

Index

SR NO	Practical	Date	Sign
1	Design a survey form for a given case study, collect the primary data and analyze it.		
2	Obtaining descriptive statistics of data, using Excel.		
3	Perform analysis of given secondary data, using Excel.		
4	Write a program for t-test comparing two means for independent samples, using Excel.		
5	Perform testing of hypothesis using chi-squared goodness- of-fit test, using Excel.		
6	Perform testing of hypothesis using chi-squared test of independence.		
7	Perform testing of hypothesis using One-way ANOVA, using Excel.		
8	Perform testing of hypothesis using Two-way ANOVA, using Excel.		
9	Perform testing of hypothesis using MANOVA, using Excel.		
10	Perform the Random sampling for the given data and analyze it, using Excel.		

Practical No 1

Aim: Design a survey form for a given case study, collect the primary data and analyze it.

Case 1:

A researcher wants to conduct a Survey in colleges on Use of ICT in higher education from Mumbai, Thane and Navi Mumbai. The survey focuses on access to and use of ICT in teaching and learning, as well as on attitudes towards the use of ICT in teaching and learning.

Design questionnaire addressed to teachers seeks information about the target class, his experience using ICT for teaching, access to ICT infrastructure, support available, ICT based activities and material used, obstacles to the use of ICT in teaching, learning activities with the target class, your skills and attitudes to ICT, and some personal background information.

Arrange question in following groups:

- 1) Information about the target class you teach
- 2) Experience with ICT for teaching
- 3) ICT access for teaching
- 4) Support to teachers for ICT use
- 5) ICT based activities and material used for teaching
- 6) Obstacles to using ICT in teaching and learning
- 7) Learning activities with the target class
- 8) Teacher skills
- 9) Teacher opinions and attitudes
- 10) Personal background information

Case 2:

A research agency wants to study the perception about App based taxi service in Mumbai, Thane and Navi Mumbai. The survey focuses on customers attitude towards app base taxi service as well as on attitudes towards regular taxi cab.

Design questionnaire seeks information about the target taxi service, his experience using taxi services, access, support available, obstacles and some personal background information, with the following objectives:

- 1) To find out the customer satisfaction towards the App based-taxi services.
- 2) To find the level of convenience and comfort with App based -taxi services.
- 3) To know their opinion about the tariff system and promptness of service.
- 4) To ascertain the customer view towards the driver behaviour and courtesy.
- 5) To provide inputs to enhance the services to delight the customers.
- 6) To examine relationship between service quality factors and taxi passenger satisfaction.
- 7) To suggest better regulations for transportation authorities regarding customer protection and effective monitoring of taxi services.

Case 3:

A popular electronic store want to conduct a survey to develop awareness of branded laptop baseline estimates and determine popularity of different company's laptop. It suggests steps to be initiated or strengthened in the field of demand in a region. The key indicators are among the general population, demand branded laptop and the problem users.

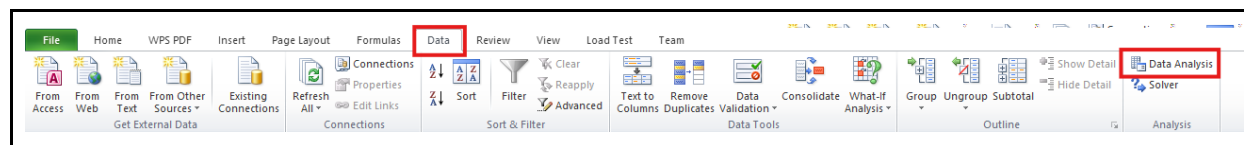
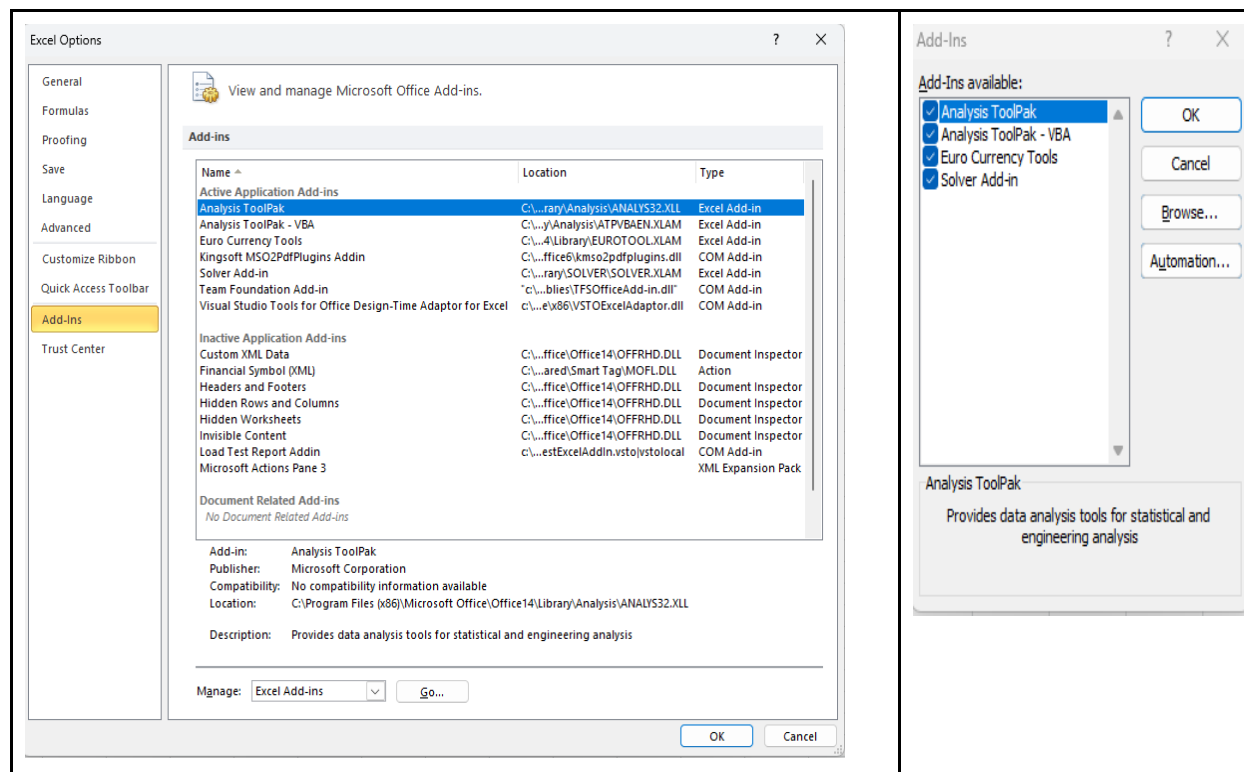
The objectives of this particular study are:-

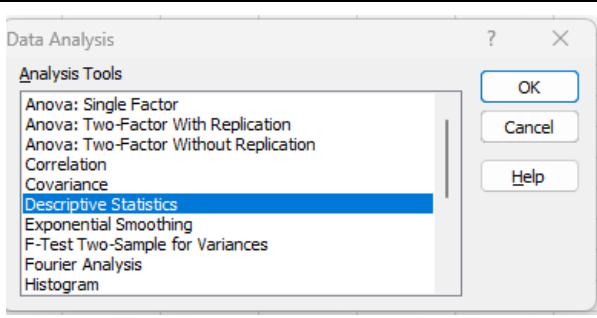
- 1) To know the preferences of different types of branded laptops by students and professionals.
- 2) To study which factor influence for choosing different types of branded laptops.
- 3) To know about the level of satisfaction towards different types of branded laptops.
- 4) To identify the perception of consumers towards the laptop positioning strategy.
- 5) To know the consumer preference towards laptop in the present era. Use the collected data for analysis.

Practical No 2

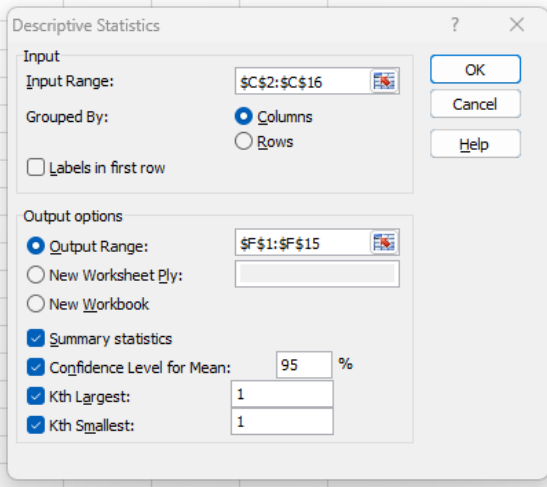
Aim: Obtaining descriptive statistics of data, using Excel.

Go to File Menu >> Options >> Add-Ins >> Select Analysis ToolPak >> Press OK





SR. NO	Name	Age	Rating
1	Pankaj	21	4.23
2	Aditi	22	5.2
3	Poornima	23	4.3
4	Vinit	24	9.15
5	Aditi	25	5.2
6	Sayali	26	5.2
7	Pankaj	27	6.2
8	Aditi	28	6.7
9	Poornima	29	8.2
10	Vinit	30	9.4
11	Aditi	31	3.2
12	Sayali	32	1.2
13	Pankaj	33	2.5
14	Aditi	34	5.6
15	Poornima	35	4.2



OUTPUT:

	A	B	C	D	E	F	G	H
1	SR. NO	Name	Age	Rating		Column1		
2	1	Pankaj	21	4.23				
3	2	Aditi	22	5.2		Mean	28	
4	3	Poornima	23	4.3		Standard Error	1.154700538	
5	4	Vinit	24	9.15		Median	28	
6	5	Aditi	25	5.2		Mode	#N/A	
7	6	Sayali	26	5.2		Standard Deviation	4.472135955	
8	7	Pankaj	27	6.2		Sample Variance	20	
9	8	Aditi	28	6.7		Kurtosis	-1.2	
10	9	Poornima	29	8.2		Skewness	-3.66008E-17	
11	10	Vinit	30	9.4		Range	14	
12	11	Aditi	31	3.2		Minimum	21	
13	12	Sayali	32	1.2		Maximum	35	
14	13	Pankaj	33	2.5		Sum	420	
15	14	Aditi	34	5.6		Count	15	
16	15	Poornima	35	4.2		Largest(1)	35	
17						Smallest(1)	21	
18						Confidence Level(95.0%)	2.476586343	
19								

Practical No 3

Aim: Perform analysis of given secondary data, using Excel.

Steps in Secondary Data Analysis

1. Determine your research question – Knowing exactly what you are looking for.
2. Locating data– Knowing what is out there and whether you can gain access to it. A quick Internet search, possibly with the help of a librarian, will reveal a wealth of options.
3. Evaluating relevance of the data – Considering things like the data's original purpose, when it was collected, population, sampling strategy/sample, data collection protocols, operationalization of concepts, questions asked, and form/shape of the data.
4. Assessing credibility of the data – Establishing the credentials of the original researchers, searching for full explication of methods including any problems encountered, determining how consistent the data is with data from other sources, and discovering whether the data has been used in any credible published research.
5. Analysis – This will generally involve a range of statistical processes. Example: Analyze the given Population Census Data for Planning and Decision Making by using the size and composition of populations

	A	B	C	D	E	F
1	AGE	MALE	FEMALE	TOTAL	MALE(%)	Female(%)
2	0-5	65,566	45,678	1,11,244		
3	6-10	57,456	98,345	1,55,801		
4	11-15	98,547	98,567	1,97,114		
5	16-20	23,567	23,876	47,443		
6	21-25	23,888	45,567	69,455		
7	26-30	90,777	34,567	1,25,344		
8	31-35	45,778	87,456	1,33,234		
9	36-40	22,333	86,988	1,09,321		
10	41-45	56,234	98,444	1,54,678		
11	46-50	78,567	43,876	1,22,443		
12	Total	5,62,713	6,63,364	12,26,077		

Put the cursor in cell B22 and click on the AutoSum and then click Enter. This will calculate the total population.

Then copy the formula in cell D22 across the row 22. To calculate the percent of males in cell E4, enter the formula $=-1*100*B4/DS\$22$. And copy the formula in cell E4 down to cell E21.

To calculate the percent of females in cell F4, enter the formula =100*C4/\$D\$22. Copy the formula in cell F4 down to cell F21

fx

Insert Function

Σ AutoSum

Recently Used

Financial

Logical

Text

Date & Time

Defined Names

Formula Auditing

Calculation Options

Calculation

F12

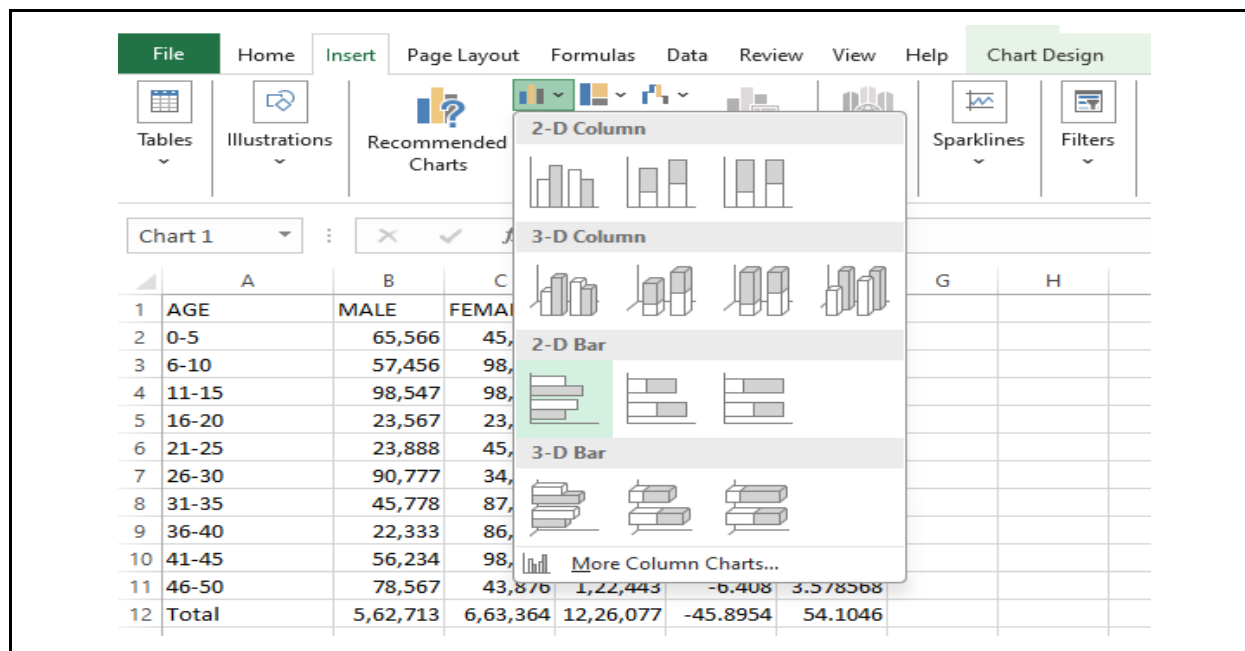
:

✕ ✓ fx

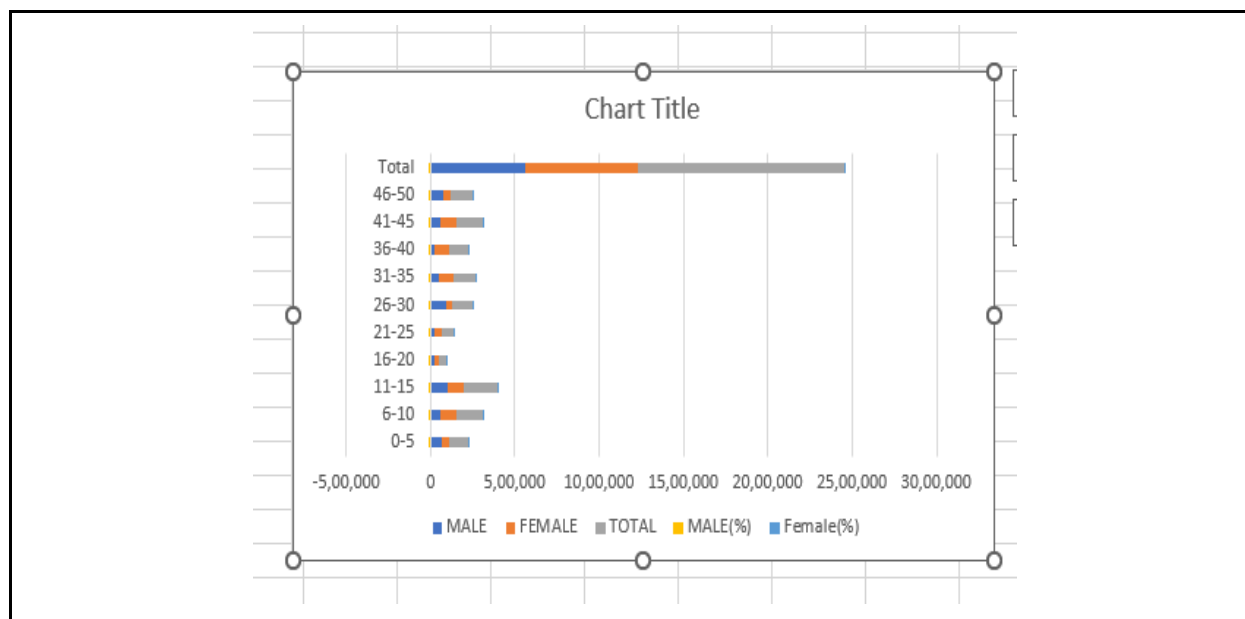
=100*C12/\$D\$12

	A	B	C	D	E	F	G	H	I	J
1	AGE	MALE	FEMALE	TOTAL	MALE(%)	Female(%)				
2	0-5	65,566	45,678	1,11,244	-5.34762	3.725541				
3	6-10	57,456	98,345	1,55,801	-4.68617	8.021111				
4	11-15	98,547	98,567	1,97,114	-8.03759	8.039218				
5	16-20	23,567	23,876	47,443	-1.92215	1.947349				
6	21-25	23,888	45,567	69,455	-1.94833	3.716488				
7	26-30	90,777	34,567	1,25,344	-7.40386	2.819317				
8	31-35	45,778	87,456	1,33,234	-3.7337	7.132994				
9	36-40	22,333	86,988	1,09,321	-1.8215	7.094824				
10	41-45	56,234	98,444	1,54,678	-4.5865	8.029186				
11	46-50	78,567	43,876	1,22,443	-6.408	3.578568				
12	Total	5,62,713	6,63,364	12,26,077	-45.8954	54.1046				

To build the population pyramid, we need to choose a horizontal bar chart with two series of data (% male and % female) and the age labels in column A as the Category X-axis labels. Highlight the range A3:A21, hold down the CTRL key and highlight the range E3:F21 Under inset tab, under horizontal bar charts select clustered bar chart



Choose Axis options tab and set the major and minor tick mark type to None, Axis labels to Low, and click OK. Click on any of the bars in your pyramid, click right and select “format data series”. Set the Overlap to 100 and Gap Width to 0. Click OK



Practical No 4

Aim: Write a program for t-test comparing two means for independent samples, using Excel.

The t distribution provides a good way to perform one sample tests on the mean when the population variance is not known provided the population is normal or the sample is sufficiently large so that the Central Limit Theorem applies.

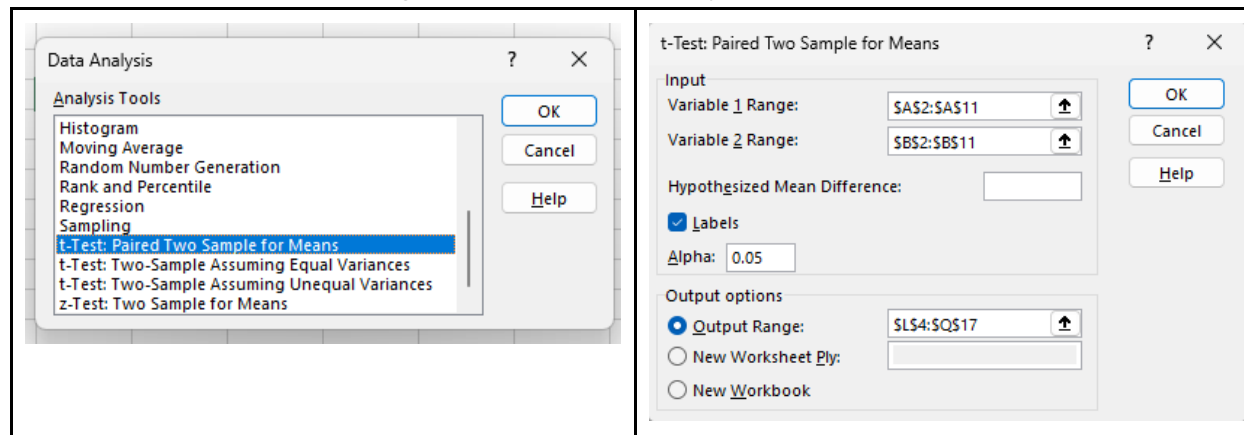
Two Sample t Test Example: A college Principal informed classroom teachers that some of their students showed unusual potential for intellectual gains. One months later the students identified to teachers a shaving potential for unusual intellectual gains showed significantly greater gains performance on a test said to measure IQ than did students who were not so identified. Below are the data for the students:

	A	B	C
1	Experimental	comparison	
2	65	8	
3	45	45	
4	41	75	
5	55	1	
6	25	5	
7	77	8	
8	55	12	
9	77	65	
10	74	5	
11	85	9	
12	59.9	23.3	mean
13	18.12429309	26.1726957	sd

Experimental Data To calculate Standard Mean go to cell A22 and type
`=SUM(A2:A21)/20` To calculate Standard Deviation go to cell A23 and type
`=STDEV(A2:A21)`

Comparison Data To calculate Standard Mean go to cell B22 and type
 $=\text{SUM}(\text{B2:B21})/20$ To calculate Standard Deviation go to cell B23 and type
 $=\text{STDEV}(\text{B2:B21})$

To find T-Test Statistics go to data Data Analysis



To calculate the T-Test square value go to cell E20 and type
 $=\text{SUM}(\text{A22-B22})/\text{SQRT}((\text{A23}*\text{A23})/\text{COUNT}(\text{A2:A21})+(\text{B23}*\text{B23})/\text{COUNT}(\text{A2:A21}))$

Now go to cell E20 and type

$=\text{IF}(\text{E20}<\text{E12}, \text{"H0 is Accepted"}, \text{"H0 is Rejected and H1 is Accepted"})$

Our calculated value is larger than the tabled value at $\alpha = .01$, so we reject the null hypothesis and accept the alternative hypothesis, namely, that the difference in gain scores is likely the result of the experimental treatment and not the result of chance variation.

Experimental	comparison			
65	8	t-Test: Paired Two Sample for Means		
45	45			
41	75		65	8
55	1	Mean	59.33333333	25
25	5	Variance	407	823.75
77	8	Observations	9	9
55	12	Pearson Correlation	-0.141186393	
77	65	Hypothesized Mean Difference	0	
74	5	df	8	
85	9	t Stat	2.758461317	
59.9	23.3	P(T<=t) one-tail	0.012366743	
18.12429309	26.1726957	t Critical one-tail	1.859548038	
		P(T<=t) two-tail	0.024733485	
		t Critical two-tail	2.306004135	
		calculate value	555.1532607	

Practical No 5

Aim: Perform testing of hypothesis using chi-squared goodness- of-fit test, using Excel.

Problem

An system administrator needsto upgrade the computersfor his division. He wants to know what sort of computer system his workers prefer. He gives three choices: Windows, Mac, or Linux. Test the hypothesis or theory that an equal percentage of the population prefers each type of computer system .

System	O	Ei	$\sum \frac{(O_i - E_i)^2}{E_i}$
Windows	20	33.33 %	
Mac	60	33.33 %	
Linux	20	33.33 %	

H0 : The population distribution of the variable is the same as the proposed distribution
HA : The distributions are different

To calculate the Chi –Squred value for Windows go to cell D2 and type =((B2-C2)*(B2-C2))/C2

To calculate the Chi –Squred value for Mac go to cell D3 and type =((B3-C3)*(B3- C3))/C3

To calculate the Chi –Squred value for Mac go to cell D3 and type =((B4-C4)*(B4- C4))/C4

Go to Cell D5 for and type=SUM(D2:D4)

To get the table value for Chi-Square for $\alpha = 0.05$ and dof = 2, go to cell D7 and type =CHIINV(0.05,2)

At cell D8 type =IF(D5>D7, "H0 Accepted","H0 Rejected")

OUTPUT:

	A	B	C	D	E	F
1	System	O	Ei	$\sum (O_i - E_i)^2 / E_i$		
2	Windows	20	33.33	5.33120012		
3	Mac	60	33.33	21.34080108		
4	Linux	20	33.33	5.33120012		
5	Total	100	100	32.00320132		
6						
7			Total Value	5.991464547		
8						
9		H0: The population distribution of the variable is the same as the proposed distribution				
10		H1:- The distributions are different				

Practical No 6

Aim: Perform testing of hypothesis using chi-squared test of independence.

In a study to understand the permormacne of M. Sc. IT Part -1 class, a college selects a random sample of 100 students. Each student was asked his grade obtained in B. Sc. IT.

The sample is as given below

	A	B	C	D	E	F	G	H
1		O	A	B	C	D	Total	$\sum (O_i - E_i)^2 / E_i$
2	Girls	11	7	5	5	11	39	6.074863267
3	Boys	30	4	3	10	14	61	6.074863267
4	Total	41	11	8	15	25	100	12.14972653
5	Ei	20.5	5.5	4	7.5	12.5	50	
6								

To calculated the expected value **Ei**

Go to Cell B5 and type =B4/2

Go to Cell C5 and type =C4/2

Go to Cell D5 and type =D4/2

Go to Cell E5 and type =E4/2

Go to Cell F5 and type =F4/2

Go to Cell S6 and calculate total girl students = SUM(B2:F2)

Go to Cell S7 and calculate total girl students = SUM(B3:F3)

Now Calculate

Go to cell H2 and type

=SUM((B2-\$B\$5)^2/\$B\$5,(C2-\$C\$5)^2/\$C\$5,(D2-\$D\$5)^2/\$D\$5,(E2-\$E\$5)^2/\$E\$5,(F2-\$F\$5)^2/\$F\$5)

Go to cell H3 and type

=SUM((B3-\$B\$5)^2/\$B\$5,(C3-\$C\$5)^2/\$C\$5,(D3-\$D\$5)^2/\$D\$5,(E3-\$E\$5)^2/\$E\$5,(F3-\$F\$5)^2/\$F\$5)

To get the table value go to cell H7 and type =CHIINV(0.05,4)

Go to cell O13 and type =IF(T8>=T11," H0 is Accepted", "H0 is Rejected")

OUTPUT:

F14								
	A	B	C	D	E	F	G	H
1		O	A	B	C	D	Total	$\sum (O_i - E_i)^2 / E_i$
2	Girls	11	7	5	5	11	39	6.074863267
3	Boys	30	4	3	10	14	61	6.074863267
4	Total	41	11	8	15	25	100	12.14972653
5	Ei	20.5	5.5	4	7.5	12.5	50	
6								
7								Critical Value of $\alpha=0.05$ for $df=(2-1) * (5-1)$
8								9.487729037
9					Decesion	H0 is Accepted		

Practical No 7

Aim: Perform testing of hypothesis using One-way ANOVA, using Excel.

ANOVA Assumptions

- The dependent variable (SAT scores in our example) should be continuous.
- The independent variables (districts in our example) should be two or more categorical groups.
- There must be different participants in each group with no participant being in more than one group. In our case, each school cannot be in more than one district.
- The dependent variable should be approximately normally distributed for each category.
- Variances of each group are approximately equal. From our data exploration, we can see that the average SAT scores are quite different for each district.

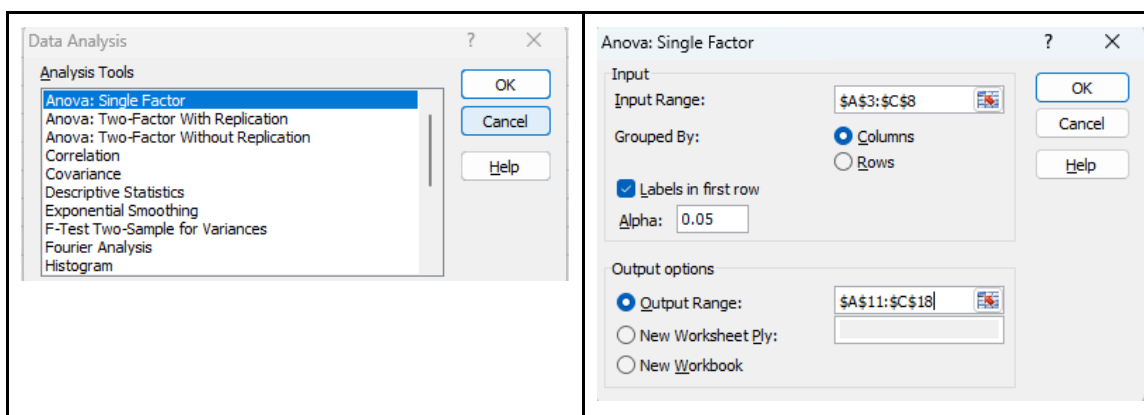
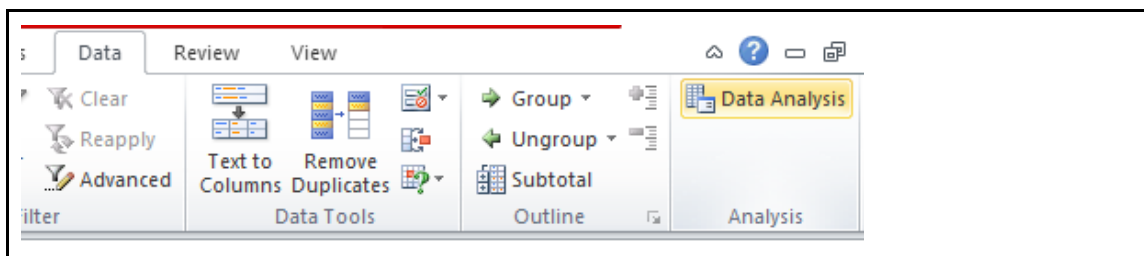
Since we have five different groups, we cannot use the t-test, use the 1-way ANOVA test anyway just to understand the concepts. If there is at least one group with a significant difference with another group, the null hypothesis will be rejected.

H0 - There are no significant differences between the Subject's mean SAT scores. $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

H1 - There is a significant difference between the Subject's mean SAT scores.

- **To perform ANOVA go to data >> Data Analysis**

	A	B	C	
1				
2	Group A	Group B	Group c	
3		5	5	7
4		6	9	5
5		7	4	8
6		8	7	6
7		9	8	9
8		5	6	4
9				



Input Range : \$A\$3:\$C\$8 (Select columns to be analyzed in group)

Output Range : \$A\$11:\$C\$18 (Can be any Range)

Output:

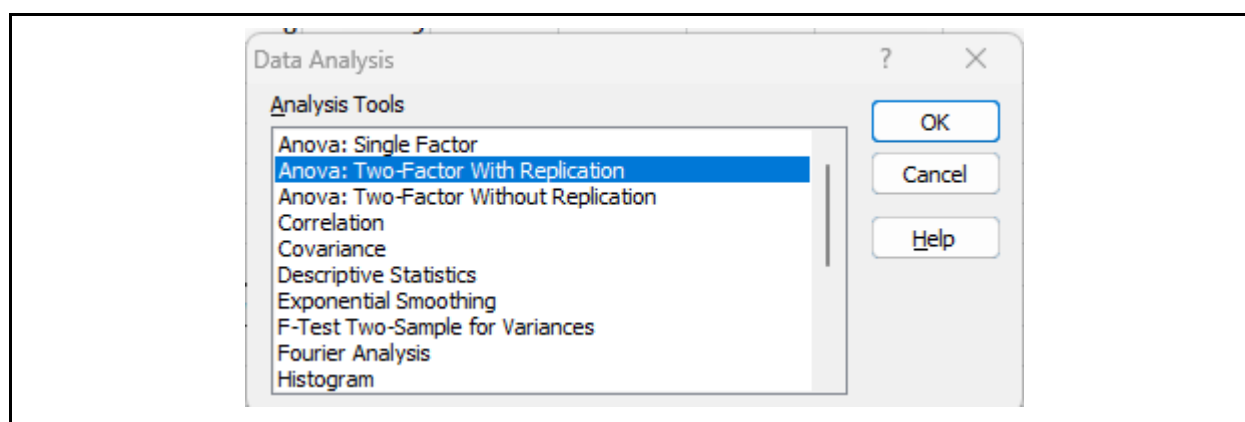
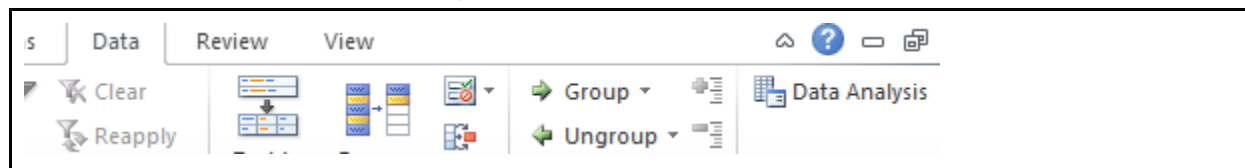
11	Anova: Single Factor						
12							
13	SUMMARY						
14	<i>Groups</i>	<i>Count</i>	<i>Sum</i>	<i>Average</i>	<i>Variance</i>		
15	5	5	35	7	2.5		
16	5	5	34	6.8	3.7		
17	7	5	32	6.4	4.3		
18							
19							
20	ANOVA						
21	<i>Source of Variation</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P-value</i>	<i>F crit</i>
22	Between Groups	0.933333	2	0.466667	0.133333	0.876452	3.885294
23	Within Groups	42	12	3.5			
24							
25	Total	42.93333	14				
26							

Since the resulting p-value is less than 0.05. The null hypothesis (H0) is rejected and conclude that there is a significant difference between the SAT scores for each subject.

Practical No 8

Aim: Perform testing of hypothesis using Two-way ANOVA, using Excel

Go to Data tab >> Data Analysis



Output:

Anova: Two-Factor With Replication			
SUMMARY	Treatment	Value	Total
A			
Count	4	4	8
Sum	5	26.5	31.5
Average	1.25	6.625	3.9375
Variance	0.25	0.695833333	8.659821429
B			
Count	4	4	8
Sum	6	27.6	33.6
Average	1.5	6.9	4.2
Variance	0.333333333	1.346666667	9.051428571

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Sample	0.475833333	2	0.237916667	0.4568	0.640449286	3.554557146
Columns	166.4266667	1	166.4266667	319.5392	6.6297E-13	4.413873419
Interaction	0.175833333	2	0.087916667	0.1688	0.845999467	3.554557146
Within	9.375	18	0.520833333			
Total	176.4533333	23				

P-value = 0.0588079 column in the ANOVA Source of Variation table at the bottom of the output. Because the p-values for both medicine dose and interaction are less than our significance level, these factors are statistically significant. On the other hand, the interaction effect is not significant because its p-value (0.0588) is greater than our significance level. Because the interaction effect is not significant, we can focus on only the main effects and not consider the interaction effect of the dose.

Practical No 9

Aim: Perform testing of hypothesis using MANOVA, using Excel.

MANOVA is the acronym for Multivariate Analysis of Variance. When analyzing data, we may encounter situations where we have there multiple response variables (dependent variables). In MANOVA there also some assumptions, like ANOVA. Before performing MANOVA we have to check the following assumptions are satisfied or not.

- The samples, while drawing, should be independent of each other.
- The dependent variables are continuous in nature and the independent variables are categorical.
- The dependent variables should follow a multivariate normal distribution.
- The population variance-covariance matrices of each group are same, i.e. groups are homogeneous.

Go to <http://www.real-statistics.com/free-download/>

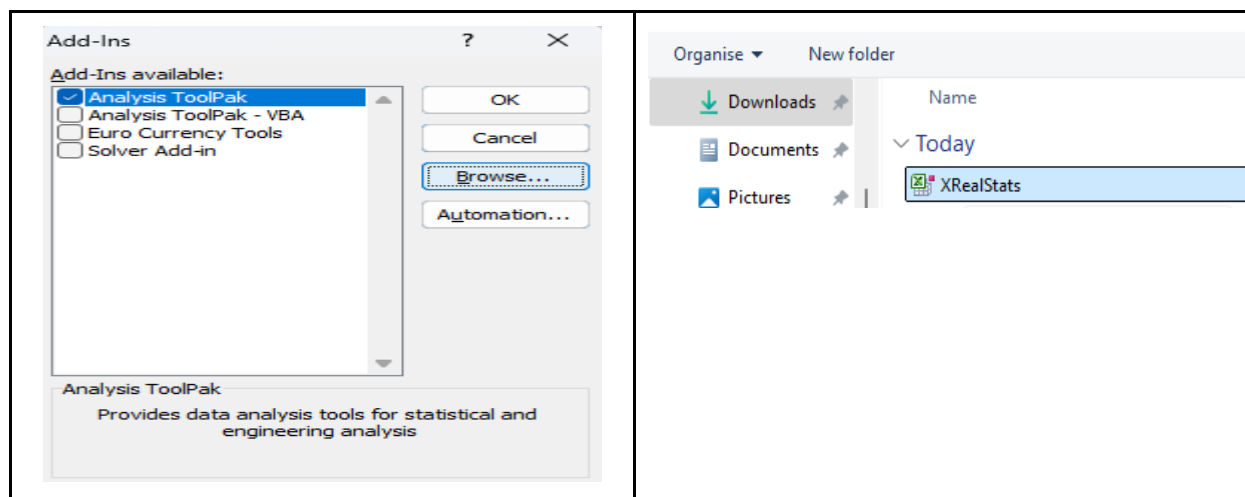
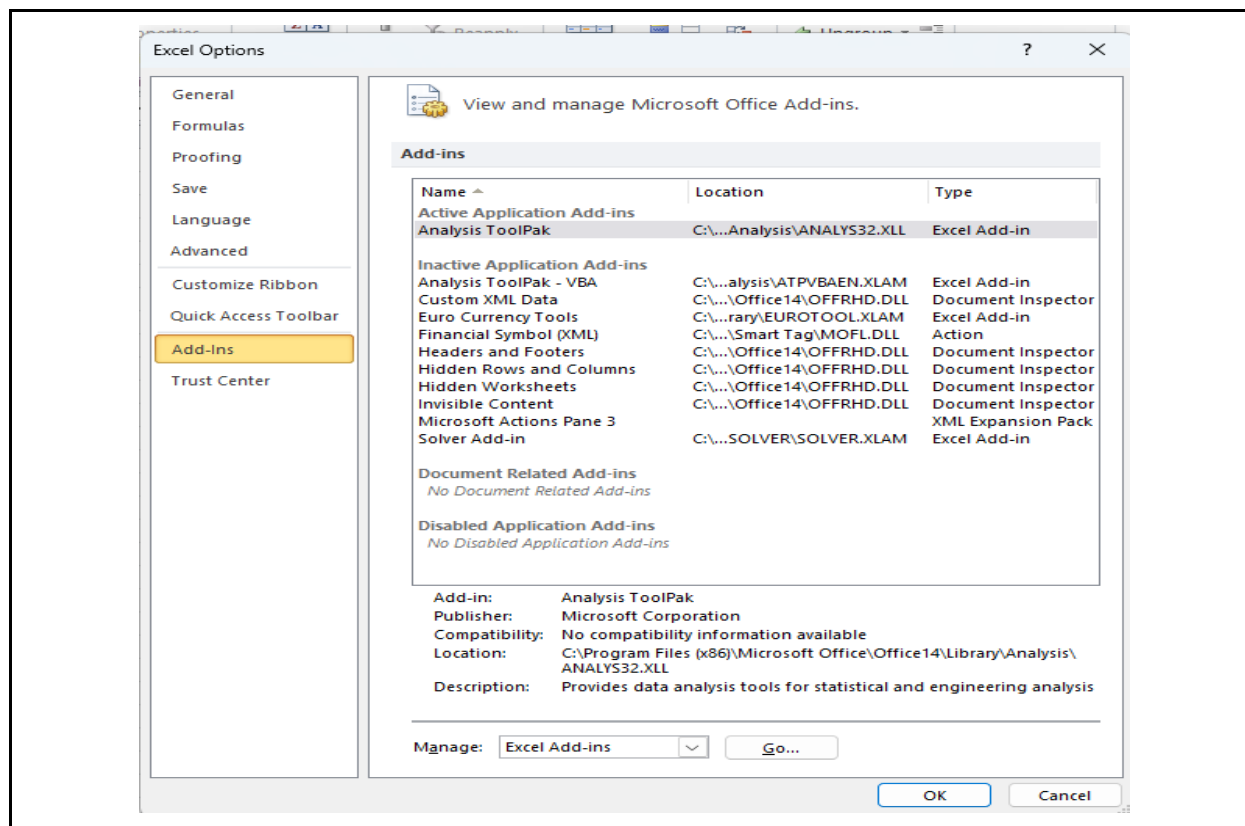
1. Download Real Statistics Resource Pack

Real Statistics Resource Pack: contains a variety of supplemental functions and data analysis tools not provided by Excel. These complement the standard Excel capabilities and make it easier for you to perform the statistical analyses described in the rest of this website.

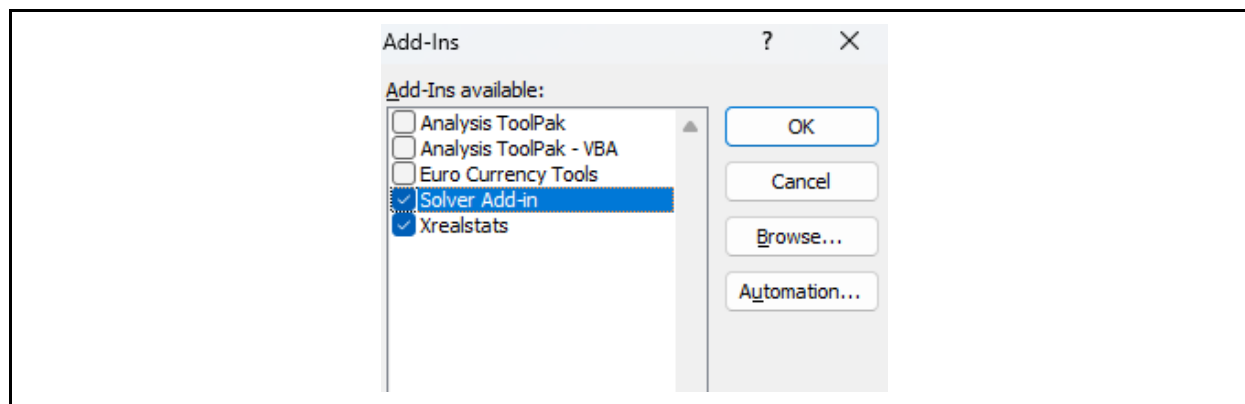
Free Download
Real Statistics Resource Pack

- **Version 1:** If you accept the **License Agreement**, click here on **Real Statistics Resource Pack for Excel 2010/2013/2016/2019/2021/365 with help** to download the latest Excel for Windows version of the resource pack that accompanies this website (Release 9.2.1). This version includes worksheet function help support and is named **XRealStats.xlam**.

Install Add-in in excel. Select File > Help|Options > Add-Ins and click on the Go button at the bottom of the window (see Figure 1). Add-ins -> Analysis Pack -> Go



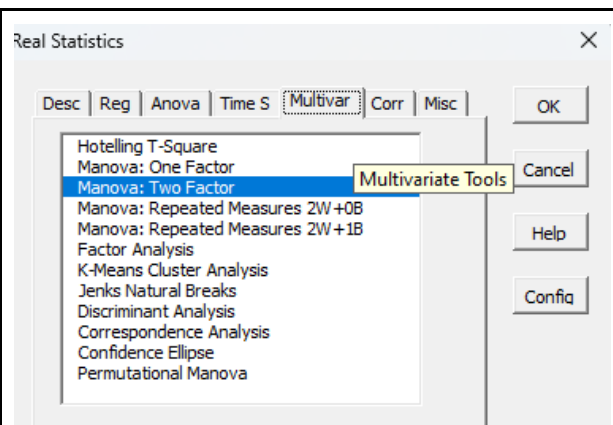
Select the following Add-Ins. Click OK.



Now create an excel sheet with following data.

A study was conducted to see the impact of social-economic class (rich, middle, poor) and gender (male, female) on kindness and optimism using on a sample of 24 people based on the data in Figure 1.

	A	B	C	D	E
1	Gender	Economic	kindnes	optimism	
2	male	wealthy	5	4	
3	male	wealthy	6	5	
4	male	middle	7	7	
5	male	poor	8	5	
6	male	poor	8	9	
7	female	wealthy	6	4	
8	female	wealthy	8	6	
9	female	middle	5	7	
10	female	poor	3	5	
11	female	poor	7	6	
12					



Press ctrl-m to open Real Statistics menu.

Select the data excluding column names. Select a cell for output.

Manova: Two Factors

Input Range: Sheet1!\$A\$4:\$D\$11

Analysis type:
☒ Regular ☐ Repeated Measures

Options:
☒ Significance Analysis
☒ Sum of Squares and Cross Product Matrices
☒ Covariance Matrices
☒ Outliers ☒ Box's Test
☒ Group Means ☐ Contrast

Alpha: 0.05

Output Range: Sheet1!\$A\$41:\$X\$53

Output:

Two-Way MANOVA							SSCP Matrices	
fact A	stat	df1	df2	F	p-value	part eta-sq	Tot	
Pillai Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	21.71429	6
Wilk's Lambda	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	6	16
Hotelling's Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!		
Roy's Largest Root	#VALUE!						Row (A)	
							6.914286	4.4
fact B	stat	df1	df2	F	p-value	part eta-sq	4.4	2.8
Pillai Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!		
Wilk's Lambda	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	Column (B)	
Hotelling's Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	2.714286	-2.5
Roy's Largest Root	#VALUE!						-2.5	3.25
fact AB	stat	df1	df2	F	p-value	part eta-sq	Interaction (AB)	
Pillai Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Wilk's Lambda	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!
Hotelling's Trace	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!	#VALUE!		
Roy's Largest Root	#VALUE!						Res	
							#VALUE!	#VALUE!
							#VALUE!	#VALUE!

Practical No 10

Aim: Perform the Random sampling for the given data and analyze it, using Excel.

Example 1: From a population of 10 women and 10 men as given in the table in Figure 1 on the left below, create a random sample of 6 people for Group 1 and a periodic sample consisting of every 3rd woman for Group 2.

You need to run the sampling data analysis tool twice, once to create Group 1 and again to create Group 2. For Group 1 you select all 20 population cells as the Input Range and Random as the Sampling Method with 6 for the Random Number of Samples. For Group 2 you select the 10 cells in the Women column as Input Range and Periodic with Period 3

Open existing excel sheet with population data Sample Sheet looks as given below:

	A	B	C	D	E	F	G	H	I	J	K
1	Sr.no	Roll No	Student'sName	Gender	Grade		Sr.no	Roll No	Student'sName	Gender	Grade
2	1	1	Aarav Sharma	Male	A		62	7	Anika Patel	Female	C
3	2	2	Anika Patel	Male	B		63	9	Rohan Mehta	Female	A
4	3	5	Rohan Mehta	Male	A		64	11	Sneha Singh	Female	B
5	4	13	Sneha Singh	Male	B		65	14	Aditi Verma	Female	A
6	5	16	Aditi Verma	Male	C		66	25	Manish Kumar	Female	B
7	6	17	Manish Kumar	Male	A		67	36	Isha Shah	Female	C
8	7	34	Isha Shah	Male	B		68	40	Neha Gupta	Female	A
9	8	35	Neha Gupta	Male	C		69	41	Pooja Reddy	Female	B
10	9	38	Pooja Reddy	Male	A		70	46	Rajesh Pillai	Female	C
11	10	42	Rajesh Pillai	Male	B		71	47	Karan Desai	Female	A
12	11	43	Karan Desai	Male	C		72	10	Priya Nair	Female	B
13	12	45	Priya Nair	Male	A		73	20	Vikram Rao	Female	C
14	13	48	Vikram Rao	Male	B		74	21	Divya Agarwal	Female	A
15	14	49	Divya Agarwal	Male	A		75	72	Ravi Joshi	Female	D
16	15	50	Ravi Joshi	Male	B		76	73	Simran Kaur	Female	D
17	16	51	Simran Kaur	Male	C		77	75	Arjun Das	Female	D
18	17	54	Arjun Das	Male	A		78	77	Sonal Tiwari	Female	D
19	18	55	Sonal Tiwari	Male	B		79	82	Rahul Chawla	Female	D
20	19	56	Rahul Chawla	Male	C		80	84	Sonal Tiwari	Female	D
21	20	59	Sonal Tiwari	Male	A		81	91	Rahul Nikam	Female	D
22	21	61	Rahul Nikam	Male	B		82	95	Aarav Sharma	Female	D
23	22	62	Aarav Sharma	Male	C		83	4	Anika Patel	Female	C
24	23	65	Anika Patel	Male	A		84	87	Aarav Sharma	Female	O

Set Cell O1 = Male and Cell O2 = Female

To generate a random sample for male students from given population go to Cell O1 and type

=INDEX(E\$2:E\$62,RANK(B2,B\$2:B\$62))

Drag the formula to the desired no of cell to select random sample.

Now, to generate a random sample for female students go to cell P1 and type

=INDEX(K\$2:K\$40,RANK(H2,H\$2:H\$40))

Drag the formula to the desired no of cell to select random sample.

Output:

O	P
Male	Female
A	C
C	D
B	D
A	D
C	D
B	D
A	A
C	C
B	B
A	A
B	C
A	B
C	A
B	C
A	B
C	A
B	C
A	B
C	A
B	A
A	C
B	A