



K. S. SCHOOL OF ENGINEERING AND MANAGEMENT

DEPARTMENT OF CIVIL ENGINEERING

BCV358A – DATA ANALYTICS WITH EXCEL

LABORATORY MANUAL

STUDENT NAME: _____

USN: _____



K S SCHOOL OF ENGINEERING AND MANAGEMENT

Holiday Village Road, Vajarahalli Village, Mallasandra, off, Kanakapura Rd,
Bengaluru, Karnataka 560109

VISION

To impart quality education in engineering and management to meet technological business and societal needs through holistic education and research.

MISSION

K. S. School of Engineering and Management shall,

- Establish state-of-art infrastructure to facilitate effective dissemination of technical and managerial knowledge.
- Provide comprehensive educational experience through a combination of curricular and experiential learning, strengthened by industry-institute interaction.
- Pursue socially relevant research and disseminate knowledge.
- Inculcate leadership skills and foster entrepreneurial spirit among students.

DEPARTMENT OF CIVIL ENGINEERING

VISION

- To emerge as one of the leading Civil Engineering Department by producing competent and quality ethical engineers with strong foot hold in the areas of Infrastructure development and research.

MISSION

- Provide industry oriented academic training with strong fundamentals and applied skills.
- Engage in research activities in Civil Engineering and allied fields and inculcate the desired perception and value system in the students.

CONTENTS

EXPT. No.	NAME OF THE EXPERIMENT	PAGE No.
01	Introduction to Data Analysis Using Spreadsheet	01
02	Perform Basic Spreadsheets Tasks	04
03	Cleaning & Wrangling Data Using Spreadsheets	09
04	Flash Fill and Text to Columns	25
05	Analyzing Data Using Spreadsheets	31
06	VLOOKUP and HLOOKUP	34
07	Final Project	37

Experiment No: 01**Date:****INTRODUCTION TO ANALYSIS USING SPREADSHEET****What is Excel?**

Excel is pronounced "Eks - sel"

It is a spreadsheet program developed by Microsoft. Excel organizes data in columns and rows and allows you to do mathematical functions. It runs on Windows, macOS, Android and iOS.

The first version was released in 1985 and has gone through several changes over the years. However, the main functionality mostly remains the same.

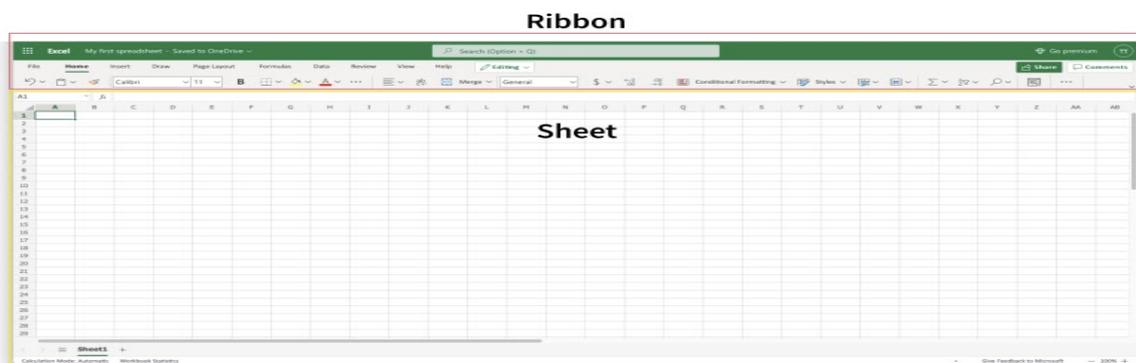
Excel is typically used for:

- Analysis
- Data entry
- Data management
- Accounting
- Budgeting
- Data analysis
- Visuals and graphs
- Programming
- Financial modeling
- And much, much more!

Why Use Excel?

- It is the most popular spreadsheet program in the world
- It is easy to learn and to get started.
- The skill ceiling is high, which means that you can do more advanced things as you become better
- It can be used with both work and in everyday life, such as to create a family budget
- It has a huge community support
- It is continuously supported by Microsoft
- Templates and frameworks can be reused by yourself and others, lowering creation costs

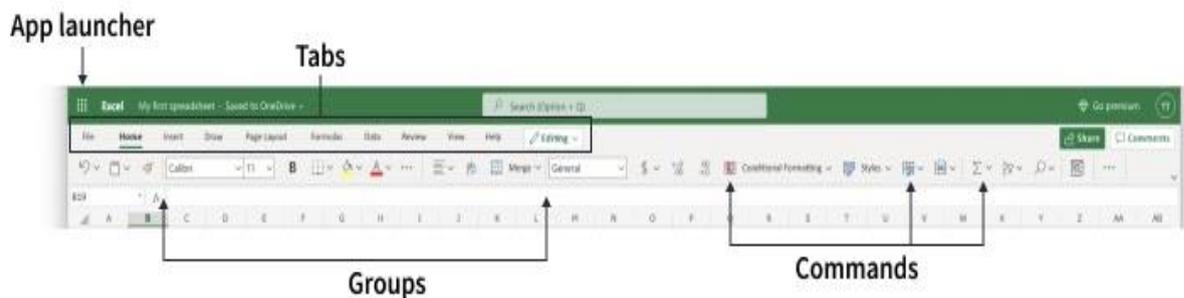
The Ribbon is marked with a red rectangle and the Sheet is marked with a yellow rectangle:



The Ribbon provides shortcuts to Excel commands. A command is an action that allows you to make something happen. This can for example be to: insert a table, change the font size, or to change the color of a cell.

The Ribbon may look crowded and hard to understand at first. Don't be scared, It will become easier to navigate and use as you learn more. Most of the time we tend to use the same functionalities over again.

The Ribbon is made up by the App launcher, Tabs, Groups and Commands. In this section we will explain the different parts of the Ribbon.



App launcher

The App launcher icon has nine dots and is called the Office 365 navigation bar. It allows you to access the different parts of the Office 365 suite, such as Word, PowerPoint and Outlook. App launcher can be used to switch seamlessly between the Office 365 applications.

Tabs

The tab is a menu with sub divisions sorted into groups. The tabs allow users to quickly navigate between options of menus which display different groups of functionality.

Groups

The groups are sets of related commands. The groups are separated by the thin vertical line break.

Commands

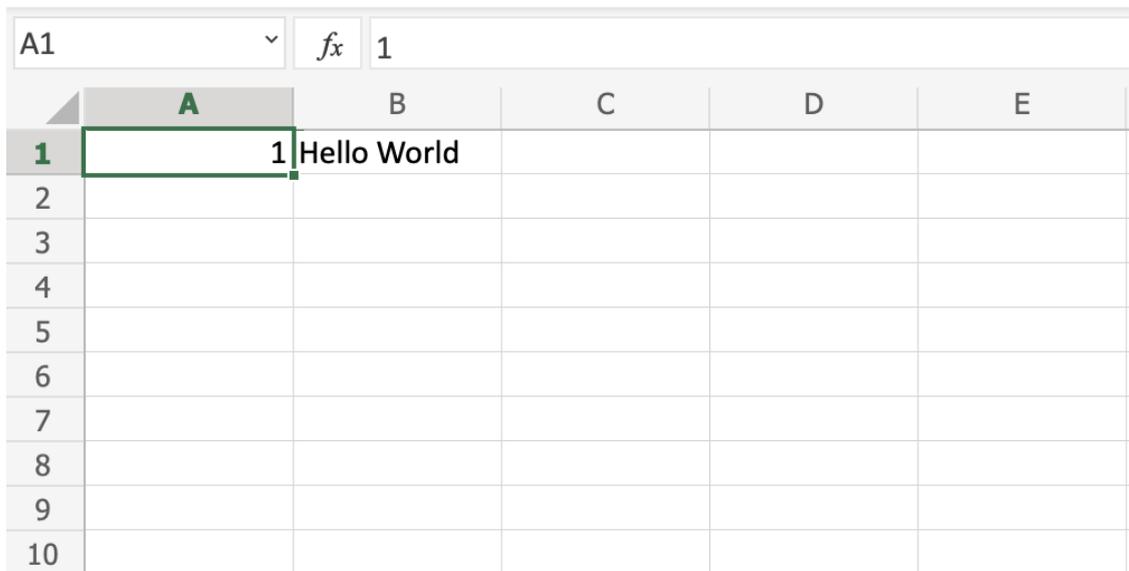
The commands are the buttons that you use to do actions.

Now, let's have a look at the Sheet. Soon you will be able to understand the relationship between the Ribbon and the Sheet, and you can make things happen.

The Sheet is a set of rows and columns. It forms the same pattern as we have in math exercise books, the rectangle boxes formed by the pattern are called cells.

Values can be typed to cells.

Values can be both numbers and letters:



The image shows a screenshot of an Excel spreadsheet. The Name Box at the top left displays 'A1', the Formula Bar shows 'fx' and '1'. The spreadsheet grid has columns A through E and rows 1 through 10. Cell A1 is selected and contains the number '1'. Cell C4 contains the text 'Hello World'.

	A	B	C	D	E
1	1				
2					
3					
4			Hello World		
5					
6					
7					
8					
9					
10					

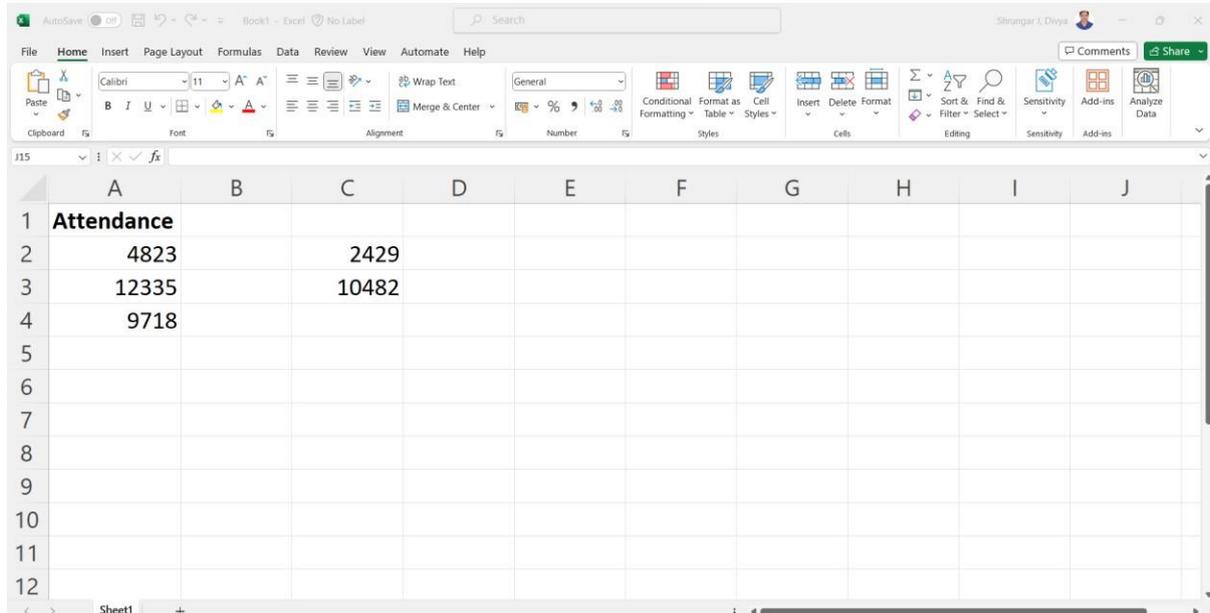
Have a look at the picture below. Hello world was typed in cell C4. The reference can be found by clicking on the relevant cell and seeing the reference in the Name Box to the left, which tells you that the cell's reference is C4.

Experiment No: 02

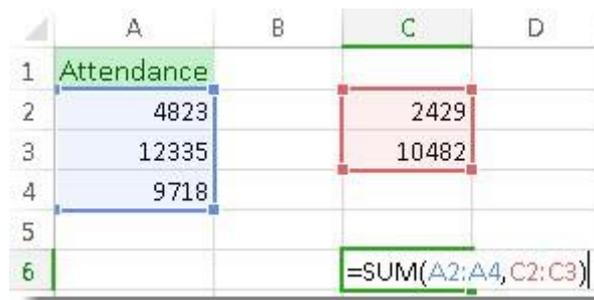
Date:

PERFORM BASIC SPREADSHEET TASKS

Exercise 1: Determine the Sum of the Values of the Given numbers.

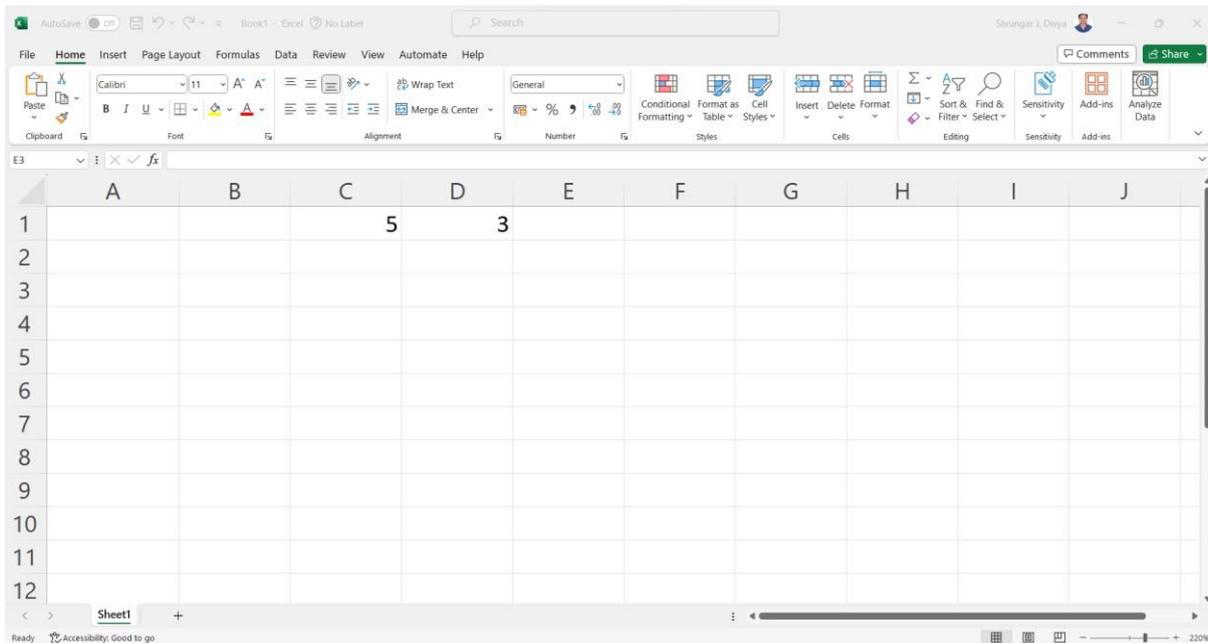


Solution :



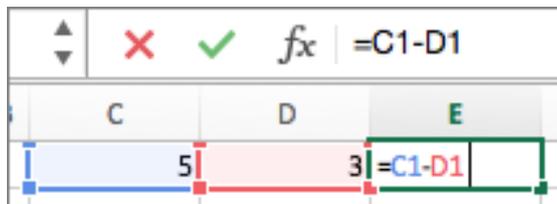
Formula that uses two cell ranges: **=SUM(A2:A4,C2:C3)** sums the numbers in ranges A2:A4 and C2:C3. Press Enter to get the total of 39787.

Exercise 2: Determine the Subtract of the Values of the Given numbers.



Solution:

1. Type a number in cells C1 and D1.
2. In cell E1, type an equal sign (=) to start the formula.
3. After the equal sign, type **C1-D1**.



Exercise 3: Determine the Average of the Values of the Given numbers.

	A	B	C	D	E	F
1	Trainer	Pokeball	Great ball	Ultraball	Master ball	Average
2	Iva	10	4	1	1	
3	Liam	12	3	0	1	
4	Jenny	15	1	3	1	
5	Iben	4	2	6	0	
6	Adora	10	4	1	1	
7	Kasper	9	2	1	0	

Solution :

	A	B	C	D	E	F	G	H
1	Trainer	Pokeball	Great ball	Ultraball	Master ball	Average		
2	Iva	10	4	1	1	=AVERAGE(
3	Liam	12	3	0	1	AVERAGE (number1; [number2]; ...)		
4	Jenny	15	1	3	1			
5	Iben	4	2	6	0			
6	Adora	10	4	1	1			
7	Kasper	9	2	1	0			

	A	B	C	D	E	F	G	H
1	Trainer	Pokeball	Great ball	Ultraball	Master ball	Average		
2	Iva	10	4	1	1	=AVERAGE(B2:E2		
3	Liam	12	3	0	1	AVERAGE (number1; [number2]; ...)		
4	Jenny	15	1	3	1			
5	Iben	4	2	6	0			
6	Adora	10	4	1	1			
7	Kasper	9	2	1	0			

	A	B	C	D	E	F	G	H
1	Trainer	Pokeball	Great ball	Ultraball	Master ball	Average		
2	Iva	10	4	1	1	4		
3	Liam	12	3	0	1			
4	Jenny	15	1	3	1			
5	Iben	4	2	6	0			
6	Adora	10	4	1	1			
7	Kasper	9	2	1	0			

F2						f_x =AVERAGE(B2:E2)			
	A	B	C	D	E	F	G	H	
1	Trainer	Pokeball	Great ball	Ultraball	Master ball	Average			
2	Iva	10	4	1	1	4			
3	Liam	12	3	0	1	4			
4	Jenny	15	1	3	1	5			
5	Iben	4	2	6	0	3			
6	Adora	10	4	1	1	4			
7	Kasper	9	2	1	0	3			

Exercise 4: Determine the Records of Apple in the given data.

2	Sales Transactions			
3				
4	Region	Sales Rep	Product	Units
5	East	Tom	Apple	6,380
6	West	Fred	Grape	5,619
7	North	Amy	Pear	4,565
8	South	Sal	Banana	5,323
9	East	Fritz	Apple	4,394
10	West	Sravan	Grape	7,195
11	North	Xi	Pear	5,231
12	South	Hector	Banana	2,427
13	East	Tom	Banana	4,213
14	West	Fred	Pear	3,239
15	North	Amy	Grape	6,420
16	South	Sal	Apple	1,310
17	East	Fritz	Banana	6,274
18	West	Sravan	Pear	4,894
19	North	Xi	Grape	7,580
20	South	Hector	Apple	9,814

Solutions:

In the following example we used the formula =FILTER(A5:D20,C5:C20=H2,"") to return all records for Apple, as selected in cell H2, and if there are no apples, return an empty string ("").

F5									
	A	B	C	D	E	F	G	H	I
1									
2	Sales Transactions						Product:	Apple	
3									
4	Region	Sales Rep	Product	Units	Region	Sales Rep	Product	Units	
5	East	Tom	Apple	6,380	East	Tom	Apple	6,380	
6	West	Fred	Grape	5,619	East	Fritz	Apple	4,394	
7	North	Amy	Pear	4,565	South	Sal	Apple	1,310	
8	South	Sal	Banana	5,323	South	Hector	Apple	9,814	
9	East	Fritz	Apple	4,394					
10	West	Sravan	Grape	7,195					
11	North	Xi	Pear	5,231					
12	South	Hector	Banana	2,427					
13	East	Tom	Banana	4,213					
14	West	Fred	Pear	3,239					
15	North	Amy	Grape	6,420					
16	South	Sal	Apple	1,310					
17	East	Fritz	Banana	6,274					
18	West	Sravan	Pear	4,894					
19	North	Xi	Grape	7,580					
20	South	Hector	Apple	9,814					

Experiment No: 03

Date:

CLEANING AND WRANGLING DATA USING SPREADSHEETS

Remove Duplicates

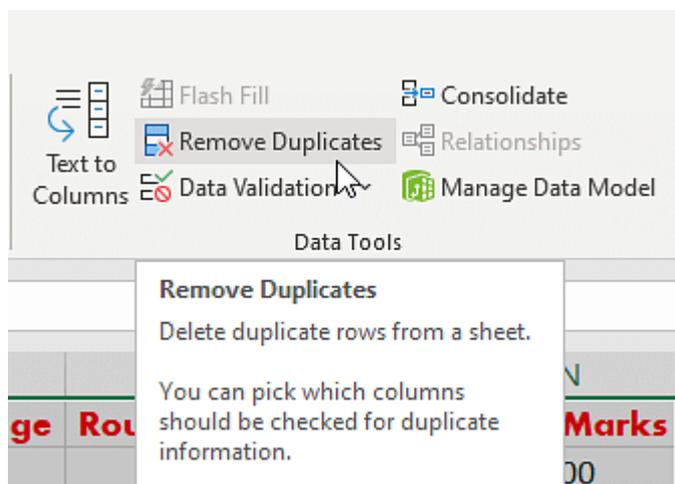
One of the easiest ways of cleaning data in Excel is to remove duplicates. There is a considerable probability that it might unintentionally duplicate the data without the user's knowledge. In such scenarios, you can eliminate duplicate values.

Here, you will consider a simple student dataset that has duplicate values. You will use [Excel's built-in function](#) to remove duplicates, as shown below.

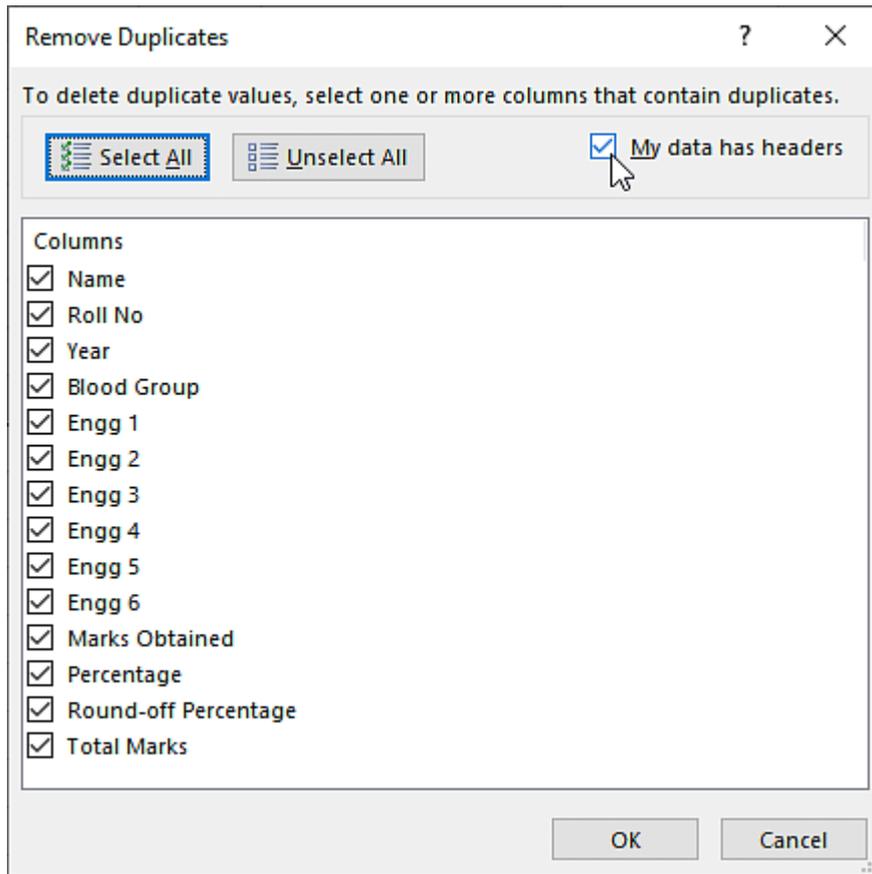
The original dataset has two rows as duplicates. To eliminate the duplicate data, you need to select the data option in the toolbar, and in the Data Tools ribbon, select the "Remove Duplicates" option. This will provide you with the new dialogue box, as shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Name	Roll No	Year	Blood Group	Engg 1	Engg 2	Engg 3	Engg 4	Engg 5	Engg 6	Marks Obtained	Percentage	Round-off Percentage	Total Marks
2	Joe	1011	IV	O+ve	32	65	86	88	79	79	429	71.5	71.5	600
3	John	2011	IV	B+ve	65	76	92	75	65	85	458	76.33333333	76.4	600
4	Mary	1028	IV	A-ve	32	87	33	33	52	65	302	50.33333333	50.4	600
5	Mark	1072	IV	AB+ve	69	87	82	65	78	85	466	77.66666667	77.7	600
6	Sunan	2874	IV	A-ve	33	90	88	90	65	54	420	70	70	600
7	Jennifer	2084	IV	B+ve	65	33	34	85	31	33	281	46.83333333	46.9	600
8	Mike	2907	IV	AB+ve	76	91	56	71	94	62	450	75	75	600
9	Tim	2917	IV	AB-ve	44	33	78	70	89	86	400	66.66666667	66.7	600
10	Jeffery	2962	IV	O-ve	77	92	98	60	82	89	498	83	83	600
11	Morgan	1972	IV	O+ve	87	84	55	62	72	91	451	75.16666667	75.2	600
12	John	2011	IV	B+ve	65	76	92	75	65	85	458	76.33333333	76.4	600
13	Mark	1072	IV	AB+ve	69	87	82	65	78	85	466	77.66666667	77.7	600
14	Mike	2807	IV	AB+ve	76	91	56	71	94	62	450	75	75	600

Here, you need to select the columns you want to compare for duplication. Another critical step is to check in the headers' option as you included the column names in the data set. Excel will automatically scan it by default.



Next, you must compare all columns, so go ahead and check all the columns as shown below.



Select Ok, and Excel performs the operations required and provides you with the data set after filtering out the duplicate data, as shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Name	Roll No	Year	Blood Group	Engg 1	Engg 2	Engg 3	Engg 4	Engg 5	Engg 6	Marks Obtained	Percentage	Round-off Percentage	Total Marks
2	Joe	1011	IV	O+ve	32	65	86	88	79	79	429	71.5	71.5	600
3	John	2011	IV	B+ve	65	76	92	75	65	85	458	76.33333333	76.4	600
4	Mary	1028	IV	A-ve	32	87	33	33	52	65	302	50.33333333	50.4	600
5	Mark	1072	IV	AB+ve	69	87	82	65	78	85	466	77.66666667	77.7	600
6	Sunan	2874	IV	A-ve	33	90	88	90	65	54	420	70	70	600
7	Jennifer	2084	IV	B+ve	65	33	34	85	31	33	281	46.83333333	46.9	600
8	Mike	2907	IV	AB+ve	76	91	56	71	94	62	450	75	75	600
9	Tim	2917	IV	AB-ve	44	33	78	70	89	86	400	66.66666667	66.7	600
10	Jeffery	2962	IV	O-ve	77	92	98	60	82	89	498	83	83	600
11	Morgan	1972	IV	O+ve	87	84	55	62	72	91	451	75.16666667	75.2	600
12	Mike	2807	IV	AB+ve	76	91	56	71	94	62	450	75	75	600
13														
14														
15														
16														
17														

In the next part of Excel Data Cleaning, you will understand data parsing from text to column.

Data Parsing from Text to Column

Sometimes, there is a possibility that one cell might have multiple data elements separated by a data delimiter like a comma. For example, consider that there is one column that stores **address** information.

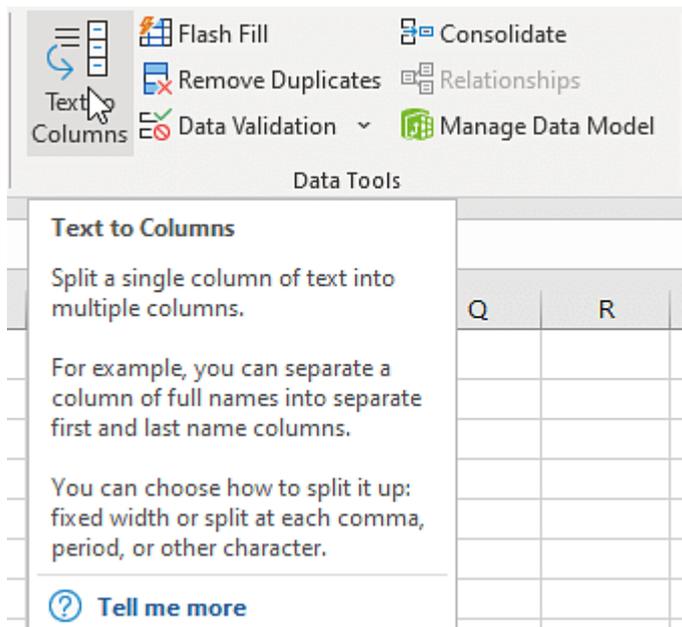
The address column stores the street, district, state, and nation. Commas separate all the data elements. You must now divide the street, district, state, and nation from the address columns into separate columns.

Excel's inbuilt functionality called "text to column" can achieve this. Now, try an example for the same.

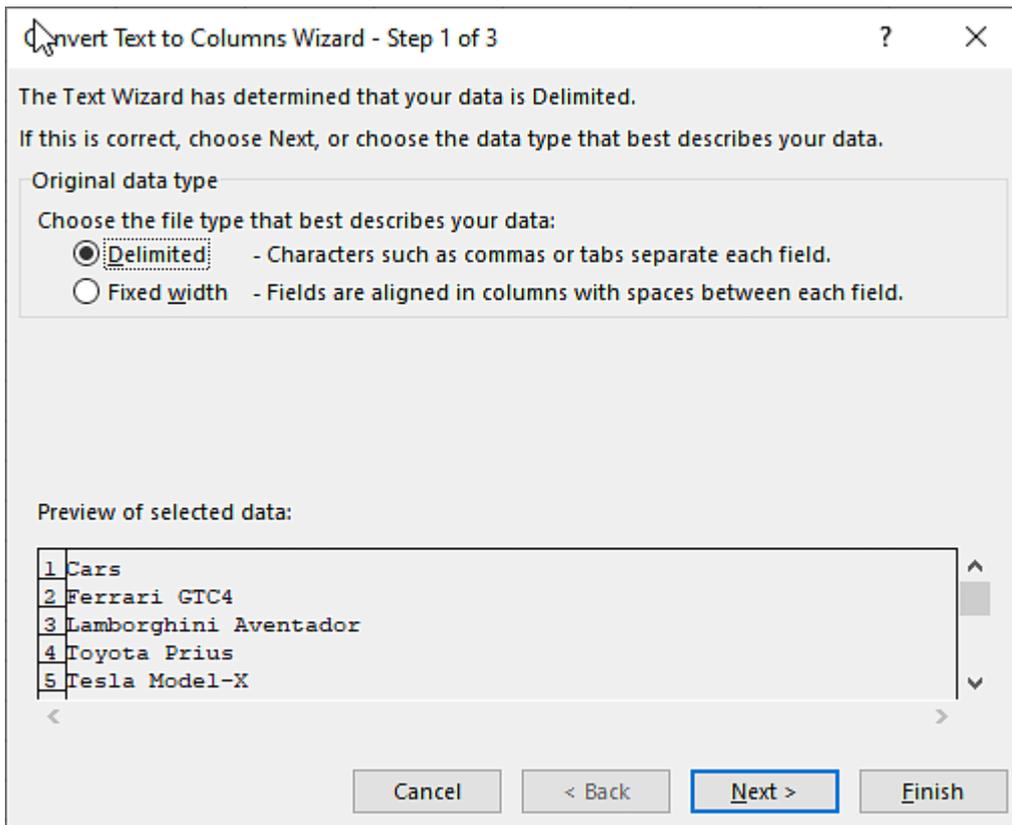
Here, you have the car manufacturer and the car model name separated by space as the data delimiter. The tabular data is shown below.

	A
1	Cars
2	Ferrari GTC4
3	Lamborghini Aventador
4	Toyota Prius
5	Tesla Model-X
6	Honda NSX
7	Ford Raptor
8	Chevrolet Corvette
9	Dodge Challenger
10	Toyota Supra
11	BMW M4
12	Mercedes 300-SL
13	Audi A6
14	Audi S8

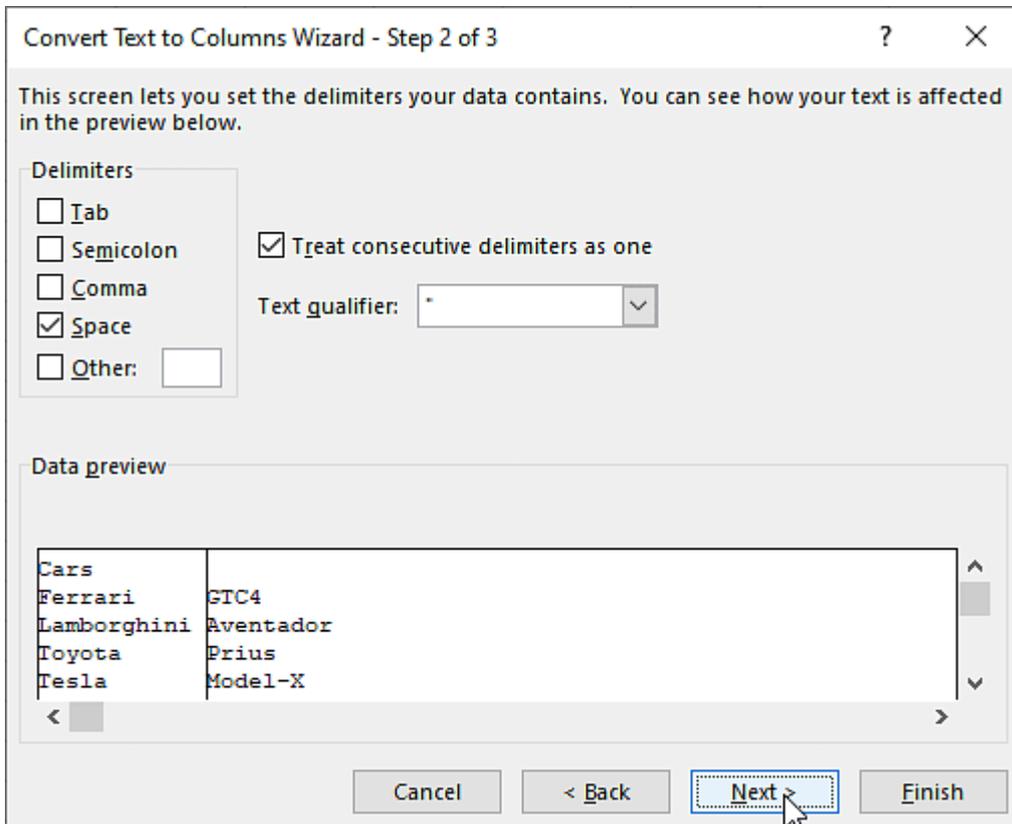
Select the data, click on the data option in the toolbar and then select "Text to Column", as shown below.



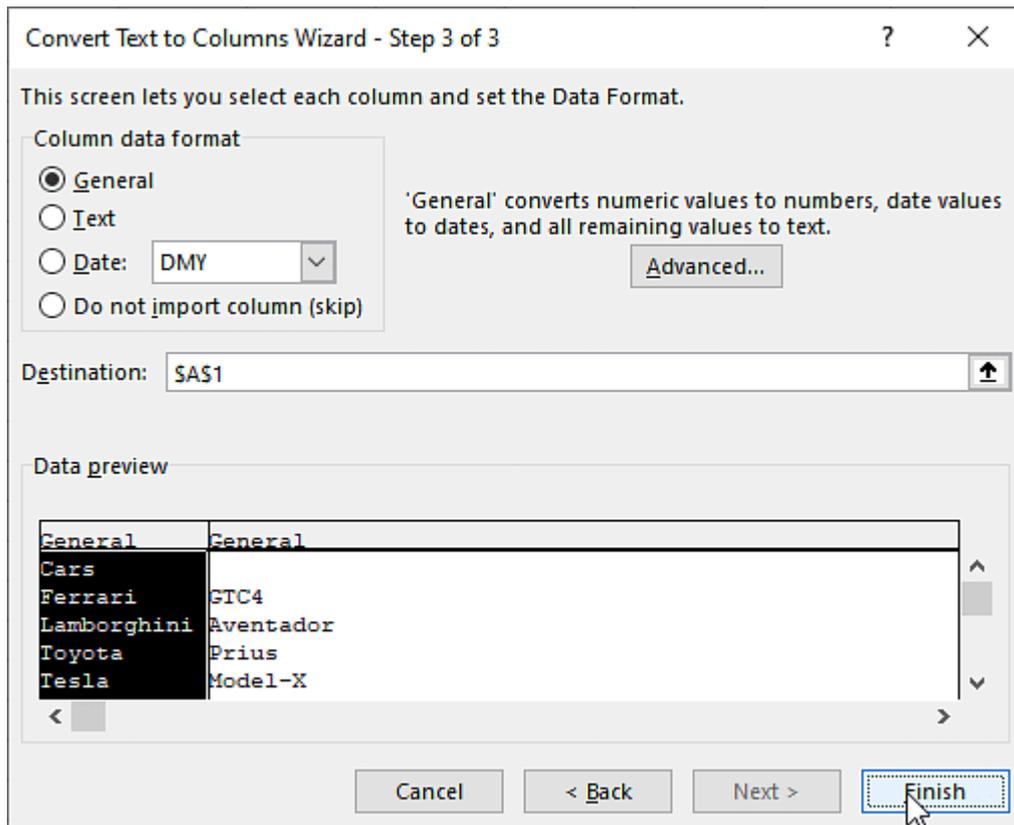
A new window will pop up on the screen, as shown below. Select the delimiter option and click on "next". In the next window, you will see another dialogue box.



In the new page dialogue box, you will see an option to select the type of delimiter your data has. In this case, you need to select the "space" as a delimiter, as shown below.



In the last dialogue box, select the column data format as "General", and the next step should be to click on the finish, as shown in the following image.



The final resultant data will be available, as shown below.

	A	B
1	Cars	Cars
2	Ferrari	GTC4
3	Lamborghini	Aventador
4	Toyota	Prius
5	Tesla	Model-X
6	Honda	NSX
7	Ford	Raptor
8	Chevrolet	Corvette
9	Dodge	Challenger
10	Toyota	Supra
11	BMW	M4
12	Mercedes	300-SL
13	Audi	A6
14	Audi	S8

Followed by Data parsing, in this tutorial about Excel Data Cleaning, you will learn how to delete all formatting.

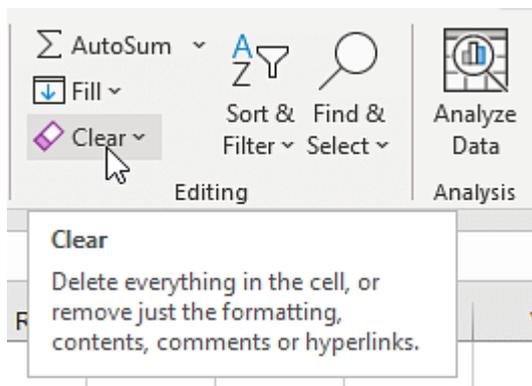
Another good way of cleaning data in excel is to ensure even formatting or, in some cases, even removing the formatting. The formatting can be as simple as coloring your cells and aligning the text in the cells. It can be a logical condition applied to your cells using [Excel's conditional formatting](#) option from the home tab.

However, in situations where you wish to remove the formatting, you can do it in the

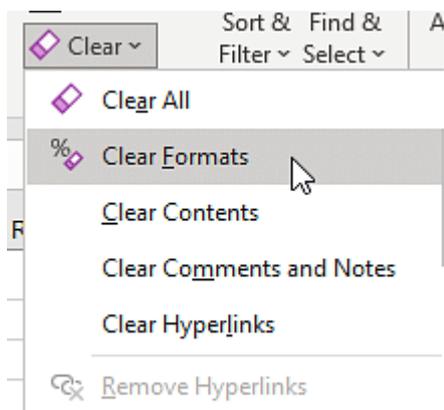
following ways. First, try to eliminate the regular formatting. In the previous example, you took the case of car manufacturers and car models data tables with heading cells colored in blue, and the text was center aligned.

	A	B
1	Cars	Cars
2	Ferrari	GTC4
3	Lamborghini	Aventador
4	Toyota	Prius
5	Tesla	Model-X
6	Honda	NSX
7	Ford	Raptor
8	Chevrolet	Corvette
9	Dodge	Challenger
10	Toyota	Supra
11	BMW	M4
12	Mercedes	300-SL
13	Audi	A6
14	Audi	S8

Now, use the clear option to remove the formats. Select the tabular data as shown below. Select the "home" option and go to the "editing" group in the ribbon. The "clear" option is available in the group, as shown below.



Select the "clear" option and click on the "clear formats" option. This will clear all the formats applied on the table.



The final data table will appear as shown below.

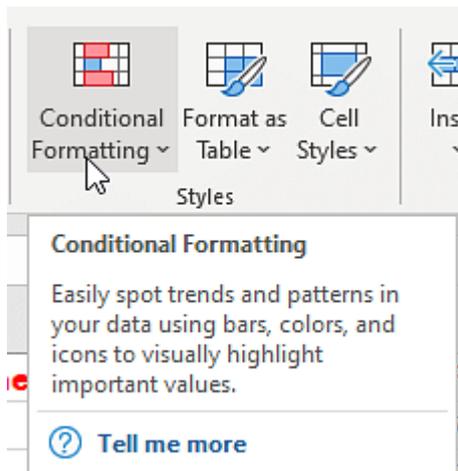
	A	B
1	Cars	Cars
2	Ferrari	GTC4
3	Lamborghini	Aventador
4	Toyota	Prius
5	Tesla	Model-X
6	Honda	NSX
7	Ford	Raptor
8	Chevrolet	Corvette
9	Dodge	Challenger
10	Toyota	Supra
11	BMW	M4
12	Mercedes	300-SL
13	Audi	A6
14	Audi	S8

Now, you must learn how to eliminate conditional formatting for cleaning data in Excel. This time, consider a different sheet. You must use the student's details sheet, which includes conditional formatting in Excel.

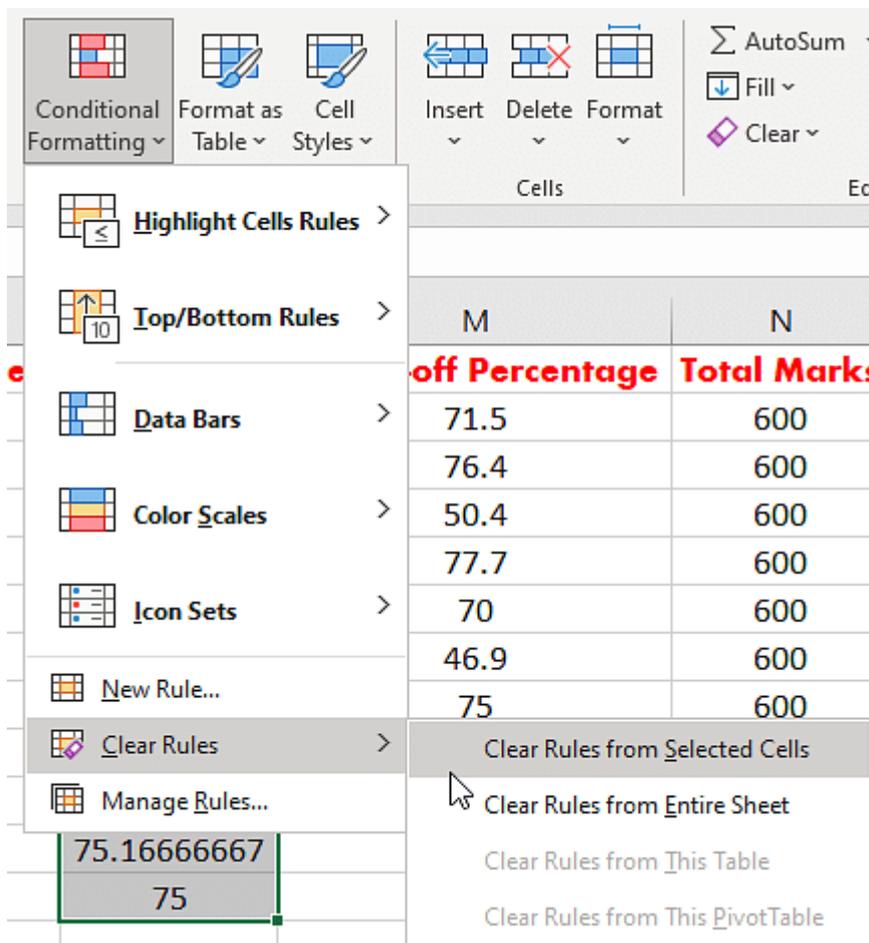
To eliminate conditional formatting in Excel, select the column or table with conditional formatting as shown below.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Name	Roll No	Year	Blood Group	Engg 1	Engg 2	Engg 3	Engg 4	Engg 5	Engg 6	Marks Obtained	Percentage	Round-off Percentage	Total Marks
2	Joe	1011	IV	O+ve	32	65	86	88	79	79	429	71.5	71.5	600
3	John	2011	IV	B+ve	65	76	92	75	65	85	458	76.33333333	76.4	600
4	Mary	1028	IV	A-ve	32	87	33	33	52	65	302	50.33333333	50.4	600
5	Mark	1072	IV	AB+ve	69	87	82	65	78	85	466	77.66666667	77.7	600
6	Sunan	2874	IV	A-ve	33	90	88	90	65	54	420	70	70	600
7	Jennifer	2084	IV	B+ve	65	33	34	85	31	33	281	46.83333333	46.9	600
8	Mike	2907	IV	AB+ve	76	91	56	71	94	62	450	75	75	600
9	Tim	2917	IV	AB-ve	44	33	78	70	89	86	400	66.66666667	66.7	600
10	Jeffery	2962	IV	O-ve	77	92	98	60	82	89	498	83	83	600
11	Morgan	1972	IV	O+ve	87	84	55	62	72	91	451	75.16666667	75.2	600
12	Mike	2807	IV	AB+ve	76	91	56	71	94	62	450	75	75	600

Then navigate to "Home", and select conditional formatting.



Then in the dialogue box, select the clear rules option. Here, you can either choose to eliminate rules only in the selected cells or eliminate rules from the entire column.



After you eliminate all conditions, the resultant table would look as follows.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Name	Roll No	Year	Blood Group	Engg 1	Engg 2	Engg 3	Engg 4	Engg 5	Engg 6	Marks Obtained	Percentage	Round-off Percentage	Total Marks
2	Joe	1011	IV	O+ve	32	65	86	88	79	79	429	71.5	71.5	600
3	John	2011	IV	B+ve	65	76	92	75	65	85	458	76.33333333	76.4	600
4	Mary	1028	IV	A-ve	32	87	33	33	52	65	302	50.33333333	50.4	600
5	Mark	1072	IV	AB+ve	69	87	82	65	78	85	466	77.66666667	77.7	600
6	Sunan	2874	IV	A-ve	33	90	88	90	65	54	420	70	70	600
7	Jennifer	2084	IV	B+ve	65	33	34	85	31	33	281	46.83333333	46.9	600
8	Mike	2907	IV	AB+ve	76	91	56	71	94	62	450	75	75	600
9	Tim	2917	IV	AB-ve	44	33	78	70	89	86	400	66.66666667	66.7	600
10	Jeffery	2962	IV	O-ve	77	92	98	60	82	89	498	83	83	600
11	Morgan	1972	IV	O+ve	87	84	55	62	72	91	451	75.16666667	75.2	600
12	Mike	2807	IV	AB+ve	76	91	56	71	94	62	450	75	75	600

You can always use a shortcut method to eliminate the conditional formatting in Excel. It is by pressing the sequential combination of the following keys as follows.

ATL + E + A + F

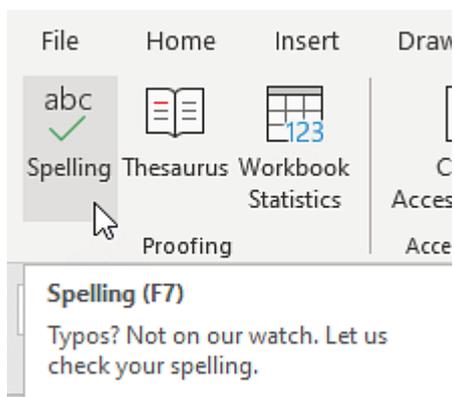
Next, in this Excel Data Cleaning tutorial, you will learn about Spell Check.

Spell Check

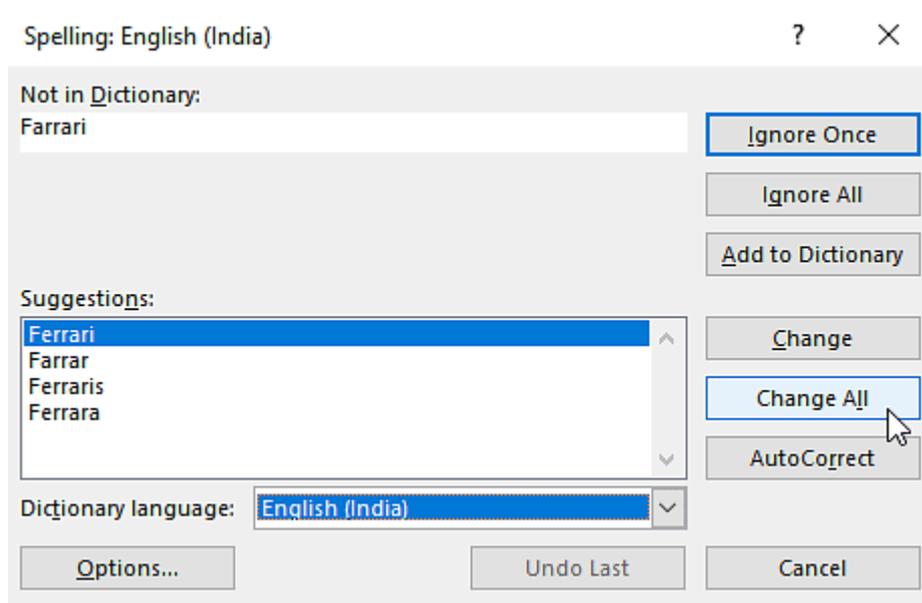
The feature of checking the spelling is available in MS Excel as well. To check the spellings of the words used in the spreadsheet, you can use the following method. Select the data cell, column, or sheet where you want to perform the spell check.

	A
1	Farrari
2	Lammorghini
3	Toyota
4	Tessla
5	Honda
6	Ford
7	Chevrolet
8	Dodge
9	Toyota
10	Audi
11	Marcedes

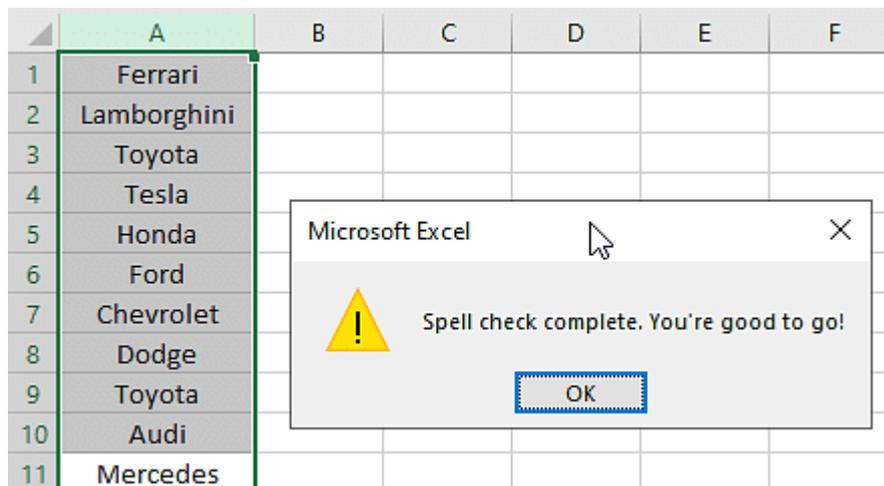
Now, go to the review option as shown below.



Microsoft Excel will automatically show the correct spelling in the dialogue box, as shown below. You can replace the words as per the requirement as shown below.



The final reviewed data table will like the one below.



In the next segment of this Excel Data Cleaning tutorial, you will learn about changing the text case.

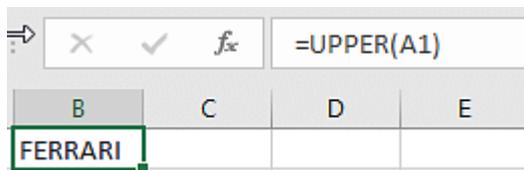
Change Case - Lower/Upper/Proper

You can manipulate the data in the [Excel worksheet](#) in terms of character cases as per the requirements. To apply case changes, you can follow the following steps.

Select the table or columns that need the case to be changed, as shown below.

	A
1	Ferrari
2	Lamborghini
3	Toyota
4	Tesla
5	Honda
6	Ford
7	Chevrolet
8	Dodge
9	Toyota
10	Audi
11	Mercedes

Select the cell next to the column and apply the formula as per the requirement, as shown below.



=UPPER(cell address) - for Upper case conversion

=LOWER(cell address) - for Lower case conversion

=PROPER(cell address) - for Sentence case conversion

Now, you can drag the cell can to the last row, as shown below.

	A	B
1	Ferrari	FERRARI
2	Lamborghini	LAMBORGHINI
3	Toyota	TOYOTA
4	Tesla	TESLA
5	Honda	HONDA
6	Ford	FORD
7	Chevrolet	CHEVROLET
8	Dodge	DODGE
9	Toyota	TOYOTA
10	Audi	AUDI
11	Mercedes	MERCEDES
12		

The final data table will appear as shown below.

	A	B
1	Ferrari	FERRARI
2	Lamborghini	LAMBORGHINI
3	Toyota	TOYOTA
4	Tesla	TESLA
5	Honda	HONDA
6	Ford	FORD
7	Chevrolet	CHEVROLET
8	Dodge	DODGE
9	Toyota	TOYOTA
10	Audi	AUDI
11	Mercedes	MERCEDES

Now that you learned spell check, in the upcoming section of Excel Data Cleaning, you will learn how to Highlight Errors in an Excel spreadsheet.

Highlight Errors

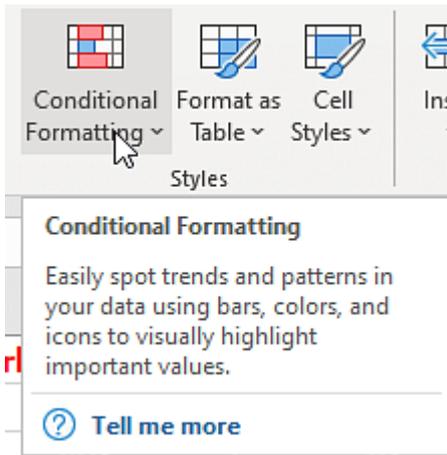
Highlighting errors in an Excel spreadsheet is helpful to find or sort out the erroneous data with ease. You can do error Highlighting with the help of conditional formatting in Excel. Here, you must consider the student data set as an example.

Imagine that you are interviewing all the students. There are eligibility criteria. You can shortlist the students if they have 60% aggregate marks. Now, apply conditional formatting and sort out the students who are eligible and not eligible.

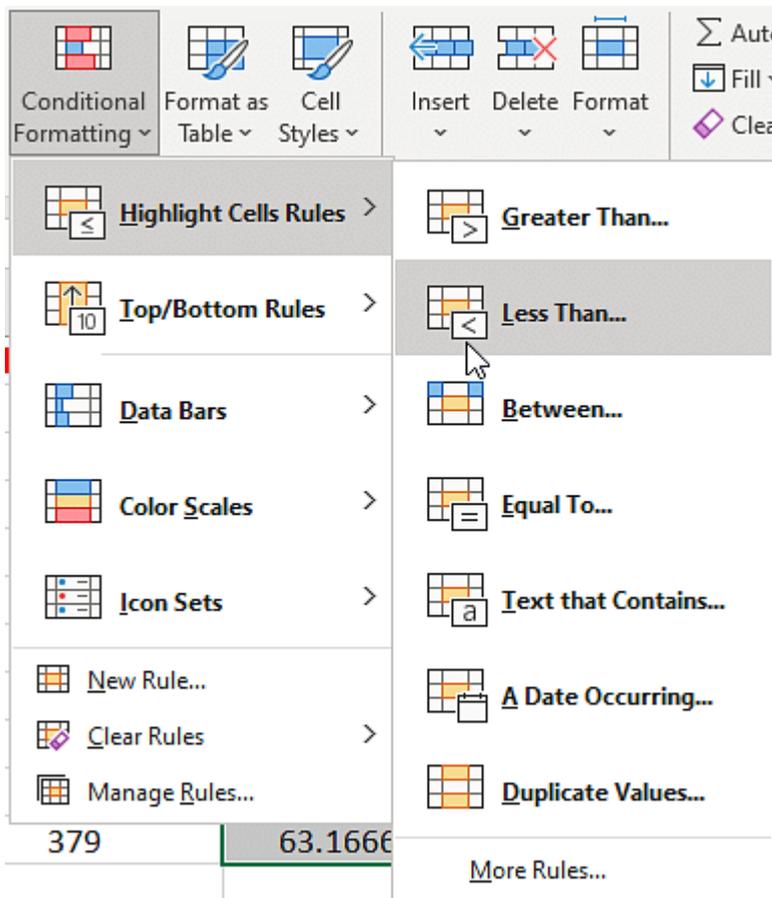
First, select the aggregate/percentage column as shown below.

M	Ro
Percentage	
71.5	
68	
59	
68.16666667	
62.83333333	
52	
57.16666667	
79.33333333	
67.66666667	
63.16666667	

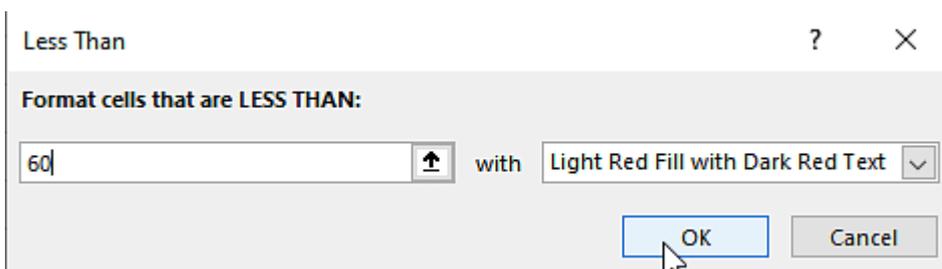
Select "Home", and in the Styles group, select conditional formatting, as shown below.



In the conditional formatting option, select the highlight option, and in the next drop-down, select the less than an option as shown below.



In the settings window, you will find a slot to provide the aggregate as "60" percent and press ok.



Excel will now select and highlight cells with an aggregate of less than 60 percent. In the next part of Excel Data Cleaning, you will understand the trim function.

TRIM Function

The TRIM function is used to eliminate excess spaces and tab spaces in the Excel worksheet cells. The excessive blank spaces and tab spaces make the data hard to understand. Using the "TRIM" function can eliminate these excessive blank spaces.

Select the data cells with excessive blank spaces and tab spaces. Now, select a new cell adjacent to the first cell.

Apply the TRIM() function and drag the cell as shown below.

	A
1	Hi, Welcome to Data Analytics
2	In Excel
3	This chapter
4	is based
5	on TRIM () Method

It shows the final data after the elimination of the excess space as follows.

B1		=TRIM(A1)	
	A		B
1	Hi, Welcome to Data Analytics		Hi, Welcome to Data Analytics
2	In Excel		In Excel
3	This chapter		This chapter
4	is based		is based
5	on TRIM () Method		on TRIM () Method

Next, in the Excel Data Cleaning tutorial, you will look at the Find and Replace function.

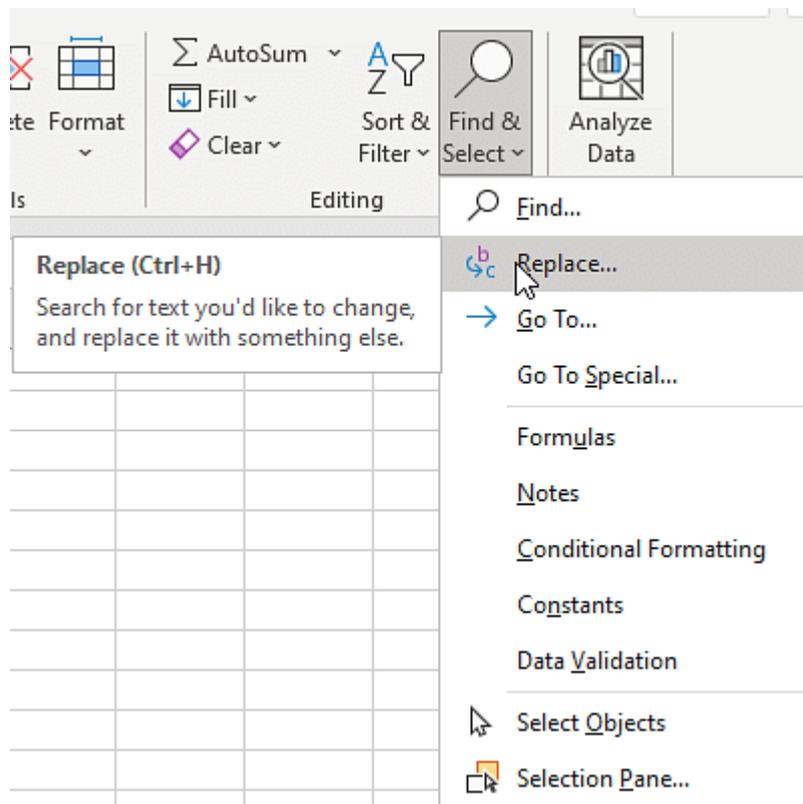
Find and Replace

Find and Replace will help you fetch and replace data in the entire worksheet to help in organizing and cleaning data in Excel. Consider the employee data example.

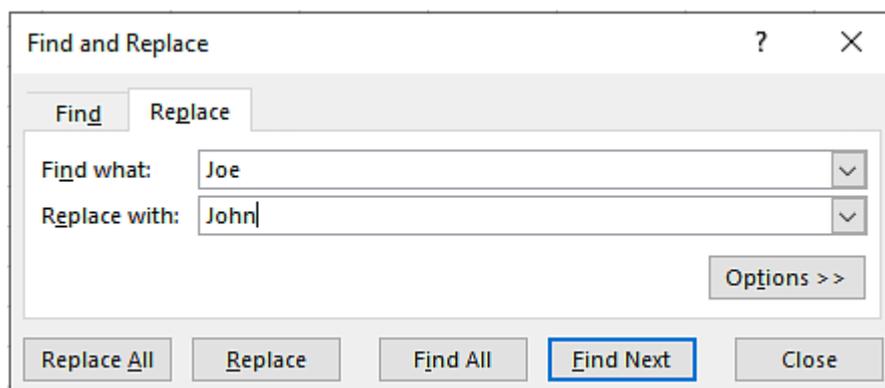
Here, try to fetch an employee with the name Joe and try to rename or replace his name with John, after changing his first name.

Employee Details									
S no	Name	EMP No	Designation	Salary	Hike %	New Salary	Blood Group	Phone number	
1	Joe	1011	CEO	100000	15%	115000	O+ve	289749782	
2	John	2011	Software Developer	15000	15%	17250	B+ve	382741987	
3	Mary	1028	Tester	19000	15%	21850	A-ve	222147868	
4	Mark	1072	Finance	20000	15%	23000	AB+ve	656398101	
5	Sunan	2874	Finance	15000	15%	17250	A-ve	164192719	
6	Jennifer	2084	Tester	29000	15%	33350	B+ve	688578990	
7	Mike	2907	Marketing	19000	15%	21850	AB+ve	157378911	
8	Tim	2917	Marketing	12000	15%	13800	AB-ve	538975791	
9	Jeffery	2962	Software Developer	10000	15%	11500	O-ve	745932616	
10	Morgan	1972	Tester	29000	15%	33350	O+ve	274729436	

The "find and replace" option is present in the home ribbon in the editing group, as shown below.



Click on the option, and a new window will open, where you can enter the data to be fetched and enter the text you need to replace, as shown below.



Click on "replace all", and it will replace the text. The final dataset will be as shown below.

	A	B	C	D	E	F	G	H	I
1		Employee Details							
2									
3	S no	Name	EMP No	Designation	Salary	Hike %	New Salary	Blood Group	Phone number
4	1	John	1011	CEO	100000	15%	115000	O+ve	289749782
5	2	John	2011	Software Developer	15000	15%	17250	B+ve	382741987
6	3	Mary	1028	Tester	19000	15%	21850	A-ve	222147868
7	4	Mark	1072	Finance	20000	15%	23000	AB+ve	656398101
8	5	Sunan	2874	Finance	15000	15%	17250	A-ve	164192719
9	6	Jenniffer	2084	Tester	29000	15%	33350	B+ve	688578990
10	7	Mike	2907	Marketing	19000	15%	21850	AB+ve	157378911
11	8	Tim	2917	Marketing	12000	15%	13800	AB-ve	538975791
12	9	Jeffery	2962	Software Developer	10000	15%	11500	O-ve	745932616
13	10	Morgan	1972	Tester	29000	15%	33350	O+ve	274729436

Experiment No: 04

Date:

FLASH FILL AND TEXT TO COLUMNS

Separate the country using Flash Fill Kicks

Participants
Ronnie Anderson, UK
Tom Boone, Canada
Sally Brook, USA
Jeremy Hill, Australia
Mattias Waldau, USA
Robert Furlan, France
David White, UK

Solutions:

The screenshot shows the Excel ribbon with the 'Data' tab selected. In the 'Data Tools' group, the 'Flash Fill' button is highlighted with a blue box and an arrow. Below the ribbon, a table shows the following data:

Participants	Country
Ronnie Anderson, UK	UK
Tom Boone, Canada	
Sally Brook, USA	
Jeremy Hill, Australia	
Mattias Waldau, USA	
Robert Furlan, France	
David White, UK	

The screenshot shows the Excel spreadsheet with the following data:

	A	B	C	D	E
1	Participants	Country			
2	Ronnie Anderson, UK	UK			
3	Tom Boone, Canada	Canada			
4	Sally Brook, USA	USA			
5	Jeremy Hill, Australia	Australia			
6	Mattias Waldau, USA	USA			
7	Robert Furlan, France	France			
8	David White, UK	UK			
9					

The 'Flash Fill Options' menu is open over the 'Canada' cell in the 'Country' column, showing the following options:

- Undo Flash Fill
- Accept suggestions
- Select all 0 blank cells
- Select all 6 changed cells

	A	B
1	Address	Zip code
2	St-Joris Weert 3051 Belgium	3051
3	Illinois, 60606, USA	60606
4	California, 92618, USA	92618
5	Madrid 28014 Spain	28014
6	San Francisco, CA, 94105, USA	94105

Experiment No: 05

Date:

ANALYZING DATA USING SPREADSHEETS

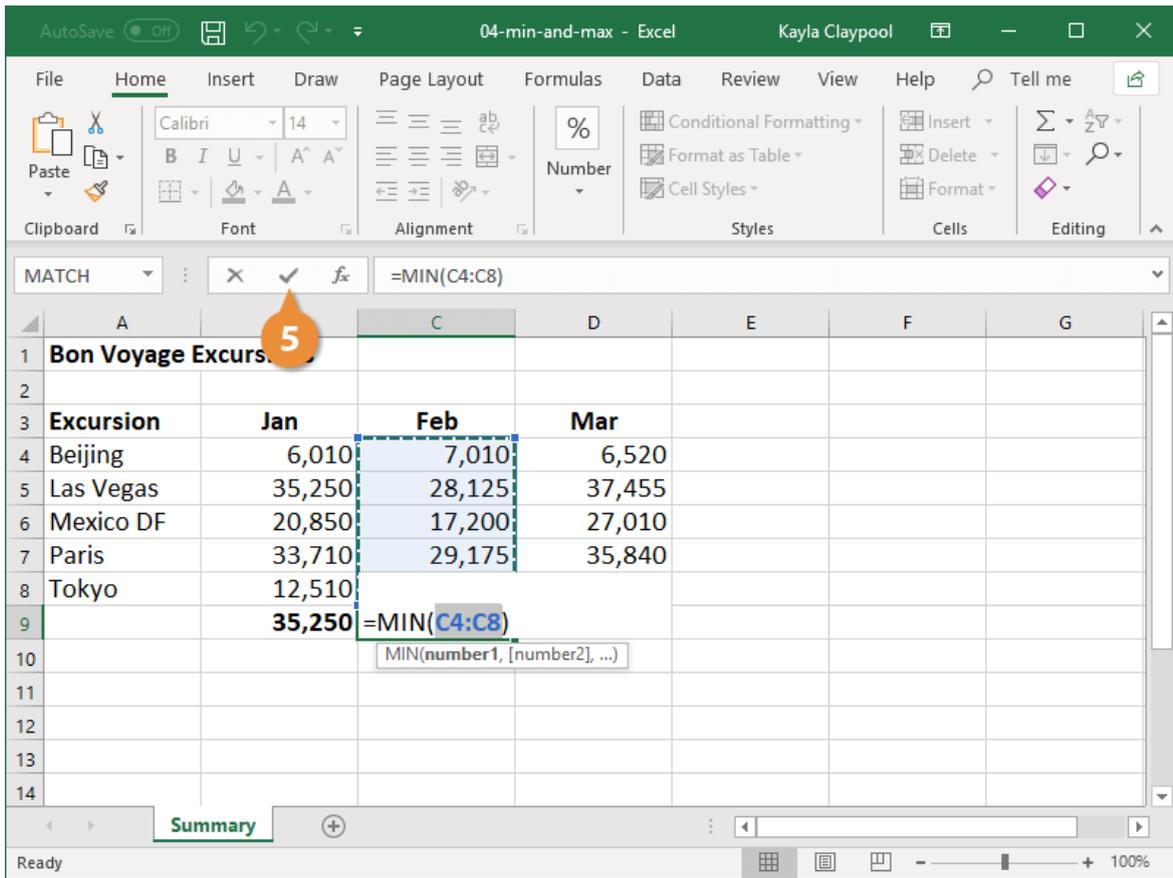
Maximum and Minimum Value

Excursion	Jan	Feb	Mar
Beijing	6,010	7,010	6,520
Las Vegas	35,250	28,125	37,455
Mexico DF	20,850	17,200	27,010
Paris	33,710	29,175	35,840
Tokyo	12,510	14,750	11,490

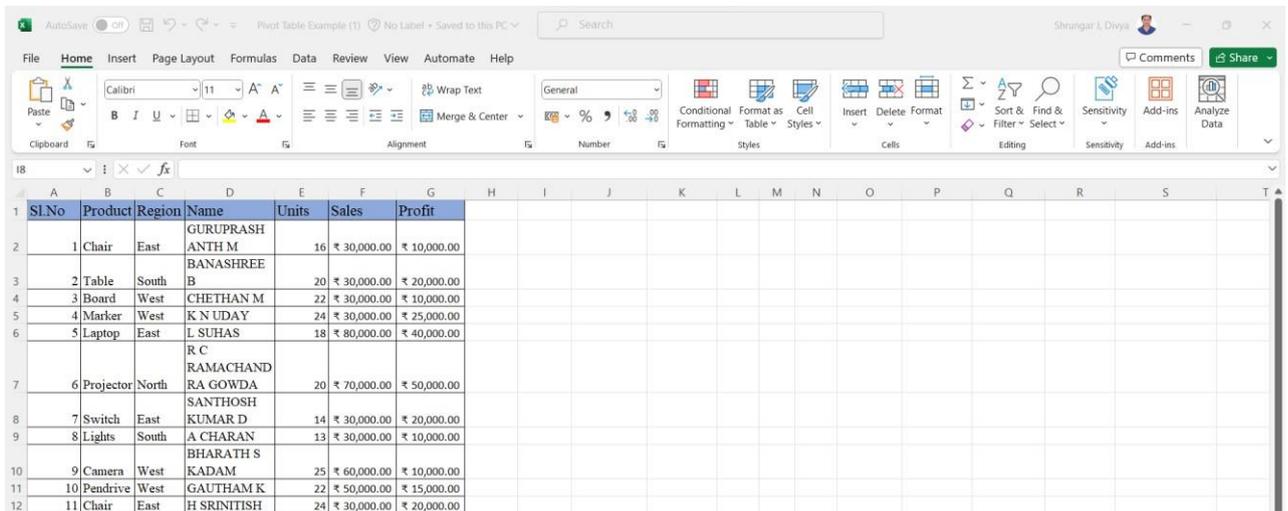
Solution :

Excursion	Jan	Feb	Mar
Beijing	6,010	7,010	6,520
Las Vegas	35,250	28,125	37,455
Mexico DF	20,850	17,200	27,010
Paris	33,710	29,175	35,840
Tokyo	12,510	14,750	11,490

Formula in B9: `=MAX(B4:B8)`



Exercise 8: Create the Pivot Tables



Solutions:

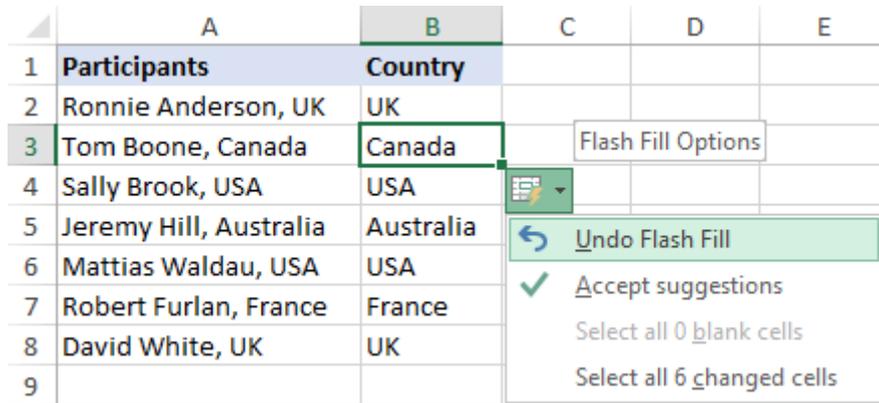
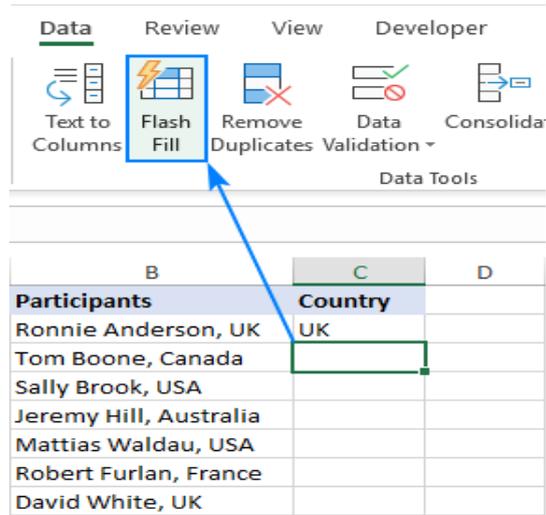
The screenshot shows an Excel spreadsheet with a PivotTable. The PivotTable is located in columns H through M and rows 8 through 15. It summarizes the data from the source table (rows 1-15, columns A-G) by 'Product' (rows) and 'Region' (columns). The PivotTable includes columns for 'Sum of Sales', 'Sum of Profit', and 'Sum of Profit2'. The PivotTable Fields task pane on the right shows the following configuration:

- Choose fields to add to report:** Name, Sales, Profit (checked); Product, Region, Units (unchecked).
- Filters:** (Empty)
- Columns:** Values
- Rows:** Name
- Values:** Sum of Sales, Sum of Profit, Sum of Profit2

Exercise 9: Separate the country using Flash Fill Kicks

Participants
Ronnie Anderson, UK
Tom Boone, Canada
Sally Brook, USA
Jeremy Hill, Australia
Mattias Waldau, USA
Robert Furlan, France
David White, UK

Solutions:



	A	B
1	Address	Zip code
2	St-Joris Weert 3051 Belgium	3051
3	Illinois, 60606, USA	60606
4	California, 92618, USA	92618
5	Madrid 28014 Spain	28014
6	San Francisco, CA, 94105, USA	94105

Experiment No: 06

Date:

VLOOKUP AND HLOOKUP

Determine the VLOOKUP in the given data.

K S SCHOOL OF ENGINEERING AND MANAGEMENT, BENGALURU - 560109						
NT DETAILS				DEPARTMENT		
OF CIVIL ENGINEERING						
SEMESTER	USN	NAME OF THE STUDENT	PARENT/GAURDIAN NAME	DOB	AGE (YR)	SEX
III	1KG22CV001	ARCHANA U				
III	1KG22CV002	BASAVASHRBE				
III	1KG22CV004	RAJESHWARI S				
III	1KG22CV005	SHASHIKALA G S				
III	1KG22CV006	TEJAS K				
III	1KG22CV007	AADITHEN B				
III	1KG22CV008	AMITH L				
III	1KG22CV009	BHARATH				
III	1KG22CV010	CHANDAN B				
III	1KG22CV011	DARSHAN R				
III	1KG22CV012	DEEPAK SHARMA				
III	1KG22CV013	FARHAN SHARIFF				
III	1KG22CV014	I SANJAY KUMAR				
III	1KG22CV015	JAMIL AKHTER				
III	1KG22CV016	KARTHIK M P				
III	1KG22CV017	LALIT KUMAR NAIDU				
III	1KG22CV018	MOHAMMED MAHEER				
III	1KG22CV018	SHARIF				
III	1KG22CV019	NAGESH R				
III	1KG22CV020	NAGENDRA PRASAD				
III	1KG22CV021	NANDISH K				
III	1KG22CV022	NAVEEN R				
III	1KG22CV023	NIKHTH R				

Solution :

The command used for to search in another sheet

=VLOOKUP(D4,Sheet2!\$D\$3:\$E\$30,2,FALSE)

The screenshot shows an Excel spreadsheet with the following data:

K S SCHOOL OF ENGINEERING AND MANAGEMENT, BENGALURU - 560109						
NT DETAILS				DEPARTMENT		
OF CIVIL ENGINEERING						
SEMESTER	USN	NAME OF THE STUDENT	PARENT/GAURDIAN NAME	DOB	AGE (YR)	SEX
III	1KG22CV001	ARCHANA U	2.FALSE)			
III	1KG22CV002	BASAVASHRRE				
III	1KG22CV004	RAJESHWARI S				
III	1KG22CV005	SHASHIKALA G S				
III	1KG22CV006	TEJAS K				
III	1KG22CV007	AADITHEN B				
III	1KG22CV008	AMITH L				
III	1KG22CV009	BHARATH				

K S SCHOOL OF ENGINEERING AND MANAGEMENT, BENGALURU - 560109						
NT DETAILS				DEPARTMENT		
OF CIVIL ENGINEERING						
SEMESTER	USN	NAME OF THE STUDENT	PARENT/GAURDIAN NAME	DOB	AGE (YR)	SEX
III	1KG22CV001	ARCHANA V	ULLURU SHETTY K T			
III	1KG22CV002	BASAVASHIRKE	KRISHNATH			
III	1KG22CV004	RAJESHWARI S	SOMANNA			
III	1KG22CV005	SHASHIKALA G S	SHIVANNA G S			
III	1KG22CV006	TEJAS K	KRISHNA V			
III	1KG22CV007	AADITHEN B	K BALASUBRAMANIAM			
III	1KG22CV008	AMITH L	S V LOKESH REDDY			
III	1KG22CV009	BHARATH	M PRAKASH			
III	1KG22CV010	CHANDAN B	BARNABAS			
III	1KG22CV011	DARSHAN R	RAMESH			
III	1KG22CV012	DEEPAK SHARMA	RAMANAND SHARMA			
III	1KG22CV013	FARHAN SHARIF	ISMAIL SHARIF			
III	1KG22CV014	I SANJAY KUMAR	A INBARAJ			
III	1KG22CV015	JAMIL AKHTER	BABU JAN ANSARI			
III	1KG22CV016	KARTHIK M P	PUTTASWAMY			
III	1KG22CV017	LALIT KUMAR NAIDU	KALIDAS NAIDU			

Exercise 6: Fetch the marks of student D in Management, in the given data using HLOOKUP .

	A	B	C	D	E	F
Student name	A	B	C	D	E	F
Accounts		75	65	70	60	59
Economics		65	72	78	89	67
Management		70	68	90	72	58
Mathematics		80	90	75	65	87

Solution :

A	B	C	D	E	F	G	H	
Student roll no	A	B	C	D	E			
Accounts	75	65	70	60	59			
Economics	65	72	78	89	67			
Management	70	68	90	72	58			
Mathematics	80	90	75	65	87			
Fetch Marks of D in Management	=Hlookup(HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])							

A	B	C	D	E	F	G	H	
Student name	A	B	C	D	E			
Accounts	75	65	70	60	59			
Economics	65	72	78	89	67			
Management	70	68	90	72	58			
Mathematics	80	90	75	65	87			
Fetch Marks of D in Management	=hlookup("D" HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])							

A	B	C	D	E	F	G	H	
Student name	A	B	C	D	E			
Accounts	75	65	70	60	59			
Economics	65	72	78	89	67			
Management	70	68	90	72	58			
Mathematics	80	90	75	65	87			
Fetch Marks of D in Management	=hlookup("D",A1:F5 HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])							

Student name	A	B	C	D	E
Accounts	75	65	70	60	59
Economics	65	72	78	89	67
Management	70	68	90	72	58
Mathematics	80	90	75	65	87

Fetch Marks of D in Management	=HLOOKUP("D",A1:F5,4
--------------------------------	----------------------

HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])

Student name	A	B	C	D	E
Accounts	75	65	70	60	59
Economics	65	72	78	89	67
Management	70	68	90	72	58
Mathematics	80	90	75	65	87

Fetch Marks of D in Management	=HLOOKUP("D",A1:F5,4,)
--------------------------------	------------------------

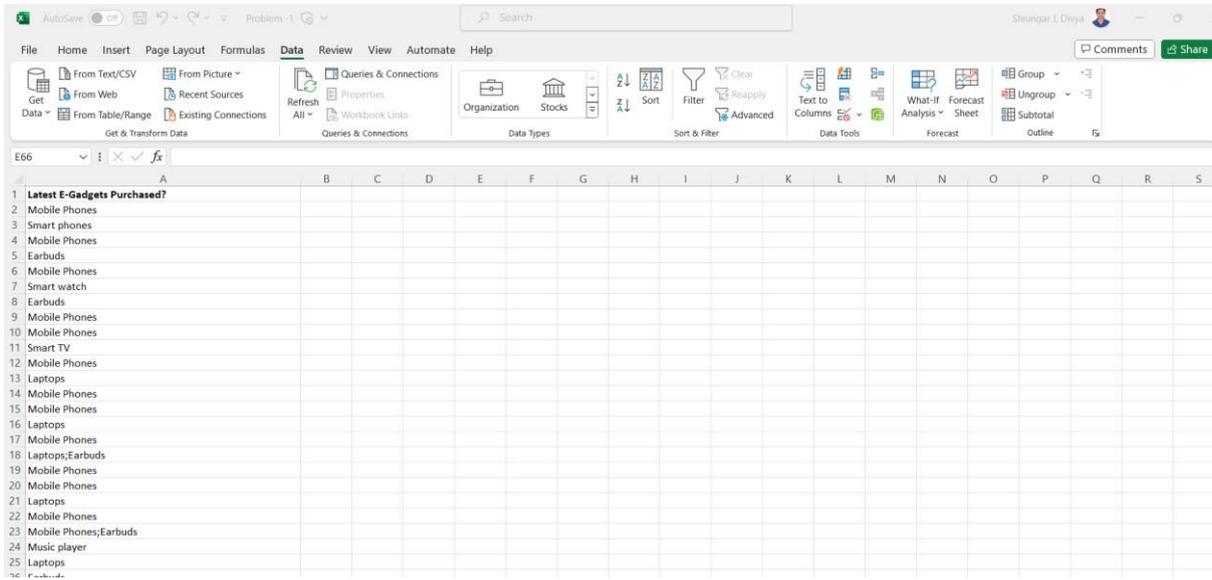
HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])

- TRUE - Approximate match
- FALSE - Exact match

Approximate match - the values in the first row of table_array must be sorted in ascending order

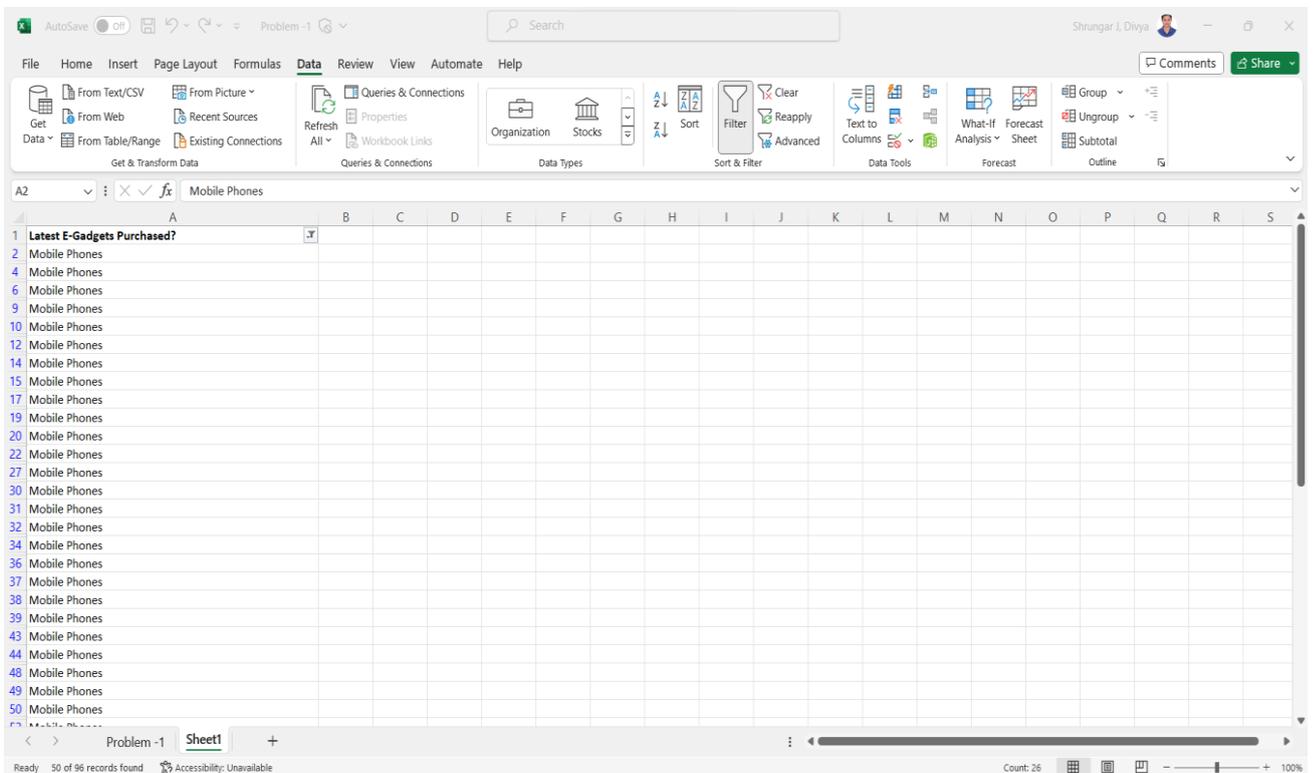
B8 fx =HLOOKUP("D",A1:F5,4,FALSE)

	A	B	C	D	E	F	G
1	Student name	A	B	C	D	E	
2	Accounts	75	65	70	60	59	
3	Economics	65	72	78	89	67	
4	Management	70	68	90	72	58	
5	Mathematics	80	90	75	65	87	
6							
7							
8	Fetch Marks of D in Management	72					
9							

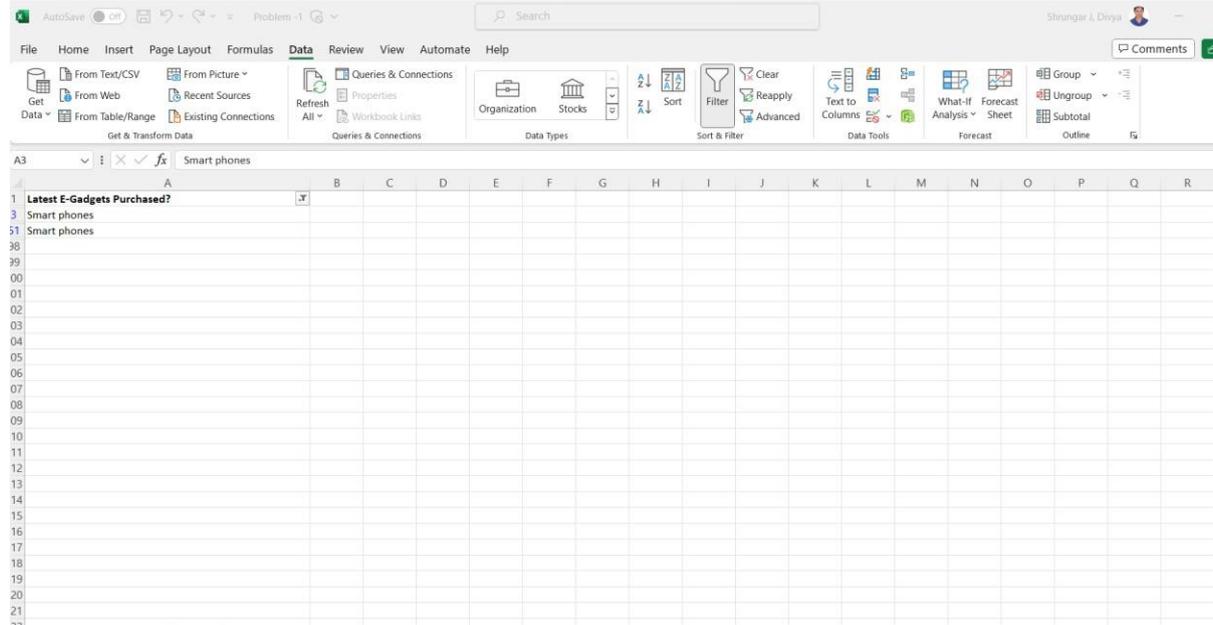
Experiment No: 07**Date:****PROJECT**

Solution:

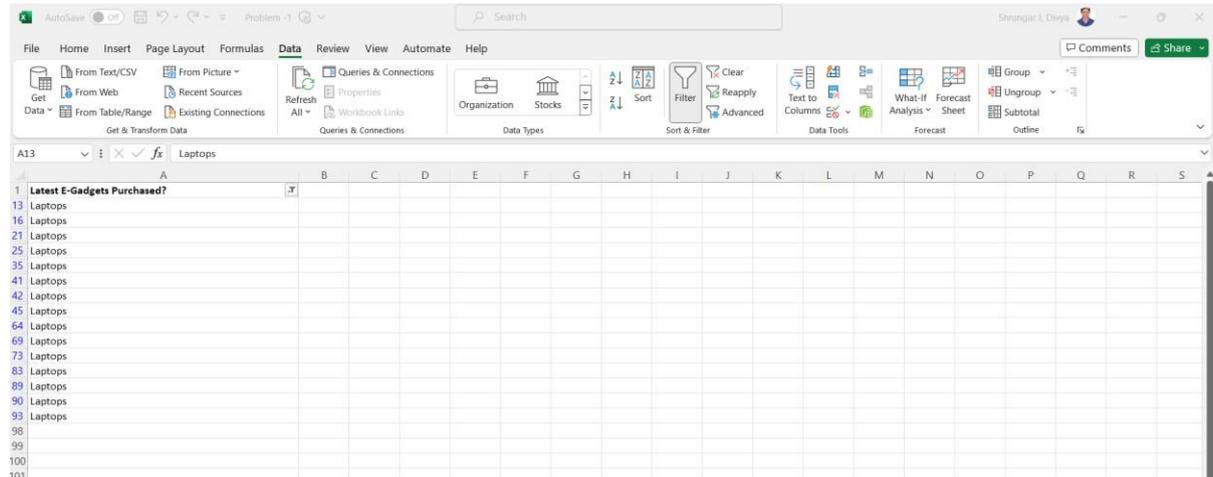
Mobile Phones: Bottom Will be present 50 Nos



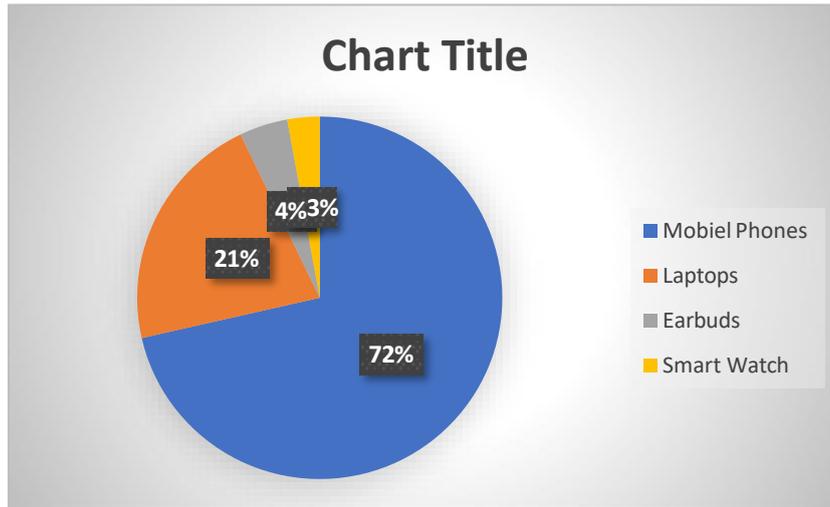
Smart Phones: 2 Nos



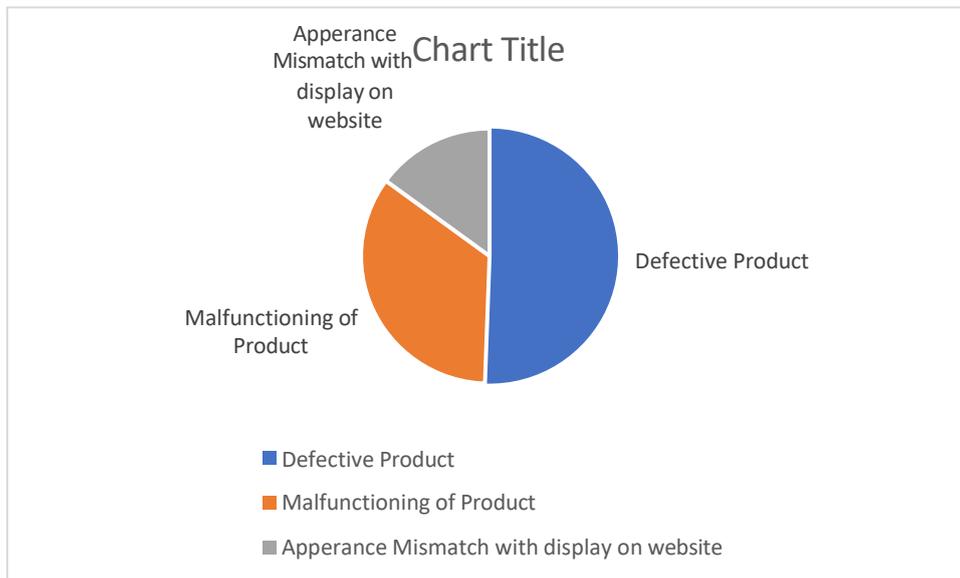
Laptops- 15



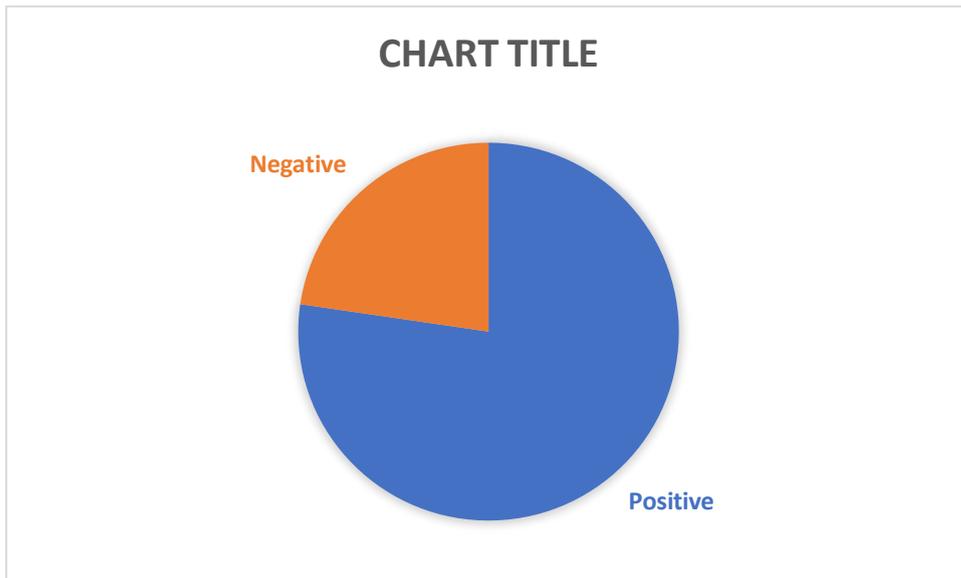
Exercise 11 : Draw the Pie Chart of E gadgets Purchased



Defective Product	44
Malfunctioning of Product	30
Appearance Mismatch with display on website	13



Positive	75
Negative	22



Replacement of Product	62
Refund of money	34

