD.)

- **Table 4: Component Matrix:** This table shows the correlation of the factor with the variables under consideration. It is helpful in the detection of the structure of the factor. A variable is said to be contained in a factor if the correlation of the variable with the factor is maximum among all the factors. In the example 8 out of 9 variables are highly correlated to 1^{st} factor as compared to second factor therefore these 8 variables (Curb weight, Engine size, Fuel capacity, Fuel efficiency, Width, Horsepower, Length, Wheelbase, Price in thousands) are said to be contained in 1^{st} factor whereas 9^{th} one (price in thousand) is said to be contained in 2^{nd} factor however the correlation of 9^{th} variable with both the factors are approximately similar and it may be contained in any of the factors. It is drawback of the component matrix and therefore the rotated component matrix is used.

EXAMPLE (CONTD.)

Component Matrix

	Component	
	1	2
Curb weight	0.923	0.039
Engine size	0.882	-0.243
Fuel capacity	0.865	0.119
Fuel efficiency	-0.845	0.106
Width	0.829	0.241
Horse power	0.771	-0.533
Length	0.732	0.512
Wheelbase	0.722	0.588
Price in thousands	0.610	-0.694

Rotated Component Matrix

	Component	
	1	2
Wheelbase	0.931	0.040
Length	0.887	0.104
Width	0.779	0.371
Fuel capacity	0.725	0.486
Curb weight	0.716	0.585
Price in thousands	-0.005	0.924
Horse power	0.221	0.911
Engine size	0.498	0.768
Fuel efficiency	-0.562	-0.641

EXAMPLE (CONTD.)

- **Table 5: Rotated Component Matrix:** This table shows the correlation of the factors retained with the variables after applying Varimax rotation. It is helpful in the detection of the structure of the factor. A variable is said to be contained in a factor if the correlation of the variable with the factor is maximum among all the factors. In the example 5 variables (Wheelbase, Length, Width, Fuel capacity, Curb weight) are highly correlated to 1st factor and are said to be contained in 1st factor. Other 4 variables (Price in thousands, Horsepower, Engine size, Fuel efficiency) are highly correlated to the 2nd factor as compared to first factor and are said to be contained in 2nd factor.

EXAMPLE (CONTD.)

- **Table 6: Component Score Coefficient Matrix:** This matrix shows the coefficients of the variable in the factor structure. It is used to obtain the value of factor for different set of observations (or individuals) under consideration. These values are used for further calculation. For our data these coefficients are computed by using Regression method.

REFERENCES

1. Anderson TW, An introduction to Multivariate Statistical Analysis, 3rd Edition, John Wiley & Sons Inc., New Jersey.
2. Lesson 12 for Course STAT505 of Penn State University available on website https://online.stat.psu.edu/stat505/book/export/html/691.
3. Malhotra NK, Birks DF, Marketing Research an Applied Approach, 4th Edition, Prentice Hall, New Delhi.
4. Johnson RA, Wichern DW, Applied Multivariate Statistical Analysis, 3rd Edition, Prentice Hall, New Delhi.
5. Morrison DF, Multivariate Statistical Methods, 2nd Edition, McGraw Hill Publication, India.
6. Rencher AC, Methods of Multivariate Analysis, 2nd Edition, Wiley Interscience, New York.
7. Everitt BS, Dunn G, Applied Multivariate Data Analysis, 2nd Edition, John Wiley & Sons, London.
8. Jobson JD, Applied Multivariate Data Analysis Vol. II, Springer–Verlag Inc. New York.