**QUESTION 1**

There are two datasets on 14 separate variables (HATCO and Hatmiss) which are examples of a segmentation study for a business-to-business situation, specifically a survey of existing customers of HATCO consist of 100 observations while Hatmiss consist of 70 observations. Three types of information were collected:

* First is the perception of HATCO on seven attributes identified in past studies as the most influential in the choice of suppliers. The respondents, purchasing managers of firms buying from HATCO, rated HATCO on each attribute.
* Second are actual purchase outcomes, either the evaluations of each respondent’s satisfaction with HATCO or the percentage of his or her product purchases from HATCO.
* Third are general characteristics of the purchasing companies (e.g., firm size, industry type).

The data provided should give HATCO a better understanding of both the characteristics of its customers and the relationships between their perceptions of HATCO and their actions toward HATCO (purchases and satisfaction). A definition of each variable and an explanation of its coding are given in the following sections.

**Perceptions of HATCO:** Each of these variables was measured on a graphic rating scale, where a ten-centimetre line was drawn between the endpoints, labelled "Poor" and "Excellent". Respondents indicated their perceptions by making a mark anywhere on the line. The mark was then measured and the distance from zero (in centimetres) was recorded. The result was a scale ranging from zero to ten, rounded to a single decimal place. The seven HATCO attributes rated by each respondent are as follows:

* X1 = Delivery speed—amount of time it takes to deliver the product once an order has been confirmed.
* X2 = Price level—perceived level of price charged by product suppliers.
* X3 = Price flexibility—perceived willingness of HATCO representatives to negotiate price on all types of purchases.
* X4 = Manufacturer’s image—overall image of the manufacturer/supplier.
* X5 = Service—overall level of service necessary for maintaining a satisfactory relationship between supplier and purchaser.
* X6 = Salesforce’s image—overall image of the manufacturer’s sales force.
* X7 = Product quality—perceived level of quality of a particular product (e.g., performance or yield).

**Purchase Outcomes:** Two specific measures were obtained that reflected the outcomes of the respondent’s purchase relationships with HATCO. These measures include:

* X9 = Usage level—how much of the firm’s total product is purchased from HATCO, measured on a 100-point percentage scale, ranging from 0 to 100 percent.
* X10 = Satisfaction level—how satisfied the purchaser is with past purchases from HATCO, measured on the same graphic rating scale as the perceptions X1 to X7

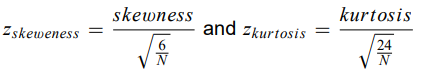
**Purchaser Characteristics:** The five characteristics of the responding firms used in the study, some metric and some nonmetric, are as follows:

* X8 = Size of firm—size of the firm relative to others in this market. This variable has two categories: 1 = large, and 0 = small.
* X11 = Specification buying-extent to which a particular purchaser evaluates each purchase separately (total value analysis) versus the use of specification buying, which details precisely the product characteristics desired. This variable has two categories:1 = employs total value analysis approach, evaluating each purchase separately, and 0 = use of specification buying.
* X12= Structure of procurement—method of procuring/purchasing products within a particular company. This variable has two categories: 1 =centralized procurement, and 0 = decentralized procurement.
* X13 =Type of industry—industry classification in which a product purchaser belongs. This variable has two categories: 1 = industry A classification, and 0 = other industries.
* X14 = Type of buying situation—type of situation facing the purchaser. This variable has three categories: 1 = new task, 2 = modified rebuy, and 3 = straight rebuy.

1. Using the data **Hatmiss**, use SPSS to produce an output for the missing value analysis (MVA) for all the variables. The analysis should include the univariate statistics, separate variance t-tests, missing patterns. Produce the output.
2. Write a brief report discussing the results of the missing value analysis and ensure that you determine whether the data is missing at random (MAR) or missing completely at random (MCAR).
3. Looking at your output, decide on a method that can be used to estimate the missing data and substantiate your answer.

**QUESTION 2**

In this question you are using the data set **HATCO.** Use SPSS to answer the following questions.

1. Is the data appropriate for factor analysis? Substantiate your answer.
2. Which variables will be appropriate for factor analysis?
3. Do a factor analysis using the principal component analysis with a promax rotation. Ensure that you include the scree plot, and under options sort the factors by size and use items with factor loadings .5 and above. Produce the output.
4. How many factors will be needed if the criterion is to explain 80% of the total variance?
5. What do you notice about the item "Price flexibility"?
6. Create a new variable "Price flexibility2" by computing the necessary transformation of "10 - Price flexibility" and redo a factor analysis using the principal component analysis with a promax rotation using "Price flexibility2" instead of "Price flexibility". Produce the output and do not suppress small coefficients in your output.
7. Interpret the results of the factor analysis (including appropriateness of the factor analysis, grouping the variables according to the factor loadings and suggesting, with justification, possible labels for the factors).
8. Show numerically that the communality estimate for Price flexibility2 is :714. Interpret this estimate.
9. Do reliability analysis of the factors obtained in (f) and give the reliability coefficient for each construct and the overall reliability analysis of the instrument, and interpret it using Jain and Angural (2017) guidelines.
10. Three new variables were created, using the factor analysis. (i) Explain how these new variables for each respondent may be obtained. (ii) Suggest one possible way that these new variables can be used.
11. Discuss the rotation method used, paying specific attention to: (i) the method used, its implications and appropriateness, (ii) the necessity of the rotation method, and (iii) the success of the rotation method.
12. Calculate the composite variables of each factor (by taking the average), and give the descriptive statistics for the composite variables and interpret them.
13. Use the distribution command in SAS JMP to test for normality of the three composite variables created. (Hint: Use histograms, boxplots, normal quantile plots and Shapiro Wilk’s test). Ensure that on the summary statistics output you add skewness and kurtosis on the analysis results.
14. Do bivariate profiling of all variables by drawing box plots for all the variables by X8 = Size of firm, X11 = Specification buying, X12 = Structure of procurement, X13 = Type of industry and X14 = Type of buying situation and comment on your plots.
15. For each of the composite variable compute and hence tests for normality.
16. Do your results in (m) and (o) agree? Explain for each composite variable
17. For those variables that are nonnormal if any, determine the transformation that makes them more nearly normal.
18. Test whether the mean of the composite variables differs by:

(i) X8 = Size of firm.

(ii) X11 = Specification buying.

(iii) X12 = Structure of procurement.

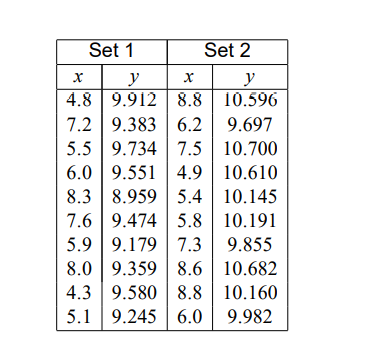
(iv) X13 = Type of industry.

(v) X14 = Type of buying situation.

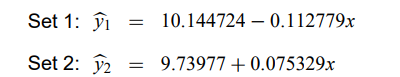
(Ensure that you draw error bars, and for variables with more than two categories you do posthoc analysis)

**QUESTION 3**

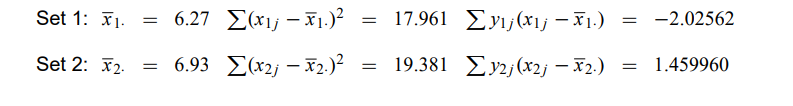
The following data was given for two sets:



1. Give a graphical representation of the data. Using this representation, would you say that an analysis of covariance model is applicable? Justify your answer.
2. Write down the appropriate model in the form  with X of full column rank for each of the following two cases. (Give the parameters explicitly.)
3. Assume the regression lines are not parallel.
4. Assume the regression lines are parallel.
5. The following regression lines were fitted to the data:



1. Test whether the slopes of the fitted regression lines are significantly different.
2. Assume equal slopes and variances, and test Given that:



1. Using any statistical package of your choice, produce the output of the analysis of covariance (where you assume equal slopes). Produce the output.