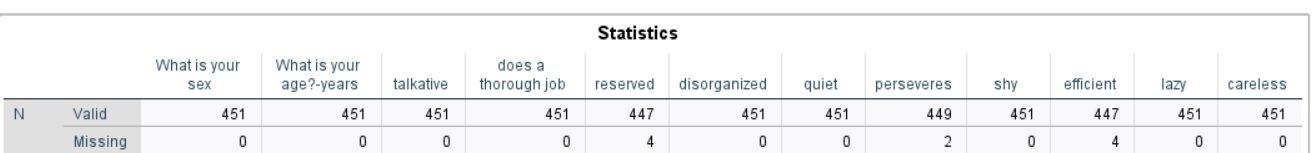
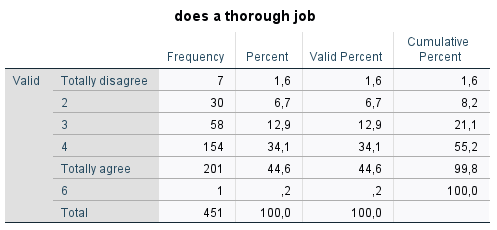
**Introduction to Statistics**

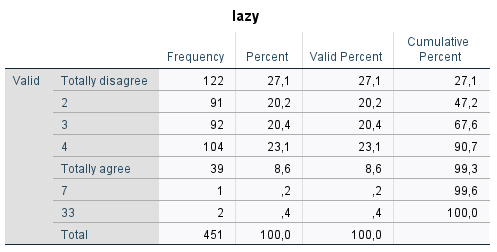
1. **DATA PREPARATION**

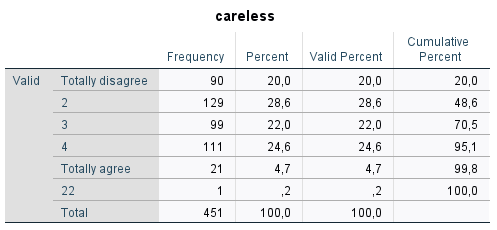
From the statistics table, we can see the presence of missing values for the variables **reserved, perseveres and efficient.** We have replaced these missing values as -1 and indicated this on SPSS for the missing value definition.

****

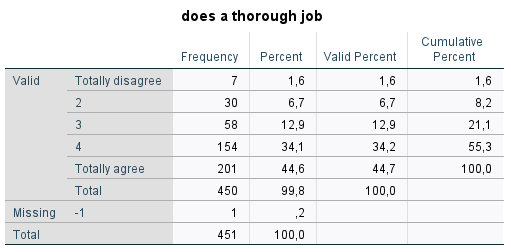
When analyzing the other tables, we can observe that the variables: **does a thorough job, lazy and careless** have wrong values that are not in the defined scale. Hence, I will consider these values as missing values.

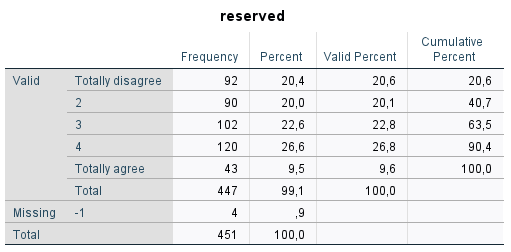
****

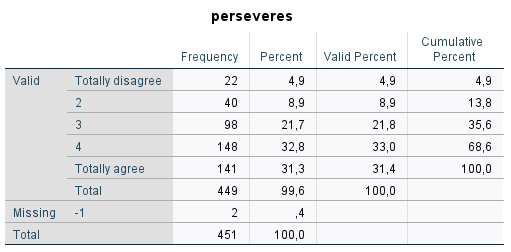
****

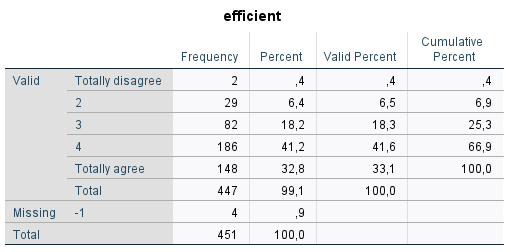
****

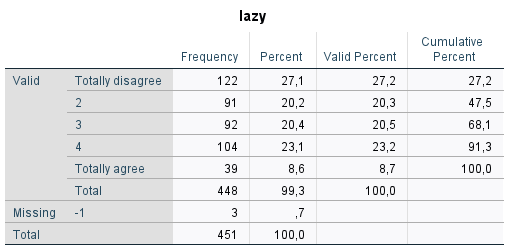
After running again the frequency of variables, we can observe that the missing values are identify by **-1** and the variables are grouped in the defined scale.

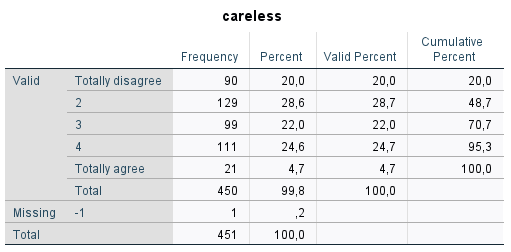






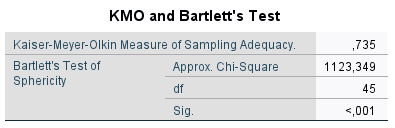




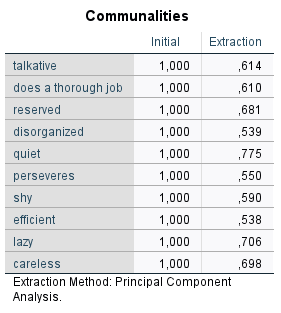
****

1. **FACTOR ANALYSIS ON PERSONALITY VARIABLES (PERS 1 → PERS 10)**

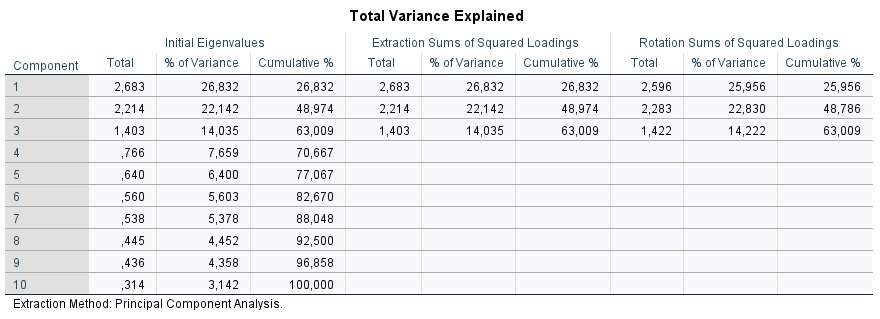
I would perform factor analysis to **reduce the variables** related to personality measures. The first thing is, we need to know if it is meaningful to perform a factor analysis. The KMO results in a value of **0.735** which is greater than 0.50 and Bartlett’s Test results in a value of **0.001** which is less than 0.05. Therefore these results indicates that it makes sense to perform the factor analysis on this data.



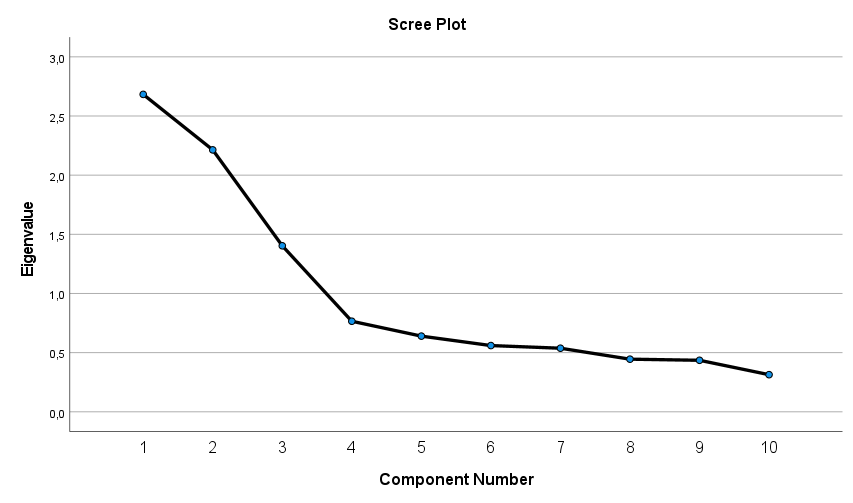
The communalities show that the proposed factors solution explain in a good level the variance for each one of the variables under analysis, in this case each variable are at least explained in **53.9%** by the factor solution.



Using the Total Variance table and scree plot, we can decide about the number of factors which we can include in our model.



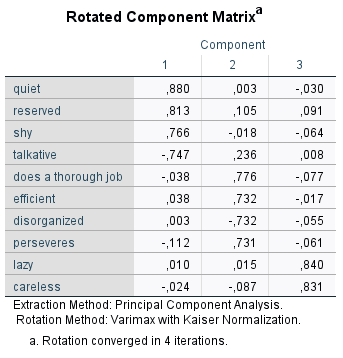
The total variance explain suggests that we should consider **three** factors, based on the initial Eigenvalues, the first three are the only ones greater than 1. These three components explain the **63.009%** of the total variance in the dataset.



From the above scree plot, we can say that the elbow is at 4, and you consider elbow – 1 factors, which means **three factors**. Both metrics shows the same conclusion about the inclusion of three factors in the analysis.

**Linking variables to factors**

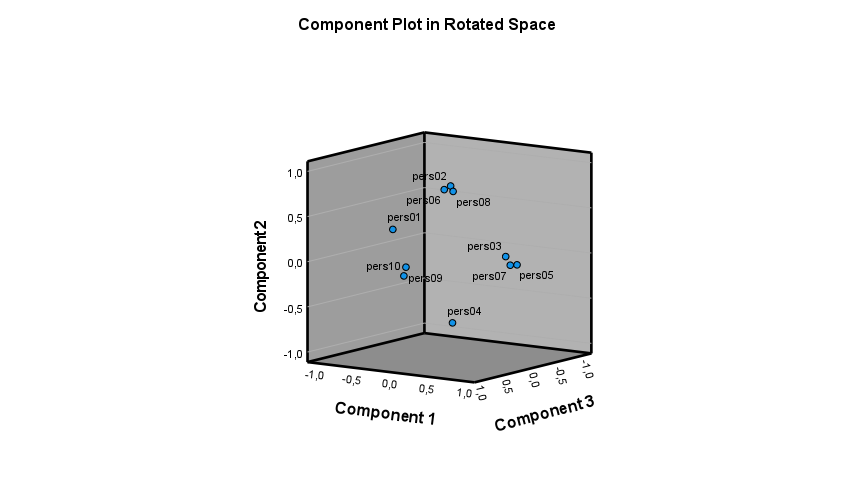
By using the Rotated Component Matrix, we can decide about which variables to put in each factor.

****

**JOB PERFORMANCE**

**COMMITMENT**

**BEHAVIOUR**



We can conclude that the variables quiet, reserved, shy and talkative are in the same factor. Variables: does a thorough job, efficient, disorganized, perseveres are in the same factor and finally lazy and careless belong to the last factor which gives us three factors in total.

I will define the variables in the following way:

**Behaviour :** This refers to how people act towards their relationship with others which include measure about being quiet, reserved, shy and talkative.

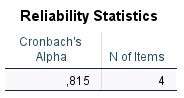
**Job performance:** It describe the variables related to how people do their job and include measures of does a thorough job, efficient; perseveres, disorganized.

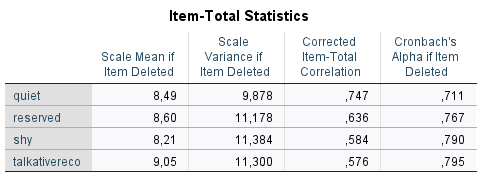
**Commitment:** This factor explains the people who are lazy or careless. It includes measures of Lazy and careless.

From the results, we observe that it is necessary to perform a recoding in the variables **talkative and disorganized** because they have the minus symbol which means that they are not in the same direction as the other variables within the factor and before creating the new variables, we need to test the reliability of the variables in each factor. We will use Cronbach’s Alpha for testing the variables inside the first two factors,

**Factor 1: Behaviour**

We run the reliability analysis for the variables inside the first factor; the Cronbach’s Alpha value is **0.815** which is greater than 0.70. This indicates that the variables inside the factor 1 are consistent.

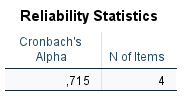
****

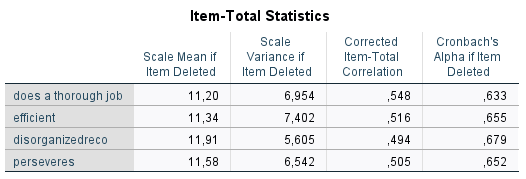
****

We can also understood from the above table that there is no improvement in Cronbach’s Alpha value if any of the variables has been removed from this factor.

**Factor 2: Job Performance**

We run the reliability analysis for the variables inside the second factor; the Cronbach’s Alpha value is **0.715** which is greater than 0.70. This indicates that the variables inside the factor 2 are consistent.

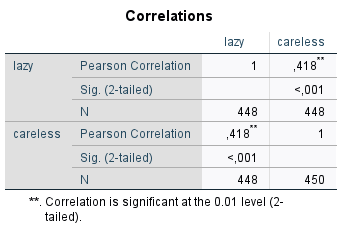
****

****

We can also understood from the above table that there is no improvement in Cronbach’s Alpha value if any of the variables has been removed from this factor.

**Factor 3: Commitment**

For the third factor we only have two variables and hence Cronbach’s alpha cannot be calculated. I will perform Pearson correlation in order to verify the consistency within variables from factor 3.

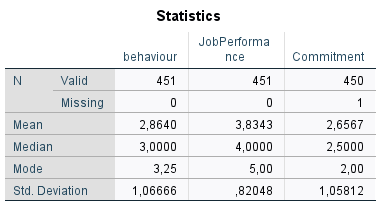


In this case, we can observe that r = **0.418**, with a significance level of p < .001. Hence, we can conclude that we can aggregate the two variables belonging to the third factor.

Now we can summate all variables because it satisfies all the **three conditions**:

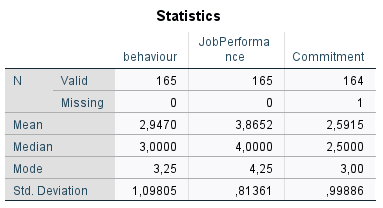
* All questions need to be measured on the **same scale**
* All questions need to be scaled in the **same direction**
* The new variable should contain only variables that measure the same construct (Check **Cronbach’sAlpha**)

After summated the three new variables, we have the following statistics for the total data set:

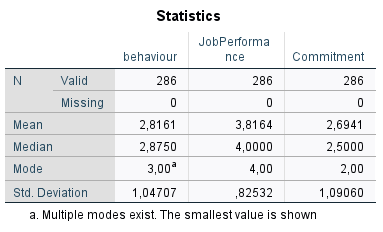
****

Splitting by age (mean=31), below are the statistics results:

**Age >= 31:**



**Age <31:**

****