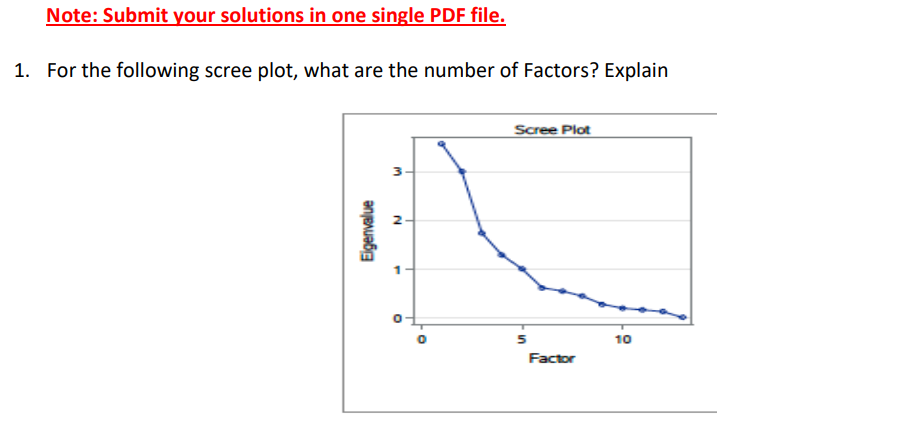
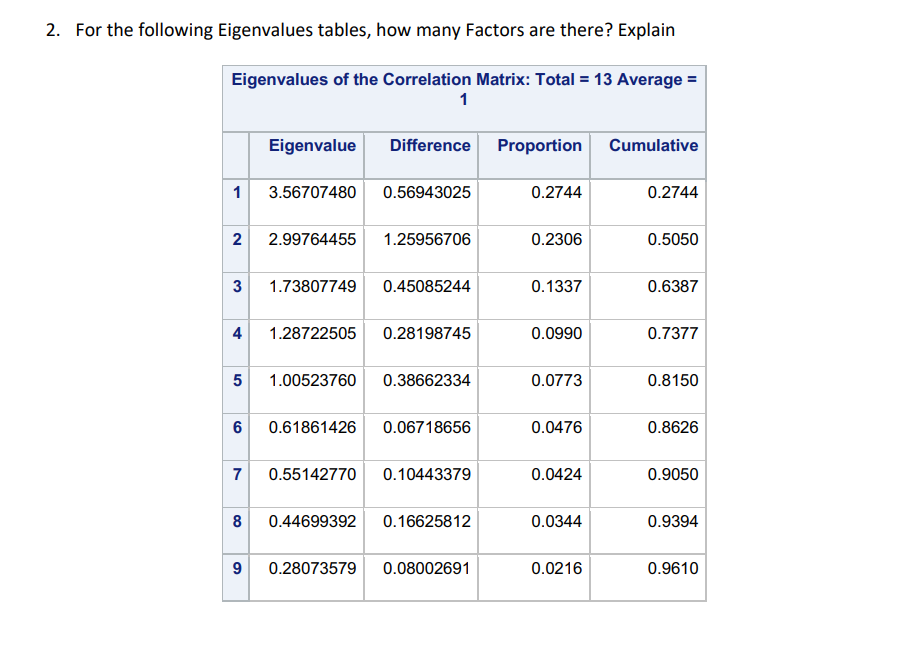
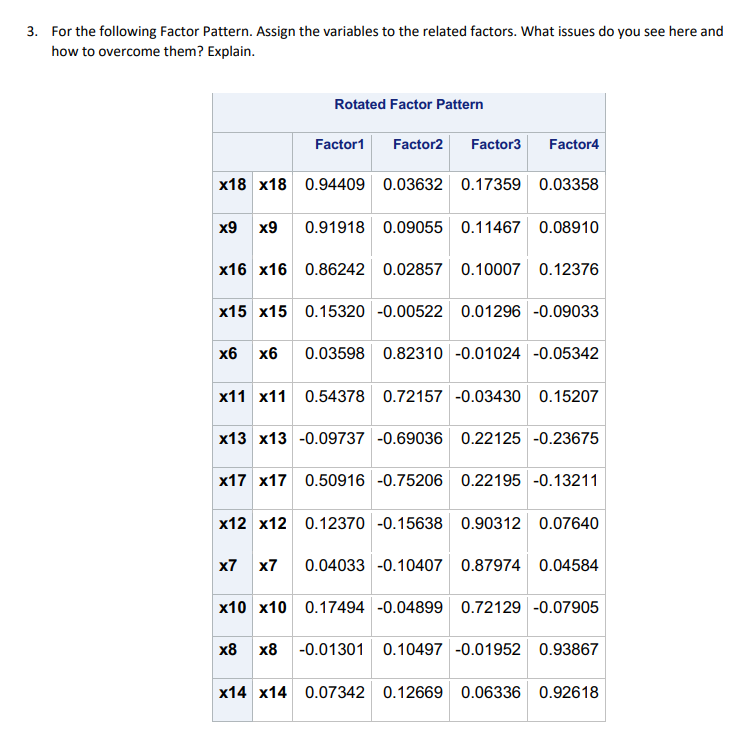
Multivariate Data Analysis 2: Factor Analysis



Solution: For the following question, the number of factors will be 5.

Reason: A scree plot is a graph drawn between factors on the x-axis and eigenvalues on the y-axis. The number of factors is determined by the point on the graph where it flattens out and the factors are 5.

Solution: In the above correlation matrix, the number of factors is 5. We determine this by considering the number of factors where eigenvalues are greater than 1. When examining the row-wise values, we find that there are 5 factors.

Keeping the loading value threshold at 0.6 and assigning variables to different factors:

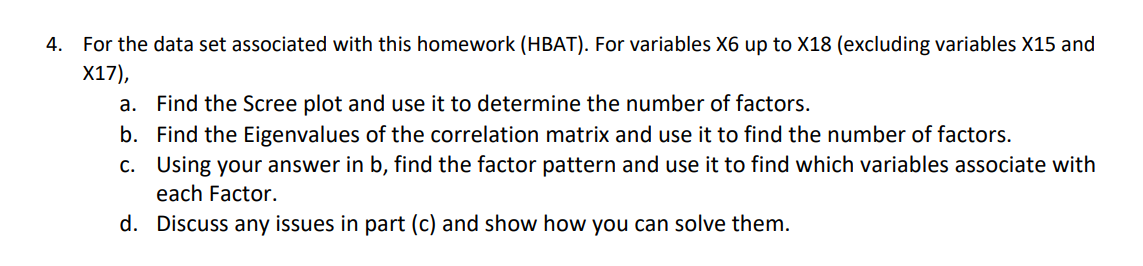
- Factor 1: X18, X9, X16

- Factor 2: X6, X11, X13, X7

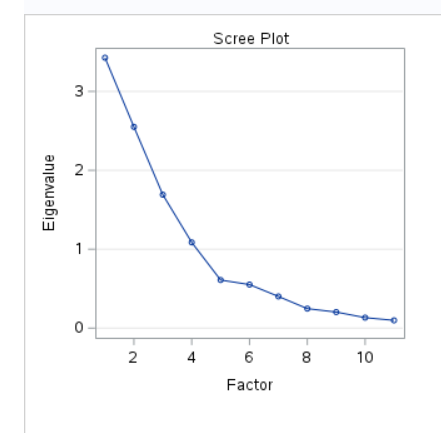
- Factor 3: X12, X7, X10

- Factor 4: X8, X14

After assigning every variable to a factor, variable X15 is not assigned to any factor because it does not have any loading values greater than 0.6. Therefore, it is not contributing significantly to any of the identified factors based on the loading value threshold.

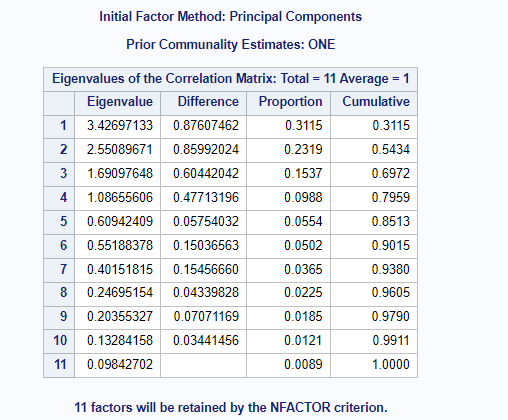


a.



This scree plot indicates that the number of factors is 4, as the graph flattens out at the 4th factor.

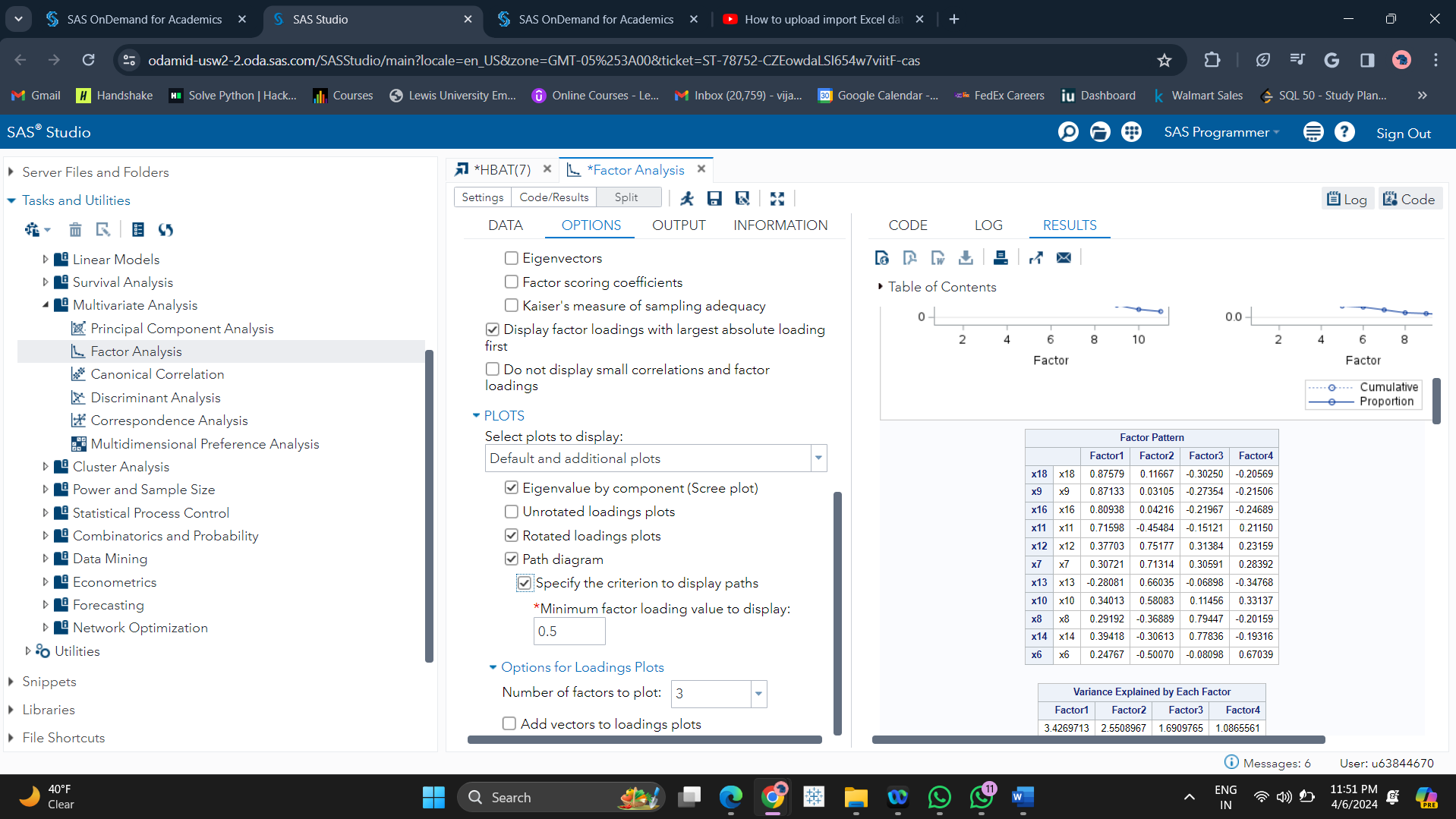
b.

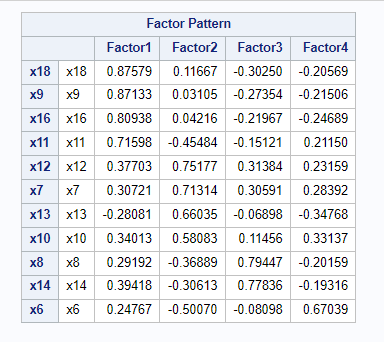


Number of factors will be 4 as eigenvalues are above 1 up to the fourth eigenvalue row-wise.

c.

By specifying the criteria to 0.5 as our threshold loading value and this is the factor pattern





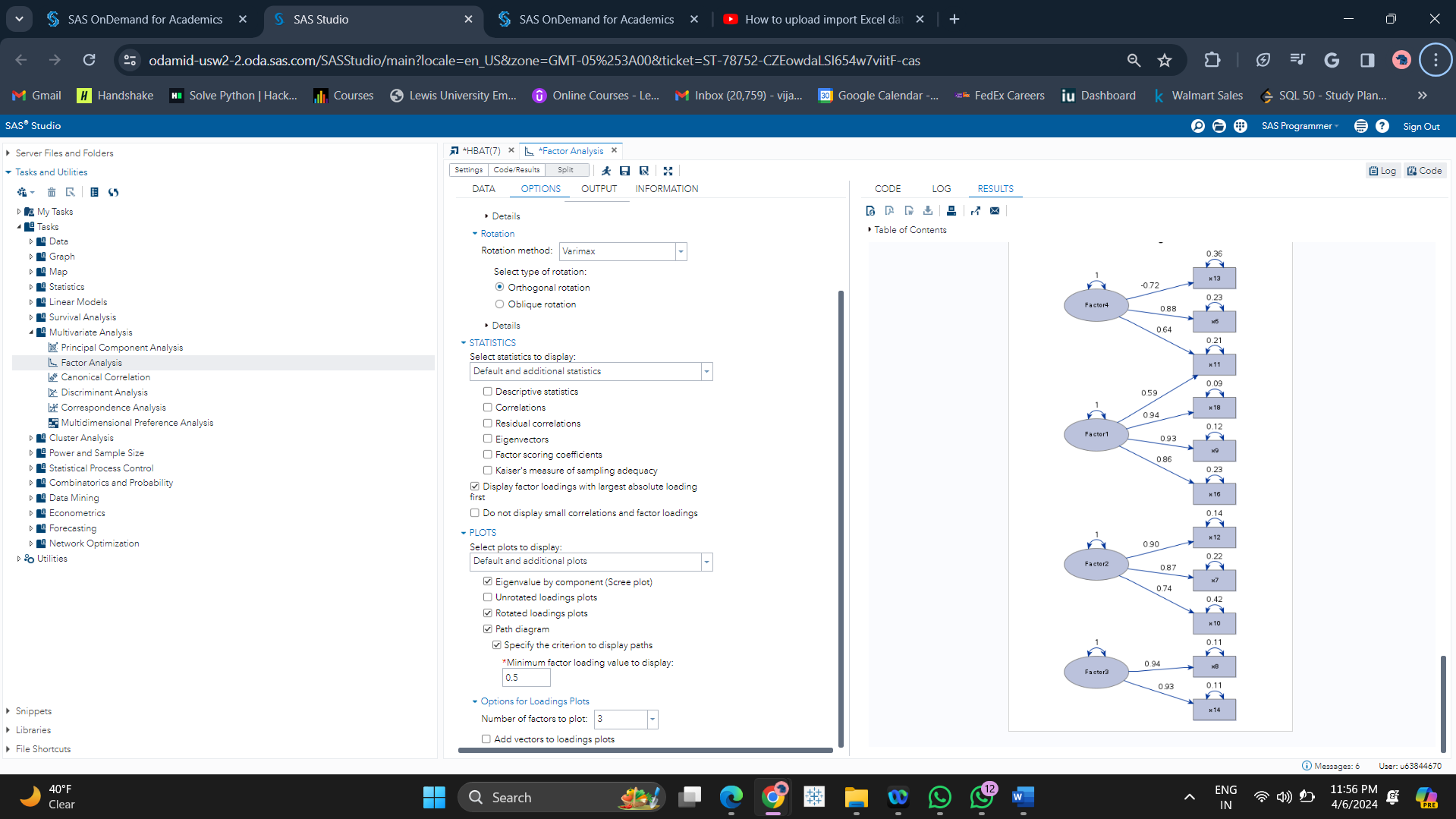
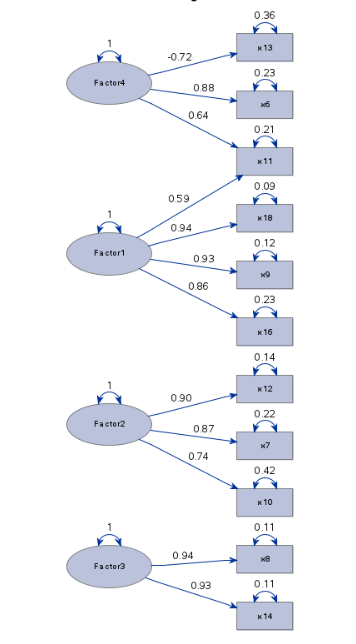
According to this Factor Pattern

Factor 1 : X18,X9,X16,X11,

Factor 2 : X12,X7,X13,X10

Factor 3 : X8,X14

Factor 4 : x6

d. 

X11 belongs to both Factor 4 and Factor 1, which is known as a cross-loading problem. In factor analysis, each variable should ideally be associated with only one factor to avoid cross-loadings. To address this issue, the variable causing cross-loadings should be removed, and the analysis should be rerun.

Additionally, Factor 3 contains only two variables assigned to it. In factor analysis, a factor should ideally have at least three variables, although in rare cases, accepting two variables for a factor may be considered.