

**ASSIGNMENT**

---

<b>Course Code</b>	<b>:</b>	<b>MS - 66</b>
<b>Course Title</b>	<b>:</b>	<b>Marketing Research</b>
<b>Assignment Code</b>	<b>:</b>	<b>MS-66/TMA/JULY/2022</b>
<b>Coverage</b>	<b>:</b>	<b>All Blocks</b>

---

**Note:** Attempt all the questions and submit this assignment to the coordinator of your study centre. **Last date of submission for July 2022 session is 31<sup>st</sup> October, 2022 and for January 2023 session is 30<sup>th</sup> April, 2023.**

1. Explain the solomon-four group design. How far does this design succeed in controlling different extraneous variables? Illustrate your answer with the help of a suitable example from marketing.
2. Briefly explain the various probability and non-probability sampling methods with their merits and demerits.
3. With the help of examples, discuss the areas where qualitative research can be used in marketing.
4. Monthly salary of 32 employees of a firm is given below. Tabulate the data after selecting a suitable class interval.

2250 1800 1650 1760 3520 5600 2450 2680  
2700 1680 3650 3240 5850 3150 1860 2425  
4520 3275 4215 3760 1950 1850 3750 2825  
4500 3800 4300 2750 4370 3350 2375 3215

5. A manufacturer of industrial supplies developed the following model for predicting the number of sales per month

$$Y = 41 + .3X_1 + .05X_2 - 7X_3 + 10X_4$$

where Y = Sales per month

X<sub>1</sub> = Number of manufacturing firms'

X<sub>2</sub> = Number of wholesale and retail firms

X<sub>3</sub> = Number of competing firms

X<sub>4</sub> = Number of full-time company sales people.

- i) Explain the correct interpretation of all estimated parameters in the equation.
- ii) If R<sup>2</sup> = 0.49, what does this figure mean, to you?
- iii) Explain how you will go about testing the validity of this multiple linear regression equation

*ASSIGNMENT REFERENCE MATERIAL (July22 to Jan23)*

*MS-66*

*Marketing Research*

**Q1. Explain the solomon-four group design. How far does this design succeed in controlling different extraneous variables? Illustrate your answer with the help of a suitable example from marketing.**

**Ans.** This design is also called the four group six study design. This test overcomes the limitations of pretest - post test control group designs. This design is infact a combination of 'Before After with one control group' and 'After only with one control Group' designs. The design is symbolically presented as:

Experimental Groups	:	R	O <sub>1</sub>	X	O <sub>2</sub>
Control Groups	:	R	O <sub>3</sub>		O <sub>4</sub>
Experimental Groups	:	R			O <sub>6</sub>
Control Groups	:	R			O <sub>6</sub>

You may note that the experimental group 2 and control group 2 do not receive any pretest measurements. The results of group 2 help us in measuring and eliminating the sensitizing testing - effect since this type of sensitizing cannot occur in measurement 05. The results of the difference of various pretest and post test measurements give us the following effects.

Experimental Groups	:	$O_2 - O_1 = \text{treatment effect} + \text{extraneous} + \text{sensitizing testing effect}$	(i)
Control Groups 1	:	$O_4 - O_3 = \text{extraneous factors}$	(ii)
Experimental Groups 2	:	$O_5 - O_1 = (\text{treatment effect} + \text{extraneous factors})$	(iii)
	:	$O_5 - O_3 = (\text{treatment effect} + \text{extraneous factors})$	(iv)
Contol Groups 2	:	$O_6 - O_1 = \text{extraneous factors}$	(v)
	:	$O_6 - O_3 = \text{extraneous factors}$	(vi)

By taking the average of equations (iii) and (iv), we get

$$O_5 - \frac{O_1 + O_3}{2} = \text{Treatment effect} + \text{Extraneous effect} \quad \text{(vii)}$$

By taking the average of equations (v) and (vi), we get

$$O_5 - \frac{O_1 + O_3}{2} = \text{Extraneous effect} \quad (\text{viii})$$

Treatment effect

$$= \left( O_5 - \frac{O_1 + O_3}{2} \right) - \left( O_4 - \frac{O_1 + O_3}{2} \right)$$

$$= O_5 - O_4 = \text{Treatment effect}$$

By subtracting equation (vii) from equation (i), we obtain the sensitizing testing effect as:

$$\text{Sensitizing testing effect} = (O_2 - O_1) - \left( O_5 - \frac{O_1 + O_3}{2} \right)$$

Please note that in order to measure the change in experiment group 2, we required an estimate of pretest measurement on O5, we assumed that groups are equal before the experimental because of the random assignments of the subjects to all four of our group I and control group 1 and control group 1 respectively. Given the just mentioned assumption, the pretest measurement of experimental group 1 and control group 1 could be taken as an estimate of pretest measurement of experiment group 2. that is what we used in the equations (iii), (iv), (v) and (vi) listed above).

This design is also referred to as "the ideal control experiment." As you have seen that we are not only able to control all extraneous variables and sensitizing testing effect, but also could get their estimates. However, this design is not very commonly used in marketing because of increase in cost, time and effort required to conduct the experiment.

The three experimental designs discussed above allow us to manipulate or control an independent variable so as to measure its effect on the dependent variable thereby enabling us to make proper casual statements. Below we will discuss some designs where we can manipulate or control more than one level of independent variable on the dependent variable.

**Q2. Briefly explain the various probability and non-probability sampling methods with their merits and demerits.**

**Ans. Probability Sampling Methods**

In probability sampling methods every element of the population has a known chance of being selected. Please note that the term known chance does not mean equal chance. Equal chance probability sampling is a special case of probability sampling, called simple random sampling. In probability sampling methods, there is no chance of arbitrary or biased selection and therefore the laws of probability apply. Therefore, it permits us to measure the sampling error which is the difference between the population value and the sample, value.

There are a number of different sampling procedures that fall under probability sampling methods. Some of these methods are listed as under:

**i) Simple Random Sampling:** Under this sampling design, each member of the population has known and equal probability of being included in the sample. For details on how to draw samples using this sampling procedure, please review unit 13 of MS-8 course. Simple random sampling is not widely used in marketing research because of the following reasons.

i) In consumer research studies, we usually select individuals, households, shops or areas as the sampling units. It may not be easy to prepare a sampling frame as it is very difficult to get lists of households, individuals and shops, although areas may be completely represented through maps.

ii) We know that an industry comprises of various firms of different sizes. If one wants to study some aspects of an industry, one might like to choose a sampling design where there is a higher probability of a larger firm being selected. If that is the case, the very concept of simple random sampling becomes inapplicable in such situations. The simple random sampling has some applications in Industrial Marketing where generally purchasing agents or companies or areas are the sampling units which are usually not very big in number. Therefore, it becomes easy to prepare a sampling frame thus facilitating the use of simple random sampling.

**ii) Systematic Sampling:** The mechanics of taking a systematic sample are very simple. Suppose the population consists of ordered  $N$  units (numbered from 1 to  $N$ ) and a sample of size  $n$  is selected from the population in such a way that  $N/n=K$  (rounded to the nearest integer). Here  $K$  is called a sample interval. Systematic sampling then consists in selecting a number at random between 1 and  $K$  (both inclusive) and then selecting every subsequent  $K$ th unit till a sample of size  $n$  is obtained.

To make the above more clear, let us assume that we have an ordered population of size  $N=500$ . Suppose it is decided to take a sample a size  $n=50$ . Therefore, our sampling interval would be  $N/n=500/50=10$ . We then select a number at random between 1 to 10 (both inclusive) Suppose it turns out to be 6. Then our sample units would be 6, 16, 26, 36 and so on.

**iii) Stratified Sampling:** In stratified sampling, the entire population is divided into various mutually exclusive and collectively exhaustive strata (groups). By mutually exclusive it is meant that if an element of a group belongs to one strata, then it doesn't belong to any other strata. By collectively exhaustive we mean that all the elements of various strata put together completely cover all the elements of the population. The groups (strata) are created on the basis of a variable (criteria) known to be correlated with the variable under study. The possible criteria for stratification of a population could be income of the individuals, age, sex, frequency of a purchase of a product, size of the household, size of the retail store, region of the country and so on. The stratification is also possible on the basis of more than one variable. This, of course, increases the number of stratum. The cost of stratification may come as a constraint in increasing the number of stratum. A variable which is considered to be good in stratification of one population may not be so in the case of other. However, one thing should be kept in mind that stratification should be done in such a way so as to minimise the variability among sampling unit within strata (more homogeneous) and maximise the variability among strata (more heterogeneous).

**iv) Cluster Sampling:** If we divide all the elements of the population into suitable, clusters; and select few clusters randomly and all the elements of the selected clusters are used, then

this method of sampling is called cluster sampling'. This method of collecting data is cheaper since collection of data from nearby units is easier, faster and more convenient than collecting data over units scattered over a region. For instance, it would not only be cheaper but also convenient to collect data on all households in a sample of few villages.(clusters) than to survey a sample of the same number of households selected randomly from a list of all households. The criteria for dividing the population into mutually exclusive and collectively exhaustive clusters is, that the elements in the clusters should be as heterogeneous as possible and elements between cluster should be as homogeneous as possible.

### **Non-Probability Sampling Methods**

These methods do not provide every item of population any known chance of being selected in, the sample. Here there is no attempt to select a representative sample. The elements of samples are selected on the convenience and/or judgement of the researcher or field interviewer. The selection process is subjective. Since the sample is not representative of population, it is not possible to make an estimate of sampling error. Also, we cannot say whether our sample estimates are correct or not.

There are many different sampling methods which fall under the umbrella of non- probability sampling methods. These are listed below:

**i) Convenience sampling:** Under convenience sampling, as the name implies, the samples are selected at the convenience of the researcher or investigator. Here, we have no way of determining the representativeness of the sample. This results into biased estimates. Therefore, it is not possible to make an estimate of sampling error as the difference between sample estimate and population parameter is unknown, both in terms of magnitude and direction. It is therefore suggested that convenience sampling should not be used in both, descriptive and causal studies as it is not possible to make any definitive statements about the results from such a sample.

This method may be quite useful in exploratory designs as a basis for generating hypotheses. The method is also useful in testing of questionnaire etc. at the pretest phase of the study. Convenience sampling is extensively used in marketing studies and otherwise.

**ii) Judgement sampling:** Judgement sampling is also called purposive sampling. Under this sampling procedure, a researcher deliberately or purposively draws a sample from the population which he thinks is a representative of the population. Needless to mention, all members of the population are not given chance to be selected in the sample. The personal bias of the investigator has a great chance of entering the sample and if the investigator chooses a sample to give results which favours his view point, the entire study may be vitiated.

However, if personal biases are avoided, then the relevant experience and the acquaintance of the investigator with the population may help to choose a relatively representative sample from the population. It is not possible to make an estimate of sampling error as we cannot determine how, precise our sample estimates are.

**iii) Quota sampling:** This is one of the most commonly used sampling method in marketing research studies. Here the sample is selected on the basis of certain basic parameters such as age, sex, income and occupation that describe the nature of a population so as to make it

representative of the population. The investigators or field workers are instructed to choose a sample that conforms to these parameters. The field workers are assigned quotas of the numbers of units satisfying the required characteristics on which data should be collected. However, before collecting data on these units the investigators are supposed to verify that the units qualify these characteristics.

The Quota sampling method has some weaknesses. These are listed below:

i) It is usually difficult to obtain an accurate and up to date proportion of respondents assigned to each cell.

ii) As the number of parameters (control characteristics) associated with the objectives of the study become large, the total number of cells increase. This makes the task of field staff difficult as it may not be easy to get a desired respondent.

iii) It is very important that all of the proper parameters (control characteristics) related to the study in question must be incorporated while taking sample. The results of the study could be misleading if any relevant parameter is omitted for one reason or the other.

**Q3. With the help of examples, discuss the areas where qualitative research can be used in marketing.**

**Ans. 1.** It is used in 'concept generation and evaluation' e.g., whenever company wants to develop new product, or modify existing product, it would conduct Qualitative Research among target consumers to understand new emerging consumer needs, or problems which consumer has with the existing product. Therefore, every time when one goes into the market place, one would find newer models of the same product with addition/deletion of features, new and better external appearance, better finish of the product etc.

2. It is used to define the problem areas more fully - in marketing research one normally starts with qualitative research which is validated further by quantitative research.

3. It is used to formulate hypothesis for further investigation/quantification.

4. It is used to obtain large amount of data about beliefs, attitudes, etc. as data input for developing questionnaires, attitude scales, which would be used as input for multivariate analysis studies.

5. It is also used to conduct post-research study i.e., to amplify or explain same points that emerge from a major study, without having to repeat on a large scale.

6. In some areas of marketing research --- especially understanding of consumer behaviour, a more flexible approach is required, rather than a rigid approach as provided by a structured questionnaire - hence qualitative research could be used.

7. In studies of distribution channels, sales, pricing strategies quantitative approach is most suitable, whereas in concept development, product development (needs of 4 consumer), advertising research -" qualitative approach is more suitable.

Ultimately, whether one uses a Qualitative Research or Quantitative Research approach, depends upon the objective of the research work, and the value of decisions that would be taken from the data generated by the research work.

Qualitative Research is taking more and more inputs from psychology, to develop latest techniques in Qualitative Research. Some of the latest techniques used are as follows:

**1) Imagery Technique:** This technique is used to determine the images of two competing brands, As an example, readers were asked to describe two competing magazines.

Magazine 1 was described as young, attractive, bubbling teenager.

Magazine 2 was described as older, less interesting, introvert lady. This information helped the owner of the magazine to reposition magazine 2.

**2) Synectics:** This technique gives the researcher an insight into the intuitive or emotional impressions of the consumer for either a brand or advertising campaign, by unleashing the creative side of the consumer. The respondents are asked to go through a series of games designed to lower their inhibitions which help in unleashing the child present in every adult. The basic idea is that creativity works at child level. With this exercise, the researcher can get original ideas for use in developing advertising campaign, as also the consumer's real feelings about products.

**3) Neuro-Linguistic Programming:** This technique provides an insight into the 'Non-Verbal Impact' of an advertisement on the target consumer. This is done by analyzing as to which of the five senses (smell, taste, touch, sight, aural) a consumer uses to recall an advertisement. As an example, the Onida advertisement campaign was analyzed using this technique, and it was able to reveal strong sensory effect of Rasping Voice and Shattering Glass. The underlying idea behind this technique is that much of the recording of experience goes on 'below the surface' and outside conscious awareness. For the purposes of marketing, there is a codification of the sensory modalities into Visual, Auditory and Kinesthetic (touch, taste and smell). The shorthand version is called as VAK segmentation. Human are emotional creatures, and a person would buy on emotion and justify the purchase on logic. Thus, to decode the emotions surrounding the product purchase and experience, researchers use the NLP technique. NLP techniques are modelled on techniques used in other fields of psychology (Ericksonian Hypnosis and Classical Conditioning), and hence only persons trained in the field of psychology can use this technique.

**Q4. Monthly salary of 32 employees of a firm is given below. Tabulate the data after selecting a suitable class interval.**

2250	1800	1650	1760	3520	5600	2450	2680
2700	1680	3650	3240	5850	3150	1860	2425
4520	3275	4215	3760	1950	1850	3750	2825
4500	3800	4300	2750	4370	3350	2375	3215

**Ans.** To select a suitable class interval for the given data, we can use the following formula:

Class interval = (Highest value - Lowest value) / Number of classes

In this case, the lowest value is 1650 and the highest value is 5850. We can choose 5 or 6 classes to group the data. Let's choose 6 classes to have a more detailed analysis.

Class interval =  $(5850 - 1650) / 6 = 700$

So, the class intervals will be:

1650 - 2349 2350 - 3049 3050 - 3749 3750 - 4449 4450 - 5149 5150 - 5850

Using these class intervals, we can tabulate the data as follows:

Salary Range	Frequency
1650 - 2349	4
2350 - 3049	7
3050 - 3749	6
3750 - 4449	8
4450 - 5149	4
5150 - 5850	3

We can interpret the data as follows:

4 employees have a monthly salary in the range of 1650 - 2349.

7 employees have a monthly salary in the range of 2350 - 3049.

6 employees have a monthly salary in the range of 3050 - 3749

8 employees have a monthly salary in the range of 3750 - 4449

4 employees have a monthly salary in the range of 4450 - 5149

3 employees have a monthly salary in the range of 5150 - 5850

This tabulation helps us to get a better understanding of the distribution of the data and identify any outliers or unusual patterns. We can also use this data to calculate measures of central tendency and dispersion to get further insights into the data.



**Q5. A manufacturer of industrial supplies developed the following model for predicting the number of sales per month**

$$Y = 41 + .3X_1 + .05X_2 - 7X_3 + 10X_4$$

where Y = Sales per month

X<sub>1</sub> = Number of manufacturing firms'

X<sub>2</sub> = Number of wholesale and retail firms

X<sub>3</sub> = Number of competing firms

X<sub>4</sub> = Number of full-time company sales people.

- i) Explain the correct interpretation of all estimated parameters in the equation.
- ii) If R<sup>2</sup> = 0.49, what does this figure mean, to you?
- iii) Explain how you will go about testing the validity of this multiple linear regression equation

**Ans. i)** The estimated parameters in the equation are the coefficients of each independent variable.

The coefficient of X<sub>1</sub> is 0.3, which means that for every one-unit increase in the number of manufacturing firms, sales are expected to increase by 0.3 units, holding all other variables constant.

The coefficient of X<sub>2</sub> is 0.05, which means that for every one-unit increase in the number of wholesale and retail firms, sales are expected to increase by 0.05 units, holding all other variables constant.

The coefficient of X<sub>3</sub> is -7, which means that for every one-unit increase in the number of competing firms, sales are expected to decrease by 7 units, holding all other variables constant.

The coefficient of X<sub>4</sub> is 10, which means that for every one-unit increase in the number of full-time companies salespeople, sales are expected to increase by 10 units, holding all other variables constant.

The intercept term (41) represents the expected level of sales when all independent variables are equal to zero.

**ii) Ans:-** R<sup>2</sup> is the coefficient of determination, which represents the proportion of the total variation in the dependent variable that is explained by the independent variables in the regression model. In this case, R<sup>2</sup> = 0.49 means that 49% of the variation in sales can be explained by the independent variables in the model, while the remaining 51% is unexplained.

**iii) Ans.** To test the validity of the multiple linear regression equation, we can use the following methods:

**Check for linearity:** We can check for linearity by examining scatter plots of the dependent variable against each independent variable. If the scatter plots show a linear pattern, then the assumption of linearity is satisfied.

**Check for normality:** We can check for normality by examining the distribution of the residuals (the differences between the observed values and the predicted values). If the residuals are normally distributed, then the assumption of normality is satisfied.

**Check for homoscedasticity:** We can check for homoscedasticity by examining a plot of the residuals against the predicted values. If the plot shows a random pattern with no trend, then the assumption of homoscedasticity is satisfied.

**Check for multicollinearity:** We can check for multicollinearity by examining the correlation matrix of the independent variables. If the correlation between any pair of independent variables is high (i.e., greater than 0.8), then multicollinearity may be a problem.

**Conduct hypothesis tests:** We can test the significance of each independent variable by conducting a t-test on the corresponding coefficient. If the p-value is less than the significance level (e.g., 0.05), then we can conclude that the independent variable is a significant predictor of the dependent variable.

By examining these assumptions and conducting hypothesis tests, we can determine whether the multiple linear regression equation is a valid and reliable model for predicting the number of sales per month.

