# MODULE 3 UNIT 3

## Activity submission

Learning outcome:

LO4: Analyse the results of factor analysis in a business context.

LO5: Reflect on the use of factor analysis.

### Name: Matlotlo Magasa

#### 1. Instructions and guidelines (Read carefully)

##### Instructions

1. Insert your name and surname in the space provided above, as well as in the **file name.** Save the file as: **First name Surname M3U3 Activity Submission** – **e.g. Lilly Smith M3U3 Activity Submission.** **NB:** *Please ensure that you use the name that appears in your student profile on the Online Campus.*

2. Write all your answers in this document. There is an instruction that says, “Start writing here” under each question. Please type your answer there.

3. Submit your assignment in **Microsoft Word only**. No other file types will be accepted.

4. You will be required to include visualisations that you have made in Tableau in this activity. To do this, export them as an image file and paste them into this document. Tableau has a page that details how to export your Tableau view as an image file.

5. Do **not delete the plagiarism declaration** or the **assignment instructions and guidelines**. They must remain in your assignment when you submit.

**PLEASE NOTE:** **Plagiarism cases will be investigated in line with the Terms and Conditions for Students.**

**IMPORTANT NOTICE:** Please ensure that you have checked your course calendar for the due date for this assignment.

##### Guidelines

1. There are 8 pages and 5 questions in this assignment.

2. This activity submission relies heavily on work done in the IDE notebook in this unit. If you have not done so already, return to the IDE notebook and complete this activity before attempting to answer the questions in this assignment.

3. You will be required to download a data file, titled LOANS.sav, in order to interact with the data in Tableau. Return to the Module 3 downloads folder and save a copy of the data file before attempting to answer the questions.

4. Make sure that you have carefully read and fully understood the questions before answering them. Answer the questions fully, but concisely, and as directly as possible. Follow all specific instructions for individual questions (e.g. “list”, “in point form”).

5. Answer all questions in your own words. Do not copy any text from the notes, readings or other sources. **The assignment must be your own work only.**

|  |
| --- |
| **Plagiarism declaration:** |
| **1. I know that plagiarism is wrong. Plagiarism is to use another’s work and pretend that it is one’s own.**  **2. This assignment is my own work.**  **3. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.**  **4. I acknowledge that copying someone else’s assignment (or part of it) is wrong and declare that my assignments are my own work.** |

#### 2. Mark allocation

Each question receives a mark allocation. However, you will only receive a final percentage mark and will not be given individual marks for each question. The mark allocation is there to show you the weighting and length of each question.

Question 1 10

Question 2 10

Question 3 10

Question 4 10

Question 5 10

**TOTAL 50**

#### 3. Questions

##### Question 1

One of the steps you need to complete before performing a factor analysis is to generate a correlation matrix between the variables in order to investigate whether multicollinearity exists. In doing this, you would have found that there is a medium-level positive correlation between income and the two different types of debt: credit card debt and other debt. The correlation coefficients from the matrix you generated are copied in Table 1 for your reference.

Table 1: Correlation coefficients between income and different types of debt.

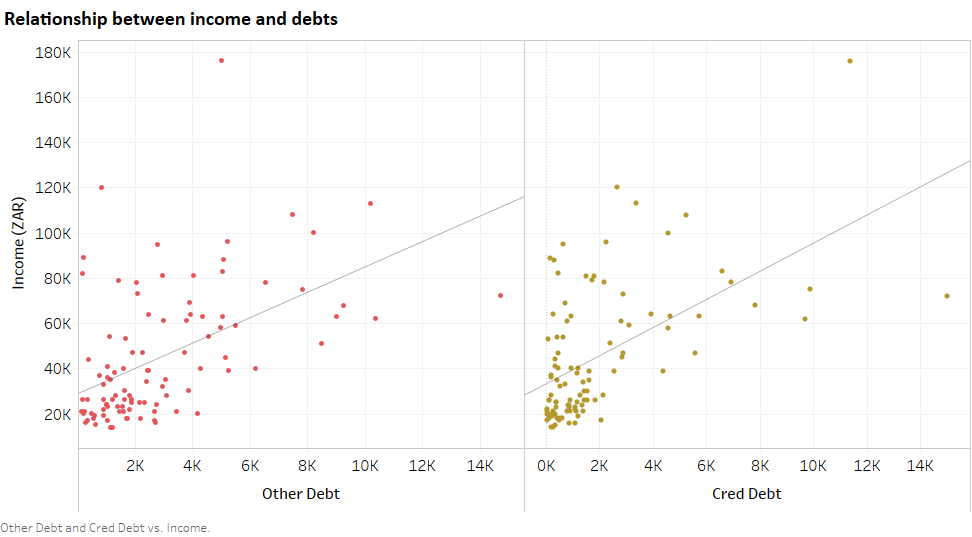
|  |  |  |
| --- | --- | --- |
|  | **Cred\_debt** | **Other\_debt** |
| **Income** | 0.5524 | 0.5113 |

Speedy Loans management is, therefore, interested in visualising this relationship between income and the types of debt customers hold. They believe this will be a more intuitive way to present the findings to the board of directors.

Before analysing the data in Tableau, you will be required to download the LOANS.sav data file from the Module 3 downloads folder. This can be imported into Tableau to allow you to interact with the data set.

To analyse the relationship between income and different types of debt, you have been asked to create a scatterplot in Tableau, with “Cred debt” and “Other debt” on the x-axis and “Income” on the y-axis. Remember to convert all the variables from “Measures” to “Dimensions” to view individual data points on the scatter plot. If desired, you can change the colours of the data points for each type of debt and include a trend line for each plot. Rename the worksheet “Relationship between income and debts”.

Paste your Tableau visualisation here:



R-Squared value for Income vs Credit Debt = 0.305

R-Squared value for Income vs Other debt = 0.305

##### Question 2

An important step in factor analysis is selecting the number of factors to include in the model. You have encountered two methods for selecting the number of factors: Kaiser’s criterion and the elbow rule. Based on the results output you generated in the previous component, answer the following questions based on these methods:

* Using Kaiser’s criterion, how many factors should be included? How did you reach this decision? How much cumulative variance do these factors explain?
* Using the elbow rule, how many factors should be included? How did you reach this decision? How much cumulative variance do these factors explain?
* If the above methods recommend different numbers of factors, how would you decide which method to use?

(Max. 200 words)

Start writing here:

Kaiser’s method suggests 2 factors because these two variables have a eigenvalues greater than 1. These 2 factors explain a cumulative variance of 77.10%.

The elbow rule suggests 3 factors, this is the point at which the plot fattens out and any factor after this point the factors explain little variation of the model. The factors on the left of this bend are those that are selected as appropriate for the model. At the 3rd factor, the cumulative variation is 85.19%.

I would use a factor of 2 because it is difficult to explain the third factor. The first two factors have a more fitting concepts for the variables. The cumulative variance for these 2 factors is above 77% which is appropriate according to "Multivariate Analysis" by Hair et al (2012) which suggests a cumulative variance of above 60% for factors selected.

##### Question 3

In the IDE notebook from the previous component, capture a screenshot of the **unrotated factor matrix** and answer the question that follows.

Paste the unrotated factor matrix screenshot here:



Using the information in the screenshot, substantiate why applying a rotation technique is appropriate for this model. Examine at least two variables in this motivation.

(Max. 150 words)

Start writing here:

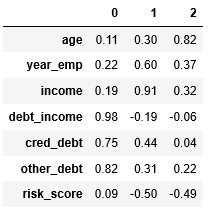
Age is the variable whose factor loading loads almost equally on factors one and three. This is unwanted because it creates a confusion on interpreting the model. Ideally a variable should load significantly high on one factor. The rotation changed the factor loading for Age to load higher on one factor, and we no longer observe cross-loading.

Credit debt has two factors that seem unclear because it loads moderately for both factor 1 and 2. Therefore it is necessary to hopefully get unique association of a variable with one factor by apply the orthogonal factor rotation technique.

##### Question 4

In the IDE notebook from the previous component, capture a screenshot of the rotated factor matrix and answer the questions that follow.

Paste the rotated factor matrix screenshot here:



What has applying this rotation technique achieved? How have the factor loadings changed between the unrotated and rotated factor matrices? Examine at least two variables in your answer.

(Max. 150 words)

Start writing here:

The rotation changed the factor loading for Age to load higher on one factor (factor 3), and we no longer observe cross-loading. However, the opposite is true for the Risk Score. Now there is moderate, and almost equal, factor loadings across two factors (factor 2 & 3) where before rotation this was not the case. These confuses the interpretation on which factor the risk score should be placed under, making interpretation a little more complex.

##### Question 5

Before performing a factor analysis, it is important to think about possible conceptual factors that might help to explain the variables in the data set. This not only allows you to determine whether factor analysis is an appropriate analysis method but assists with factor naming.

Given the results of your analysis and the number of factors you have selected, discuss the following:

* Has factor analysis achieved the goal of reducing the number of variables to a smaller number of conceptual factors?
* Does the grouping of variables under each factor make sense?
* Conceptually, what could you name each factor, considering the variables that load highly onto each factor? Explain your reasoning.

(Max. 200 words)

Start writing here:

Yes, FA achieved a reduction of 7 variables into less than 3 factors which will make business analysis of customers simpler and easier to understand. These factors can provide insight into how the company can provide loans to its potential customers based on the results from analysing these factors.

The 3 factors in this case demonstrate why it may be better to only take on 2 Factors (Using the Kaiser method instead of the Elbow method). Age is the only variable with its own factor which may be unnecessary as it can be placed under factor 2.

Naming the Factors:

FACTOR 1: 'Debt Records'

FACTOR 2: 'Financial Profile'

\*\*FACTOR 3: 'Personal Profile' It may be appropriate to use only 2 factors and place the Age under factor 2 and rename it "Personal Financial Profile”.

**4. Rubric**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **No Submission** | **Poor** | **Good** | **Excellent** |
| **Question 1**  *The image contains the correct graphs and adheres to the instructions given.* *The graphs contain the correct variables.* | No submission or attempt. (0) | The student attempts to answer the question, but the image contains the incorrect graphs, or the incorrect variables are used. (4) | The image contains the correct graphs using the correct variables, but not all instructions have been adhered to. (6) | The image contains the correct graphs using the correct variables, and all instructions have been adhered to. (10) |
| **Question 2**  *The submission correctly identifies the number of factors to be included for each rule, and clearly explains how each decision was reached.* | No submission or attempt. (0) | The student correctly identifies the number of factors to be included in at least one instance, but was unable to explain how this decision was reached. (4) | The student correctly identifies the number of factors to be included in at least one instance and adequately explains how this decision was reached. (6) | The student correctly identifies the number of factors to be included for each rule and clearly explains how each decision was reached. (10) |
| **Question 3**  *The submission contains a screenshot of the unrotated factor matrix and clearly substantiates the reasoning for applying a rotation technique.* | No submission or attempt. (0) | The student provides a screenshot of the unrotated factor matrix, but does not clearly or adequately substantiate the reasoning for applying a rotation technique. (4) | The student provides a screenshot of the unrotated factor matrix and adequately substantiates the reason for applying a rotation technique, but does not examine at least two variables. (6) | The student provides a screenshot of the unrotated factor matrix, clearly substantiates the reason for applying a rotation technique, and examines two or more variables. (10) |
| **Question 4**  *The submission contains a screenshot of the rotated factor matrix and clearly explains the effects of the rotation technique.* | No submission or attempt. (0) | The student provides a screenshot of the rotated factor matrix, but does not clearly or adequately explain the effect of applying the rotation technique. (4) | The student provides a screenshot of the rotated factor matrix, adequately explains the effect of the factor rotation, but does not examine at least two variables. (6) | The student provides a screenshot of the rotated factor matrix, clearly explains the effect of the factor rotation, and examines two or more variables. (10) |
| **Question 5**  *The submission critically engages with the use of factor analysis in this situation,* *and successfully states the names for the factors.* | No submission or attempt. (0) | The student attempts to critically engage with the use of factor analysis in this situation, but does not successfully state the names for the factors. (4) | The student addresses the grouping of variables, adequately reflects on the use of factor analysis, and attempts to name the factors. (6) | The student clearly addresses the grouping of variables, reflects on the use of factor analysis, and provides adequate names for the factors. (10) |

**Total:** 50 marks