



Assignment - Advanced Excel

Submitted by :-

Name :- Devashish

Kaushik

Roll no.: 2402208016

Course – BCom(H) ACCA

ASSIGNMENT – 1

Advance Excel

Topic : Numerical Function (Count,CountA,CountIF,SUMIF)

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 5th August 2025

Item	Region	Sales
grapes	North	250
apples	South	155
grapes	West	130
lemon	North	255
apples	North	160
grapes	South	280
lemon	east	170
apples	east	285
apples	West	110

Q1	How Many Rows are there in the dataset?		
A1		10	COUNTA(A1:A10)
Q2	Find the number of rows where region is north?		
A2		3	COUNTIF(B1:B10,B2)
Q3	Find the number of row item is grapes and region is south?		
A3		1	COUNTIFS(A1:A10,A2,B1:B10,B3)
Q4	Find the total sales if the region is west		
A4		240	SUMIF(B1:B10,B4,C1:C10)
Q5	Find the total sales if the item is apple and region is north		
A5		160	SUMIFS(C1:C10,A1:A10,A3,B1:B10,B2)
Q6	Find the number of rows where item is lemon region is north and sales is more than 250?		
A6		1	COUNTIFS(A1:A10,A5,B1:B10,B2,C1:C10,">250")

ASSIGNMENT – 2

Advance Excel

Topic : Numerical Function (Count, CountA, Count IF, SUMIF, Sumproduct, Cell Referencing)

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 12th August 2025

DataSet – 1

Student Name	Marks
Deepak	85
Binil	96
Suman	78
Shilpa	65
Sunil	89

Q1. Find the total marks of the student ?

413 SUM(B3:B7)

Q2. Find the Max. Marks?

96 MAX(B3:B7)

Q3. Find the Min.Marks?

65 MIN(B3:B7)

Q4. Find the average marks?

82.6 AVERAGE(B3:B7)

Other Method

82.6 $\text{SUM(B3:B7)}/\text{COUNT(B3:B7)}$

DataSet-2

Student Name	Marks
Deepak	85
Binil	G
Suman	78
Shilpa	65
Sunil	89

Count Function	4
CountA Function	5

DataSet-3

Student Name	Marks
Deepak	85
Binil	
Suman	78
Shilpa	65
Sunil	89

Count	4
CountA	4
CountBlan	1

DataSet-3

Student Name	Marks
Deepak	85
Binil	
Suman	78
Shilpa	65
Sunil	89

Count	4
CountA	5
CountBlan	0

DataSet-4

Student Name	Gender
Deepak	F
Binil	M
Suman	F
Shilpa	F
Sunil	M

Refrencing				
M	2	COUNTIF(\$I\$27:\$I\$31,\$K28)		
F	3			

DataSet-5

Column1	Column2	Column3	Column4	Column5	Column6	Column7	Column8	Column9	Column10
1	2	3	4	5	6	7	8	9	10
2	4	6	8	10	12	14	16	18	20
3	6	9	12	15	18	21	24	27	30
4	8	12	16	20	24	28	32	36	40
5	10	15	20	25	30	35	40	45	50
6	12	18	24	30	36	42	48	54	60
7	14	21	28	35	42	49	56	63	70
8	16	24	32	40	48	56	64	72	80
9	18	27	36	45	54	63	72	81	90
10	20	30	40	50	60	70	80	90	100

DataSet-6

Products	Quantity	Cost	Total	SumProduct
P1	150	10	1500	27500
P2	200	15	3000	
P3	250	20	5000	
P4	300	25	7500	
P5	350	30	10500	
Total Sum	27500			

ASSIGNMENT – 3

Advance Excel

Topic : Left Function, Right Function, Mid Function, Length Function, Right Function, Replace Function, Substitute Function.

Name:

Devashish

Kaushik

Roll no.:-

2402208016

Date: - 19th

August 2025

Left Function: - The left function in excel extract a specified number of characters from the beginning (left side) of the text string.

Syntax= Left (text, num characters)

Right Function: - The right function in excel extract a specified number of characters from the beginning (right side) of the text string.

Syntax= Right (text, num characters)

Mid Function: - The mid function extracts a sub-string of a specified length from a middle of the text and it takes 3 arguments.

- The text string
- The starting position of the sub-string.
- The number of characters to extract.

Length Function (Len): - The length function returns number of characters in a text string including space and punctuation.

Find Function: - Find function locates the starting position of one text string within another and returns that position with a number. (Default Definition)

Syntax: - Find (Find Text, within text, start number)

Find function locates the starting position of one text string within another and returns that position with a number and the third argument states the starting position of the given text. (Complete)

Replace Function: - The replace function replaces a question of a string with another string.

Syntax: - Replace(Old text,start num, num of Chars, new text)

Substitute Function: - Substitute function replaces a specific sub- string within a new sub-string =Substitute (text, old text, new text)

Difference between replace and substitute?

In excel both replace and substitute replaces the string but they differ in the target the text to be changed.

Substitute replaces specific text apparenecies while replaced modifies the position of len

ASSIGNMENT-4

Advance Excel

Topic : Random (Rand) and Random Between, IF function

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 6th September

2025

Random Number Generation: - To generate a random number between 0 and 1. We use the function **Rand()**. It generates a random number between 0 and 1.

0.82164495 | RAND()

To generate a random number between 0 and 100 we will use the function **Rand()*100**.

28.95733692 | RAND()*100

To generate a random value between 20 and 100 we will use the function **20+80*rand()**.

75.40694052 | 20+80*RAND()

To generate a random number between a certain range we can use the function **Randbetween()** Or
A+rand()*(B-A)

42 RANDBETWEEN(20,100)

85.67261009 20+RAND()*100-20

Name	Excel	Python	SQL	Tableau
A	48	66	43	86
B	59	65	76	67
C	86	46	83	50
D	51	65	53	59
E	68	71	61	73
F	40	90	51	85

We use if function when we are checking for some condition.

IF has 3 arguments: -

- Condition
- Message if(condition is true)
- Message if(condition is false)

Name	Sales	Target	Status	Bonus	Q1. If sales is more than target status is target achieved or not achieved ?
Jasmine	580	640	Not Achieved	FALSE	Q2. If sales is more than the target bonus is 30% of the target?
Karan	1200	1000	Target Achieved	300	
Kent	710	710	Not Achieved	FALSE	
Formula Used					
IF(B19>C19,"Target Achieved","Not Achieved")					
IF(B19>C19,C19*30%,FALSE)					

Name	Salary	DA(15% of HRA(35% BASIC + Convene Taxable In PF(12%))	Tax	Net Income
A	4900	735	1715	490 7840 588 1568 5684 Income to be deducted as follow
B	3626	543.9	1269.1	362.6 5801.6 435.12 580.16 4786.32 No tax if taxable value is less than 3000
C	2450	367.5	857.5	245 3920 294 392 3234 10% tax if between 3000 and 6000
D	2842	426.3	994.7	284.2 4547.2 341.04 454.72 3751.44 20% tax if 6000 and 9000
E	12000	1800	4200	1200 19200 1440 5760 12000 30% tax if more than 9000
F	3525	528.75	1233.75	352.5 5640 423 564 4653
G	18350	2752.5	6422.5	1835 29360 2202 8808 18350
H	2842	426.3	994.7	284.2 4547.2 341.04 454.72 3751.44

ASSIGNMENT- 5

Advance Excel

Topic : IF function, IFS function, Date Conversion, Today and Now Function

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 9th September

2025

Q1) Find the commission based on the rule given below

ITEM CODE	UNITS	PRICE	SALES
100	1000	5.5	5500
101	1200	6.75	8100
102	1300	2	2600
103	800	4	3200
104	1200	5	6000
105	1800	0.5	900
106	145	2	290
107	1900	3.5	6650
108	1550	2.25	3487
109	2000	2	4000
110	2200	3	6600

SALES	COMMISSION
LESS THAN 1000	NO COMMISSION
1000 TO 5000	2% OF SALES
ABOVE 5000	5% OF SALES

Solution: -

ITEM CODE	UNITS	PRICE	SALES	COMMISSION BY IF	COMMISSION BY IFS
100	1000	5.5	5500	275	275
101	1200	6.75	8100	405	405
102	1300	2	2600	52	52
103	800	4	3200	64	64
104	1200	5	6000	300	300
105	1800	0.5	900	NO COMMISSION	NO COMMISSION
106	145	2	290	NO COMMISSION	NO COMMISSION
107	1900	3.5	6650	332.5	332.5
108	1550	2.25	3487	69.74	69.74
109	2000	2	4000	80	80
110	2200	3	6600	330	330

Find the Grade of the following Data

Sales	Grade
Less Than 1000	POOR
1000 to 5000	GOOD
Above 5000	Excellent

ITEM CODE	UNITS	PRICE	SALES	COMMISSION BY IF	COMMISSION BY IFS	GRADE
100	1000	5.5	5500	275	275	Excellent
101	1200	6.75	8100	405	405	Excellent
102	1300	2	2600	52	52	Good
103	800	4	3200	64	64	Good
104	1200	5	6000	300	300	Excellent
105	1800	0.5	900	NO COMMISSION	NO COMMISSION	Poor
106	145	2	290	NO COMMISSION	NO COMMISSION	Poor
107	1900	3.5	6650	332.5	332.5	Excellent
108	1550	2.25	3487	69.74	69.74	Good
109	2000	2	4000	80	80	Good
110	2200	3	6600	330	330	Excellent

Used If and IFS

IF is used for one logical test and IFS is used for many arguments.

Date Conversion

Find the date if date month year is given

DATE	MONTH	YEAR	DAY
09-09-2025		9	2025

DAY	MONTH	YEAR	DATE
5		6	2010 05-06-2010
9		9	2010 09-09-2010

We can find day month year by the date.

Now function gives current date and time.

For the date find year month date hour minute second

NOW	DAY	Month	YEAR	Hour	Minute	Seconds
09-09-2025 10:47	9	9	2025	10	47	14

Difference between today and now function

Today	Now
Today function only gives date	Now function gives date and time

ASSIGNMENT– 6

Advance Excel

Topic : Average, Minimum and Maximum, Median, Quartiles, Interquartile Range.

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 16th September 2025

mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
21	6	160	110	3.9	2.62	16.46	0	1	4	4
21	6	160	110	3.9	2.875	17.02	0	1	4	4
22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
18.7	8	360	175	3.15	3.44	17.02	0	0	3	2
18.1	6	225	105	2.76	3.46	20.22	1	0	3	1
14.3	8	360	245	3.21	3.57	15.84	0	0	3	4
24.4	4	146.7	62	3.69	3.19	20	1	0	4	2
22.8	4	140.8	95	3.92	3.15	22.9	1	0	4	2
19.2	6	167.6	123	3.92	3.44	18.3	1	0	4	4
17.8	6	167.6	123	3.92	3.44	18.9	1	0	4	4
16.4	8	275.8	180	3.07	4.07	17.4	0	0	3	3
17.3	8	275.8	180	3.07	3.73	17.6	0	0	3	3
15.2	8	275.8	180	3.07	3.78	18	0	0	3	3
10.4	8	472	205	2.93	5.25	17.98	0	0	3	4
10.4	8	460	215	3	5.424	17.82	0	0	3	4
14.7	8	440	230	3.23	5.345	17.42	0	0	3	4
32.4	4	78.7	66	4.08	2.2	19.47	1	1	4	1
30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
33.9	4	71.1	65	4.22	1.835	19.9	1	1	4	1
21.5	4	120.1	97	3.7	2.465	20.01	1	0	3	1
15.5	8	318	150	2.76	3.52	16.87	0	0	3	2
15.2	8	304	150	3.15	3.435	17.3	0	0	3	2
13.3	8	350	245	3.73	3.84	15.41	0	0	3	4
19.2	8	400	175	3.08	3.845	17.05	0	0	3	2

27.3	4	79	66	4.08	1.935	18.9	1	1	4	1
26	4	120.3	91	4.43	2.14	16.7	0	1	5	2
30.4	4	95.1	113	3.77	1.513	16.9	1	1	5	2
15.8	8	351	264	4.22	3.17	14.5	0	1	5	4
19.7	6	145	175	3.62	2.77	15.5	0	1	5	6
15	8	301	335	3.54	3.57	14.6	0	1	5	8
21.4	4	121	109	4.11	2.78	18.6	1	1	4	2

Q) what is the average mpg of all the cars?

=AVERAGE(A2:A33)

20.090625

Q) which car has the maximum horsepower (hp)?

=MAX(D2:D33)

335

Q) what is the minimum and maximum width of cars?

=MAX(A2:K33) 472

=MIN(A2:K33) 0

Q) find the median displacement

=MEDIAN(A2:K33)

4

Q) find the first second third quartiles of the column mpg

First Quartile =QUARTILE.EXC(A2:A33,1) = 15.275

Second Quartile =QUARTILE.EXC(A2:A33,2) = 19.2

Third Quartile =QUARTILE.EXC(A2:A33,3) = 22.8

Quartiles are the values which divide data into 4 equal parts

Q1 is below 25% and above 75%

Q2 is also known as median below which 50% and above is also 50% Q3 is also known as third quartile below which 75%.

IQR - it is the difference between first and third quartile i.e. = Q3-Q1

If no of values in the data set is given, then median is the average of two middle values.

$$=\text{SUM(P17:P18)}/2 = 196.3$$

Q) find the number of cars which has 4 cylinders

cyl
6
6
4
6
8
6
8
4
4
6
6
8
8
8
8
8
8
4
4
4
4
8
8
8
8
4
4
4
4
8
6
8
4

11 Cylinders

Formula used - COUNTIF(B2:B33,4) ASSIGNMENT – 7

Advance Excel

Topic : Correlation, Scatterplot, Pie and Bar Chart, Slope, Intercept, Percentile.

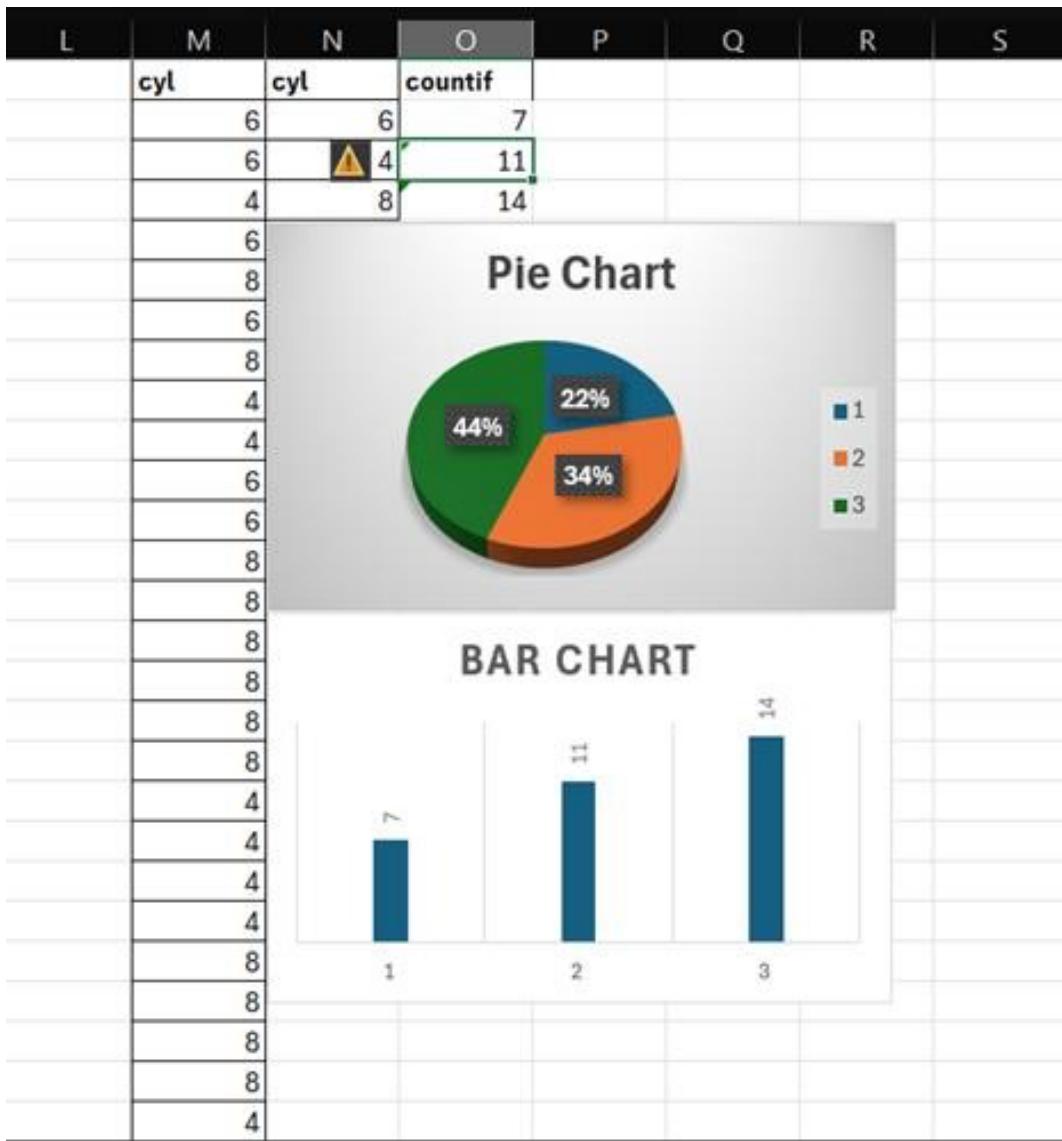
Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 23rd September 2025

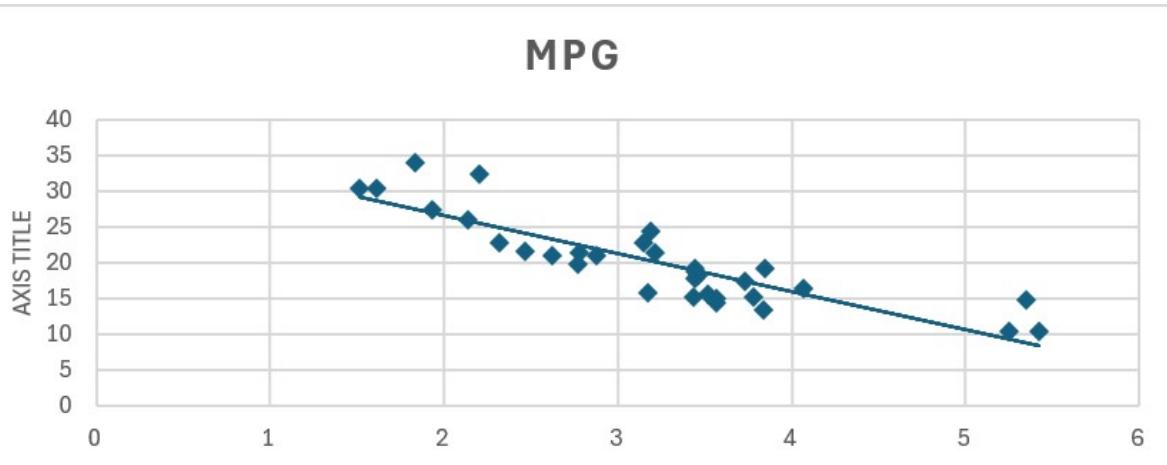
	A	B	C	D	E	F	G	H	I	J	K
1	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
2	21	6	160	110	3.9	2.62	16.46	0	1	4	4
3	21	6	160	110	3.9	2.875	17.02	0	1	4	4
4	22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
5	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
6	18.7	8	360	175	3.15	3.44	17.02	0	0	3	2
7	18.1	6	225	105	2.76	3.46	20.22	1	0	3	1
8	14.3	8	360	245	3.21	3.57	15.84	0	0	3	4
9	24.4	4	146.7	62	3.69	3.19	20	1	0	4	2
10	22.8	4	140.8	95	3.92	3.15	22.9	1	0	4	2
11	19.2	6	167.6	123	3.92	3.44	18.3	1	0	4	4
12	17.8	6	167.6	123	3.92	3.44	18.9	1	0	4	4
13	16.4	8	275.8	180	3.07	4.07	17.4	0	0	3	3
14	17.3	8	275.8	180	3.07	3.73	17.6	0	0	3	3
15	15.2	8	275.8	180	3.07	3.78	18	0	0	3	3
16	10.4	8	472	205	2.93	5.25	17.98	0	0	3	4
17	10.4	8	460	215	3	5.424	17.82	0	0	3	4
18	14.7	8	440	230	3.23	5.345	17.42	0	0	3	4
19	32.4	4	78.7	66	4.08	2.2	19.47	1	1	4	1
20	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
21	33.9	4	71.1	65	4.22	1.835	19.9	1	1	4	1
22	21.5	4	120.1	97	3.7	2.465	20.01	1	0	3	1
23	15.5	8	318	150	2.76	3.52	16.87	0	0	3	2
24	15.2	8	304	150	3.15	3.435	17.3	0	0	3	2
25	13.3	8	350	245	3.73	3.84	15.41	0	0	3	4
26	19.2	8	400	175	3.08	3.845	17.05	0	0	3	2
27	27.3	4	79	66	4.08	1.935	18.9	1	1	4	1

Q) Find the number of cars which has 8 cylinders, also find the frequency distribution on no. of cars based on the number of cylinders also make the bar and the pie chart.



Q) find the correlation between mpg and wt also make the scatter plot between mpg and wt and show the trendline the scatterplot, find the value of mpg if wt is 4.2.

T	U
wt	mpg
2.62	21
2.875	21
2.32	22.8
3.215	21.4
3.44	18.7
3.46	18.1
3.57	14.3
3.19	24.4
3.15	22.8
3.44	19.2
3.44	17.8
4.07	16.4
3.73	17.3
3.78	15.2
5.25	10.4
5.424	10.4
5.345	14.7
2.2	32.4
1.615	30.4
1.835	33.9
2.465	21.5
3.52	15.5
3.435	15.2
3.84	13.3
3.845	19.2
1.935	27.3
2.14	26
1.513	30.4
3.17	15.8
2.77	19.7
3.57	15
2.78	21.4



We make scatter plot between two variables to show the linear relationship between them

Trendline is given by $y=mx+c$ where y = dependent variable = mpg, x = independent variable = wt, m = slope, c = constant or intercept
 Formula for slope
 $=SLOPE(\text{Sheet1!A1:A33}, \text{Sheet1!F1:F33})$

-5.344471573

Formula for intercept =INTERCEPT(Sheet1!A1:A33,Sheet1!F1:F33)

37.28512617

The trendline is given by mpg =-5.344471573*wt+37.28512617

What is the correlation between mpg and wt?

Formula used =CORREL(Sheet1!B:B,Sheet1!G:G)

-0.591242074

Q) What is the interquartile range for the column wt ?

Formula used =QUARTILE.EXC(F1:F33,1)

2.50375, Q1

3.325,Q2

3.69,Q3

1.18625,IQR (Q3-Q1)

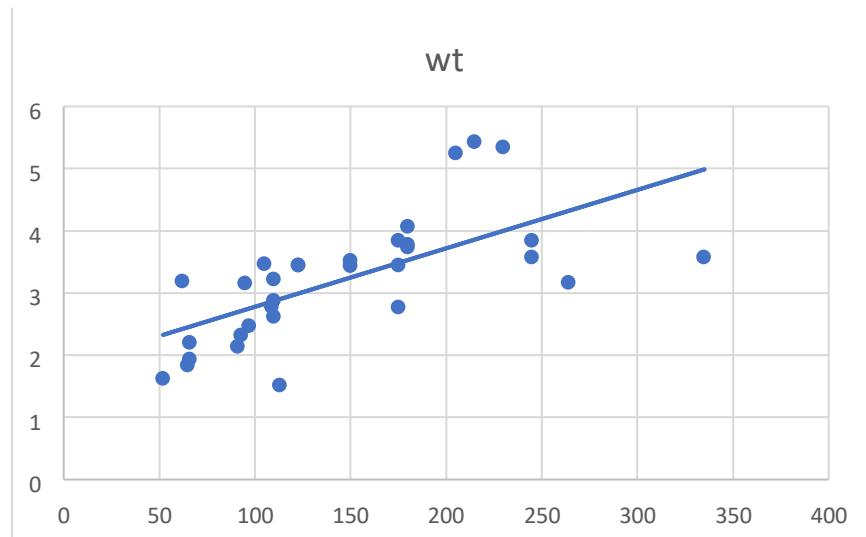
Q) What is the 90th percentile for the hp?

Formula used =PERCENTILE.EXC(D1:D33,0.9)

245,HP

Q) make the scatterplot between hp and wt, where weight is the independent variable also find the trendline between weight and hp

hp	wt
110	2.62
110	2.875
93	2.32
110	3.215
175	3.44
105	3.46
245	3.57
62	3.19
95	3.15
123	3.44
123	3.44
180	4.07
180	3.73
180	3.78
205	5.25
215	5.424
230	5.345
66	2.2
52	1.615
65	1.835
97	2.465
150	3.52
150	3.435
245	3.84
175	3.845
66	1.935
91	2.14
113	1.513
264	3.17
175	2.77
335	3.57
109	2.78



Q) find the no. of cars which has 4 gears, 4 cylinders, and mpg more than avg mpg.

	A
1	mpg
2	21
3	21
4	22.8
5	21.4
6	18.7
7	18.1
8	14.3
9	24.4
10	22.8
11	19.2
12	17.8
13	16.4
14	17.3
15	15.2
16	10.4
17	10.4
18	14.7
19	32.4
20	30.4
21	33.9
22	21.5
23	15.5
24	15.2
25	13.3
26	19.2
27	27.3
28	26
29	30.4
30	15.8
31	19.7
32	15
33	21.4

Formula used =AVERAGE(A1:A33)

20.090625

Formula used =COUNTIFS(J1:J33,4, B1:B33,4, A1:A33,>"&W58)

8

Assignment -8

Advance Excel

Topic : Numerical Functions, Average, Quartiles, Unique Values, Median.

Name: Devashish Kaushik

Roll no.: 2402208016

Date: - 30th September 2025

1.What is the total number of transactions?

Total Transaction

6146 COUNT(N2:N6147)

2.What is the total transaction amount?

Total Transaction Amount

88179698 SUM(N2:N6147)

3.What is the average transaction amount?

Average Transaction Amount

14347.49 AVERAGE(N2:N6147)

4.what is the maximum transaction amount?

Max Transaction Limit

151223 MAX(N2:N6147)

5. what is the minimum transaction amount

Min Transaction Limit

0 MIN(N2:N6147)

6.what is the avg transaction amount per year

Average Transaction of Per Year		Average	
Row Labels	Sum of Replace		
2002	21340	4641037	AVERAGE(R24:R42)
2003	138756		
2004	144116		
2005	635331		
2006	1185342		
2007	2725680		
2008	4389541		
2009	6055199		
2010	10541065		
2011	12028881		
2012	6718157		
2013	6669010		
2014	5933882		
2015	4598449		
2016	3275638		
2017	2237044		
2018	2280623		
2019	1698816		
2020	16902828		
Grand Total	88179698		

7.how many transaction occurred in 2020

Transaction occurred in 2020	
Row Labels	Count of Replace
2002	2
2003	11
2004	12
2005	36
2006	73
2007	177
2008	303
2009	456
2010	712
2011	809
2012	477
2013	465
2014	397
2015	341
2016	234
2017	162
2018	156
2019	115
2020	1208
Grand Total	6146

8. what is the total amount of card type

Total Number of card Type	
Debit	3511
Credit	2057
Debit (Prepaid)	578

9. what is the proportion of credit card vs debit card

Proportion of Credit and Debit Card		
Total	5568	
Debit	3511	63%
Credit	2057	37%

10 what is the avg debit amt vs credit amt

Average Debit Vs Credit	
Row Labels	Average of Replace
Credit	11174.38017
Debit	18557.88864
Debit (Prepaid)	64.44809689
Grand Total	14347.49398

11 identify the top five highest transactions

Top 5 Transaction

151223
141391
137669
132439
130971

12 identify the bottom five transactions

Bottom 5
0

13 what is the median transaction amt

Median Transaction
12592.5

14 how many transaction are above avg amt

15 how many transactions are below first quartile

15	Transaction below 1st Quartile	7026
16	Q1	7026
	Q2	12591
	Q3	19140.75
	IQR	12114.75

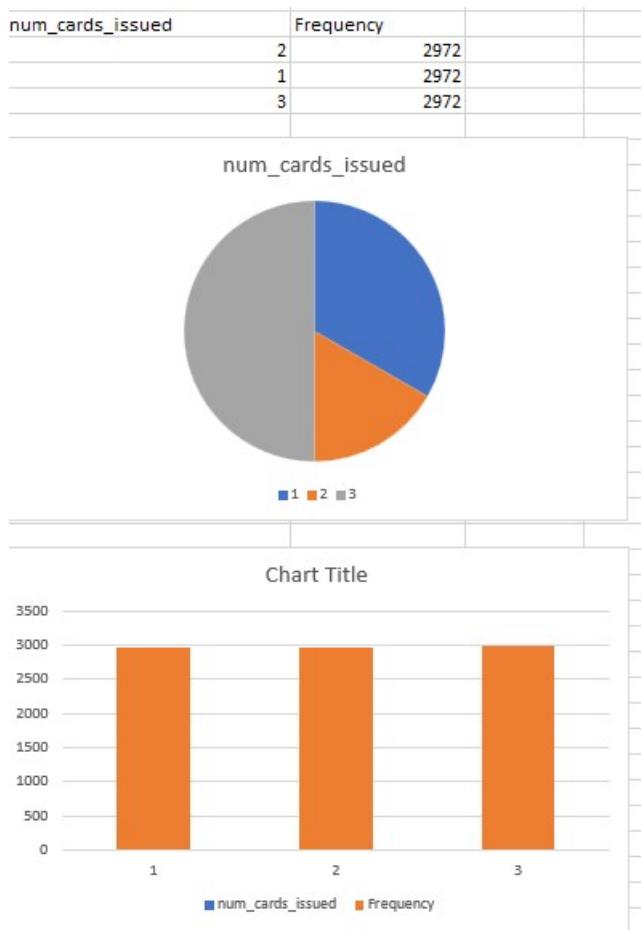
16 find the quartiles of transaction amt and also find iqr also find the number of transaction which we can consider as outliers

Q1	7026
Q2	12591
Q3	19140.75
IQR	12114.75
55485 Outliers	

17 how many cards are of debit rank, mastercard, issued year 2019 , limit is more thn avg limit.

18 find the number of cards which expeired in 2023

19 find the unique values of the column number of cards issued also find frequence distribution of each category also make bar and the pie chart. Find the percentage of customer who has 2cards.



20.in which year the total credit limit is maximum

REPLACE	151223	<input type="button" value="▼"/>
Row Labels		<input type="button" value="▼"/>
2009		

Assignment – 9

Advance Excel

Topic : Rank and Percentile.

Name :Devashish Kaushik

Roll no.: 2402208016

Date: - 7th October 2025

Rank Vs Percentile

Difference Between Percentage and Percentile?

Percentage	Percentile
It is calculated out of the particular observation	It is calculated on the overall observation

Consider the source in a school examination

Sr No	Score of CBSE Exam	Point	Score of CBSE Exam	Rank	Percent
1	354	4	575	1	100.00%
2	265	5	555	2	83.30%
3	463	7	476	3	66.60%
4	575	6	465	4	50.00%
5	555	3	463	5	33.30%
6	465	1	354	6	16.60%
7	476	2	265	7	0.00%

Naresh is in the HR Department of a Software Company ABC Soft. Total 30 Candidate appear for the interview ABC soft is recruiting 5 candidates in this intake based on 2 rounds of assessment. The 1st round is a Written test where top 15 scoring candidate are alone shortlisted and the next round is a personal interview where candidates are screen and the top 5 candidates are selected for recruitment. If 2 people score the same they are shortlisted and accommodated even if its exceed the target of 15 candidates at the end of 1st round. The 2nd

round is used to shortlist them to 5 candidates. Naresh is faced with a decision on how to assess these candidates he concerns her colleagues who is from the analytics department to help him in fact his colleague almost accompany Naresh for recruitment to help him with these works, He decides to feed the scores in excel spreadsheet to access them Naresh now decides

Candidates Number	Score of 1sst Round	Marks of Second Round	Point	Score of 1sst Round	Rank	Percent
1	78		26	95	1	100.00%
2	56		18	83	2	96.50%
3	57		16	80	3	93.10%
4	47		7	79	4	82.70%
5	68		11	79	4	82.70%
6	58		15	79	4	82.70%
7	79		1	78	7	72.40%
8	68		17	78	7	72.40%
9	69		30	78	7	72.40%
10	58		14	70	10	62.00%
11	79		20	70	10	62.00%
12	40		25	70	10	62.00%
13	50		9	69	13	51.70%
14	70		19	69	13	51.70%
15	79		23	69	13	51.70%
16	80		5	68	16	44.80%
17	78		8	68	16	44.80%
18	83		29	64	18	41.30%
19	69		28	60	19	37.90%
20	70		27	59	20	34.40%
21	40		6	58	21	27.50%
22	30		10	58	21	27.50%
23	69		3	57	23	24.10%
24	48		2	56	24	20.60%
25	70		13	50	25	17.20%
26	95		24	48	26	13.70%
27	59		4	47	27	10.30%
28	60		12	40	28	3.40%
29	64		21	40	28	3.40%
30	78		22	30	30	0.00%

Since we need to shortlist top 15 candidate out of 30 candidate which scores are above 50% are selected								
Point	Score of 1sst Round	Rank	Percent	Marks of Second Round	Point	Marks of Second Round	Rank	Percent
26	95	1	100.00%	85	1	85	1	100.00%
18	83	2	96.50%	76	11	79	2	92.80%
16	80	3	93.10%	58	2	76	3	85.70%
7	79	4	82.70%	67	6	69	4	78.50%
11	79	4	82.70%	57	7	68	5	64.20%
15	79	4	82.70%	69	14	68	5	64.20%
1	78	7	72.40%	68	4	67	7	57.10%
17	78	7	72.40%	47	3	58	8	42.80%
30	78	7	72.40%	57	10	58	8	42.80%
14	70	10	62.00%	58	5	57	10	28.50%
20	70	10	62.00%	79	9	57	10	28.50%
25	70	10	62.00%	40	13	56	12	21.40%
9	69	13	51.70%	56	8	47	13	14.20%
19	69	13	51.70%	68	12	40	14	7.10%
23	69	13	51.70%	35	15	35	15	0.00%
Point	Marks of Second Round	Rank	Percent					
1	85	1	100.00%					
11	79	2	92.80%					
2	76	3	85.70%					
6	69	4	78.50%					

Consider the score of the contest Rank them and percentile score

<i>Score of Student in a Contest</i>	<i>Point</i>	<i>Score of Student in a Contest</i>	<i>Rank</i>	<i>Percent</i>
320	7	482	1	100.00%
215	6	475	2	83.30%
223	4	472	3	66.60%
472	5	421	4	50.00%
421	1	320	5	33.30%
475	3	223	6	16.60%
482	2	215	7	0.00%

Advance Excel

Topic : Random and Sampling

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 14th September 2025

Random and Sampling

Sampling is the process of selecting a sub-set (sample) from a larger group of population together in a group and draw conclusion about whole population that you want to study is called population. Sampling refers to a part of population selecting for analysis. Sampling frame refers to a list or data base from which a sample is drawn.

Importance - it saves time and cost, compare to study the whole population. It enables faster decision making. It allows researcher to make reference about whole population. It reduces data overload and makes analysis more manageable.

Types - probability and non-probability sampling .

Probability sampling - every unique in the population has a known and non-zero.

- 1) simple random sampling - every element has an equal chance of being selected.
- 2) systematic sampling - select every Kth element after a random start.
- 3) stratified sampling - population/into groups (strata) and random sample taken from each strata.
- 4) Cluster Sampling - population divided into clusters and some clusters are randomly selected.

2) Non probability Sampling - selection is based on judgement, convenience or other non-random factors.

- 1) Convenience sampling - choosing readily available correspondence.

- 2) Judgemental (purposive sampling) - select correspondence based on researchers' expertise.
- 3) Quota Sampling - setting quotas for sub-groups.

4) Snowfall Sampling - existing respond ended refer other potential responded.

Characteristics of a good sample design –

- 1) representativeness - accurately represents the research.
- 2) Adequate size - large enough to ensure liability.
- 3) practibility - feasible in terms of time and cost.
- 4) lack of bias - free from personal or selection bias.
- 5) Homogeneity - captures the diversity within the population.

Steps –

Define the target population.

Determine the sampling frame.

Select the sampling technique. Design
the sample size.

Execute the sampling process.

Evaluate the sampling error and bias.

Questions -

Q1) Consider the population of customers served by a small electronic service company on each day. Find the simple random sample and periodic sample of size 5.

SOL1) No. of customers per day	Random Sampling
12	47
33	10
23	45
34	47
43	
12	
32	
54	
55	
15	
14	
24	
34	
54	
23	
40	
35	
25	
27	
46	

Q2) Ajay is the head of the inspection department of an automated company ABC auto. His job is to make sure that the parts produced adhere to the specifications required. Every day he meets his subordinates and gives them instruction on how to inspect the products. While inspecting it is a general custom that the entire population of products is not inspected. Instead, a sample of the population is inspected, and the results are analysed to arrive at the population characteristics. On the previous day he was assigned a work of inspecting 20 crankshafts. He decided to analyse a sample of 8 crankshafts and arrive at the results for the population. All the crankshafts in the population are assigned and equal probability of being selected. To select a sample randomly he uses the simple random sampling. He can also sample every second sample or unit so that he has 10 samples. He decided to use both techniques and compare this research. Use simple random sampling to take sample for inspection of parts. Use systematic random sampling to take sample for inspection of parts.

SOL2) Part No.	Length of the parts	Random Sampling	Systematic Sampling
1	45	18	10
2	44	44	20
3	43	44	46
4	47	19	47
5	46	9	
6	44	46	
7	45	14	
8	43	47	
9	45	45	
10	46	43	
11	47		
12	45		
13	44		
14	43		
15	45		
16	43		
17	44		
18	45		
19	44		
20	47		

Assignment – 11

Advance Excel

Topic : Regression

Name :- Devashish Kaushik

Roll no.: - 2402208016

Date: - 4th November 2025

Regression

Regression is a statistical technique where we have a dependent and one or more independent variables where we are trying to predict variable with the help of independent variable.

When we have 1 independent and one dependent variable it is called simple linear regression. When we have dependent(y) and independent (x). Simple linear regression is given by $y = mx + c + e$, where y is the dependent variable and x is an independent variable, m is slope, c is intercept and e is error/residual.

C= Intercept which is equal to minimum value of 5 which is equal to 0.

M= Slope which is equal to regression coefficient which is change in y and when there is unit change in x.

M and c are also called parameters.

We have 2 types of models: -

- Deterministic
- Stochastic

O

L

S

To find the parameter of linear regression model we use OLS method (ordinary least square method)

Error or residual is a difference between actual y and predicted y

SLR = 2 parameters

MLS = (n+1) parameter

Slope can be negative or positive. If the slope is negative then both the variable x and y will move in same direction. If slope is negative then both the variable will move in opposite direction.

Multiple linear regression is given by the

$$Y = c + m_1x_1 + m_2x_2 + \dots + m_nx_n + e$$

C is the intercept which is the minimum value of y when all xis are 0.

M1, M2, M3... are called as parameter or regression coefficient.

M1 is the regression coefficient of X1 which is the change in y when there is a unit change in x1 keeping all other variable constant.

$$\text{Fare} = 1000 + 10 * \text{distance} + 200 * \text{toll} + e \quad Y_1 = 10;$$

$$X_2 = 2 \quad Y = ?$$

$$Y_1 = 11; \quad X_2 = 2; \quad Y = ?$$

1500	$1000 + 10 * 10 + 200 * 2$
1510	$1000 + 10 * 11 + 200 * 2$

The change in Between 1510 and 1500 is of 10 and it would be the M1 of the Data.

The goal of MR is to maximize the correlation between dependent and independent variable and minimize the correlation between the procedure. A procedure used to find the regression coefficient is minimizing the sum of square Alan which is called ordinary least square method.

R=

R^2 = dependent between 0 to 1 and dependent and independent variable
 Coefficient of determination it tells how good the model is for prediction. High the value of the R^2 .

MSE (Mean square error): - it is the mean of square errors. MSE should be minimum

Q) For the given data find if there is a relationship between quantity sold, price and advertising.

Qty sold	Price	Advertising
8500	2	2800
4700	5	200
5800	3	400
7400	2	500
6200	5	3200
7300	3	1800
5600	4	1900

Q) Find the qty sold if the price is 4 and advertising is 3000

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.980681431							
R Square	0.961736068							
Adjusted R Square	0.942604102							
Standard Error	310.5239249							
Observations	7							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	2	9694299.568	4847149.784	50.26854403	0.001464128			
Residual	4	385700.4318	96425.10794					
Total	6	10080000						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	8536.213882	386.9117478	22.06243137	2.49812E-05	7461.974654	9610.453111	7461.974654	9610.453111
Price	-835.722351	99.65304469	-8.386320297	0.001106064	-1112.40356	-559.041143	-1112.40356	-559.041143
Advertising	0.592228496	0.104346803	5.675578729	0.004755309	0.302515325	0.881941666	0.302515325	0.881941666

For the given data R is .98 (+ve correlation between x and y)

For the given data R square is 0.96 which is way too high the model is good

C = 8536.21

M1 = Coefficient of price = -835.72

M2= Coefficient of Advertising= .59

It is linear in parameter not in variables.

Sum of the errors is always zero

If r is zero which we cannot use predict