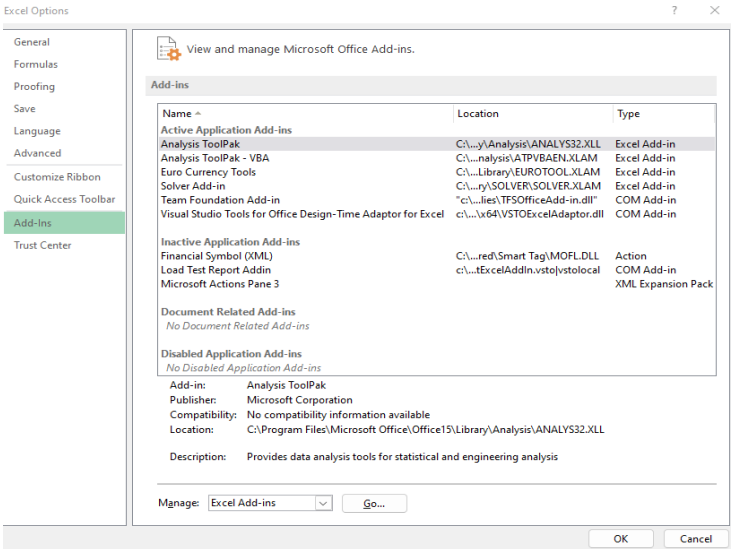
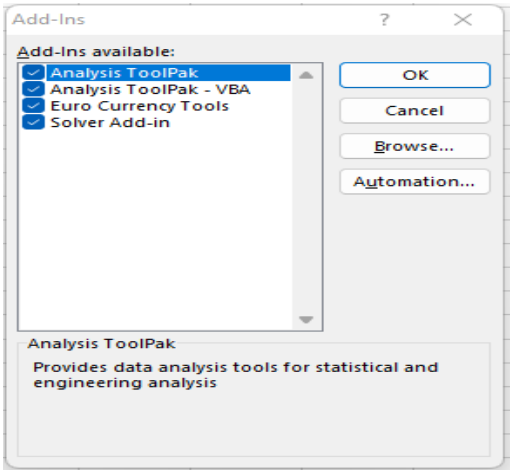
RESEARCH IN COMPUTING

PRACTICAL-1

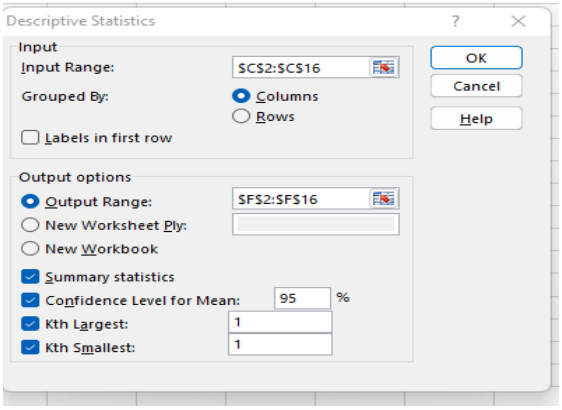
Aim: Write a program for obtaining descriptive statistics of data.

Step1: Program/Steps to obtain descriptive statistics of data. Using Excel Go to File Menu -> Options ->Add-Ins -> Select Analysis Tool Pak -> Press OK

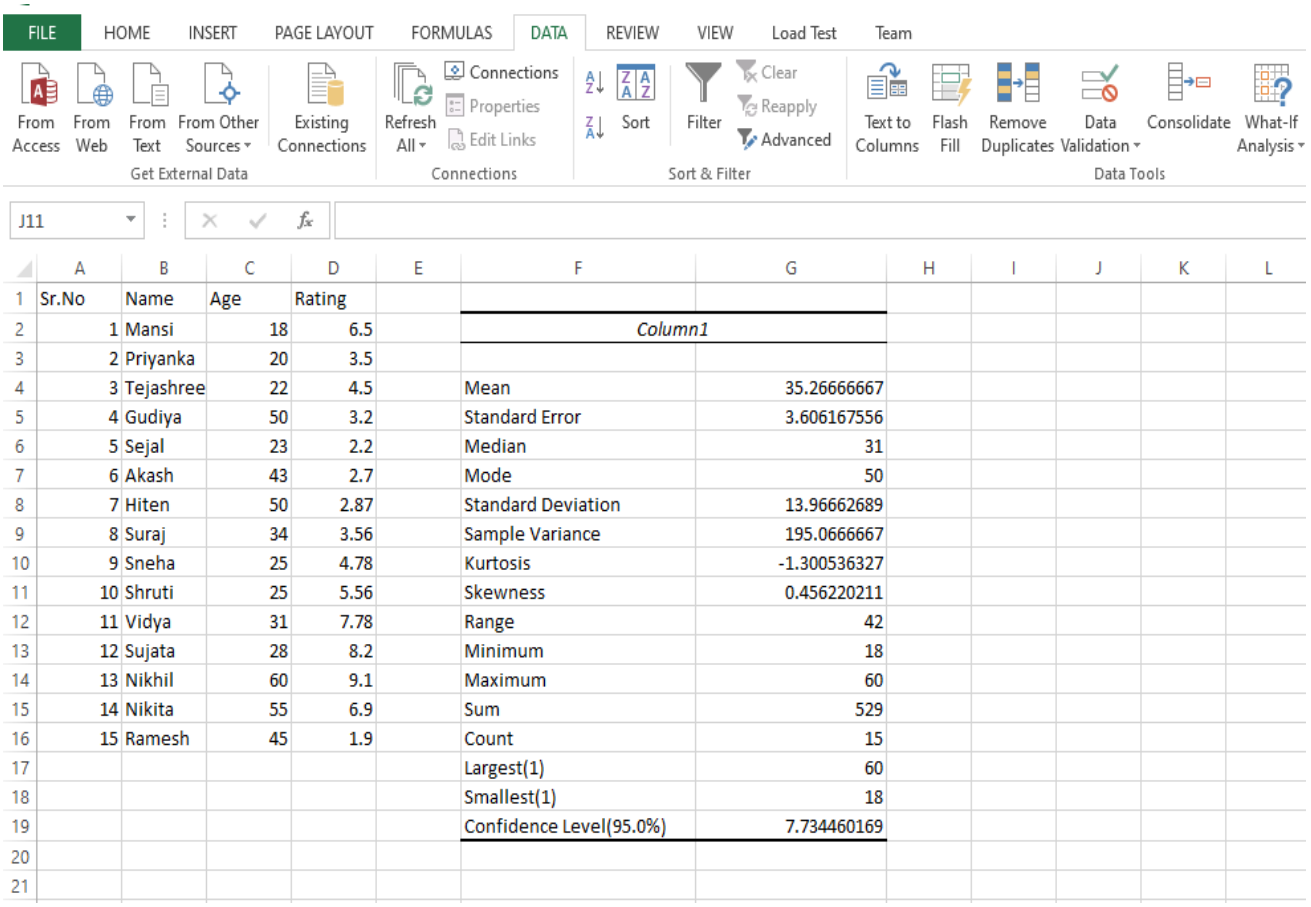




Step2: Go to Data Menu -> Data Analysis -> Descriptive Statistics ->and Select the Age Range



Output



	A	B	C	D	E	F	G	H	I	J	K	L
1	Sr.No	Name	Age	Rating								
2	1	Mansi	18	6.5		*Column1*						
3	2	Priyanka	20	3.5								
4	3	Tejashree	22	4.5		Mean	35.26666667					
5	4	Gudiya	50	3.2		Standard Error	3.606167556					
6	5	Sejal	23	2.2		Median	31					
7	6	Akash	43	2.7		Mode	50					
8	7	Hiten	50	2.87		Standard Deviation	13.96662689					
9	8	Suraj	34	3.56		Sample Variance	195.0666667					
10	9	Sneha	25	4.78		Kurtosis	-1.300536327					
11	10	Shruti	25	5.56		Skewness	0.456220211					
12	11	Vidya	31	7.78		Range	42					
13	12	Sujata	28	8.2		Minimum	18					
14	13	Nikhil	60	9.1		Maximum	60					
15	14	Nikita	55	6.9		Sum	529					
16	15	Ramesh	45	1.9		Count	15					
17						Largest(1)	60					
18						Smallest(1)	18					
19						Confidence Level(95.0%)	7.734460169					
20												
21												

PRACTICAL -2

Aim: Import data from different data sources (from Excel, csv, mysql, sql server, oracle to R/Python/Excel)

- **READ DATA FROM MYSQL**

#Install Python-MySQL connector(Download from Resource folder

Click Start> MySQL CLI> Login > Create database > create table > Insert records in table

```
mysql> Create database mydb1;
Query OK, 1 row affected (0.00 sec)

mysql> use mydb1;
Database changed
mysql> Create table test(name varchar(10), eid  int);
Query OK, 0 rows affected (0.02 sec)

mysql> insert into test values('Hiten','1');
Query OK, 1 row affected (0.00 sec)

mysql> insert into test values('Sky','2');
Query OK, 1 row affected (0.02 sec)

mysql> insert into test values('Abhishek','3');
Query OK, 1 row affected (0.00 sec)

mysql> insert into test values('Chiku','4');
Query OK, 1 row affected (0.00 sec)

mysql> insert into test values('Rustom','5');
Query OK, 1 row affected (0.00 sec)

mysql>
```

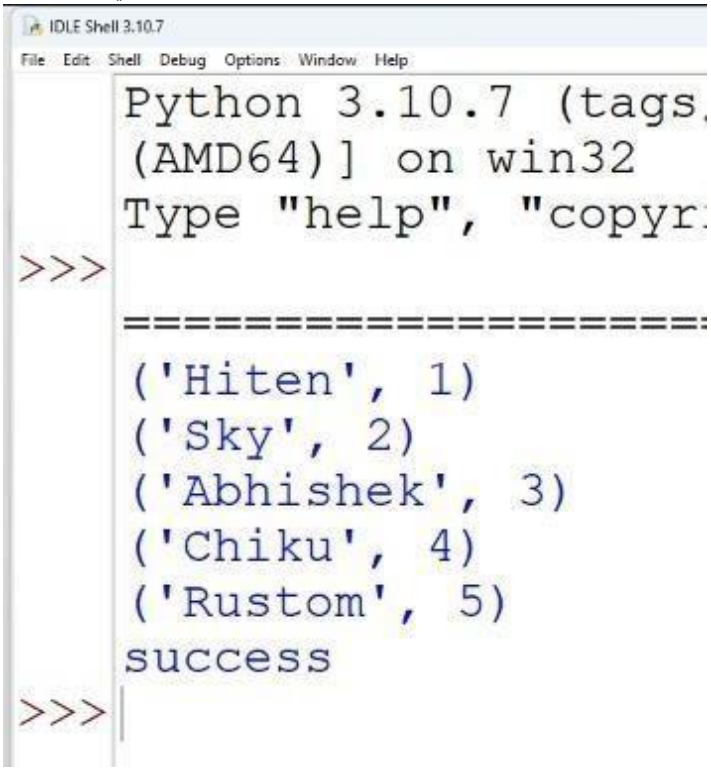
#To install pymysql package run pip command on cmd

#pip3 install pymysql

CODE:

```
import pymysql
db = pymysql.connect(host='localhost',user='root',password='root',database='mydb1')
cursor = db.cursor()
sql = """SELECT * from test"""
try:
    cursor.execute(sql)
    results = cursor.fetchall()
    for row in results:
        print(row)
```

```
    print("success")
    db.commit()
except:
    db.rollback()
db.close()
```



PRACTICAL-3

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

Using the collected data for analysis

	A	B	C	D	E	F	G	H	I
1	Timestamp	Username	Name	Phone number	1.Age	2.Gender	3.Subject	4.Teaching hours per week	1. Including this year, how long have you using ICT?
2	2022/11/24 7:07:34 PM GMT+5:30	siddhibhosale08@gmail.com	Siddhi Vinayak Bhosale	8208228407	21-30	Female	Science	Fewer than 10hours	1-3 years
3	2022/11/24 7:09:48 PM GMT+5:30	pawarprerana9702@gmail.com	PRERANA TANAJI PAWAR	8108876282	21-30	Female	Technology	10-15 hours	3-6 years
4	2022/11/24 7:21:26 PM GMT+5:30	snehapasi2001@gmail.com	Pasi sneha ramkevel	9372622367	21-30	Female	Science	Fewer than 10hours	1-3 years
5	2022/11/24 9:41:07 PM GMT+5:30	pratikarolkar2.model@gmail.com	Pratik Arolkar	9.1808E+11	21-30	Male	Technology	Fewer than 10hours	1-3 years
6	2022/11/24 10:00:30 PM GMT+5:30	sandipdhangar1997@gmail.com	Sandip Dhangar	7776969640	21-30	Male	Technology	Fewer than 10hours	1-3 years
7	2022/11/24 10:01:59 PM GMT+5:30	jayshreenaidu67@gmail.com	Jayshree	9930092706	40-50	Female	Science	Fewer than 10hours	3-6 years
8	2022/11/25 7:51:07 AM GMT+5:30	jaydeeppagar37617@gmail.com	Jaydeep Pagar	8369747784	21-30	Male	Science	Fewer than 10hours	1-3 years
9	2022/11/25 10:38:10 AM GMT+5:30	sharvarigavhane19@gmail.com	Sharvari Gavhane	9820188487	21-30	Female	Technology	Fewer than 10hours	1-3 years
10	2022/11/25 11:17:06 AM GMT+5:30	anujfivezero@gmail.com	Anuj Gandhat	7977936516	21-30	Male	Science	Fewer than 10hours	1-3 years
11	2022/11/25 11:25:00 AM GMT+5:30	shardulikaa@gmail.com	Shardulika Agrahari	9137519993	21-30	Female	Other	Fewer than 10hours	1-3 years
12	2022/11/25 11:38:06 AM GMT+5:30	priyankakondhavle01@gmail.com	Priyanka Kondhavle	9152698805	21-30	Female	Science	Fewer than 10hours	1-3 years
13	2022/11/25 11:54:23 AM GMT+5:30	durguderasika10@gmail.com	Rasika	865241852	21-30	Female	Other	Fewer than 10hours	1-3 years
14	2022/11/25 12:55:53 PM GMT+5:30	aakashachale20@gmail.com	Aakash Achale	7666453334	21-30	Male	Technology	Fewer than 10hours	1-3 years
15	2022/11/25 3:15:47 PM GMT+5:30	yashu020@icloud.com	Yashu	9136288011	21-30	Male	Science	16-20 hours	3-6 years
16	2022/11/25 4:22:16 PM GMT+5:30	qfaizal755@gmail.com	Faizal	9702982558	21-30	Male	Other	Fewer than 10hours	1-3 years
17	2022/11/25 4:31:56 PM GMT+5:30	hraot07@gmail.com	Hiten	7385890876	21-30	Male	Science	Fewer than 10hours	1-3 years
18	2022/11/25 8:39:16 PM GMT+5:30	nbalajinadar@gmail.com	Balaji Nadar	9820945498	21-30	Male	Other	Fewer than 10hours	1-3 years
19	2022/11/25 9:34:56 PM GMT+5:30	pallanaidu01@gmail.com	Eshika Naidu	9819356960	30-40	Female	Mathematics	10-15 hours	3-6 years
20	2022/11/25 9:39:42 PM GMT+5:30	naiduhemakshi1111@gmail.com	Hemakshi M Naidu	7875754291	40-50	Female	Mathematics	10-15 hours	6-10 years
21	2022/11/26 3:06:08 PM GMT+5:30	malitejashri46@gmail.com	Tejashri prakash mali	7304350488	21-30	Female	Technology	Fewer than 10hours	1-3 years
22	2022/11/27 10:19:20 AM GMT+5:30	abhishekdas22222@gmail.com	Abhishek Das	8878219561	21-30	Male	Technology	10-15 hours	1-3 years
23	2022/11/27 10:30:28 AM GMT+5:30	cvishakha018@gmail.com	Vishakha Vijay Chavan	9011653459	21-30	Female	Technology	Fewer than 10hours	1-3 years
24	2022/11/27 10:31:13 AM GMT+5:30	pranitawaingankar20@gmail.com	Shruti Waingankar	8369521630	21-30	Female	Science	Fewer than 10hours	1-3 years
25	2022/11/27 10:36:41 AM GMT+5:30	surajgupta37@gmail.com	Suraj Gupta	8108899101	21-30	Male	Science	Fewer than 10hours	1-3 years
26	2022/11/27 10:42:18 AM GMT+5:30	santoshyadav2397@gmail.com	Santosh Yadav	9005063716	21-30	Male	Technology	Fewer than 10hours	1-3 years
27									

J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA
2. How often do you use computers or the internet in your classes?	1. What a	2. Does yo	1. Is partic	2. Who pr	3. Have yo	1. Which c	2. What a	1. What o	1. To what extent are you confident using ICT?								
All of the time	Desktop c	Yes	Yes	A more ex	Introducto	Electronic	Communi	Insufficier	Use a spreadsheet to plot a graph								
All of the time	Digital car	Yes	Yes	A more ex	Advanced	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Downloa								
All of the time	Computer	Yes	Yes	A more ex	Introducto	Existing or	Create yo	Insufficier	Create a presentation with video or audio clips movies;Capture and edit digital photos, movies or o								
Sometimes	Digital car	No	Yes	A more ex	Advanced	Existing or	Use ICT to	Insufficier	Create a presentation with video or audio clips movies;Capture and edit digital photos, movies or o								
Sometimes	Digital car	Yes	Yes	A more ex	Introducto	Material c	Post hom	Insufficier	Use a spreadsheet to plot a graph								
Sometimes	Computer	No	Yes	A more ex	Introducto	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Create or								
Sometimes	Digital car	No	No	An online	Other pro	Existing or	Communi	Most of th	Create a presentation with video or audio clips movies;Capture and edit digital photos, movies or o								
All of the time	Computer	Yes	Yes	A more ex	Advanced	Existing or	Use ICT to	Insufficier	Create a presentation with video or audio clips movies								
All of the time	Desktop c	Yes	Yes	A more ex	Subject-sp	Existing or	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
All of the time	Desktop c	Yes	Yes	School ICT	Introducto	Material c	Create yo	Using ICT	Create a presentation with video or audio clips movies								
Sometimes	Desktop c	Yes	Yes	School ICT	Introducto	Existing or	Use ICT to	Insufficier	Create a presentation with video or audio clips movies;Download and install software on a system								
All of the time	Digital car	Yes	Yes	A more ex	Introducto	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
Rarely	Computer	Yes	Yes	A more ex	Introducto	Material c	Use ICT to	Insufficier	Create a presentation with video or audio clips movies;Capture and edit digital photos, movies or o								
All of the time	Computer	No	Yes	A more ex	Advanced	Existing or	Browse/se	Most of th	Use a spreadsheet to plot a graph								
Never	Digital car	No	No	Expert fro	Other pro	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Capture and edit digital photos, movies or other images								
Sometimes	Computer	No	Yes	School ICT	Introducto	Existing or	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
Rarely	Computer	No	No	A more ex	Course on	Material c	Create yo	Insufficier	Create a presentation with video or audio clips movies;Capture and edit digital photos, movies or o								
All of the time	Digital car	Yes	Yes	School ICT	Advanced	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
Sometimes	Computer	Yes	Yes	A more ex	Subject-sp	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Download and install software on a system								
All of the time	Desktop c	Yes	No	A more ex	Introducto	Existing or	Browse/se	Insufficier	Use a spreadsheet to plot a graph								
All of the time	Desktop c	Yes	No	A more ex	Introducto	Existing or	Create yo	Most of th	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
Sometimes	Computer	Yes	Yes	School ICT	Introducto	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Create a presentation with video or audio clips movies;Capture a								
Rarely	Computer	Yes	Yes	School ICT	Introducto	Material c	Communi	Insufficier	Use a spreadsheet to plot a graph;Capture and edit digital photos, movies or other images								
Sometimes	Computer	No	No	School ICT	Introducto	Existing or	Communi	Insufficient number of system;Using ICT in teaching and learning not being a goal in your school									
Rarely	Computer	Yes	Yes	A more ex	Advanced	Material c	Post hom	Most of th	Create a presentation with video or audio clips movies								

	Question1	Question2	Question3	Question4	Question5	Question6	Question7	Question8	Question9	Question10	Question11	Question12
COUNT(N)	25	25	25	25	25	25	25	25	25	25	25	24
Not Answer	0	0	0	0	0	0	0	0	0	0	0	0
TOTAL:	25	25	25	25	25	25	25	25	25	25	25	24
DISAGREE:	8	6										
Agree:	17	19										
TOTAL:	25	25										
PERCENTAGE:												
NO(%): 133%												
YES(%): 89%												
TOTAL: 223%												

FORMULA FOR OPERATION:

count(N) = =COUNTA(H2:H26)

Not Answer = =COUNT(H2:H26)

Total =SUM(E29:E30)

YES =COUNTIF(L$2:L$26,$M24)

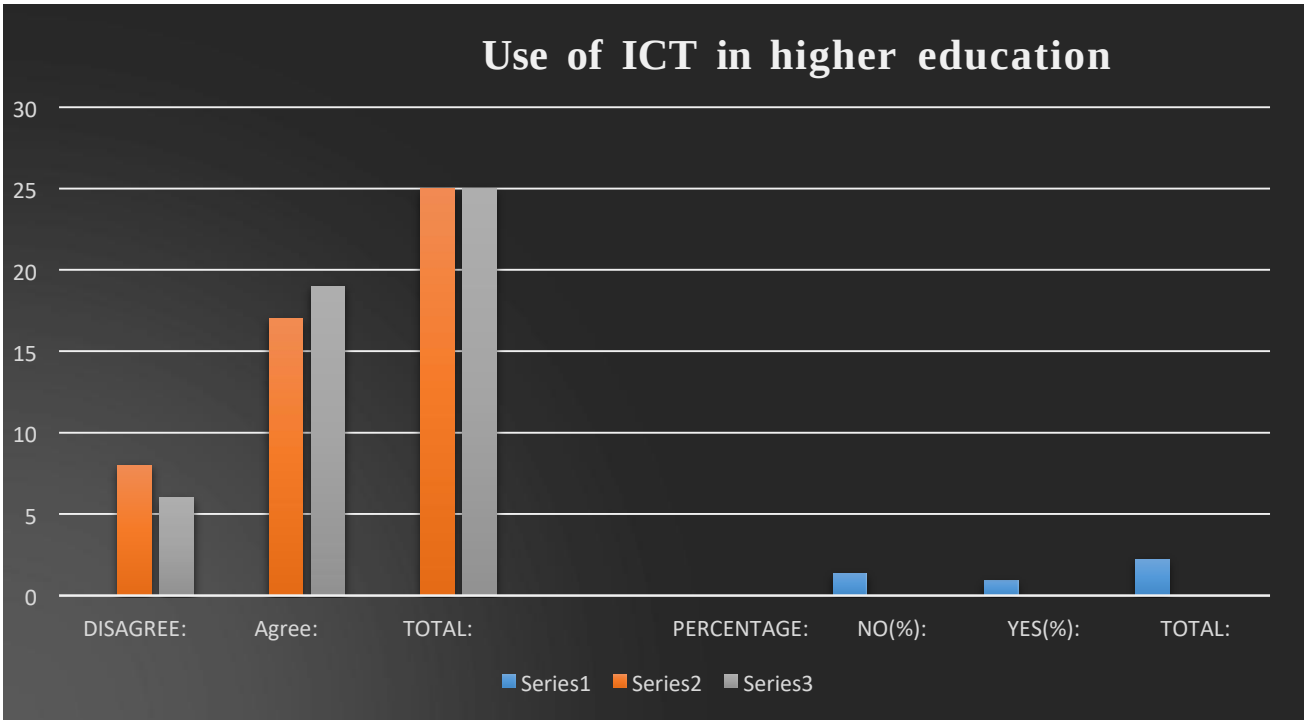
NO= COUNTIF(L$2:L$26,$M25)

PERCENTAGE:

YES= E34/F34

NO= E33/F33

TOTAL=SUM(D38:D39)



OUTPUT:

PRATICAL 4

Aim: Perform analysis of given secondary data.

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.

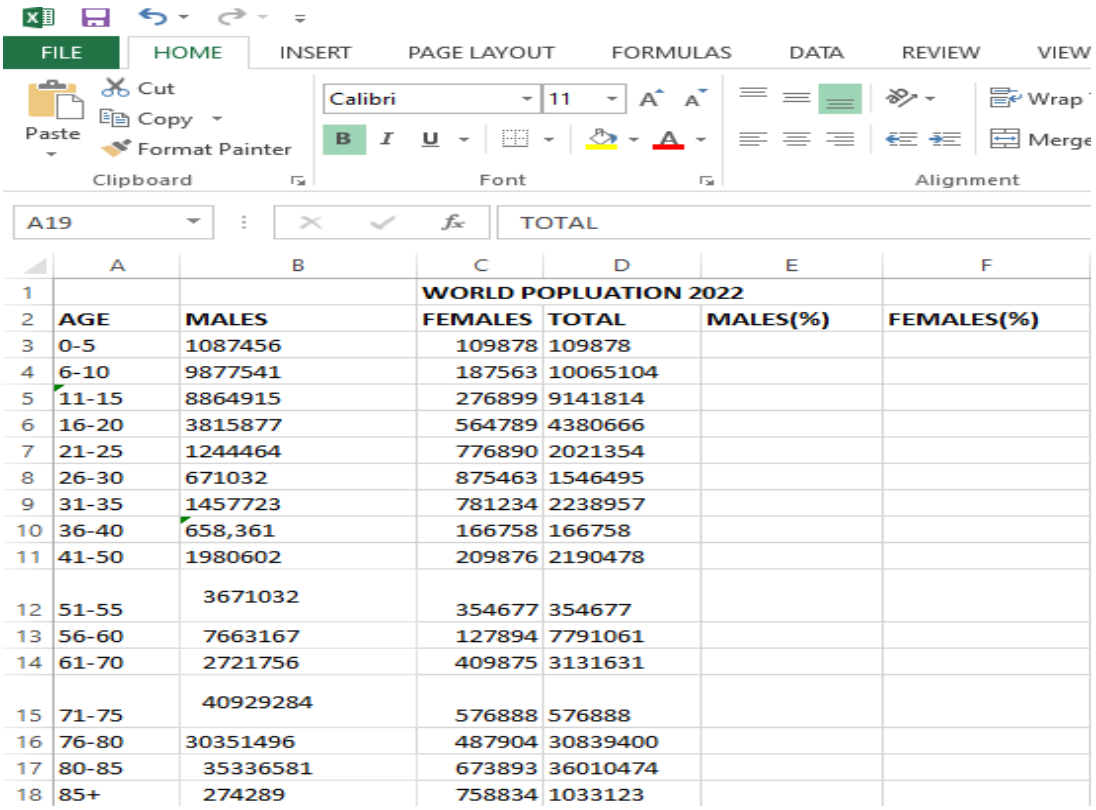
3. **Evaluating relevance of the data** – Considering things like the data's original purpose, when it was collected, population, sampling strategy/sample, data collection protocols, operationalization of concepts, questions asked, and form/shape of the data.

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: Analyse the given Population Census Data for Planning and Decision Making by using the size and composition of populations

Example: Analyse the given Population Census Data for Planning and Decision Making by using the size and composition of populations

Output:-



	A	B	C	D	E	F
1			WORLD POPLUATION 2022			
2	AGE	MALES	FEMALES	TOTAL	MALES(%)	FEMALES(%)
3	0-5	1087456	109878	109878		
4	6-10	9877541	187563	10065104		
5	11-15	8864915	276899	9141814		
6	16-20	3815877	564789	4380666		
7	21-25	1244464	776890	2021354		
8	26-30	671032	875463	1546495		
9	31-35	1457723	781234	2238957		
10	36-40	658,361	166758	166758		
11	41-50	1980602	209876	2190478		
12	51-55	3671032	354677	354677		
13	56-60	7663167	127894	7791061		
14	61-70	2721756	409875	3131631		
15	71-75	40929284	576888	576888		
16	76-80	30351496	487904	30839400		
17	80-85	35336581	673893	36010474		
18	85+	274289	758834	1033123		

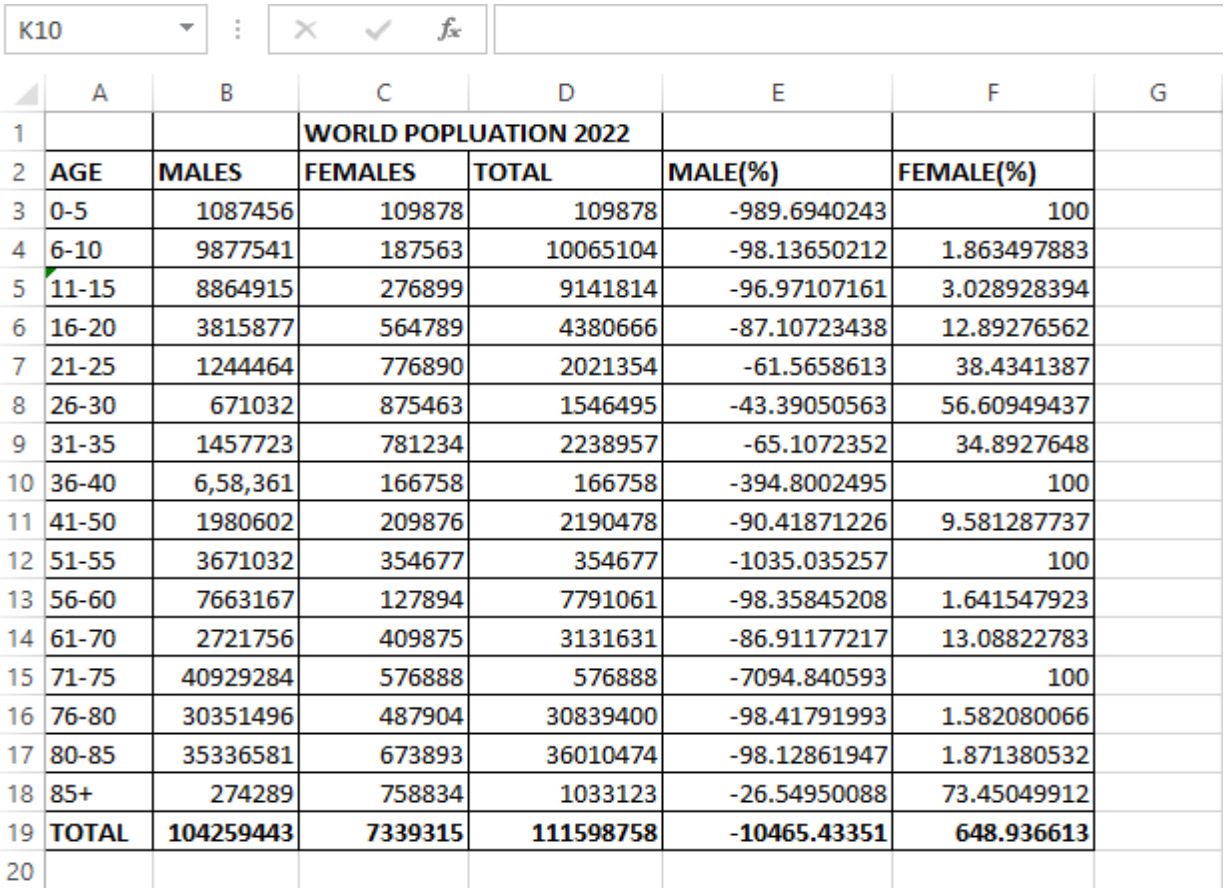
Put the cursor in cell **B19** and click on the **AutoSum** and then click **Enter**. This will calculate the total population. Then copy the formula in cell **D19** across the row **19. (Total_population)**
To calculate the percent of males in cell **E3**, enter the formula

[-1*100*Male_count*Total_population]
=-1*100*B3/$D$19

And copy the formula in cell **E3** down to cell **E18.**
To calculate the percent of females in cell **F3**, enter the formula

[1*100*Female_count*Total_population]
=**100*C3/$D$19**. Copy the formula in cell **F3** down to cell **F18.**
This gives percentage in +ve for female and –ve for male.

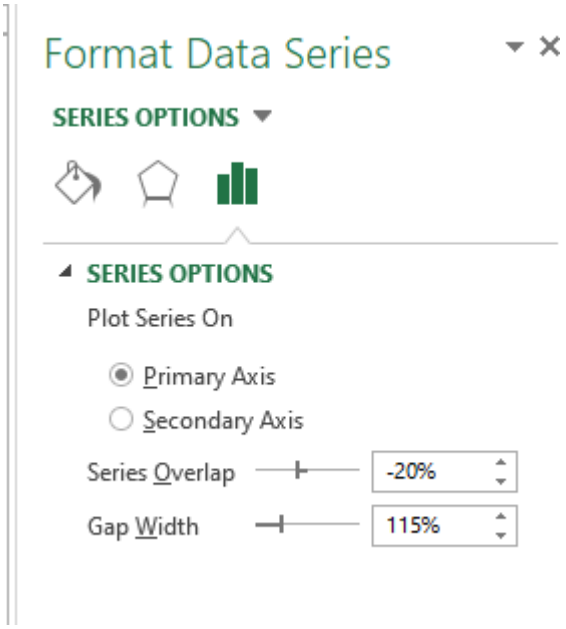
To build the population pyramid, we need to choose a horizontal bar chart with two series of data (% male and % female) and the age labels in column A as the **Category X-axis** labels. Highlight the range **A2:A18**, hold down the CTRL key and highlight the range **E2:F19**

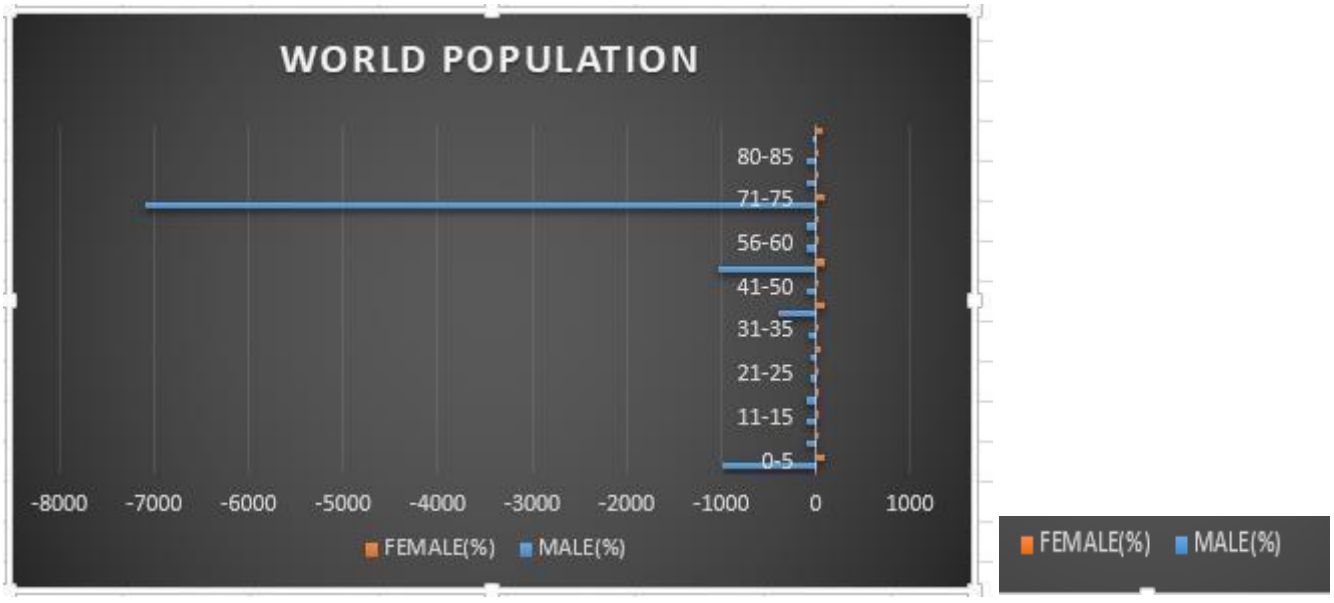


	A	B	C	D	E	F	G
1			WORLD POPLUATION 2022				
2	AGE	MALES	FEMALES	TOTAL	MALE(%)	FEMALE(%)	
3	0-5	1087456	109878	109878	-989.6940243	100	
4	6-10	9877541	187563	10065104	-98.13650212	1.863497883	
5	11-15	8864915	276899	9141814	-96.97107161	3.028928394	
6	16-20	3815877	564789	4380666	-87.10723438	12.89276562	
7	21-25	1244464	776890	2021354	-61.5658613	38.4341387	
8	26-30	671032	875463	1546495	-43.39050563	56.60949437	
9	31-35	1457723	781234	2238957	-65.1072352	34.8927648	
10	36-40	6,58,361	166758	166758	-394.8002495	100	
11	41-50	1980602	209876	2190478	-90.41871226	9.581287737	
12	51-55	3671032	354677	354677	-1035.035257	100	
13	56-60	7663167	127894	7791061	-98.35845208	1.641547923	
14	61-70	2721756	409875	3131631	-86.91177217	13.08822783	
15	71-75	40929284	576888	576888	-7094.840593	100	
16	76-80	30351496	487904	30839400	-98.41791993	1.582080066	
17	80-85	35336581	673893	36010474	-98.12861947	1.871380532	
18	85+	274289	758834	1033123	-26.54950088	73.45049912	
19	TOTAL	104259443	7339315	111598758	-10465.43351	648.936613	
20							

Under **inset** tab, under horizontal bar charts select **clustered bar chart.**

Put the tip of your mouse arrow on the **Y-axis** (vertical axis) so it says “Category Axis”, right click and chose **Format Axis**
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**.





PRATICAL 5

Aim: Perform testing of hypothesis using one sample t-test.

One sample t-test : The One Sample t Test determines whether the sample mean is statistically different from a known or hypothesised population mean. The One Sample t Test is a parametric test.

H0: Mean age of given sample is 30.

H1: Mean age of given sample is not 30

#pip3 install scipy

#pip3 install numpy

```
from scipy.stats import ttest_1samp
import numpy as np
ages = np.genfromtxt('ages.csv')
print(ages)
ages_mean = np.mean(ages)
print("Mean age:",ages_mean)
print("Test 1: m=30")
tset, pval = ttest_1samp(ages, 30)
 print('p-values - ',pval)
if pval< 0.05:
        print("we reject null hypothesis")
else:
        print("we fail to reject null hypothesis"
```

OUTPUT:

```
==================== RESTART: C:\Mansi\RIC PRACTICAL 5.PY ====================
[20. 30. 25. 13. 16. 17. 34. 35. 38. 43. 45. 48. 49. 50. 51. 54. 55. 56.
 59. 61. 62. 18. 22. 29.]
Mean age: 38.75
Test 1: m=30
p-values -  0.01333239479255858
we reject null hypothesis
```

#TEST 2

H0: Mean age of given sample is 38.
H1: Mean age of given sample is not

```
38. from scipy.stats import ttest_1samp
import numpy as np
ages = np.genfromtxt('ages.csv')
print(ages)
ages_mean = np.mean(ages)
print("Mean age:",ages_mean)
print("Test 2: m=38")
tset, pval = ttest_1samp(ages, 38)
print('p-values - ',pval)
if pval< 0.05:
        print("we reject null hypothesis")
else:
        print("we fail to reject null hypothesis")
```

OUTPUT

```
====================== RESTART: C:\Mansi\RIC PRACTICAL 5.PY ======================
[20. 30. 25. 13. 16. 17. 34. 35. 38. 43. 45. 48. 49. 50. 51. 54. 55. 56.
 59. 61. 62. 18. 22. 29.]
Mean age: 38.75
Test 1: m=38
p-values -  0.8202593087020069
we fail to reject null hypothesis
```

Practical 6

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

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 as having 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	Experiment	Comaparison	
2	35	2	
3	40	27	
4	12	38	
5	15	31	
6	21	1	
7	14	19	
8	46	1	
9	10	34	
10	28	3	
11	48	1	
12	16	2	
13	30	3	
14	32	2	
15	48	1	
16	31	2	
17	22	1	
18	12	3	
19	39	29	
20	19	37	
21	25	2	
22	27.15	11.95	Mean
23	12.50799744	14.24245414	SD

Experimental Data

To calculate Standard Mean go to cell A22 and type =SUM(A2:A21)/20

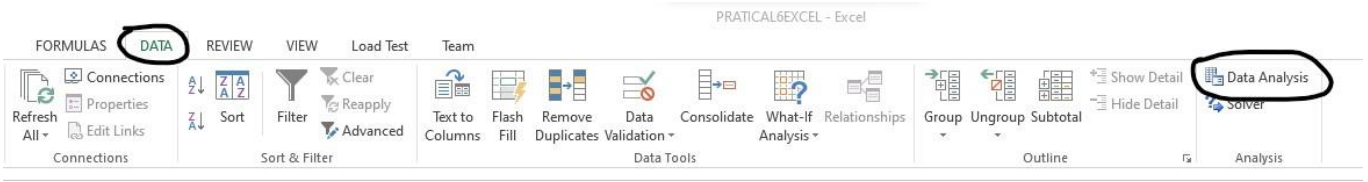
To calculate Standard Deviation go to cell A23 and type =STDEV(A2:A21)

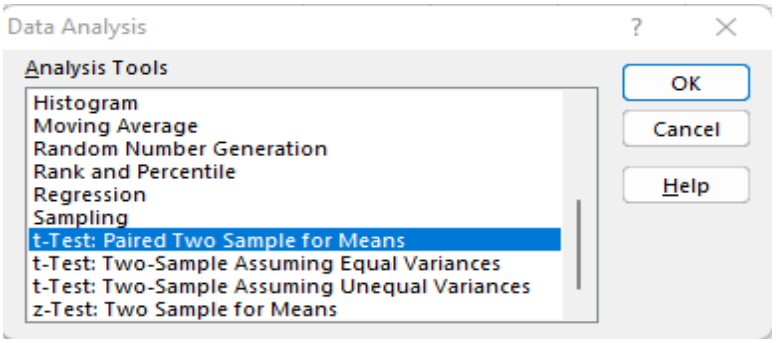
Comparison Data

To calculate Standard Mean go to cell B22 and type =SUM(B2:B21)/20

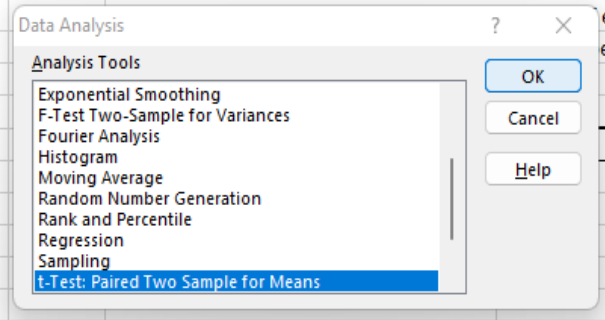
To calculate Standard Deviation go to cell B23 and type =STDEV(B2:B21)

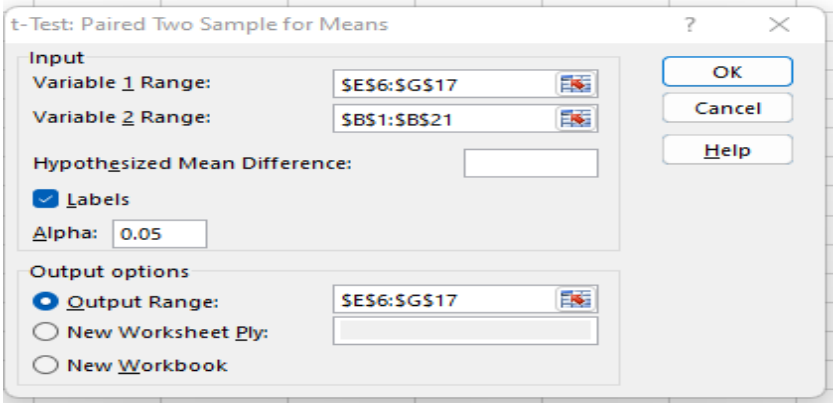
To find T-Test Statistics go to data Data Analysis





	A	B	C	D	E	F
1	Experiment	Comaparison				
2	35	2				
3	40	27				
4	12	38				
5	15	31				
6	21	1				
7	14	19				
8	46	1				
9	10	34				
10	28	3				
11	48	1				
12	16	2				
13	30	3				
14	32	2				
15	48	1				
16	31	2				
17	22	1				
18	12	3				
19	39	29				
20	19	37				
21	25	2				





To caluculate the T-Test square value go to cell E20 and type
=(A22-B22)/SQRT((A23*A23)/COUNT(A2:A21)+(B23*B23)/COUNT(A2:A21))
Now go to cell E20 and type
=IF(E20<E12,"H0 is Accepted", "H0 is Rejected and H1 is Accepted")
Our calculated value is larger than the tabled value at alpha = .01, so we reject the null hypothesisand accept the alternative hypothesis, namely, that the difference in gain scores is likely the resultof the experimental treatment and not the result of chance variation.

	Experiment	*Comaparison*
Mean	27.15	11.95
Variance	156.45	213.5236842
Observations	20	20
Pearson Correlation	-0.395904927	
Hypothesized Mean Difference	0	
df	19	
t Stat	2.996289153	
P(T<=t) one-tail	0.003711226	
t Critical one-tail	1.729132812	
P(T<=t) two-tail	0.007422452	
t Critical two-tail	2.093024054	

OUTPUT:

	A	B	C	D	E	F	G
1	Experiment	Comaparison					
2	35	2			H0 -Difference in gain score is not likely the result of experiment		
3	40	27			H1-Difference in gain score is likely the result of experimental		
4	12	38			t-Test: Paired Two Sample for Means		
5	15	31			t-Test: Paired Two Sample for Means		
6	21	1				*Experiment*	*Comaparison*
7	14	19			Mean	27.15	11.95
8	46	1			Variance	156.45	213.5236842
9	10	34			Observations	20	20
10	28	3			Pearson Correlation	-0.395904927	
11	48	1			Hypothesized Mean Difference	0	
12	16	2			df	19	
13	30	3			t Stat	2.996289153	
14	32	2			P(T<=t) one-tail	0.003711226	
15	48	1			t Critical one-tail	1.729132812	
16	31	2			P(T<=t) two-tail	0.007422452	
17	22	1			t Critical two-tail	2.093024054	
18	12	3					
19	39	29					
20	19	37			3.586175078	Calculated Value	
21	25	2					
22	27.15	11.95	Mean			H0 is Rejected and H1 is accepted	
23	12.50799744	14.24245414	SD				
24							

Practical 7

Aim: Perform testing of hypothesis using paired t-test.

The paired sample t-test is also called dependent sample t-test. It's an univariate test that tests for a significant difference between 2 related variables. An example of this is if you where to collect the blood pressure for an individual before and after some treatment, condition, or time point. The data set contains blood pressure readings before and after an intervention. These are variables "bp_before" and "bp_after".

The hypothesis being test is:
- H0 - The mean difference between sample 1 and sample 2 is equal to 0.
- **H1** - The mean difference between sample 1 and sample 2 is not equal to 0

```
from scipy import stats
import matplotlib.pyplot as plt
import pandas as pd
df = pd.read_csv("blood_pressure.csv")
print(df[['bp_before','bp_after']].describe())
tst,pval=stats.ttest_rel(df['bp_before'], df['bp_after'])
if pval< 0.05:
        print("we reject null hypothesis")
else:
        print("we fail to reject null hypothesis")
```

OUTPUT:

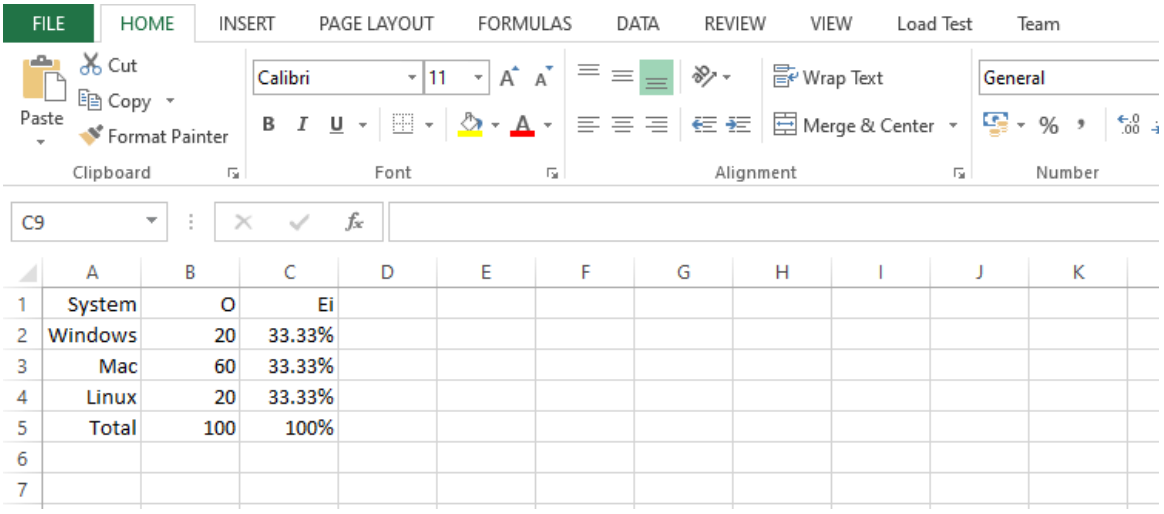
```
========================= RESTART: C:\Mansi
        bp_before    bp_after
count    5.000000    5.000000
mean   156.000000  137.800000
std     33.615473   13.198485
min    120.000000  120.000000
25%    130.000000  129.000000
50%    150.000000  140.000000
75%    180.000000  150.000000
max    200.000000  150.000000
we fail to reject null hypothesis
>>>
```

Practical 8

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

Problem in system administrator needs to upgrade the computers for 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.

Step 1: Creation of the data.



	A	B	C	D	E	F	G	H	I	J	K
1	System	O	Ei								
2	Windows	20	33.33%								
3	Mac	60	33.33%								
4	Linux	20	33.33%								
5	Total	100	100%								
6											
7											

Step 2: Setting the value of H0 and HA

H0: The population distribution of the variable is the same as the proposed distribution

HA: The distributions are different to calculate the Chi – Square value for Windows go to cell D2 and type

Step: 3

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

Step: 4

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

Step: 5

To calculate the Chi – Square value for Mac go to cell D3 and type

= ((B4-C4)*(B4-C4))/C4

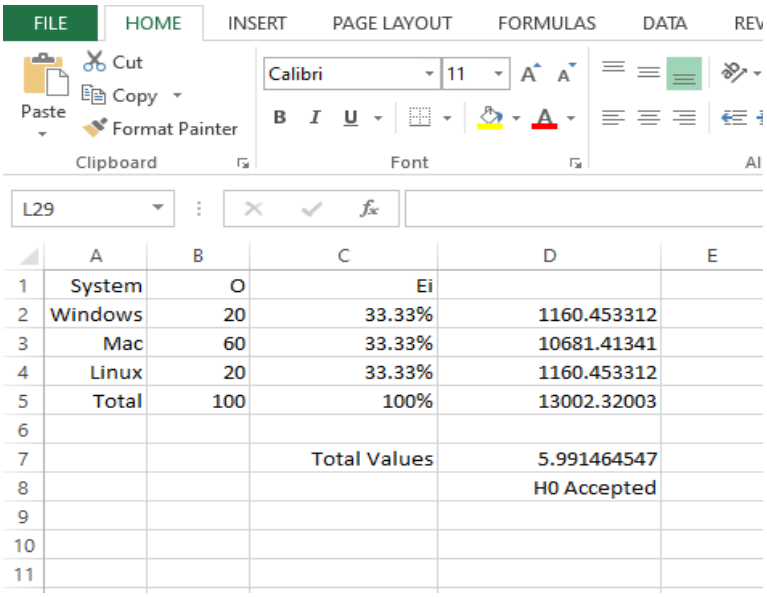
Step: 6

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,

Step: 7

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
1	System	O	Ei		
2	Windows	20	33.33%	1160.453312	
3	Mac	60	33.33%	10681.41341	
4	Linux	20	33.33%	1160.453312	
5	Total	100	100%	13002.32003	
6					
7			Total Values	5.991464547	
8				H0 Accepted	
9					
10					
11					