SECTION 5 SAMPLING – CASE STUDY

Structure **Page Nos.**

5.0 INTRODUCTION

In our previous sessions you learned about various tools, techniques and practices required to perform the statistical analysis of data. In general, it is observed that the outcome of such analysis (sample estimates)varies from the population values. A common reason for this lies in the size of sample data collected. Thus, the question is "How much data is sufficient?" and the answer relates to sampling. Recall earlier discussion on the population and the sample in Section 1.

It is impractical to study the entire population, so it is devised to study proportionate part of the population, for reliable interpretation of the population behavior. Thus the need of sampling techniques was identified. There are various methodologies devised to perform sampling of identified population, we will try to put light on most of the methodologies, which are relevant at your level of study. Further, you should keep following facts in your mind, that sampling is simply not a formula but in practice it is the Science combined with the Art of calculations and presentation.

You are advised to refer to Block 4 of BCS 040, and study the details about sampling and its related techniques. However, we will discuss about its practical aspects and how to attempt it as a practice, by using Excel and its tools.

5.1 OBJECTIVES

After going through this unit you will be able to :

- use Data Analysis Toolpak for Descriptive statistics of sampled data;
- use Descriptive statistics to calculate sample size through Excel; and
- use Excel functions for Sampling.

5.2 SAMPLING – SAMPLE SIZE

Since Computer is a discipline, that facilitates the working of our day to day life, we choose an example/Case from our daily life in this context and use it to interpret the respective results. Through the case study, we would like to emphasize that there are situations when the basic assumptions made while computing the sample size by using the formulas are not met in practice. At other times, there are factors which are influential in increasing or decreasing the sample size obtained through the use of formula. It should be kept in mind that sample size determination is the blend of using formulae, experience of similar studies, time and budget constraints and a few other elements such as output or analysis requirements etc.

Through the case, we are going to concentrate on the utilization and implementation of the sample size calculation formula for continuous or interval-scaled variables.

However, for other details you are advised to refer to Block 4 of BCS 040, and apply the skills learned in these practical sessions.

CASE DESCRIPTION :

- As a pilot project, a private company involved in hardware and maintenance related activity planned to launch its own brand of assembled computers. The company identified a small colony having 12 families in all. Each of the family is having members of different age group. For the sake of their brand performance related study, they decided to issue one computer system to each family as a single unit. They got the system installed with several applications suitable for each age group viz., Computer Games, Movies, Application Softwares, E-Books, Internet, and Security related software. The company wants to study the application usage pattern and the satisfaction level of the families, so that they can customize their product accordingly. In order to monitor the system usage time, they installed an automatic clock with the system, which records the usage time. This will help them to analyze the questionnaire feedback, which they wish to get filled, at the end of their study.

On the basis of an appraisal of the case cited above, you may identify the objectives of the study to be as under:

- What is the customer satisfaction level?
- Which software utility is performed exhaustively in most of the families?

In view of these objectives, you are required to draft a Questionnaire, where the responses to the questions are either direct or indirect, as desired. They can be categorized into one among the following scales, viz.,

- Nominal
- Ordinal
- Quantitative.

A sample questionnaire is given below, this will let you understand the Excel operations, to be performed at the later stages

QUESTIONNAIRE

Family No.:
What is the Average System Utilization time(minutes) ? :

15 30 45 60 75 90 105 120 135 150 165 180

What are the Activities Performed by your family while utilizing the system?

a) Computer Games, b) Internet, c) E-Books,

d) Application Software's, e)Movies f) Security related software

What is the level of satisfaction, achieved by your family ?

-2 = Very dissatisfied, 1 = dissatisfied, 0 = indifference,

1 = satisfied, 2 = Very satisfied.

Would your family like to purchase the product?

Yes No

We got the questionnaire filled from respective families, finally we have the following data from the computer savvy families. The collected data is through following variables

- *Family* (Nominal scale)is the observation/serial number of the questionnaire. In the present case, one questionnaire per household is filled by each family.
- *Time* (Quantitative scale)is a quantitative data type which is measured in minutes. It is the software utilization time for a family.
- *System Utilization Activity* (Nominal scale)is a nominal data type consisting of 6 choices of activity in the park:

    1. Computer Games,
    2. Internet,
    3. E-Books,
    4. Application Software's,
    5. Movies
    6. Security related software

    To this question, there may be multiple choices that for each family, one corresponding to each of the several activities they perform.

- *Product Satisfaction* (Ordinal scale) It measures family satisfaction toward the Product. It is measured in the ordinal scale with 5 values as options listed below:

    -2 = Very dissatisfied,
    1 = dissatisfied,
    0 = indifference,
    1 = satisfied,
    2 = Very satisfied.

- *Product Purchase* (Nominal scale)is measured in the nominal scale (Yes or No) regarding the final decision about purchase of the product.

Based on the responses to the questions framed in the questionnaire, collected data is presented in the following table:

Raw Data

Family	*System Utilization Time*	*System Utilization Activity*	*Product Satisfaction*	*Product Purchase*
1	30	1, 2, 3	0	N
2	30	4,6	1	Y
3	60	1, 2	2	Y
4	45	5	-1	N
5	30	6	1	N
6	60	2	2	Y
7	30	4	1	N
8	45	3, 4	-1	N
9	15	6	1	Y
10	60	2	2	Y
11	180	1, 2, 3, 4	2	Y
12	120	1,2,4	2	Y

Raw data presented in the table above is coded into the following form, which is amenable for further statistical analysis using spreadsheet software.

Raw Data coded for analysis in Formatted Excel Sheet

Family	*System Utilization Time*	*System Utilization Activity*						*Product Satisfaction*	*Product Purchase*
		1	2	3	4	5	6		
1	30	1	1	1	0	0	0	0	0
2	30	0	0	0	1	0	1	1	1
3	60	1	1	0	0	0	0	2	1
4	45	0	0	0	0	1	0	-1	0
5	30	0	0	0	0	0	1	1	0
6	60	0	1	0	0	0	0	2	1
7	30	0	0	0	1	0	0	1	0
8	45	0	0	1	1	0	0	-1	0
9	15	0	0	0	0	0	1	1	1
10	60	0	1	0	0	0	0	2	1
11	180	1	1	1	1	0	0	2	1
12	120	1	1	0	1	0	0	2	1

After the excel file with data in the format shown above has been created from the Raw Data table, notice that the responses under the head "*Product Purchase*" are marked 0 and 1, which represents the codes for responses "Y" and "N", respectively. The same also applies to the responses under "*System Utilization Activity*" head of Raw Data table.

Please note that Family, System Utilization Activity, Product Purchase, are variables measured in NOMINAL SCALE. So, we should NOT calculate sum, average, variance etc. Instead the best is to calculate frequency, percentage, plot graphs only. But here we are going to demonstrate that how you can use the tool of data analysis tool pack to get the desired answers, related to the marked objectives of your study. Now to get the answer for the stated objectives, we use the Descriptive Statistics Tool of Data Analysis ToolPak.

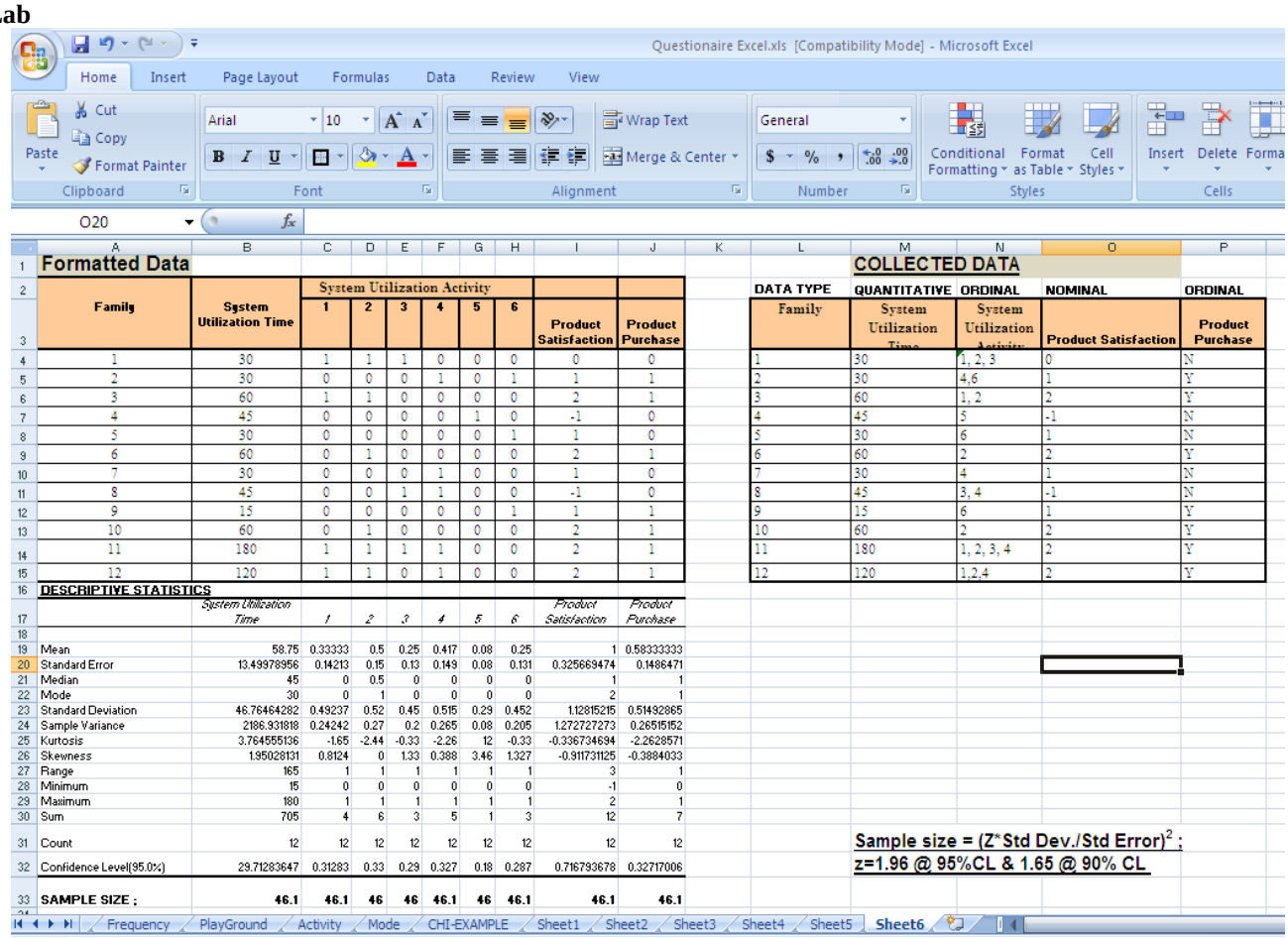
Data analysis through Excel:

Perform Following Steps:

1. Tabulate the collected data as given above in Excel Spreadsheet.
2. Click DATA TAB → DATA ANALYSIS → DESCRIPTIVE STATISTICS → OK

Note:"For activation and usage of Data Analysis Toolpak, refer to the earlier section1 & section Correlation and Regression - The snap shots are readily available there" .

We present below some of the relevant screenshots in the present case.



Questionaire Excel.xls [Compatibility Mode] - Microsoft Excel

Formatted Data

Family	System Utilization Time	System Utilization Activity 1	2	3	4	5	6	Product Satisfaction	Product Purchase
1	30	1	1	1	0	0	0	0	0
2	30	0	0	0	1	0	1	1	1
3	60	1	1	0	0	0	0	2	1
4	45	0	0	0	0	1	0	-1	0
5	30	0	0	0	0	0	1	1	0
6	60	0	1	0	0	0	0	2	1
7	30	0	0	0	1	0	0	1	0
8	45	0	0	1	1	0	0	-1	0
9	15	0	0	0	0	0	1	1	1
10	60	0	1	0	0	0	0	2	1
11	180	1	1	1	1	0	0	2	1
12	120	1	1	0	1	0	0	2	1

COLLECTED DATA

DATA TYPE	QUANTITATIVE	ORDINAL	NOMINAL	ORDINAL
Family	System Utilization Time	System Utilization Activity	Product Satisfaction	Product Purchase
1	30	1, 2, 3	0	N
2	30	4,6	1	Y
3	60	1, 2	2	Y
4	45	5	-1	N
5	30	6	1	N
6	60	2	2	Y
7	30	4	1	N
8	45	3, 4	-1	N
9	15	6	1	Y
10	60	2	2	Y
11	180	1, 2, 3, 4	2	Y
12	120	1,2,4	2	Y

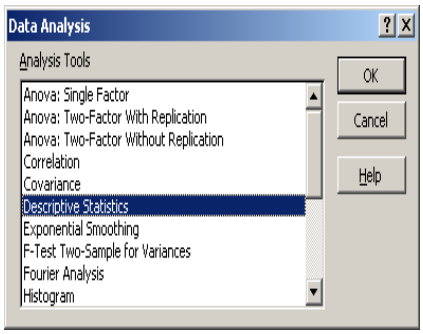
DESCRIPTIVE STATISTICS

	System Utilization Time	1	2	3	4	5	6	Product Satisfaction	Product Purchase
Mean	58.75	0.33333	0.5	0.25	0.417	0.08	0.25	1	0.58333333
Standard Error	13.49978956	0.14213	0.15	0.13	0.149	0.08	0.131	0.325669474	0.1486471
Median	45	0	0.5	0	0	0	0	1	1
Mode	30	0	1	0	0	0	0	2	1
Standard Deviation	46.76464282	0.49237	0.52	0.45	0.515	0.29	0.452	1.12815215	0.51492865
Sample Variance	2186.931818	0.24242	0.27	0.2	0.265	0.08	0.205	1.272727273	0.26515152
Kurtosis	3.764555136	-1.65	-2.44	-0.33	-2.26	12	-0.33	-0.336734694	-2.2628571
Skewness	1.95028131	0.8124	0	1.33	0.388	3.46	1.327	-0.911731125	-0.3884033
Range	165	1	1	1	1	1	1	3	1
Minimum	15	0	0	0	0	0	0	-1	0
Maximum	180	1	1	1	1	1	1	2	1
Sum	705	4	6	3	5	1	3	12	7
Count	12	12	12	12	12	12	12	12	12
Confidence Level(95.0%)	29.71283647	0.31283	0.33	0.29	0.327	0.18	0.287	0.716793678	0.32717006
SAMPLE SIZE :	46.1	46.1	46	46	46.1	46	46.1	46.1	46.1

Sample size = $(Z^*Std\ Dev./Std\ Error)^2$;
z=1.96 @ 95%CL & 1.65 @ 90% CL

Sheet tabs: Frequency | PlayGround | Activity | Mode | CHI-EXAMPLE | Sheet1 | Sheet2 | Sheet3 | Sheet4 | Sheet5 | Sheet6

Results of DESCRIPTIVE STATISTICS Procedure



Statistics	System Utilization Time	System Utilization Activity 1	2	3	4	5	6	Product Satisfaction	Product Purchase
Mean	58.75	0.33	0.50	0.25	0.42	0.08	0.25	1	0.58
Standard Error	13.50	0.14	0.15	0.13	0.15	0.08	0.13	0.33	0.15
Median	45	0	0.5	0	0	0	0	1	1
Mode	30	0	1	0	0	0	0	2	1
Standard Deviation	46.76	0.49	0.52	0.45	0.51	0.29	0.45	1.13	0.51
Sample Variance	2186.93	0.24	0.27	0.20	0.27	0.08	0.20	1.27	0.27
Kurtosis	3.8	-1.65	-2.44	-0.33	-2.26	12	-0.33	-0.34	-2.26
Skewness	2.0	0.81	0	1.33	0.39	3.46	1.33	-0.91	-0.39
Range	165	1	1	1	1	1	1	3	1
Minimum	15	0	0	0	0	0	0	-1	0
Maximum	180	1	1	1	1	1	1	2	1
Sum	705	4	6	3	5	1	3	12	7
Count	12	12	12	12	12	12	12	12	12
Sample size = $(Z^*Std\ Dev./Std\ Error)^2$	46.0992	46.099	46.099	46.099	46.099	46.099	46.099	46.099	46.0992
Z=1.96 @ 95%CL & 1.65 @ 90% CL						46.099	46.099	46.0992	46.0992

DATA INTERPRETATION

Now, let us interpret the results of descriptive statistics and analyse, what they say about our objectives viz.

- What is the customer satisfaction level?
- Which software utility is performed exhaustively in most of the families?

So far as *customer satisfaction level is concerned, the tabulated data under head of Mean and Mode gives the result as 1 and 2 respectively* i.e.,on an average the customers are satisfied and in fact they are "most frequently" - very satisfied. To interpret, we are to refer our data coding standards. As in this case they are :

- *Product Satisfaction* is an ordinal scale with 5 values:

  -2 = Very dissatisfied,
  1 = dissatisfied,
  0 = indifference,
  1 = satisfied,
  2 = Very satisfied.

Similarly, in order to study the objective "*Which software utility is performed exhaustively in most of the families? Refer to the results under the head SUM for the* ***System Utilization Activity***. *Here, the activity 2 is identified to be performed the most* i.e. *Families are mostly using the Computer system for internet related activities.*

To interpret, we are to refer our data coding standards. As in this case, they are:

- *System Utilization Activity* is a nominal data type consist of 6 choices of activity in the park:

  1. Computer Games,
  2. Internet,
  3. E-Books,
  4. Application Software's,
  5. Movies
  6. Security related software

Now, you might be wondering that the purpose of this session is to talk about sampling and we are discussing about the descriptive statistics. This is intentional, we want you to realize that sampling is not simply a formula but it is related to various factors, as in this case the number of families are limited in the target area i.e. target population is limited. By using the formula for Sample size calculation i.e. *Sample size = (Z*Std Dev. / Std Error)^2 ; we can calculate the respective sample size as*

$$\textbf{Sample size} = (Z * \text{Std Dev.} / \text{Std Error})^2$$

You might have not located this expression *Sample size=(Z*StdDev./Std Error)^2* in the Units12,13,14 of BCS 040. But, for the sake of understanding its just another formula for calculating the sample size, when standard deviation is known. However when standard deviation is not known, we use another formula,we are not going to discuss that, here we are going to concentrate on *Sample size=(Z*StdDev./Std Error)^2*

We calculated the sample size to be 46 families but our study is considering only 12. So for the reliability of the study the collected data is insufficient and thus we cannot conclude that the interpretation of the results related to our objectives is fine enough.

You might be in the position to realize the importance of sampling now, because from the above case, you identified that by limited data you might find that the objectives of the study are fulfilled but which might not be the fact as the sample size is inappropriate.

After Sample size calculation let us understand how to use Excel for Random Sampling. You might have studied various sampling methods in block 4 of BCS 040, but we are confining our discussion to only Random sampling, you can implement the rest of the sampling techniques by applying the understanding of BCS 040 and Excel concepts, learned in these sessions.

5.3 RANDOM SAMPLING

Let us extend our discussion from past session, where we considered the sample size of 12 families and based on the descriptive statistics results we analyzed that the sample size is not large enough for the study. But if we assume that the population itself is only 12 families and sample size turns out to be only 4 families, and we are going to opt for random sampling. Then among the 12 families, how to choose 4 families at random, this is the task we are going to demonstrate in this session. Among these 12 families 4 families are to be chosen at Random so we firstly Randomize the entire population data by using Rand() function. We, simply write Rand() in the subsequent column and apply it for entire population data. But, prior to that we should understand that the sampling mechanism is random sampling, so, to draw the sample by Simple Ransom Sample discussed in section 12.4/ Unit-12 of BCS 040, these probabilities have to be all equal, and we have to draw only 4 random numbers to select the required 4 sample units. Thus, we need to determine the equal probabilities of being chosen, which turns out to be (1/12) where 12 is total number of families, the result 0.08 is the equal probability assigned to each family, as shown below.

You understood the theoretical way of doing it, in our block 4 of BCS 040. Here, we are going to use Excel as the tool for the mentioned purpose. We are going to explore, following functions to perform the task RAND(), SMALL(), MATCH() and INDEX() functions to perform the above framed task i.e. to do Random Sampling.

Say the population data of 12 families under consideration is as follows

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	*Family*	1	2	3	4	5	6	7	8	9	10	11	12
2	*Surname*	Sharma	kumar	Sastry	Venugopal	Gupta	Nagpal	Dhiman	Kaushik	Bhardwaj	Garg	Giri	Iyengar
3	*Probability* ***(1/12=0.08)***	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
4	*Random#*	0.88	0.03	0.43	0.22	0.40	0.82	0.96	0.44	0.55	0.66	0.04	0.86

Task:
From among these 12 families, 4 families are to be chosen at Random.

Steps:

1. At ever draw equal probabilities have to be assigned to the units (see 3rd paragraph/section 12.4/Unit-12).These probabilities are $\frac{1}{12}$=0.08 as shown in C3..C14
2. Use excel function RAND( ) to generate random numbers in the range [0,1].
3. Calculate cumulative probabilities as shown in D3..D14.
    a. Value 0 in D2 is marked for the calculation of cumulative frequency
4. Consider a discrete random variable X having probability mass function $P(X = x_i) = p_i$, $i = 0,1,2,$ ... and $\sum_i p_i = 1$. Then we generate an uniform random number U and the corresponding observation $X = x_i$, if

$$\sum_{k=0}^{i-1} p_k \leq U < \sum_{k=0}^{i} p_k$$

5. Repeat the above step $n = 4$ times to draw the required random sample. This is shown in columns E, F, G and H.

	A	B	C	D
1	*Family*	*Surname*	*Probability*	*Cumulative Probability*
2				0
3	1	Sharma	0.08	0.08
4	2	Kumar	0.08	0.17
5	3	Sastry	0.08	0.25
6	4	Venugopal	0.08	0.33
7	5	Gupta	0.08	0.42
8	6	Nagpal	0.08	0.50
9	7	Dhiman	0.08	0.58
10	8	Kaushik	0.08	0.67
11	9	Bhardwaj	0.08	0.75
12	10	Garg	0.08	0.83
13	11	Giri	0.08	0.92
14	12	Iyengar	0.08	1.00

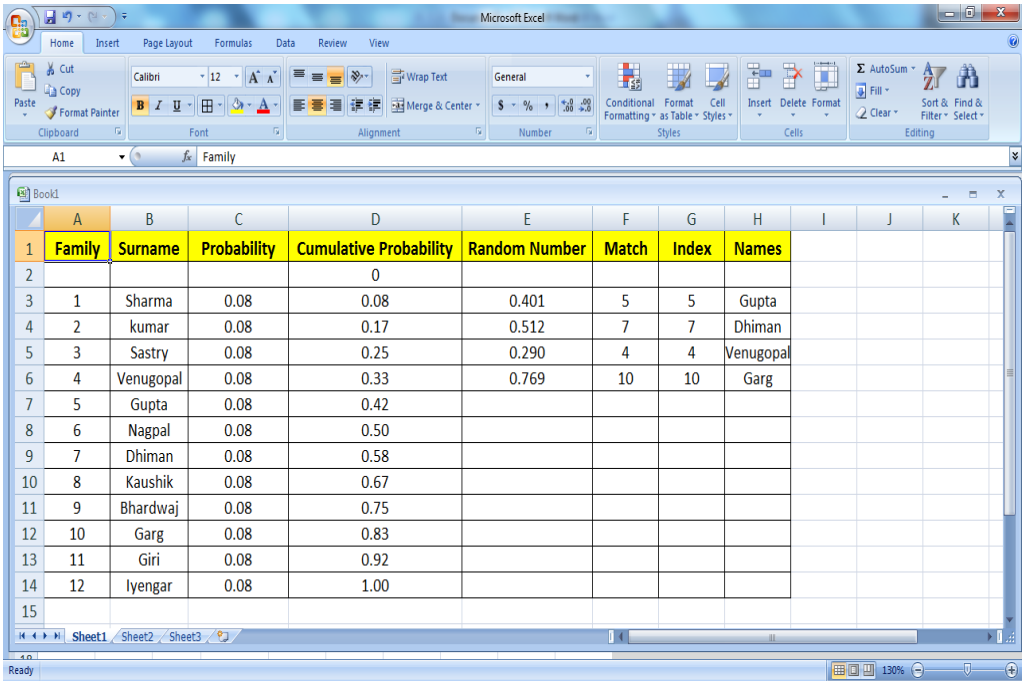
Formula :

The function Rand() generates the random numbers from 0 to 1, from the above results it is observed that $B$3:$M$3 is the range of generated random numbers.

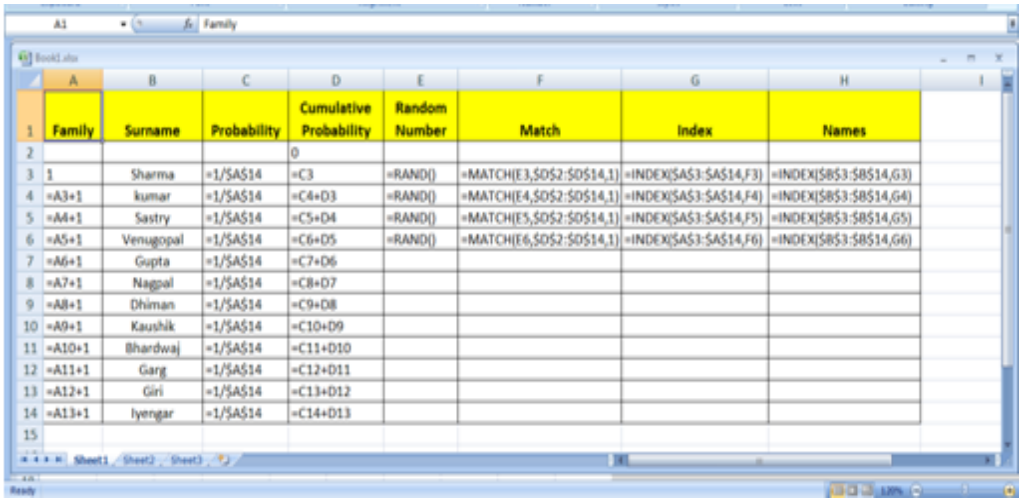
match(lookup value, lookup array), i.e. =MATCH(E3,$D$2:$D$14) :returns the relative position of an item in an array that matches a specified value in a specified order.

To be specific to the application, by using match we can determine the Index value (which is the Family number in our case). Everytime you refresh the Random# column the entire status of Match column will change randomly. Now our sample is of size 4 families, thus either we can directly refer to first 4 entries under match# as the random families to be opted OR we can use the formula Index(). Lets use index to clearly identify the Number and Surname of the Randomly sampled families.

index(reference, row num[column num]), i.e. =INDEX($A$3:$A$14,F3) returns the Randomly sampled family number and =INDEX($B$3:$B$14,G3) returns the Randomly sampled family surname. One of the Randomly Sampled Data is shown in screen shot below.



	A	B	C	D	E	F	G	H
1	Family	Surname	Probability	Cumulative Probability	Random Number	Match	Index	Names
2				0				
3	1	Sharma	0.08	0.08	0.401	5	5	Gupta
4	2	kumar	0.08	0.17	0.512	7	7	Dhiman
5	3	Sastry	0.08	0.25	0.290	4	4	Venugopal
6	4	Venugopal	0.08	0.33	0.769	10	10	Garg
7	5	Gupta	0.08	0.42				
8	6	Nagpal	0.08	0.50				
9	7	Dhiman	0.08	0.58				
10	8	Kaushik	0.08	0.67				
11	9	Bhardwaj	0.08	0.75				
12	10	Garg	0.08	0.83				
13	11	Giri	0.08	0.92				
14	12	Iyengar	0.08	1.00				
15								



	A	B	C	D	E	F	G	H
1	Family	Surname	Probability	Cumulative Probability	Random Number	Match	Index	Names
2				0				
3	1	Sharma	=1/$A$14	=C3	=RAND()	=MATCH(E3,$D$2:$D$14,1)	=INDEX($A$3:$A$14,F3)	=INDEX($B$3:$B$14,G3)
4	=A3+1	kumar	=1/$A$14	=C4+D3	=RAND()	=MATCH(E4,$D$2:$D$14,1)	=INDEX($A$3:$A$14,F4)	=INDEX($B$3:$B$14,G4)
5	=A4+1	Sastry	=1/$A$14	=C5+D4	=RAND()	=MATCH(E5,$D$2:$D$14,1)	=INDEX($A$3:$A$14,F5)	=INDEX($B$3:$B$14,G5)
6	=A5+1	Venugopal	=1/$A$14	=C6+D5	=RAND()	=MATCH(E6,$D$2:$D$14,1)	=INDEX($A$3:$A$14,F6)	=INDEX($B$3:$B$14,G6)
7	=A6+1	Gupta	=1/$A$14	=C7+D6				
8	=A7+1	Nagpal	=1/$A$14	=C8+D7				
9	=A8+1	Dhiman	=1/$A$14	=C9+D8				
10	=A9+1	Kaushik	=1/$A$14	=C10+D9				
11	=A10+1	Bhardwaj	=1/$A$14	=C11+D10				
12	=A11+1	Garg	=1/$A$14	=C12+D11				
13	=A12+1	Giri	=1/$A$14	=C13+D12				
14	=A13+1	Iyengar	=1/$A$14	=C14+D13				
15								

☞ Lab Sessions 9 and 10

1) Analyze the Case given below and answer the subsequent questions
" *A park, maintained by an NGO, the management of the Park wants to make a study which will help them to identify the satisfaction level of the visitor families, the park has facility for every family member as if the youngsters may use playground for sports and others may use it for picnic or reading or meditation or walk or jogging, but they should adhere to the norms laid by the park management. As usual the existence of playground is always an issue(generation gap), so the management want to decide that the availability of playground leads to family satisfaction or it should be removed or transformed for some other activity. Further, the management of conveyance and over crowdedness in the park are also to be addressed. The management identified that different activities are performed for different duration of time, they wish to identify the activity performed the most and its respective duration, and the mode or vehicles used by families to visit the park, based on the outcome they may decide that which activities should be commercialize and do they need to focus on parking*"

1. Identify the objectives of the study.
2. Draft a Questionnaire to study the identified objectives.
3. Tabulate the collected data and digitize it in excel spreadsheet.
4. Use the Descriptive statistics option of Data analysis toolpak, and generate the results for collected data.
5. Analyze the generated results for the identified objectives, and comment accordingly.
6. Use the results from descriptive statistics to calculate the sample size, comment whether the sample size is sufficient/ insufficient for the conducted study.

2) Develop an Excel Application that can perform Random sampling for the data tabulated below, and choose 3 villages randomly for study.

VILLAGE	**A**	**B**	**C**	**D**	**E**	**F**	**G**
IT SERVICES	115 231	267	98	155		321	144

5.4 SUMMARY

In this session you learned about the basic formulation of questionnaire for an assigned task, and learned about the various type of scaling viz. nominal, ordinal etc, for data collection. Further, you learned about the tabulation of the data gathered through questionnaire, and its digitization in excel. We also demonstrated the use of data analysis tool pack for data interpretation. The session, also enlightened the concept of random sampling through spread sheets, where you learned to use various formulas viz. rand(), match(), index() etc. After going through this session you might be in the position to practically apply the statistical concepts at some elementary level.

5.5 ANSWERS TO CHECK YOUR PROGRESS

Correlation
Check Your Progress-1

1) a) Scatter plot shows that data is quite correlated with positive magnitude and value or r is approximately close to 0.9 or so
   b) r= 0.96 thus the variables are synchronized , thus variation in one variable directly affects the other.
   c) Yes, we can use the collected data for the forecasting purpose because r~0.96 i.e. good correlation

Regression
Check Your Progress-2

1) a) Both factors are identified to be highly interdependent
   b) Unit increase in RAM size will lead to 0.69 times increase in systems performance
   c) Unit improvement in systems performance requires 0.54 times of alteration in RAM

d) 25 times of RAM change will lead to improve system performance by 0.69 X 25 times

Multiple Regression

Check Your Progress-3

Q1 a)~70.6% of variation of Y is explained by X1 and X2
b)-5.51 – 39.33 X1 + 6.295 X2

ANOVA

Check Your Progress-1

1. Alpha =0.05 i.e.5% , so level of confidence is 95%
2. Critical value of Factor F (Fcrit) = 3.238872
3. Because F= 0.290072 and Fcrit = 3.238872 i.e. F <Fcrit , so accept the NULL Hypothesis.
4. P=0.831921, which is more than the desired significance level, so accept the null hypothesis

Check Your Progress-2

1. Alpha =0.05 i.e. 5%, thus level of confidence = 95%
2. Critical value of F(Fcrit) = 3.49 and 3.25 respectively
3. F=8.015 ; Fcrit = 3.49 ; F >Fcrit , so Non Acceptance of Null Hypothesis
4. F=6.291 ; Fcrit = 3.259 ; F >Fcrit , so Non Acceptance of Null Hypothesis
5. P value > significant value, thus , Non Acceptance of Null Hypothesis