ACADEMIC TASK-2 MTH-302

(PROBABILITY AND STATISTICS)

Lovely Professional University, Punjab Phagwara

Name**: Shivam Modi**

Registration Number: **12304496**

Roll Number: **49**

Section: **K23TB**

Submitted to**: Dr. Rakesh Yadav**

BTECH

Name Of School: SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

Question 1.

Draw the histogram using the following data set and find bins.

Note -Ensure no two students use the same dataset; violations will result in UMC.

First data set (X)-

$$X = Random\ number\left(\frac{Last\ two\ digist\ of\ your\ registration\ number}{15} + Your\ roll\ no\right) + 15$$

$$Total\ no\ of\ observations = 11^2 + Last\ digit\ o\ your\ roll\ number$$

$$Y = Random\ number\left(\frac{Last\ three\ digist\ of\ your\ registration\ number}{24} + Your\ roll\ no\right) + 16$$

$$Total\ no\ of\ observations = 12^2 + Sum\ of\ Last\ two\ digits\ of\ your\ roll\ numbe$$

Tasks:

1. Visualize the data on the histogram
2. Write the Used Excel commands to find bins
3. Attach the screen print with dialogue box showing the commands.

Ans:

X = Random number ((96/15) + 49) + 15
X = Random number (6.4+ 49) + 15

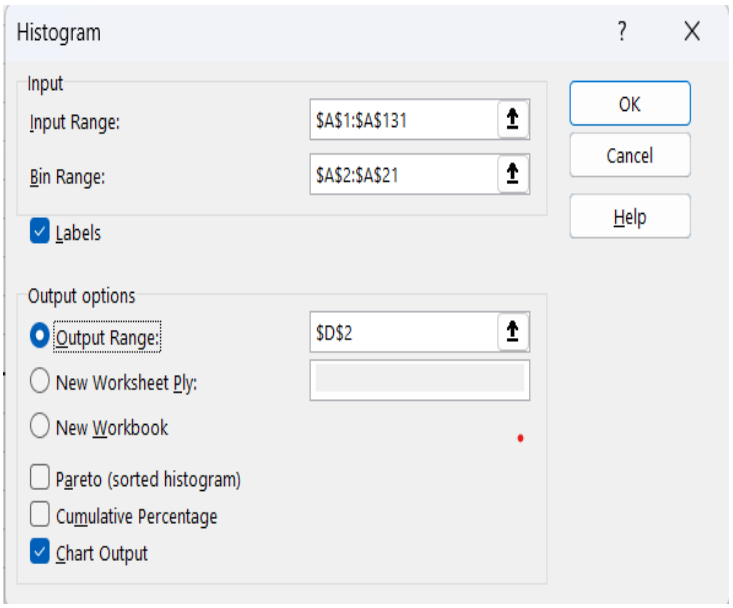
X => RAND()*55.4+15

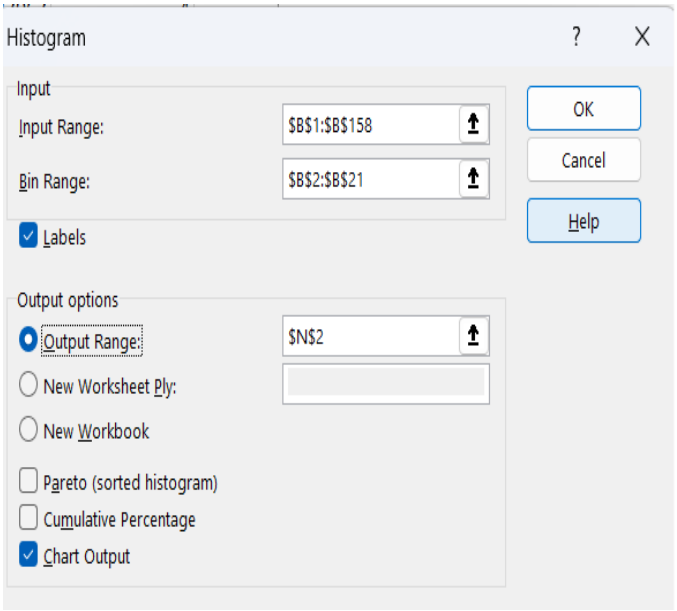
Total Observation = 11^2 + 9 = 130

Y = Random ((496/24) + 49) + 16
Y = Random (20.67+ 49) + 16

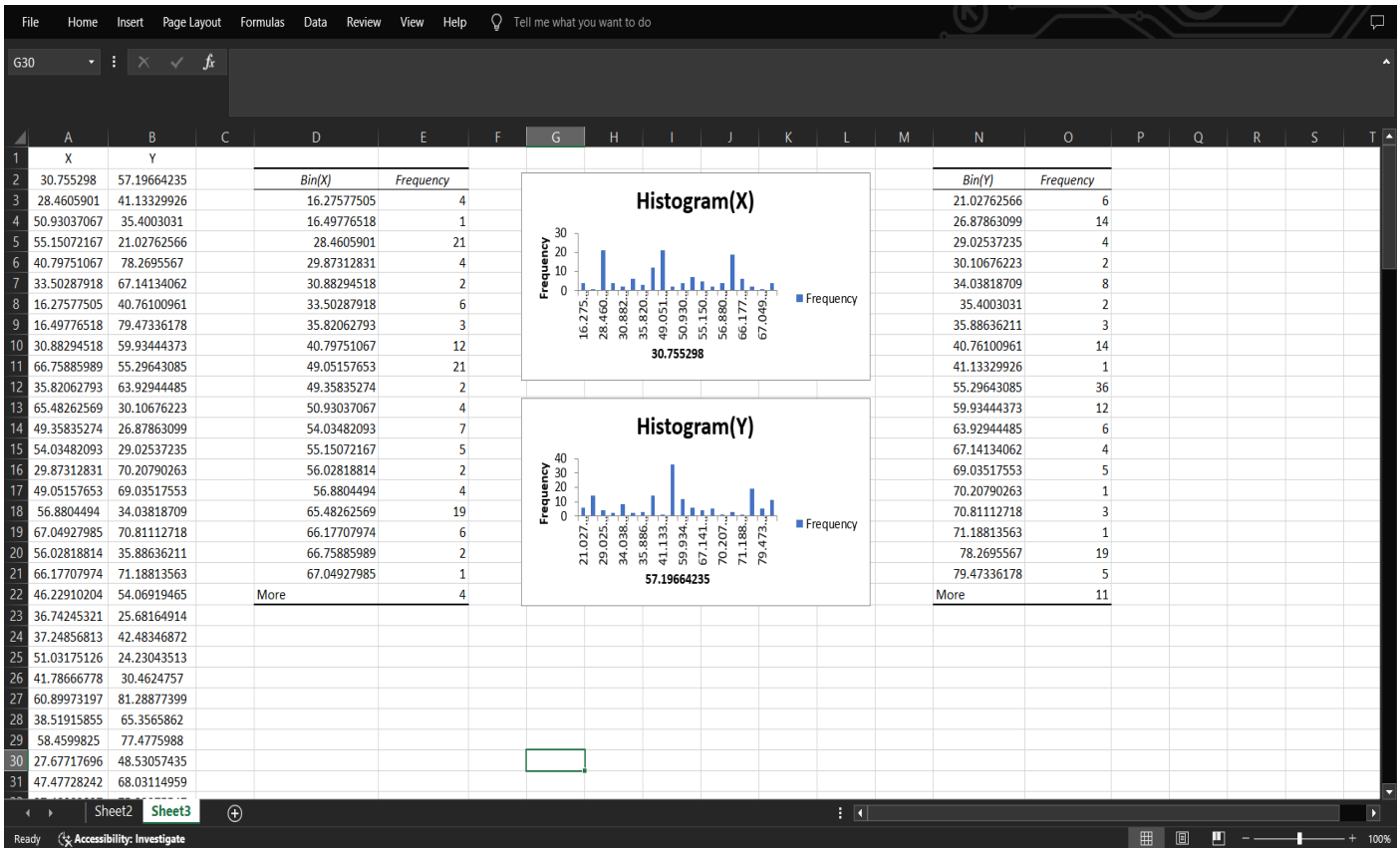
Y => RAND() * 69.67 + 16

Total Observation = 12^2 + 13 = 157





To find freq -> =FREQUENCY(data_array, bins_array)



	A	B	D	E	N	O
1	X	Y				
2	30.755298	57.19664235	*Bin(X)*	*Frequency*	*Bin(Y)*	*Frequency*
3	28.4605901	41.13329926	16.27577505	4	21.02762566	6
4	50.93037067	35.4003031	16.49776518	1	26.87863099	14
5	55.15072167	21.02762566	28.4605901	21	29.02537235	4
6	40.79751067	78.2695567	29.87312831	4	30.10676223	2
7	33.50287918	67.14134062	30.88294518	2	34.03818709	8
8	16.27577505	40.76100961	33.50287918	6	35.4003031	2
9	16.49776518	79.47336178	35.82062793	3	35.88636211	3
10	30.88294518	59.93444373	40.79751067	12	40.76100961	14
11	66.75885989	55.29643085	49.05157653	21	41.13329926	1
12	35.82062793	63.92944485	49.35835274	2	55.29643085	36
13	65.48262569	30.10676223	50.93037067	4	59.93444373	12
14	49.35835274	26.87863099	54.03482093	7	63.92944485	6
15	54.03482093	29.02537235	55.15072167	5	67.14134062	4
16	29.87312831	70.20790263	56.02818814	2	69.03517553	5
17	49.05157653	69.03517553	56.8804494	4	70.20790263	1
18	56.8804494	34.03818709	65.48262569	19	70.81112718	3
19	67.04927985	70.81112718	66.17707974	6	71.18813563	1
20	56.02818814	35.88636211	66.75885989	2	78.2695567	19
21	66.17707974	71.18813563	67.04927985	1	79.47336178	5
22	46.22910204	54.06919465	More	4	More	11
23	36.74245321	25.68164914				
24	37.24856813	42.48346872				
25	51.03175126	24.23043513				
26	41.78666778	30.4624757				
27	60.89973197	81.28877399				
28	38.51915855	65.3565862				
29	58.4599825	77.4775988				
30	27.67717696	48.53057435				
31	47.47728242	68.03114959				

Question 2.

Find Descriptive Statistics of following Dataset

Analyze central tendencies (mean, median, mode) and measures of dispersion (variance, standard deviation).

Note -Ensure no two students use the same dataset; violations will result in UMC.

First data set (X)-

$$X = Random\ number\left(\frac{Last\ two\ digist\ of\ your\ registration\ number}{15} + Your\ roll\ no\right) + 10$$

$$Total\ no\ of\ observations = 11^2 + Last\ digit\ o\ your\ roll\ number$$

$$Y = Random\ number\left(\frac{Last\ three\ digist\ of\ your\ registration\ number}{20} + Your\ roll\ no\right) + 13$$

$$Total\ no\ of\ observations = 14^2 + Sum\ of\ \ Last\ two\ digits\ of\ your\ roll\ numbe$$

Tasks:

4. Calculate mean, median, and mode using both conventional and automated processes.

5. Comment on the central tendency analysis.

Use Excel commands to find maximum, minimum, sum, and count.

X = Random number ((96/15) + 49) + 10
X = Random number (65.4)

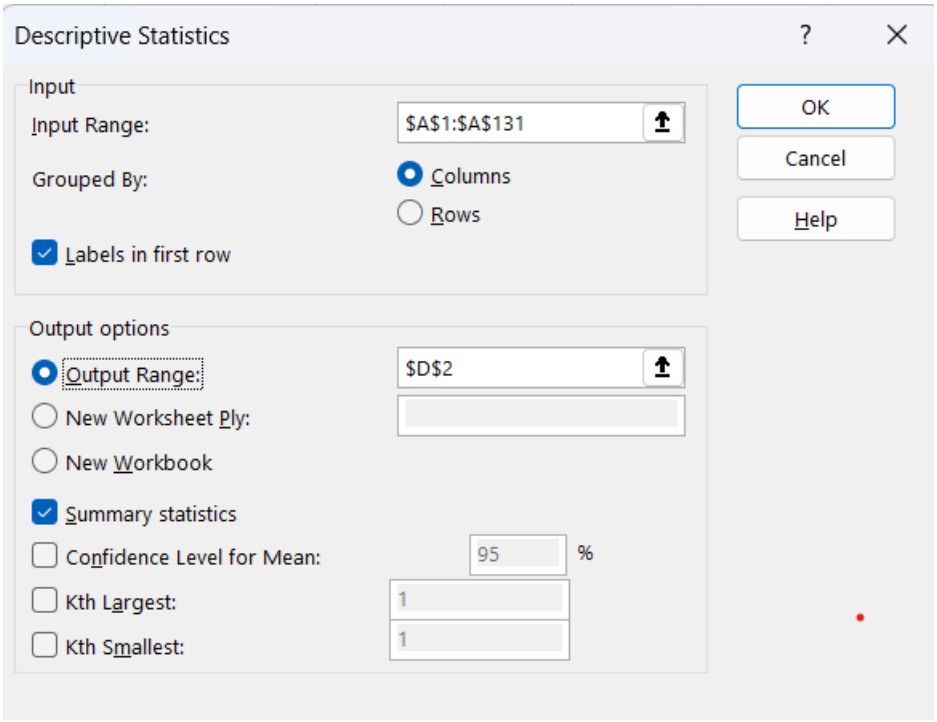
X => 0,65.4

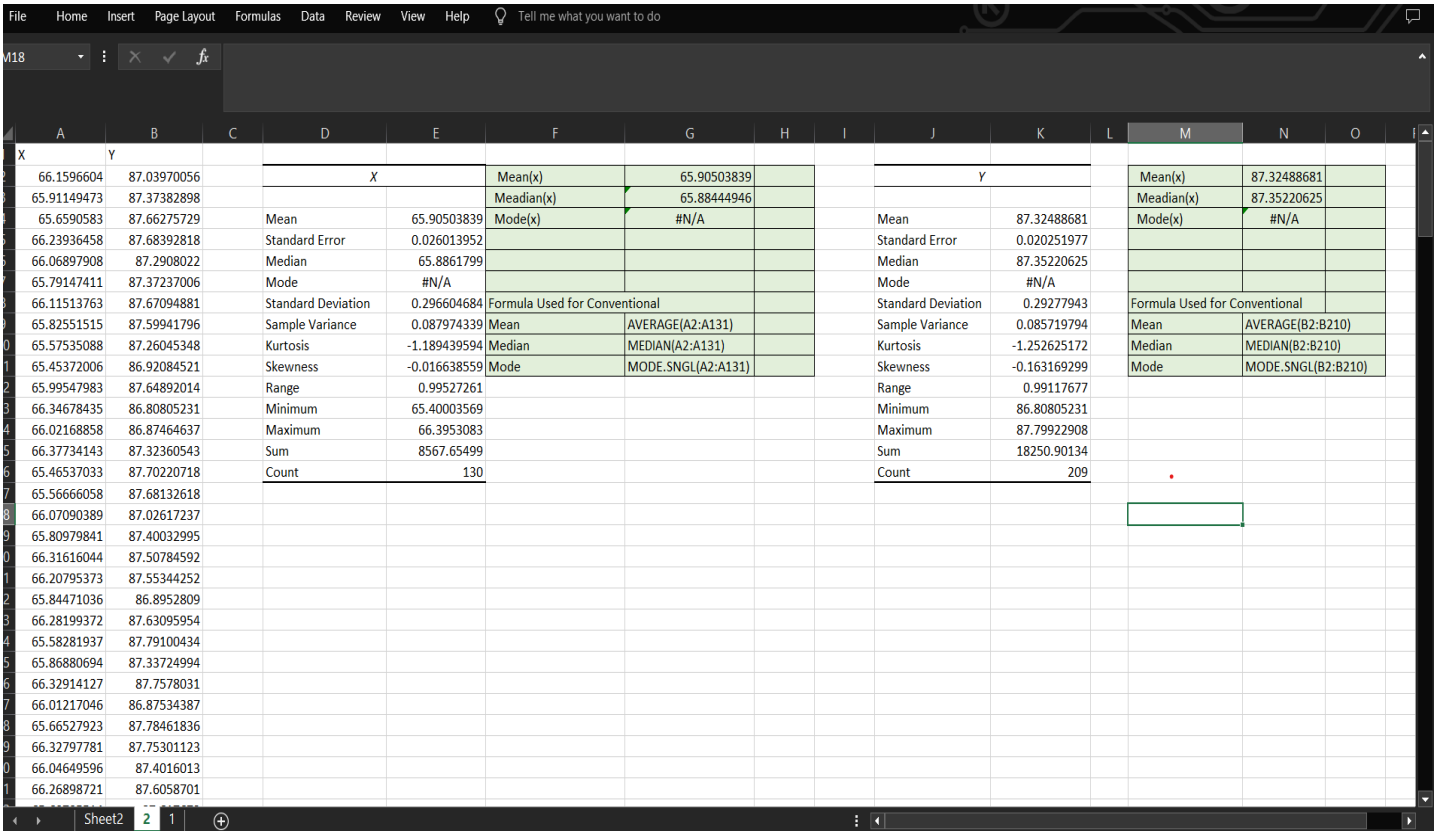
Total Observation = 11^2 + 9 = 130

Y = Random ((496/20) + 49) + 13
Y = Random (24.8+49)

Y => (0, 86.8)

Total Observation = 14^2 + 19 = 209





X	Y														
66.1596604	87.03970056		X		Mean(x)	65.90503839			Y			Mean(x)	87.32488681		
65.91149473	87.37382898				Meadian(x)	65.88444946						Meadian(x)	87.35220625		
65.6590583	87.66275729	Mean	65.90503839		Mode(x)	#N/A			Mean	87.32488681		Mode(x)	#N/A		
66.23936458	87.68392818	Standard Error	0.026013952						Standard Error	0.020251977					
66.06897908	87.2908022	Median	65.8861799						Median	87.35220625					
65.79147411	87.37237006	Mode	#N/A						Mode	#N/A					
66.11513763	87.67094881	Standard Deviation	0.296604684	Formula Used for Conventional					Standard Deviation	0.29277943		Formula Used for Conventional			
65.82551515	87.59941796	Sample Variance	0.087974339	Mean	AVERAGE(A2:A131)				Sample Variance	0.085719794		Mean	AVERAGE(B2:B210)		
65.57535088	87.26045348	Kurtosis	-1.189439594	Median	MEDIAN(A2:A131)				Kurtosis	-1.252625172		Median	MEDIAN(B2:B210)		
65.45372006	86.92084521	Skewness	-0.016638559	Mode	MODE.SNGL(A2:A131)				Skewness	-0.163169299		Mode	MODE.SNGL(B2:B210)		
65.99547983	87.64892014	Range	0.99527261						Range	0.99117677					
66.34678435	86.80805231	Minimum	65.40003569						Minimum	86.80805231					
66.02168858	86.87464637	Maximum	66.3953083						Maximum	87.79922908					
66.37734143	87.32360543	Sum	8567.65499						Sum	18250.90134					
65.46537033	87.70220718	Count	130						Count	209					
65.56666058	87.68132618														
66.07090389	87.02617237														
65.80979841	87.40032995														
66.31616044	87.50784592														
66.20795373	87.55344252														
65.84471036	86.8952809														
66.28199372	87.63095954														
65.58281937	87.79100434														
65.86880694	87.33724994														
66.32914127	87.7578031														
66.01217046	86.87534387														
65.66527923	87.78461836														
66.32797781	87.75301123														
66.04649596	87.4016013														
66.26898721	87.6058701														

Question 3.

- *Find regression statistics and analyze correlation coefficients also plot the required graph*

Note -Ensure no two students use the same dataset; violations will result in UMC.

First data set -

$$\boldsymbol{Study\ hours\ (X) = Genrate\ random\ number\ between\ 2\ \ to\ 10 + 7th\ digit\ of\ your\ registration\ no\ \ number}$$

$$\boldsymbol{Marks\ (Y) = Genrate\ random\ number\ \ beteen\ 30\ to\ \ 80 + last\ digit\ of\ your\ registration\ number}$$

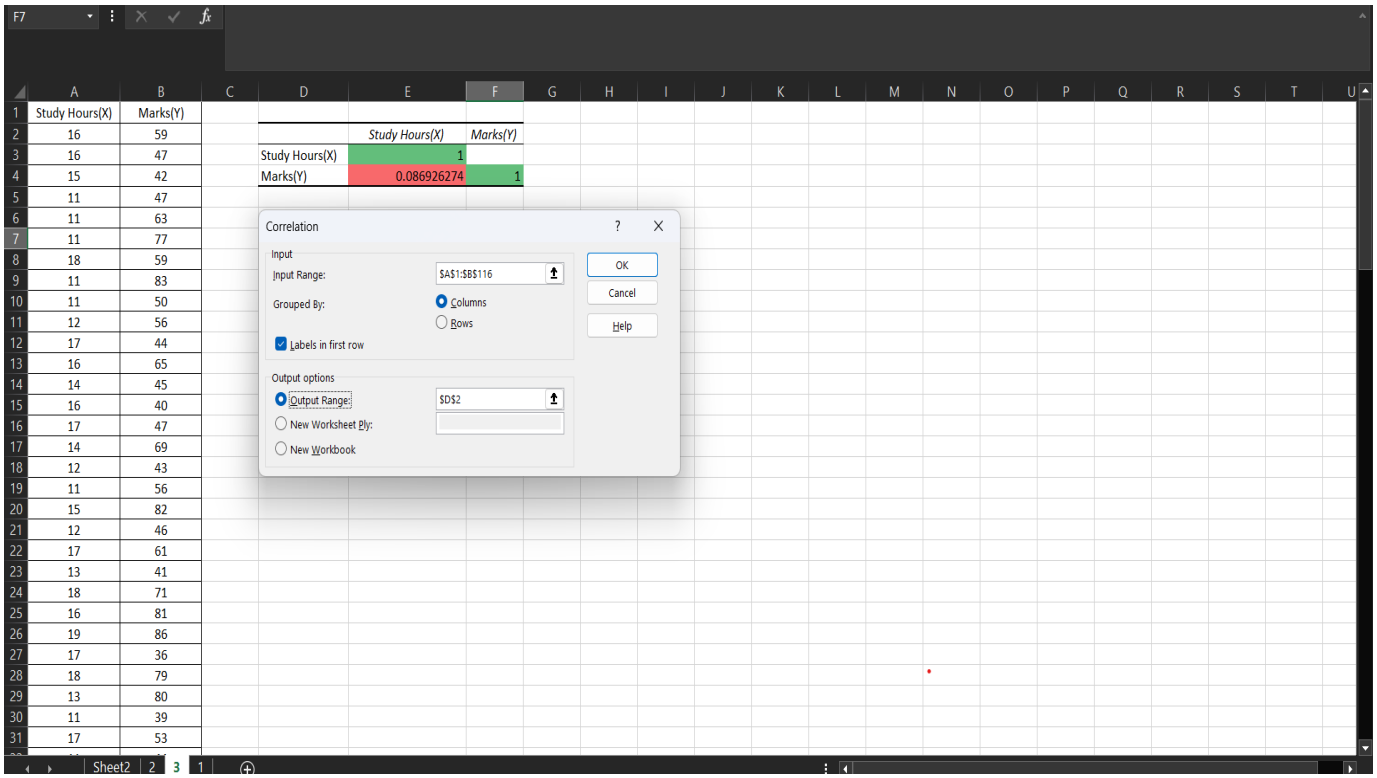
$$\boldsymbol{Total\ no\ of\ observations = 100 + sum\ of\ loas\ two\ digis\ of\ your\ registration\ number}$$

Note Attach the screen print with dialog box for every step.

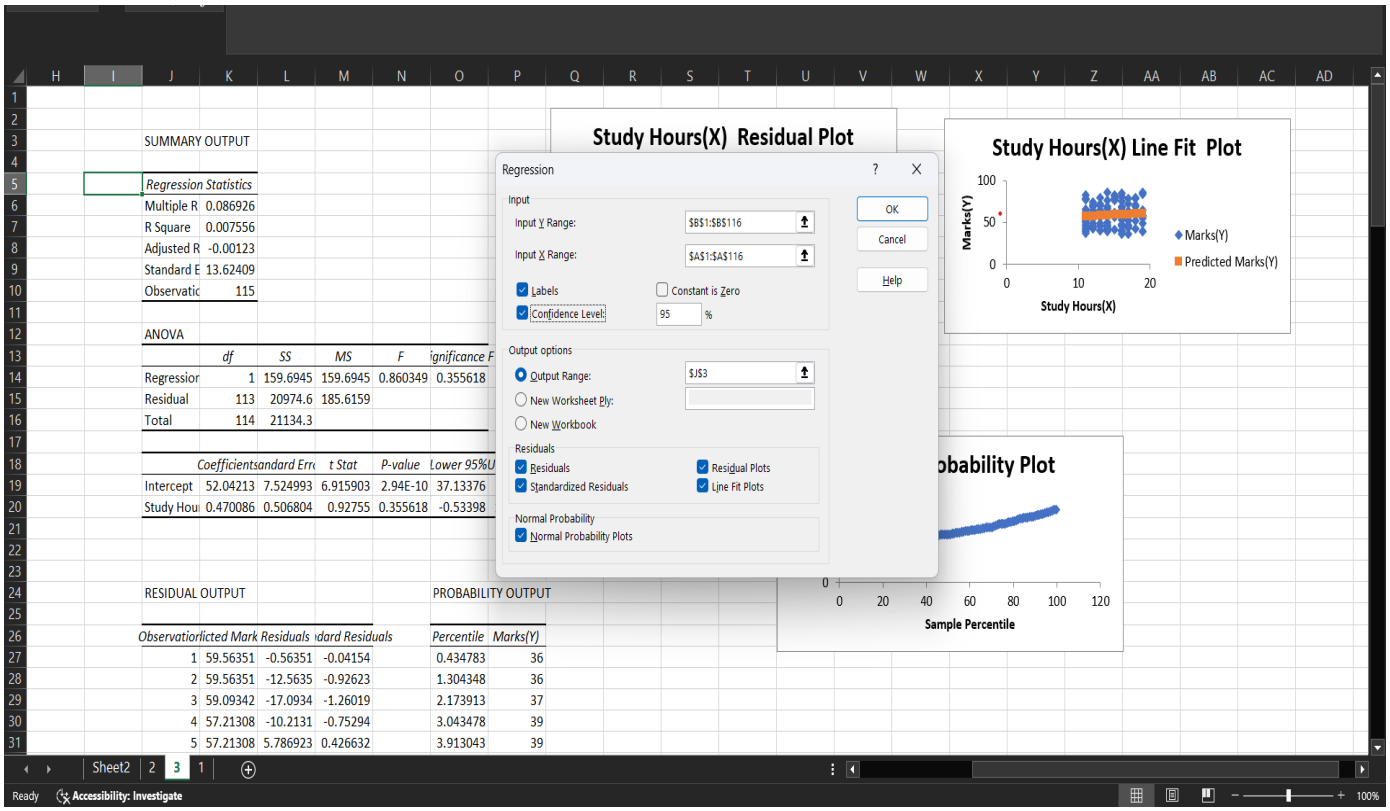
Total number of observations = 100 + 15 = 115

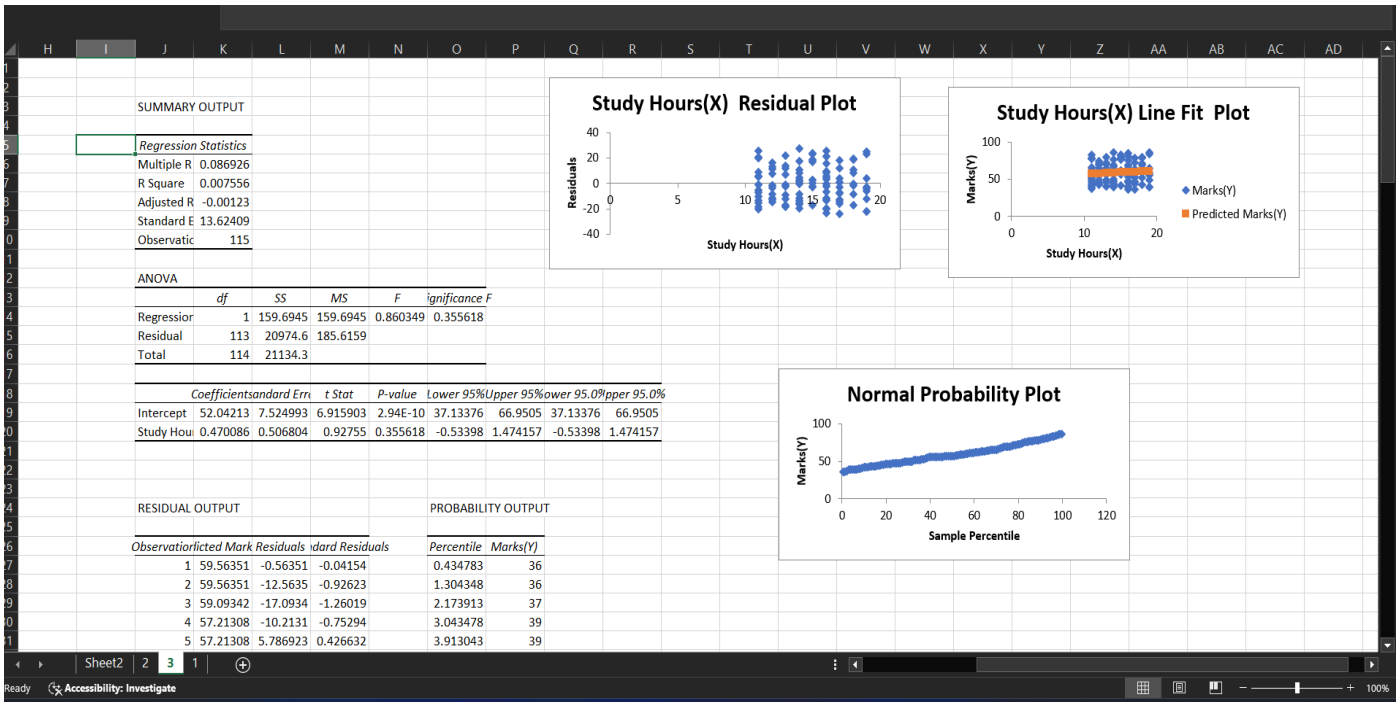
Study Hours (X) =RANDBETWEEN(2,10)+9 Marks (Y) =RANDBETWEEN(30,80)+6

Correlation->



Regression->


SUMMARY OUTPUT
Regression Statistics
Multiple R 0.086926
R Square 0.007556
Adjusted R -0.00123
Standard E 13.62409
Observatic 115
ANOVA
df SS MS F ignificance F
Regressior 1 159.6945 159.6945 0.860349 0.355618
Residual 113 20974.6 185.6159
Total 114 21134.3
Coefficients tandard Erri t Stat P-value Lower 95%U
Intercept 52.04213 7.524993 6.915903 2.94E-10 37.13376
Study Hou 0.470086 0.506804 0.92755 0.355618 -0.53398
RESIDUAL OUTPUT
PROBABILITY OUTPUT
Observatio licted Mark Residuals dard Residuals Percentile Marks(Y)
1 59.56351 -0.56351 -0.04154 0.434783 36
2 59.56351 -12.5635 -0.92623 1.304348 36
3 59.09342 -17.0934 -1.26019 2.173913 37
4 57.21308 -10.2131 -0.75294 3.043478 39
5 57.21308 5.786923 0.426632 3.913043 39
Study Hours(X) Residual Plot
Residuals
Study Hours(X)
Study Hours(X) Line Fit Plot
Marks(Y)
Predicted Marks(Y)
Study Hours(X)
Regression
Input
Input Y Range: $B$1:$B$116
Input X Range: $A$1:$A$116
Labels
Constant is Zero
Confidence Level: 95 %
Output options
Output Range: $J$3
New Worksheet Ply:
New Workbook
Residuals
Residuals
Residual Plots
Standardized Residuals
Line Fit Plots
Normal Probability
Normal Probability Plots
OK
Cancel
Help
obability Plot
Sample Percentile
Sheet2
Ready
Accessibility: Investigate


SUMMARY OUTPUT
Regression Statistics
Multiple R 0.086926
R Square 0.007556
Adjusted R -0.00123
Standard E 13.62409
Observatic 115
ANOVA
df SS MS F ignificance F
Regressior 1 159.6945 159.6945 0.860349 0.355618
Residual 113 20974.6 185.6159
Total 114 21134.3
Coefficients tandard Erri t Stat P-value Lower 95% Upper 95% ower 95.0% pper 95.0%
Intercept 52.04213 7.524993 6.915903 2.94E-10 37.13376 66.9505 37.13376 66.9505
Study Hou 0.470086 0.506804 0.92755 0.355618 -0.53398 1.474157 -0.53398 1.474157
RESIDUAL OUTPUT
PROBABILITY OUTPUT
Observatio licted Mark Residuals dard Residuals Percentile Marks(Y)
1 59.56351 -0.56351 -0.04154 0.434783 36
2 59.56351 -12.5635 -0.92623 1.304348 36
3 59.09342 -17.0934 -1.26019 2.173913 37
4 57.21308 -10.2131 -0.75294 3.043478 39
5 57.21308 5.786923 0.426632 3.913043 39
Study Hours(X) Residual Plot
Residuals
Study Hours(X)
Study Hours(X) Line Fit Plot
Marks(Y)
Predicted Marks(Y)
Study Hours(X)
Normal Probability Plot
Marks(Y)
Sample Percentile
Sheet2
Ready
Accessibility: Investigate

Question 4.

Find the pmf and cdf for Binomial and Poisson distribution also plot the required graph

Note -Ensure no two students use the same dataset; violations will result in UMC.

Data set -

X=Generate random number +7th digit of your registration no number

N (Number of trials)=25 +last two digit of your registration number

p=Probability of success=1/(sum of last two digits of your registraion number)

Mean=λ=Your roll number

Note - Attach the screen print with dialog box for every step.

Solution ->

X = rand() + 9

Number of trails = 25 + 96 = 121

P = 1/(9+6)

P= 1/15

Formula is used ->

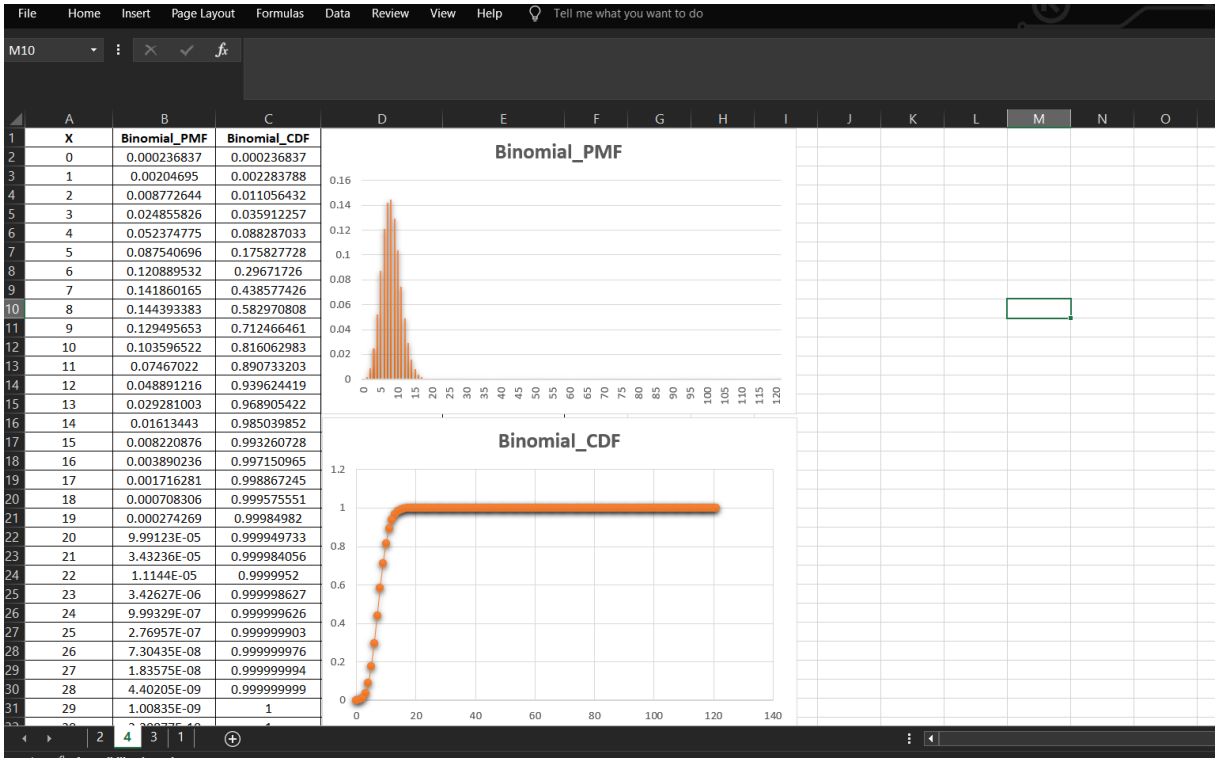
. Binomial PMF = BINOM.DIST(A2, 121, 1/15, FALSE)

. Binomial CDF = BINOM.DIST(A2, 121, 1/15, TRUE)

. Poisson PMF = POISSON.DIST(A2, 49, FALSE)

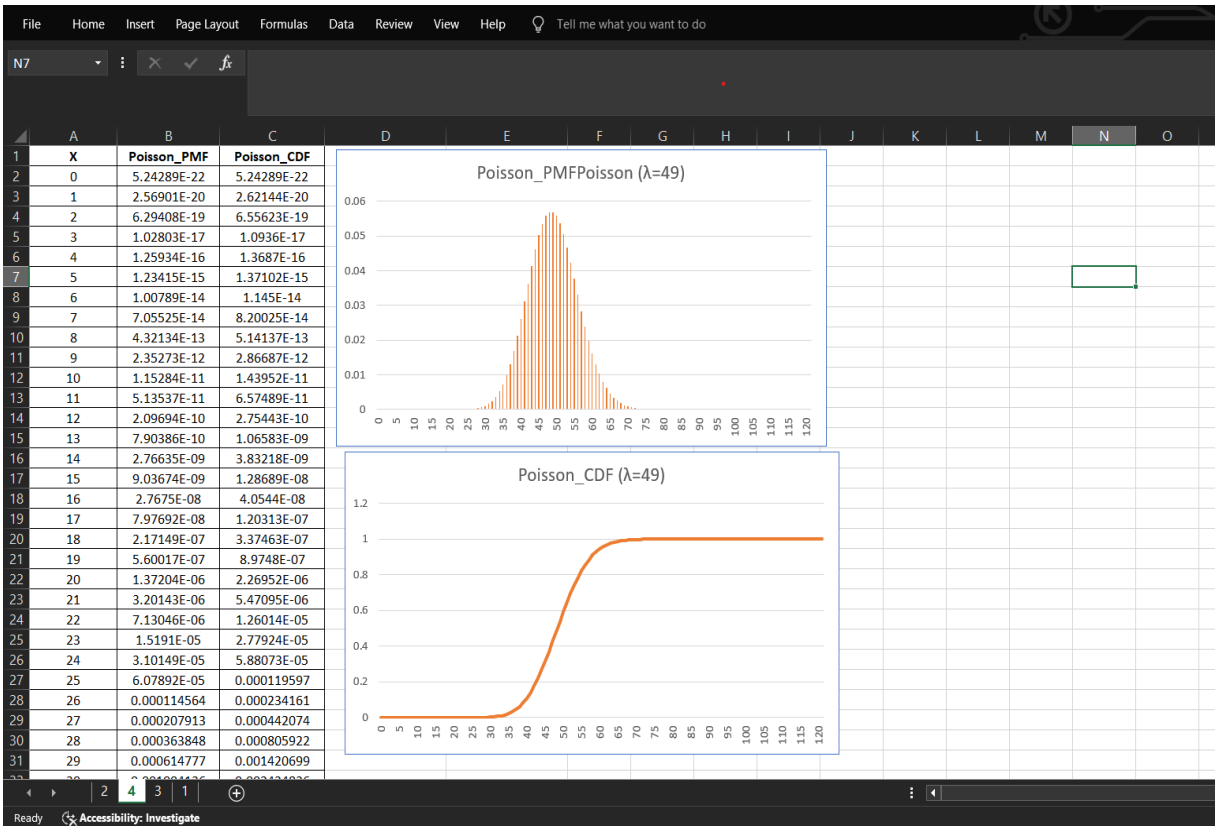
. Poisson CDF = POISSON.DIST(A2, 49, TRUE)

Binomial_PMF And Binomial_CDF ->



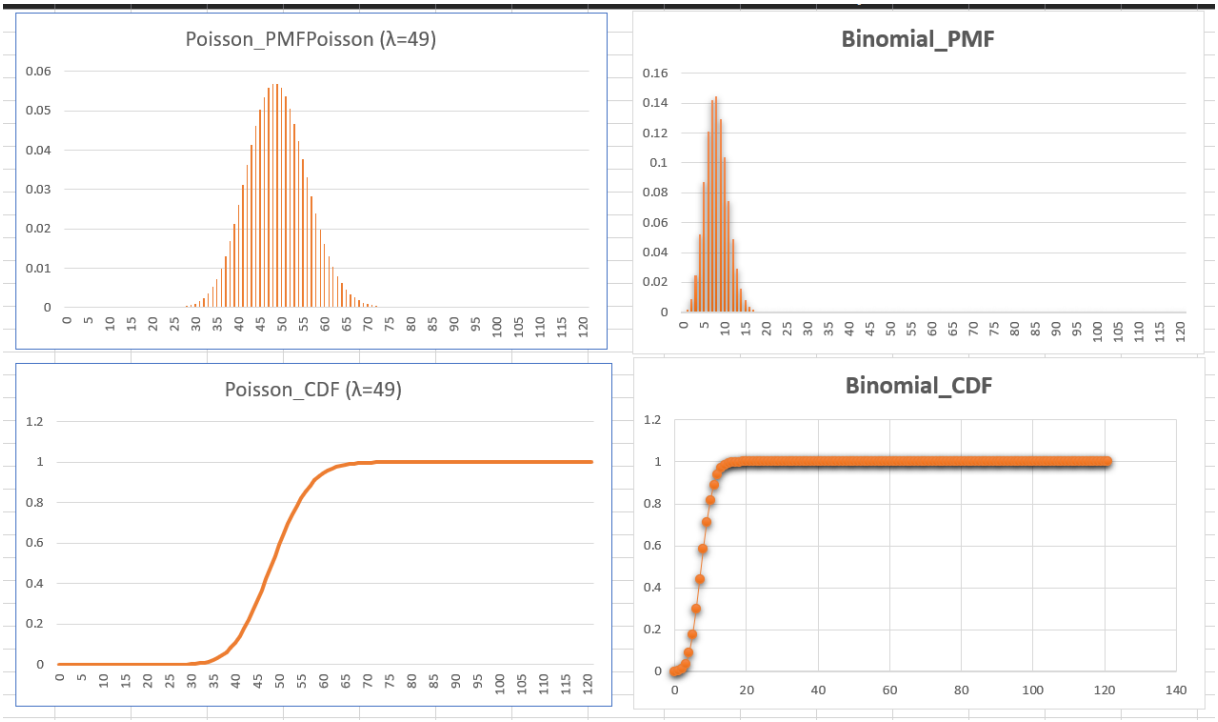
X	Binomial_PMF	Binomial_CDF
0	0.000236837	0.000236837
1	0.00204695	0.002283788
2	0.008772644	0.011056432
3	0.024855826	0.035912257
4	0.052374775	0.088287033
5	0.087540696	0.175827728
6	0.120889532	0.29671726
7	0.141860165	0.438577426
8	0.144393383	0.582970808
9	0.129495653	0.712466461
10	0.103596522	0.816062983
11	0.07467022	0.890733203
12	0.048891216	0.939624419
13	0.029281003	0.968905422
14	0.01613443	0.985039852
15	0.008220876	0.993260728
16	0.003890236	0.997150965
17	0.001716281	0.998867245
18	0.000708306	0.999575551
19	0.000274269	0.99984982
20	9.99123E-05	0.999949733
21	3.43236E-05	0.999984056
22	1.1144E-05	0.9999952
23	3.42627E-06	0.999998627
24	9.99329E-07	0.999999626
25	2.76957E-07	0.999999903
26	7.30435E-08	0.999999976
27	1.83575E-08	0.999999994
28	4.40205E-09	0.999999999
29	1.00835E-09	1

Possion_PMF and Possion_CDF:



X	Poisson_PMF	Poisson_CDF
0	5.24289E-22	5.24289E-22
1	2.56901E-20	2.62144E-20
2	6.29408E-19	6.55623E-19
3	1.02803E-17	1.0936E-17
4	1.25934E-16	1.3687E-16
5	1.23415E-15	1.37102E-15
6	1.00789E-14	1.145E-14
7	7.05525E-14	8.20025E-14
8	4.32134E-13	5.14137E-13
9	2.35273E-12	2.86687E-12
10	1.15284E-11	1.43952E-11
11	5.13537E-11	6.57489E-11
12	2.09694E-10	2.75443E-10
13	7.90386E-10	1.06583E-09
14	2.76635E-09	3.83218E-09
15	9.03674E-09	1.28689E-08
16	2.7675E-08	4.0544E-08
17	7.97692E-08	1.20313E-07
18	2.17149E-07	3.37463E-07
19	5.60017E-07	8.9748E-07
20	1.37204E-06	2.26952E-06
21	3.20143E-06	5.47095E-06
22	7.13046E-06	1.26014E-05
23	1.5191E-05	2.77924E-05
24	3.10149E-05	5.88073E-05
25	6.07892E-05	0.000119597
26	0.000114564	0.000234161
27	0.000207913	0.000442074
28	0.000363848	0.000805922
29	0.000614777	0.001420699

All Four Graphs ->



Question 5.

Solution 5.->

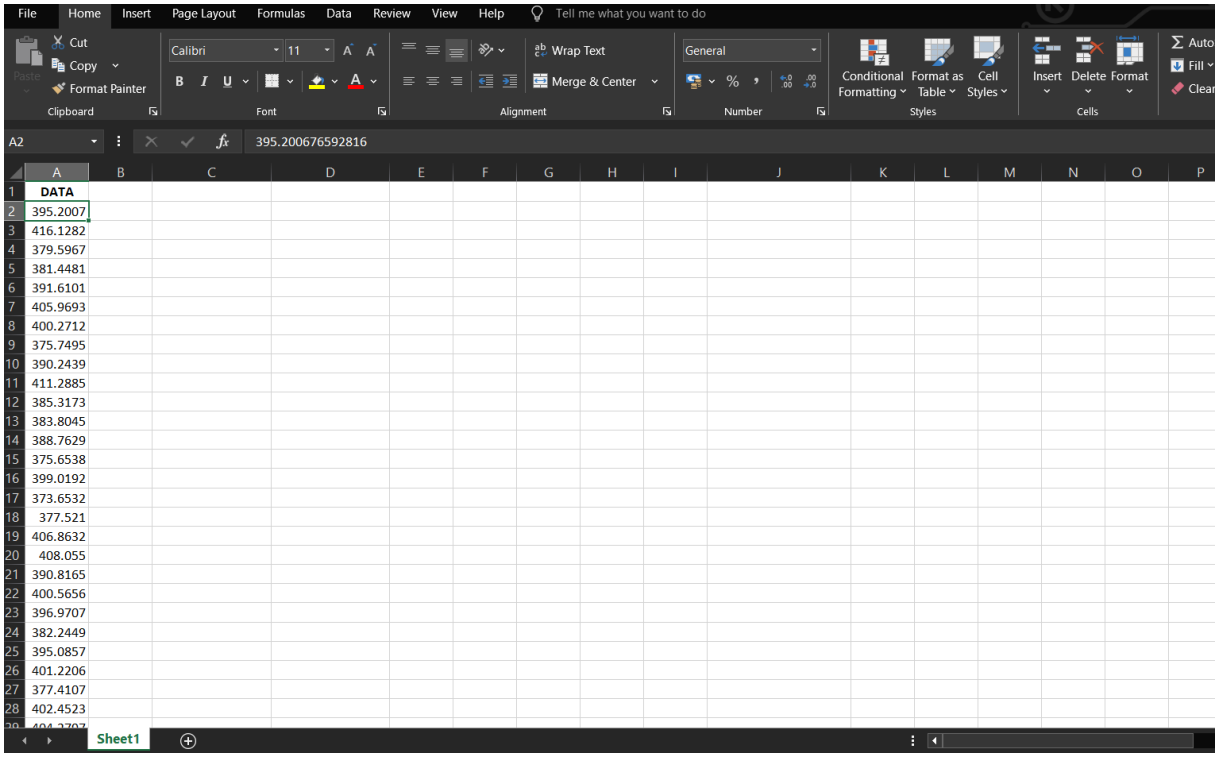
1. Total Data Points (N):
N=1000 + (23×200) + 49 = 1000 + 4600 + 49 = 5649

2. Mean (μ):
μ = (23×15) + 49 = 345 + 49 = 394

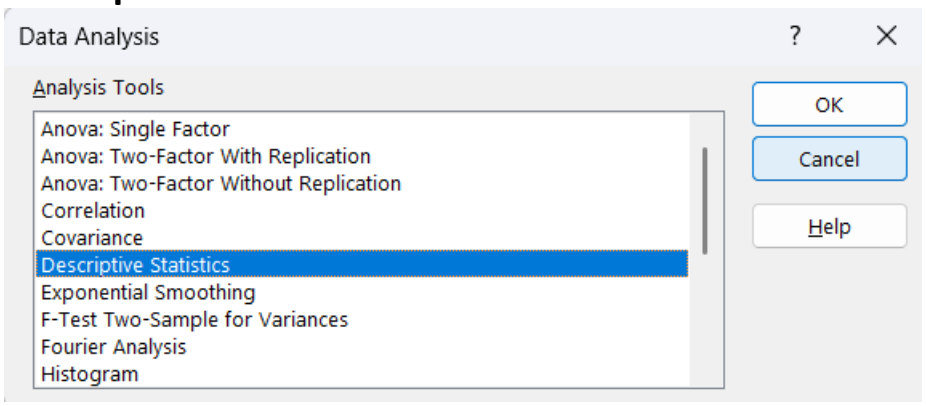
3. Standard Deviation (σ):
σ = (49×0.25) + (23×0.1) = 12.25+2.3 = 14.55

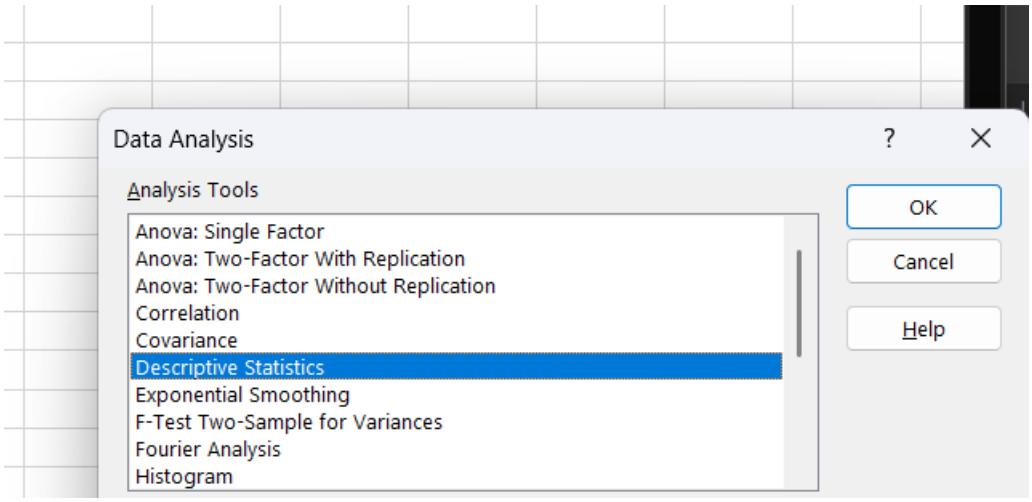
Formula To Generate Random Normal Dataset ->

. =NORM.INV(RAND(), 394, 14.55)



Descriptive Statistics:

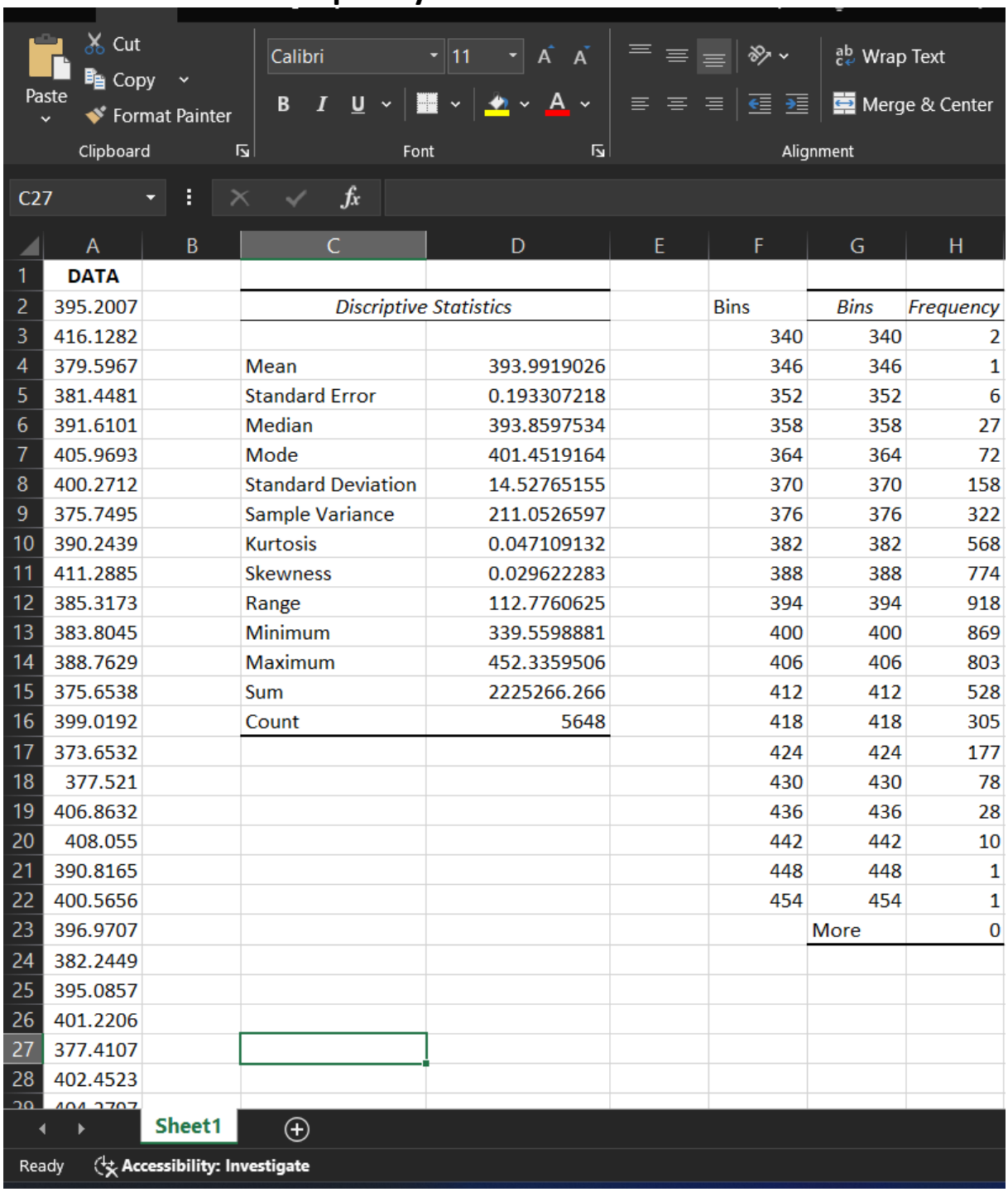




C27

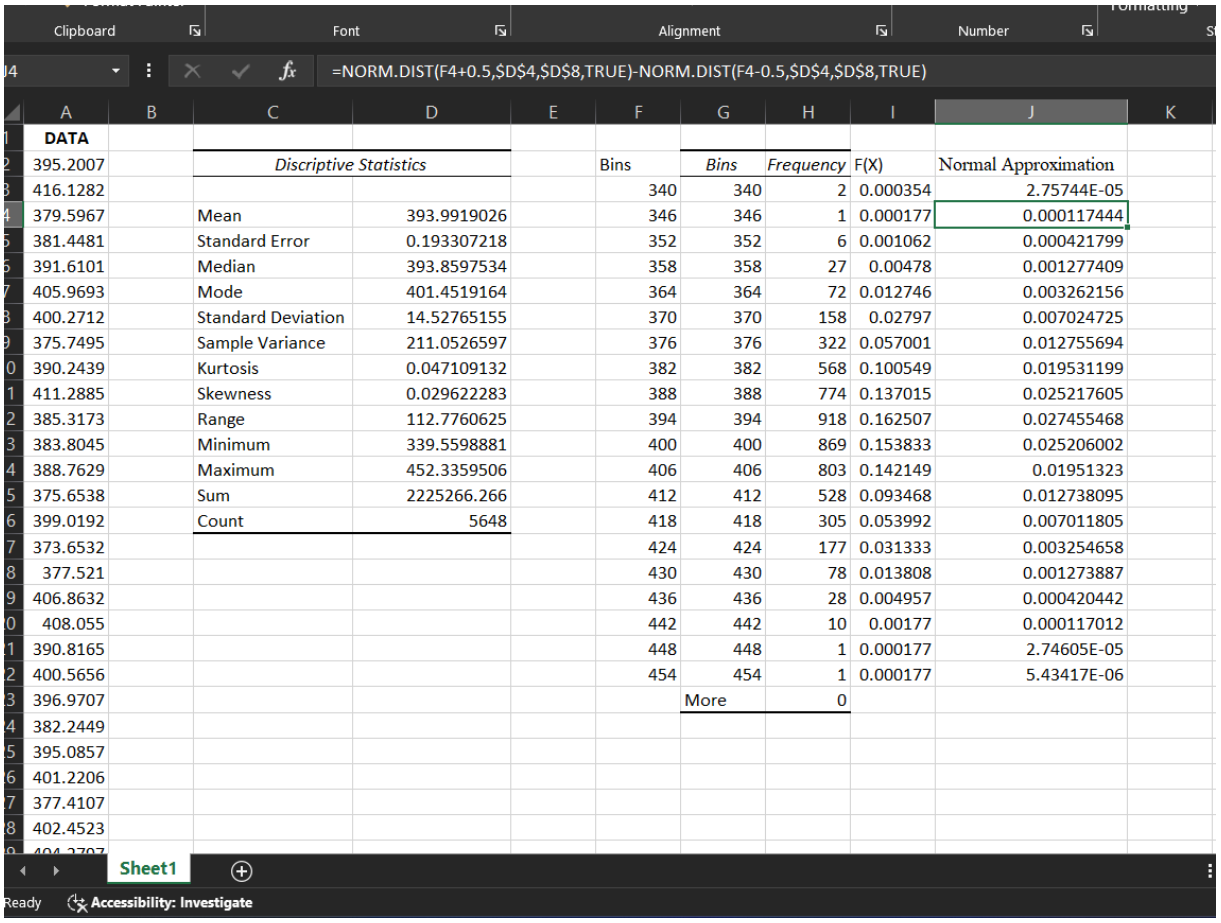
	A	B	C	D	E
1	**DATA**				
2	395.2007		*Discriptive Statistics*		
3	416.1282				
4	379.5967		Mean	393.9919026	
5	381.4481		Standard Error	0.193307218	
6	391.6101		Median	393.8597534	
7	405.9693		Mode	401.4519164	
8	400.2712		Standard Deviation	14.52765155	
9	375.7495		Sample Variance	211.0526597	
10	390.2439		Kurtosis	0.047109132	
11	411.2885		Skewness	0.029622283	
12	385.3173		Range	112.7760625	
13	383.8045		Minimum	339.5598881	
14	388.7629		Maximum	452.3359506	
15	375.6538		Sum	2225266.266	
16	399.0192		Count	5648	
17	373.6532				
18	377.521				
19	406.8632				
20	408.055				
21	390.8165				
22	400.5656				
23	396.9707				
24	382.2449				
25	395.0857				
26	401.2206				
27	377.4107				
28	402.4523				

bins and construct frequency table :



	A	B	C	D	E	F	G	H
1	**DATA**							
2	395.2007		*Discriptive Statistics*			Bins	*Bins*	*Frequency*
3	416.1282					340	340	2
4	379.5967		Mean	393.9919026		346	346	1
5	381.4481		Standard Error	0.193307218		352	352	6
6	391.6101		Median	393.8597534		358	358	27
7	405.9693		Mode	401.4519164		364	364	72
8	400.2712		Standard Deviation	14.52765155		370	370	158
9	375.7495		Sample Variance	211.0526597		376	376	322
10	390.2439		Kurtosis	0.047109132		382	382	568
11	411.2885		Skewness	0.029622283		388	388	774
12	385.3173		Range	112.7760625		394	394	918
13	383.8045		Minimum	339.5598881		400	400	869
14	388.7629		Maximum	452.3359506		406	406	803
15	375.6538		Sum	2225266.266		412	412	528
16	399.0192		Count	5648		418	418	305
17	373.6532					424	424	177
18	377.521					430	430	78
19	406.8632					436	436	28
20	408.055					442	442	10
21	390.8165					448	448	1
22	400.5656					454	454	1
23	396.9707						More	0
24	382.2449							
25	395.0857							
26	401.2206							
27	377.4107							
28	402.4523							

Normal Approximation and Graph->



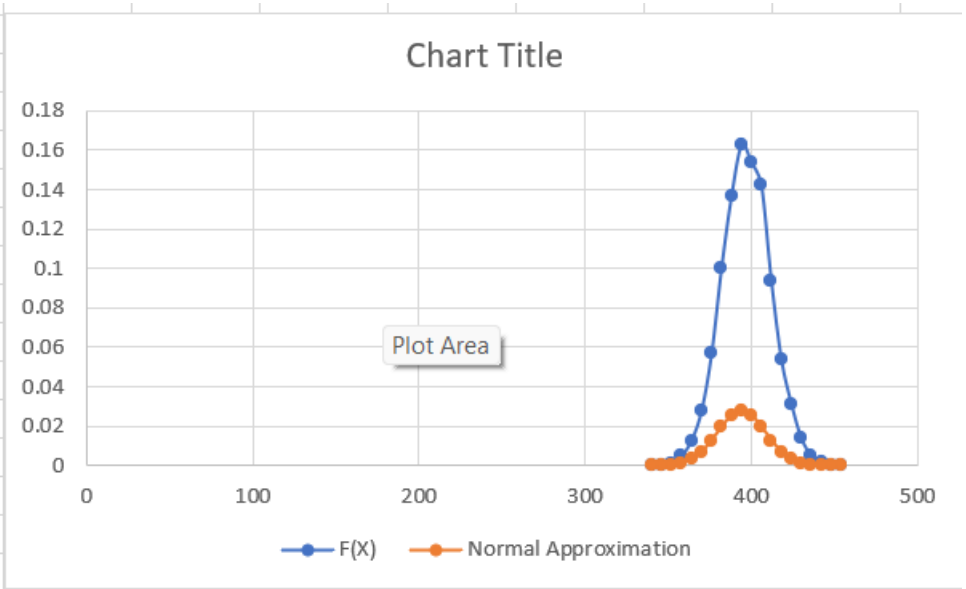
Clipboard | Font | Alignment | Number

=NORM.DIST(F4+0.5,$D$4,$D$8,TRUE)-NORM.DIST(F4-0.5,$D$4,$D$8,TRUE)

	A	B	C	D	E	F	G	H	I	J	K
1	DATA										
2	395.2007		*Discriptive Statistics*			Bins	*Bins*	*Frequency*	F(X)	Normal Approximation	
3	416.1282					340	340	2	0.000354	2.75744E-05	
4	379.5967		Mean	393.9919026		346	346	1	0.000177	0.000117444	
5	381.4481		Standard Error	0.193307218		352	352	6	0.001062	0.000421799	
6	391.6101		Median	393.8597534		358	358	27	0.00478	0.001277409	
7	405.9693		Mode	401.4519164		364	364	72	0.012746	0.003262156	
8	400.2712		Standard Deviation	14.52765155		370	370	158	0.02797	0.007024725	
9	375.7495		Sample Variance	211.0526597		376	376	322	0.057001	0.012755694	
10	390.2439		Kurtosis	0.047109132		382	382	568	0.100549	0.019531199	
11	411.2885		Skewness	0.029622283		388	388	774	0.137015	0.025217605	
12	385.3173		Range	112.7760625		394	394	918	0.162507	0.027455468	
13	383.8045		Minimum	339.5598881		400	400	869	0.153833	0.025206002	
14	388.7629		Maximum	452.3359506		406	406	803	0.142149	0.01951323	
15	375.6538		Sum	2225266.266		412	412	528	0.093468	0.012738095	
16	399.0192		Count	5648		418	418	305	0.053992	0.007011805	
17	373.6532					424	424	177	0.031333	0.003254658	
18	377.521					430	430	78	0.013808	0.001273887	
19	406.8632					436	436	28	0.004957	0.000420442	
20	408.055					442	442	10	0.00177	0.000117012	
21	390.8165					448	448	1	0.000177	2.74605E-05	
22	400.5656					454	454	1	0.000177	5.43417E-06	
23	396.9707						More	0			
24	382.2449										
25	395.0857										
26	401.2206										
27	377.4107										
28	402.4523										

Sheet1

Ready | Accessibility: Investigate



Summary:

- In this activity, I generated a unique random normal distribution dataset in Excel using the provided formulas for N, μ , and σ based on my section and roll number. I used the Excel NORMINV (or NORM.INV) function to create random values from a normal distribution. I captured screenshots of the formula used, as well as the top and bottom of the generated dataset. I then calculated descriptive statistics (mean, standard deviation, minimum, maximum, etc.) using the Data Analysis tool and formulas like AVERAGE, STDEV.S, MIN, and MAX.

- Following that, I constructed a frequency table by selecting appropriate bins based on the descriptive statistics. I computed the empirical probability distribution by dividing the frequency of each bin by the total number of observations. Screenshots of the formulas used were captured at each step.

- I then applied the normal approximation using continuity correction by calculating the probabilities for each bin range using the normal distribution formula in Excel. Screenshots of the applied formulas and the resulting probabilities were also taken.

- Finally, I plotted a scatter graph with smooth lines and markers, comparing the empirical probability distribution and the normal approximation. The complete Excel workbook was saved with all steps properly documented through screenshots.

Conclusion:

- This activity helped me understand the process of generating random normal data and analyzing it statistically. It strengthened my skills in Excel functions related to probability distributions and descriptive statistics. Additionally, the comparison between empirical and theoretical (normal approximation) probabilities demonstrated how closely real-world data can fit the ideal normal distribution under certain conditions. The graph provided a visual confirmation of the alignment between empirical data and the theoretical curve.

Question 6:

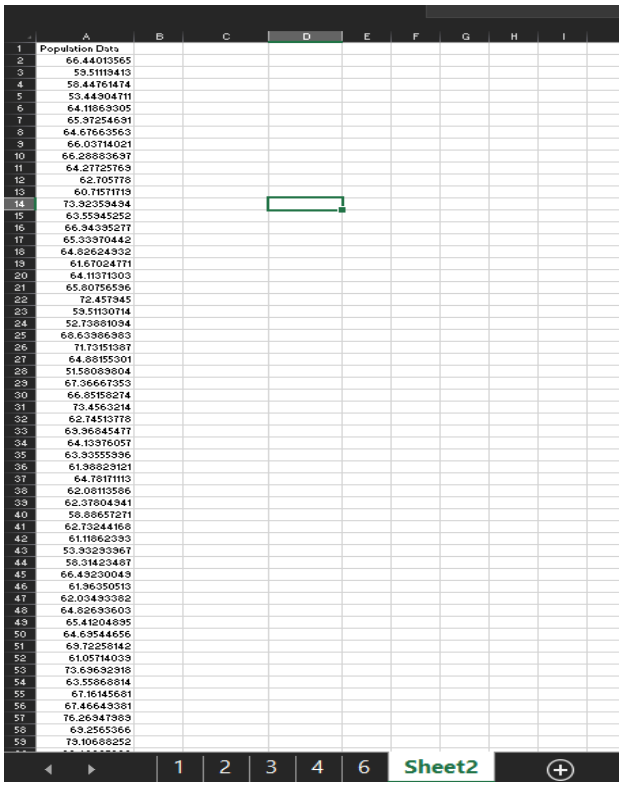
Solution:

- **N (Population Size) = 1100 + (Section number × 100) + Roll number**
  N = 1100 + (23 × 100) + 49
  = 1100 + 2300 + 49
  = 3449

- **n (Sample Size) = 500 + Roll number**
  n = 500 + 49
  = 549

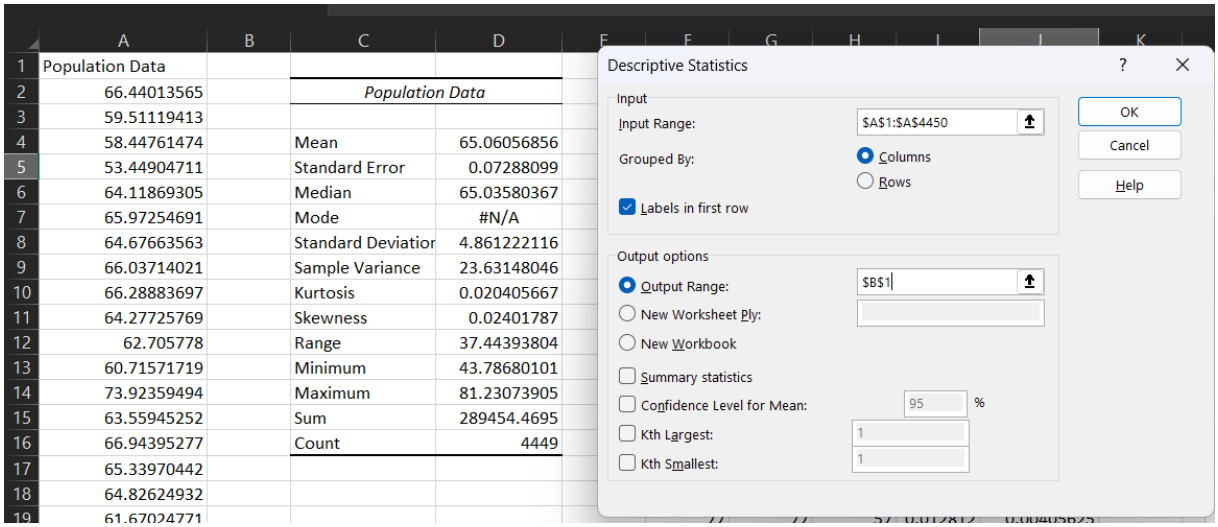
- **Total no. of samples = 100 + Roll number**
  Total samples = 100 + 49
  = 149

Normal Random Data Screenshots : = NORM.INV(RAND() ,65,5)



	A
1	Population Data
2	66.44013565
3	59.51119413
4	58.44761474
5	53.44304711
6	64.11863305
7	65.37254631
8	64.6766356
9	66.03714021
10	66.28883697
11	64.27725769
12	62.705778
13	60.71571719
14	73.32353434
15	63.5394252
16	66.94395277
17	65.33970442
18	64.82624932
19	61.67024771
20	64.11371303
21	65.80756536
22	72.457945
23	59.51130714
24	52.73881094
25	68.63386383
26	71.73151387
27	64.88155301
28	51.58083804
29	67.36667353
30	66.85158274
31	73.4563214
32	62.74513778
33	63.96845477
34	64.13976057
35	63.33555936
36	61.38623121
37	64.78171113
38	62.08113586
39	62.37804941
40	58.88657271
41	62.73244168
42	61.11862393
43	53.33293367
44	58.31423487
45	66.43230043
46	61.96350513
47	62.03493382
48	64.82693603
49	65.41204835
50	64.69544656
51	63.72258142
52	61.05714039
53	73.63693918
54	63.55868814
55	67.16145681
56	67.46643381
57	76.26947989
58	63.2565366
59	73.10688252

Descriptive Statistics Screenshots :



	A	B	C	D
1	Population Data			
2	66.44013565		*Population Data*	
3	59.51119413			
4	58.44761474		Mean	65.06056856
5	53.44904711		Standard Error	0.07288099
6	64.11869305		Median	65.03580367
7	65.97254691		Mode	#N/A
8	64.67663563		Standard Deviatior	4.861222116
9	66.03714021		Sample Variance	23.63148046
10	66.28883697		Kurtosis	0.020405667
11	64.27725769		Skewness	0.02401787
12	62.705778		Range	37.44393804
13	60.71571719		Minimum	43.78680101
14	73.92359494		Maximum	81.23073905
15	63.55945252		Sum	289454.4695
16	66.94395277		Count	4449
17	65.33970442			
18	64.82624932			
19	61.67024771			

Descriptive Statistics

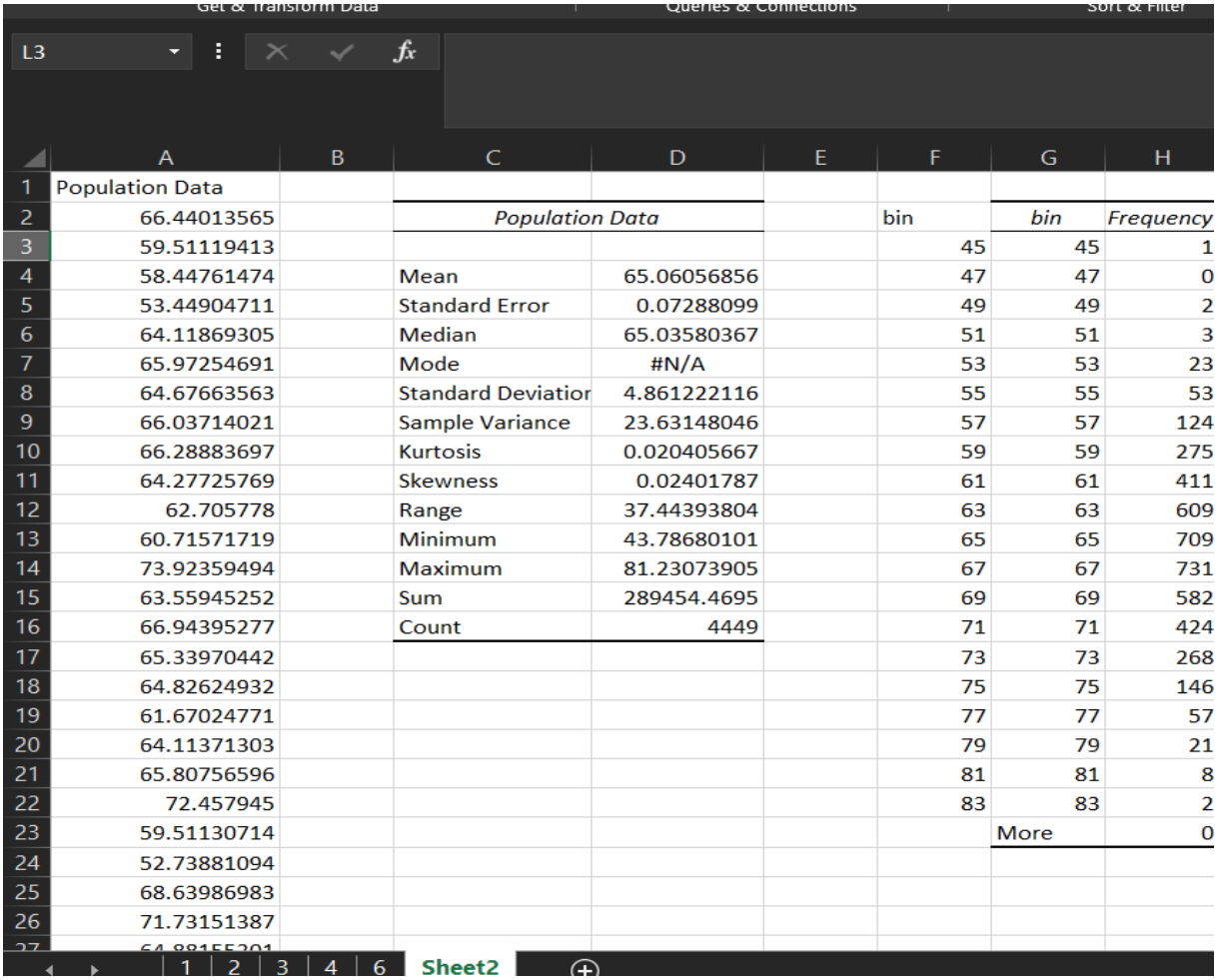
Input
Input Range: $A$1:$A$4450
Grouped By: Columns / Rows
Labels in first row

Output options
Output Range: $B$1
New Worksheet Ply:
New Workbook
Summary statistics
Confidence Level for Mean: 95 %
Kth Largest: 1
Kth Smallest: 1

OK
Cancel
Help

Bin and Frequency table:

- **Bin = MIN(Data) + (n * Bin Width)**
- **=FREQUENCY(Data Range, Bin Range)**

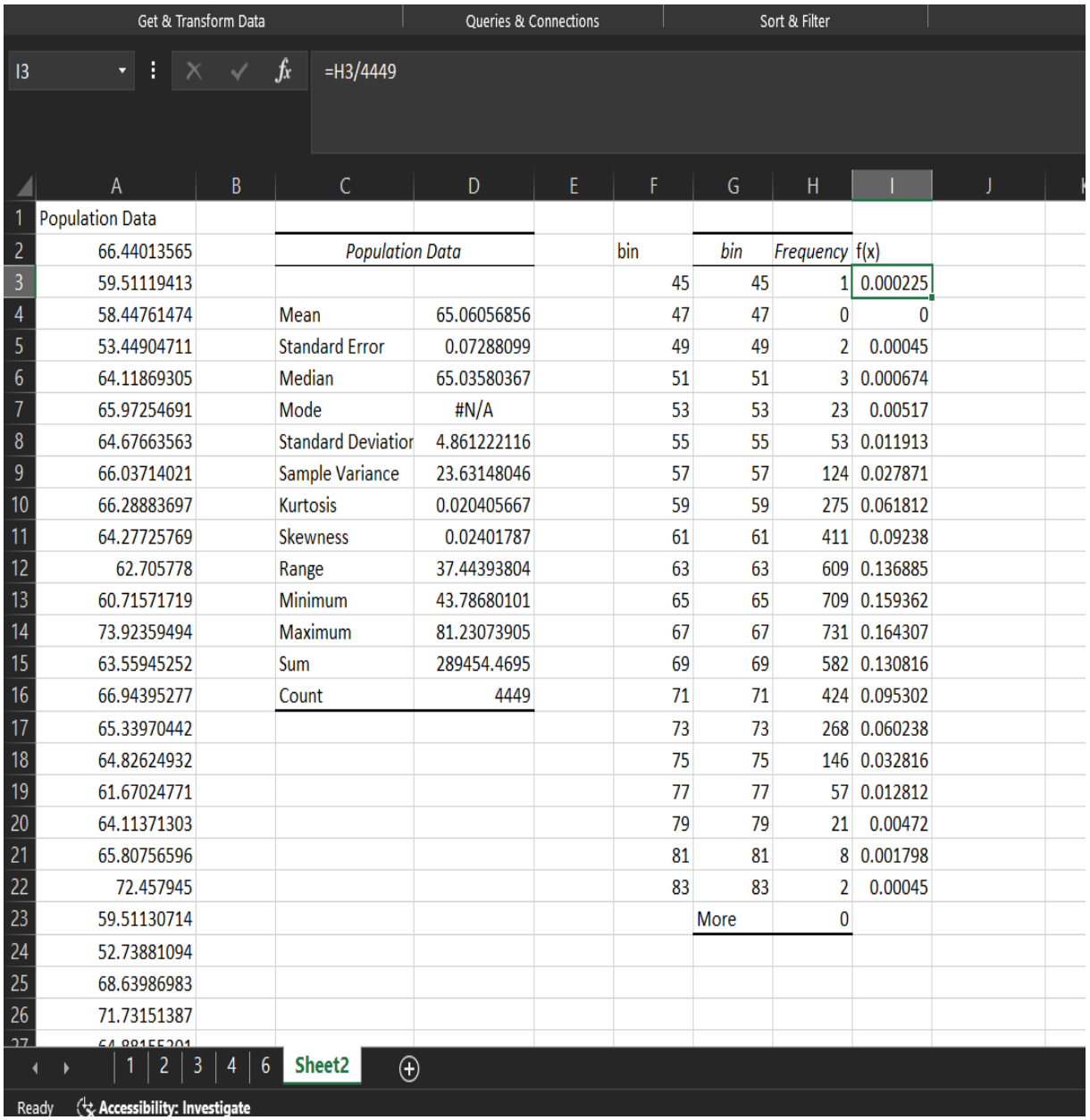


	A	B	C	D	E	F	G	H
1	Population Data							
2	66.44013565		*Population Data*			bin	*bin*	*Frequency*
3	59.51119413					45	45	1
4	58.44761474		Mean	65.06056856		47	47	0
5	53.44904711		Standard Error	0.07288099		49	49	2
6	64.11869305		Median	65.03580367		51	51	3
7	65.97254691		Mode	#N/A		53	53	23
8	64.67663563		Standard Deviatior	4.861222116		55	55	53
9	66.03714021		Sample Variance	23.63148046		57	57	124
10	66.28883697		Kurtosis	0.020405667		59	59	275
11	64.27725769		Skewness	0.02401787		61	61	411
12	62.705778		Range	37.44393804		63	63	609
13	60.71571719		Minimum	43.78680101		65	65	709
14	73.92359494		Maximum	81.23073905		67	67	731
15	63.55945252		Sum	289454.4695		69	69	582
16	66.94395277		Count	4449		71	71	424
17	65.33970442					73	73	268
18	64.82624932					75	75	146
19	61.67024771					77	77	57
20	64.11371303					79	79	21
21	65.80756596					81	81	8
22	72.457945					83	83	2
23	59.51130714						More	0
24	52.73881094							
25	68.63986983							
26	71.73151387							
27	64.88155201							

L3

Sheet2

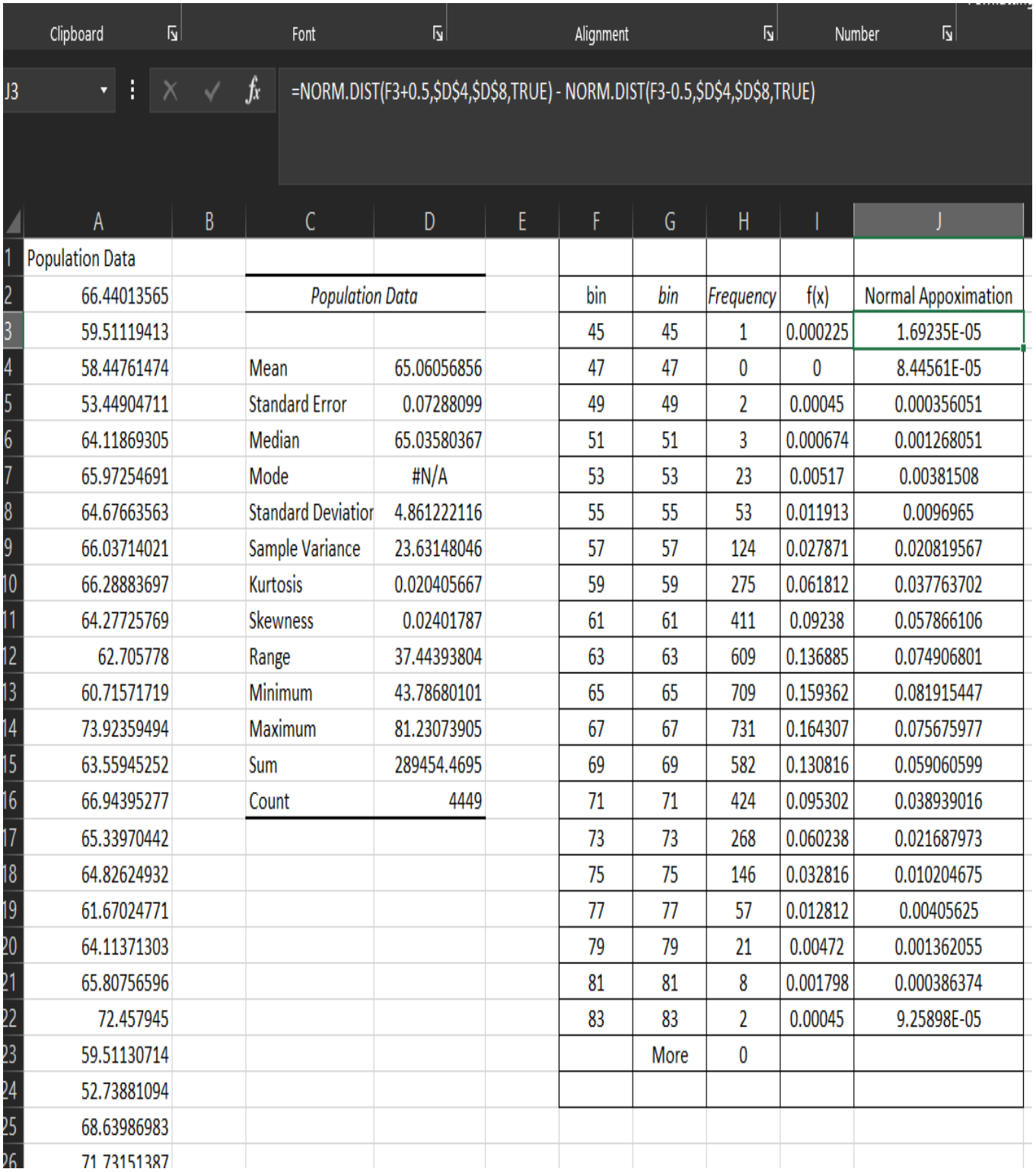
Empirical Probability: = **frequency_cell/total_N**



	A	B	C	D	E	F	G	H	I	J
1	Population Data									
2	66.44013565		*Population Data*			bin	*bin*	*Frequency*	f(x)	
3	59.51119413					45	45	1	0.000225	
4	58.44761474		Mean	65.06056856		47	47	0	0	
5	53.44904711		Standard Error	0.07288099		49	49	2	0.00045	
6	64.11869305		Median	65.03580367		51	51	3	0.000674	
7	65.97254691		Mode	#N/A		53	53	23	0.00517	
8	64.67663563		Standard Deviatior	4.861222116		55	55	53	0.011913	
9	66.03714021		Sample Variance	23.63148046		57	57	124	0.027871	
10	66.28883697		Kurtosis	0.020405667		59	59	275	0.061812	
11	64.27725769		Skewness	0.02401787		61	61	411	0.09238	
12	62.705778		Range	37.44393804		63	63	609	0.136885	
13	60.71571719		Minimum	43.78680101		65	65	709	0.159362	
14	73.92359494		Maximum	81.23073905		67	67	731	0.164307	
15	63.55945252		Sum	289454.4695		69	69	582	0.130816	
16	66.94395277		Count	4449		71	71	424	0.095302	
17	65.33970442					73	73	268	0.060238	
18	64.82624932					75	75	146	0.032816	
19	61.67024771					77	77	57	0.012812	
20	64.11371303					79	79	21	0.00472	
21	65.80756596					81	81	8	0.001798	
22	72.457945					83	83	2	0.00045	
23	59.51130714						More	0		
24	52.73881094									
25	68.63986983									
26	71.73151387									

Normal Approximation Probability with Continuity Correction:

=NORM.DIST(bin+0.5, mean, standard deviation, TRUE) - NORM.DIST(bin-0.5, mean, standard deviation, TRUE)

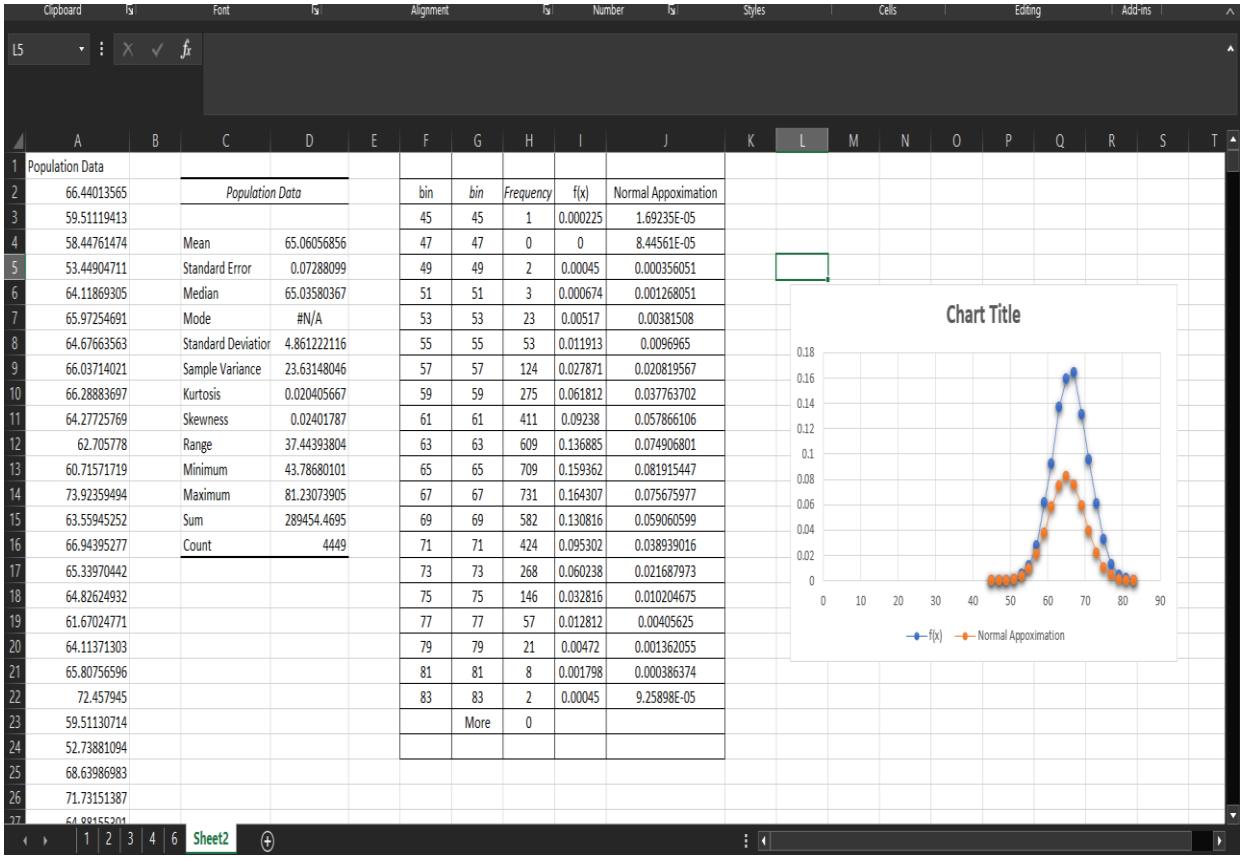


Clipboard | Font | Alignment | Number

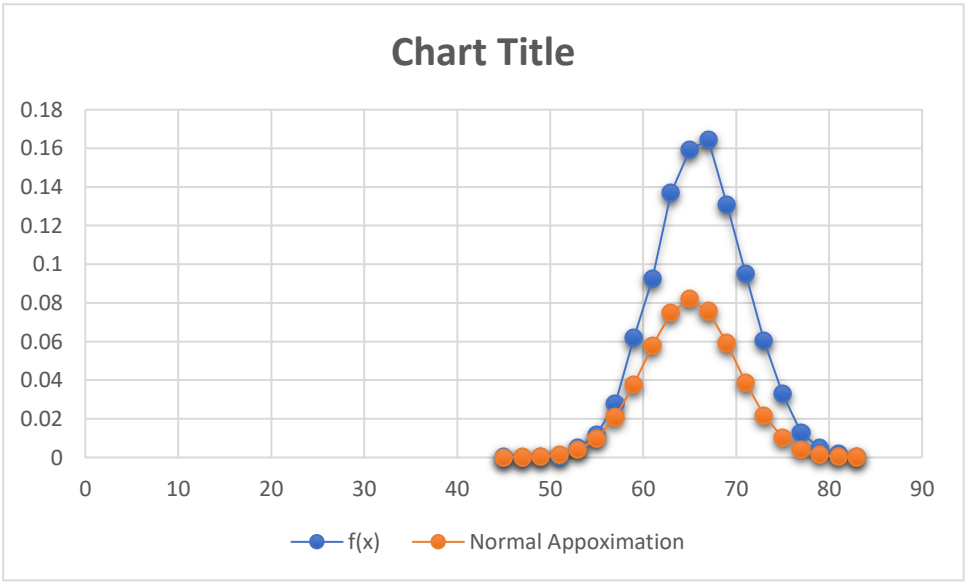
J3 | =NORM.DIST(F3+0.5,$D$4,$D$8,TRUE) - NORM.DIST(F3-0.5,$D$4,$D$8,TRUE)

	A	B	C	D	E	F	G	H	I	J
1	Population Data									
2	66.44013565		*Population Data*			bin	*bin*	*Frequency*	f(x)	Normal Appoximation
3	59.51119413					45	45	1	0.000225	1.69235E-05
4	58.44761474		Mean	65.06056856		47	47	0	0	8.44561E-05
5	53.44904711		Standard Error	0.07288099		49	49	2	0.00045	0.000356051
6	64.11869305		Median	65.03580367		51	51	3	0.000674	0.001268051
7	65.97254691		Mode	#N/A		53	53	23	0.00517	0.00381508
8	64.67663563		Standard Deviatior	4.861222116		55	55	53	0.011913	0.0096965
9	66.03714021		Sample Variance	23.63148046		57	57	124	0.027871	0.020819567
10	66.28883697		Kurtosis	0.020405667		59	59	275	0.061812	0.037763702
11	64.27725769		Skewness	0.02401787		61	61	411	0.09238	0.057866106
12	62.705778		Range	37.44393804		63	63	609	0.136885	0.074906801
13	60.71571719		Minimum	43.78680101		65	65	709	0.159362	0.081915447
14	73.92359494		Maximum	81.23073905		67	67	731	0.164307	0.075675977
15	63.55945252		Sum	289454.4695		69	69	582	0.130816	0.059060599
16	66.94395277		Count	4449		71	71	424	0.095302	0.038939016
17	65.33970442					73	73	268	0.060238	0.021687973
18	64.82624932					75	75	146	0.032816	0.010204675
19	61.67024771					77	77	57	0.012812	0.00405625
20	64.11371303					79	79	21	0.00472	0.001362055
21	65.80756596					81	81	8	0.001798	0.000386374
22	72.457945					83	83	2	0.00045	9.25898E-05
23	59.51130714						More	0		
24	52.73881094									
25	68.63986983									
26	71.73151387									

Graph Screenshot:



Population Data									
66.44013565		*Population Data*		bin	*bin*	*Frequency*	f(x)	Normal Appoximation
59.51119413				45	45	1	0.000225	1.69235E-05
58.44761474		Mean	65.06056856	47	47	0	0	8.44561E-05
53.44904711		Standard Error	0.07288099	49	49	2	0.00045	0.000356051
64.11869305		Median	65.03580367	51	51	3	0.000674	0.001268051
65.97254691		Mode	#N/A	53	53	23	0.00517	0.00381508
64.67663563		Standard Deviatior	4.861222116	55	55	53	0.011913	0.0096965
66.03714021		Sample Variance	23.63148046	57	57	124	0.027871	0.020819567
66.28883697		Kurtosis	0.020405667	59	59	275	0.061812	0.037763702
64.27725769		Skewness	0.02401787	61	61	411	0.09238	0.057866106
62.705778		Range	37.44393804	63	63	609	0.136885	0.074906801
60.71571719		Minimum	43.78680101	65	65	709	0.159362	0.081915447
73.92359494		Maximum	81.23073905	67	67	731	0.164307	0.075675977
63.55945252		Sum	289454.4695	69	69	582	0.130816	0.059060599
66.94395277		Count	4449	71	71	424	0.095302	0.038939016
65.33970442				73	73	268	0.060238	0.021687973
64.82624932				75	75	146	0.032816	0.010204675
61.67024771				77	77	57	0.012812	0.00405625
64.11371303				79	79	21	0.00472	0.001362055
65.80756596				81	81	8	0.001798	0.000386374
72.457945				83	83	2	0.00045	9.25898E-05
59.51130714					More	0		
52.73881094								
68.63986983								
71.73151387								



Summary And Conclusion:

- In this assignment, we calculated descriptive statistics like mean, standard deviation, and range.
- We created class intervals (bins) and calculated frequencies using formulas, not the histogram tool.
- Then, we plotted the actual frequency curve and compared it with the normal approximation curve.
- The graph shows that the data is approximately normally distributed, with minor variations.
- Thus, we conclude that real-world data often follows a normal distribution with slight deviations.