

# Thank you for your order!

Your order is complete. If you need a receipt, you can print this page.

Arlene Gonzales  
Paris corne Vatican St., Ciudad Grande  
Ciudad Grande  
Pasig City  
Metro Manila  
1605  
Philippines

**Order Number:**

EDX-50849487

**Payment Method:**

Visa 418359XXXXXX9107

**Order Date:**

June 18, 2022

## Order Information

Quantity	Description	Item Price
1	Seat in Analyzing Data with Excel with verified certificate IBM	\$99.00
<hr/>		
	<b>Subtotal</b>	<b>\$99.00</b>
	<b>Total</b>	<b>\$99.00</b>

[Go to dashboard](#)

[Find more courses](#)

# Analyzing Data with Excel

## General Information

This course is designed to provide you with basic working knowledge for using Excel spreadsheets for Data Analysis. It covers some of the first steps for working with spreadsheets and their usage in the process of analyzing data. It includes plenty of videos, demos, and examples for you to learn, followed by step-by-step instructions for you to apply and practice on a live spreadsheet.

Excel is an essential tool for working with data - whether for business, marketing, data analytics, or research. This course is suitable for those aspiring to take up Data Analysis or Data Science as a profession, as well as those who just want to use Excel for data analysis in their own domains. You will gain valuable experience in cleansing and wrangling data using functions and then analyze your data using techniques like filtering, sorting and creating pivot tables.

This course starts with an introduction to spreadsheets like Microsoft Excel and Google Sheets and loading data from multiple formats. With this introduction you will then learn to perform some basic level data wrangling and cleansing tasks and continue to expand your knowledge of analyzing data through the use of filtering, sorting, and using pivot tables within the spreadsheet. By performing these tasks throughout the course, it will give you an understanding of how spreadsheets can be used as a data analysis tool and understand its limitations.

There is a strong focus on practice and applied learning in this course. With each lab, you will gain hands-on experience in manipulating data and begin to understand the important role of spreadsheets. Clean and analyze your data faster by understanding functions in the formatting of data. You will then convert your data to a pivot table and learn its features to make your data organized and readable. The final project enables you to show off your newly acquired data analysis skills. By the end of this course you will have worked with several data sets and spreadsheets and demonstrated the basics of cleaning and analyzing data all without having to learn any code.

Getting started with Excel is made easy in this course. It does not require any prior experience with spreadsheets or coding. Nor does it require downloads or installation of any software. All you need is a device with a modern web browser, and ability to create a Microsoft account to access Excel online at no-cost. However if you already have a desktop version of Excel, you can follow along quite easily as well.

### Who Should Take This Course

This course benefits anyone looking to sharpen their skills using Excel spreadsheets and expand their knowledge on the fundamentals of analyzing data.

### Pre-requisite

This program does not require any pre-requisites, and is suitable for learners with or without college degrees. All you need to get started is basic computer literacy, a comfort working with numbers, a willingness to self-learn online, and the desire to enrich your profile with valuable skills.

# Syllabus

## Course Syllabus

### Module 1 - Introduction to Data Analysis Using Spreadsheets

- Introduction to Spreadsheets
- Spreadsheet Basics
- Using Spreadsheets as a Data Analysis Tool

### Module 2 - Getting Started with Using Excel Spreadsheets

- Viewing, Entering, and Editing Data
- Copying, Filling, and Formatting Cells and Data
- The Basics of Formulas
- Intro to Functions
- Referencing Data in Formulas

### Module 3 - Basics of Data Quality and Privacy

- Overview of the Data Analyst Ecosystem
- Types of Data
- Understanding Different Types of File Formats
- Sources of Data Using Service Bindings
- Languages for Data Professionals

### Module 4 - Cleaning Data

- Removing Duplicated or Inaccurate Data and Empty Rows
- Dealing with Inconsistencies in Data
- More Excel Features for Cleaning Data

### Module 5 - Data Analysis Basics, Filtering and Sorting Data

- Intro to Analyzing Data Using Spreadsheets
- Filtering and Sorting Data in Excel
- Useful Functions for Data Analysis
- Using VLOOKUP and HLOOKUP Functions

### Module 6 - Using Pivot Table

- Introduction to Creating Pivot Tables in Excel
- Pivot Table Features

### Final Project - Part 1 and Part 2

# Grading Scheme

## GRADING SCHEME

This section contains information for those earning a certificate. Those auditing the course can skip this section and click next.

1. This course contains 6 Graded Quizzes and 1 Final Assignment. There is 1 Graded Quiz per module. Your total grade at 100% is weighted as follows:
  - o Each of the 6 Graded Quizzes carries an equal weight totaling 80% of your total grade.
  - o Final Assignment carries a weight of 20% of your total grade.
2. The minimum passing mark for the **course** is 70%.
3. Permitted attempts are per **question**:
  - o One attempt - for True/False questions
  - o Two attempts - for any question other than True/False
4. There are no penalties for incorrect attempts.
5. Clicking the "**Final Check**" button when it appears, means your submission is **FINAL**. You will **NOT** be able to resubmit your answer for that question again.
6. Check your grades in the course at any time by clicking on the "Progress" tab.

## **Module Introduction**

In this module, you will learn about the fundamentals of spreadsheet applications, and you will be introduced to the Excel interface and learn how to navigate your way around a worksheet and workbook.

### **Module 1 - Introduction to Data Analysis Using Spreadsheets**

- Introduction to Spreadsheets
- Spreadsheet Basics
- Using Spreadsheets as a Data Analysis Tool

### **Learning Objectives**

**After completing this module, you will be able to:**

- Describe the fundamentals of a spreadsheet application.
- Get access to Excel on the Web.
- Navigate around an Excel worksheet and workbook.

# Course Introduction

Do you want to learn how to use spreadsheets and start analyzing data using Excel? This course from IBM is designed to help you work with Excel and gives you a good grounding in the cleaning and analyzing of data which are important parts of the skill set required to become a data analyst. You will not only learn data analysis techniques using spreadsheets, but also practice using multiple hands-on labs throughout the course.

In **module one** you will learn about the basics of spreadsheets, including spreadsheet terminology, the interface and navigating around worksheets and workbooks.

In **module two** you will learn about selecting data, entering and editing data, copying and auto filling data, formatting data, and using functions and formulas.

In **module three** you will learn about cleaning and wrangling data using a spreadsheet, including the fundamentals of data quality and data privacy, removing duplicated and inaccurate data, removing empty rows, removing data inconsistencies and white spaces, and using the flash fill and text to columns features.

In **module four** you will learn about analyzing data using spreadsheets, including filtering data, sorting data, using common data analysis functions, creating and using pivot tables, and creating and using slicers and timelines.

At the end of this course in **module five**, you will complete a series of hands-on labs which will guide you on how to create your first deliverable as a data analyst. This will involve you understanding what the business scenario is, cleaning and preparing your data, and analyzing your data.

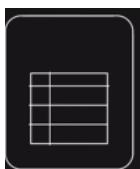
You will follow two different business scenarios throughout the course, with each using their own data set. These different scenarios and data sets will be used in the lesson videos and in the hands-on labs. After completing this course, you will be able to understand how spreadsheets can be used as a data analysis tool; understand when to use spreadsheets as a data analysis tool and their limitations; create a spreadsheet and explain its basic functionality; perform data wrangling and data cleaning tasks using Excel; analyze data using filter sort and pivot table features within Excel spreadsheets. You will also perform some intermediate level data wrangling and data analysis tasks to address a business scenario.

The course team and other peers are available to help in the course discussion forums in case you require any assistance.

Let's get started with your next video where you will get an introduction to spreadsheets.

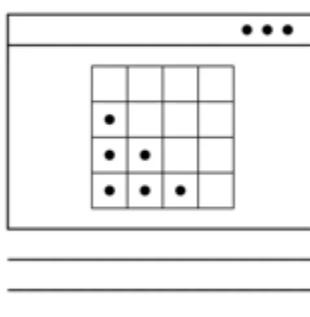
# Course introduction

This course from IBM is designed to help you work with Excel and gives you a good grounding in the cleaning and analyzing of data which are important parts of the skill set required to become a data analyst.



## Course Overview

### Module 1



- Basics of Spreadsheets
- Spreadsheet Terminology
- The Interface - Navigating Around a Spreadsheet

### Module 2



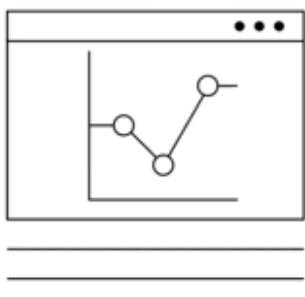
- Selecting, Entering and Editing Data
- Copying and Auto-Filling Data
- Formatting Data
- Using Functions and Formulas

## Module 3



- Cleaning and Wrangling Data Using a Spreadsheet
- The Fundamentals of Data Quality and Data Privacy
- Removing Duplicated and Inaccurate Data
- Removing Empty Rows
- Removing Data Inconsistencies and Whitespaces
- Using the 'Flash Fill' and 'Text to Columns' Features

## Module 4

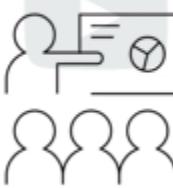


- Analyzing Data Using Spreadsheets
- Filtering and Sorting Data
- Using Common Data Analysis Functions
- Creating and Using Pivot Tables
- Creating and Using Slicers and Timelines

At the end of the course you will complete a series of hands-on labs which will guide you on how to create your first deliverable as a data analyst; this will involve you understanding what the business scenario is, cleaning and preparing your data, and analyzing your data.



**Hands-on Labs**



**Business Scenarios**



**Clean, Prepare,  
& Analyze Data**

At the end of this course you will:

- ✓ Understand how spreadsheets can be used as a data analysis tool
- ✓ Understand when to use spreadsheets as a data analysis tool and their limitations
- ✓ Create a spreadsheet and explain its basic functionality
- ✓ Perform data wrangling and data cleaning tasks using Excel
- ✓ Analyze data using filter sort and pivot table features within Excel spreadsheets
- ✓ Perform some intermediate level data wrangling and data analysis tasks to address a business scenario.

# Introduction to Spreadsheets

In this first video of the course, we will list some of the **common spreadsheet applications** available, learn about the key capabilities of spreadsheets, and discuss why spreadsheets might be a useful tool for a Data Analyst.

There are several spreadsheet applications available in the marketplace; some of them are more widely known and used than others, and some are free, while others need to be paid for.

By far the most commonly used spreadsheet application, and the most fully featured of them all is Microsoft **1. Excel**.

- The desktop version comes in a paid form as part of the Office suite and some Microsoft 365 subscriptions,
- but there is also a web-based cut-down version called Excel for the web, also known as Excel Online.
- The online version is free to users with a Microsoft account, but does not offer all the advanced features that the desktop version provides.

The next most popular is **2. Google Sheets**,

- which offers a lot, though not all of the features that Excel provides, and is free with a Google account.
- This is a web-based application and it integrates nicely with other Google apps, such as Google Forms, Google Analytics, and Google Data Studio.

Then there is **3. LibreOffice Calc**,

- a totally free and open source desktop spreadsheet application that offers more basic functionality than Excel or Google Sheets,
- but still has a lot of the tools you need for data analysis, such as charts, conditional formatting, and pivot tables.

Other spreadsheet apps include **Zoho Sheet** (a fully-featured web-based application that is comparable with Google Sheets), **OpenOffice Calc**, **Quip from Salesforce**, **Smartsheet** (which is predominantly for project management), and **Apple Numbers**, (which is included with Apple devices such as Mac computers and is also available on the App Store for other Apple devices).

So, there are many spreadsheet application options open to you, from fully-featured to basic, from cloud-based to desktop apps, from paid-for to free versions. It's up to you to decide which one best fits your needs and your budget.

Spreadsheets provide several advantages over manual calculation methods.

- For example, once you have your formulas correctly written, you can be assured that your calculations are accurate, and that the calculations will be performed automatically for you.
- Spreadsheets also help keep your data organized and easily accessible.
- Your data can be easily formatted, filtered, and sorted to suit your needs.
- If you do make mistakes in your data entry or your calculations you can easily edit them, undo them, or use error-checking tools to help remedy those mistakes.
- And lastly, you can analyze data in spreadsheets, and create charts, graphs, and reports to help visualize your data analysis.

Since spreadsheet software for personal computers first appeared on the market in the 1970s, with VisiCalc on the Apple II PC, spreadsheets have come a long way in terms of the capabilities and features they now offer businesses, from uncomplicated tables and relatively simple computations to powerful tools for the analysis, management, and visualization of enormous sets of data.

The most common business uses for spreadsheet applications include the following:

- Data Entry and Storage
- Comparing Large Datasets
- Modelling and Planning, Charting
- Identifying Trends, Flowcharts for Business Processes
- Tracking Business Sales
- Financial Forecasting
- Statistical Analysis
- Profit and Loss Accounting
- Budgeting
- Forensic Auditing
- Payroll and Tax Reporting
- Invoicing
- Scheduling

And away from the business side of things, other typical uses include

- Personal Expenses
- Household Budgeting Recipe library
- Fitness Tracking
- Calorie Counting & Weight Monitoring
- Sports Leagues such as Fantasy Football
- Cataloging Music Libraries
- Contact Lists
- Shopping Lists
- Christmas Card Lists

As a Data Analyst, you can use spreadsheets as a tool for your data analysis tasks, including:

- Collecting and harvesting data from one or more distributed and different sources.

- Cleaning data to remove duplicates, inaccuracies, errors, and resolve missing values to improve the quality of the data.
- Analyzing data by filtering, sorting, and interpreting it to determine what useful information can be gleaned from it.
- And visualizing data, to help you tell a story about your data analysis findings to key business stakeholders and any other interested parties within your organization.

In this video, we had an introduction to spreadsheets. We learned about some common spreadsheet applications, what the main capabilities of spreadsheets are, and why spreadsheets might be a useful tool for a Data Analyst.

## Introduction to Spreadsheets

Looking at:

- common spreadsheet applications available
- learn about the key capabilities of spreadsheets
- discuss why spreadsheets might be a useful tool for a Data Analyst.
- How data analyst can use spreadsheets



### Spreadsheet Applications

1.



- Most commonly used
- Part of the Office suite
- Web-based version called Excel for the web

2.



- Free with a Google account
- Web-based application
- Integrates with other Google apps

3.



- Free and open source desktop spreadsheet application

4.

Other spreadsheet apps include:



Zoho Sheet (a fully-featured web-based application that is comparable with Google Sheets),



- OpenOffice Calc, Quip from Salesforce, Smartsheet (which is predominantly for project management),
- Apple Numbers, (which is included with Apple devices such as Mac computers and is also available on the App Store for other Apple devices).

# Advantages of Using Spreadsheets

- Accurate calculations
- Automatic calculations
- Organize and access data
- Format, filter, and sort data
- Edit, Undo, and Error-check
- Analyze data
- Create charts, graphs, and reports

ID	Filename	Model	Type ID	Width	Height	Mtr	Colors	Resolution (inches)	Length	VGA	AV	S-Video	AV Out	AV In	Image	Create Rendition	Create Preview	Thumbnail	Asset ID	File ID
1	max000101.xls	00	813	480	4	15	20	1024x768	22	11								1100001	541162	
2	max000102.xls	00	813	480	1	14	30	1024x768	22	11								1100002	541163	
3	max000103.xls	00	813	480	1	14	30	1024x768	22	11								1100003	541164	
4	max000104.xls	00	813	480	1	14	30	1024x768	22	11								1100004	541165	
5	max000105.xls	00	813	480	4	17	30	1024x768	22	11								1100005	541166	
6	max000106.xls	00	813	480	4	17	30	1024x768	22	11								1100006	541167	
7	max000107.xls	00	813	480	1	11	30	1024x768	22	11								1100007	541168	
8	max000108.xls	00	813	480	1	17	30	1024x768	22	11								1100008	541169	
9	max000109.xls	00	813	480	1	17	30	1024x768	22	11								1100009	541170	
10	max000110.xls	00	813	480	1	17	30	1024x768	22	11								1100010	541171	
11	max000111.xls	00	813	480	1	17	30	1024x768	22	11								1100011	541172	
12	max000112.xls	00	813	480	1	17	30	1024x768	22	11								1100012	541173	
13	max000113.xls	00	813	480	1	17	30	1024x768	22	11								1100013	541174	
14	max000114.xls	00	813	480	1	17	30	1024x768	22	11								1100014	541175	
15	max000115.xls	00	813	480	1	17	30	1024x768	22	11								1100015	541176	
16	max000116.xls	00	813	480	1	17	30	1024x768	22	11								1100016	541177	
17	max000117.xls	00	813	480	1	17	30	1024x768	22	11								1100017	541178	
18	max000118.xls	00	813	480	1	17	30	1024x768	22	11								1100018	541179	
19	max000119.xls	00	813	480	1	17	30	1024x768	22	11								1100019	541180	
20	max000120.xls	00	813	480	1	17	30	1024x768	22	11								1100020	541181	
21	max000121.xls	00	813	480	1	17	30	1024x768	22	11								1100021	541182	
22	max000122.xls	00	813	480	1	17	30	1024x768	22	11								1100022	541183	
23	max000123.xls	00	813	480	1	17	30	1024x768	22	11								1100023	541184	
24	max000124.xls	00	813	480	1	17	30	1024x768	22	11								1100024	541185	
25	max000125.xls	00	813	480	1	17	30	1024x768	22	11								1100025	541186	
26	max000126.xls	00	813	480	1	17	30	1024x768	22	11								1100026	541187	
27	max000127.xls	00	813	480	1	17	30	1024x768	22	11								1100027	541188	
28	max000128.xls	00	813	480	1	17	30	1024x768	22	11								1100028	541189	
29	max000129.xls	00	813	480	1	17	30	1024x768	22	11								1100029	541190	
30	max000130.xls	00	813	480	1	17	30	1024x768	22	11								1100030	541191	
31	max000131.xls	00	813	480	1	17	30	1024x768	22	11								1100031	541192	
32	max000132.xls	00	813	480	1	17	30	1024x768	22	11								1100032	541193	
33	max000133.xls	00	813	480	1	17	30	1024x768	22	11								1100033	541194	
34	max000134.xls	00	813	480	1	17	30	1024x768	22	11								1100034	541195	
35	max000135.xls	00	813	480	1	17	30	1024x768	22	11								1100035	541196	
36	max000136.xls	00	813	480	1	17	30	1024x768	22	11								1100036	541197	
37	max000137.xls	00	813	480	1	17	30	1024x768	22	11								1100037	541198	
38	max000138.xls	00	813	480	1	17	30	1024x768	22	11								1100038	541199	
39	max000139.xls	00	813	480	1	17	30	1024x768	22	11								1100039	541200	
40	max000140.xls	00	813	480	1	17	30	1024x768	22	11								1100040	541201	
41	max000141.xls	00	813	480	1	17	30	1024x768	22	11								1100041	541202	
42	max000142.xls	00	813	480	1	17	30	1024x768	22	11								1100042	541203	
43	max000143.xls	00	813	480	1	17	30	1024x768	22	11								1100043	541204	
44	max000144.xls	00	813	480	1	17	30	1024x768	22	11								1100044	541205	
45	max000145.xls	00	813	480	1	17	30	1024x768	22	11								1100045	541206	
46	max000146.xls	00	813	480	1	17	30	1024x768	22	11								1100046	541207	
47	max000147.xls	00	813	480	1	17	30	1024x768	22	11								1100047	541208	
48	max000148.xls	00	813	480	1	17	30	1024x768	22	11								1100048	541209	
49	max000149.xls	00	813	480	1	17	30	1024x768	22	11								1100049	541210	
50	max000150.xls	00	813	480	1	17	30	1024x768	22	11								1100050	541211	
51	max000151.xls	00	813	480	1	17	30	1024x768	22	11								1100051	541212	
52	max000152.xls	00	813	480	1	17	30	1024x768	22	11								1100052	541213	
53	max000153.xls	00	813	480	1	17	30	1024x768	22	11								1100053	541214	
54	max000154.xls	00	813	480	1	17	30	1024x768	22	11								1100054	541215	
55	max000155.xls	00	813	480	1	17	30	1024x768	22	11								1100055	541216	
56	max000156.xls	00	813	480	1	17	30	1024x768	22	11								1100056	541217	
57	max000157.xls	00	813	480	1	17	30	1024x768	22	11								1100057	541218	
58	max000158.xls	00	813	480	1	17	30	1024x768	22	11								1100058	541219	
59	max000159.xls	00	813	480	1	17	30	1024x768	22	11								1100059	541220	
60	max000160.xls	00	813	480	1	17	30	1024x768	22	11								1100060	541221	
61	max000161.xls	00	813	480	1	17	30	1024x768	22	11								1100061	541222	
62	max000162.xls	00	813	480	1	17	30	1024x768	22	11								1100062	541223	
63	max000163.xls	00	813	480	1	17	30	1024x768	22	11								1100063	541224	
64	max000164.xls	00	813	480	1	17	30	1024x768	22	11								1100064	541225	
65	max000165.xls	00	813	480	1	17	30	1024x768	22	11								1100065	541226	
66	max000166.xls	00	813	480	1	17	30	1024x768	22	11								1100066	541227	
67	max000167.xls	00	813	480	1	17	30	1024x768	22	11								1100067	541228	
68	max000168.xls	00	813	480	1	17	30	1024x768	22	11								1100068	541229	
69	max000169.xls	00	813	480	1	17	30	1024x768	22	11								1100069	541230	
70	max000170.xls	00	813	480	1	17	30	1024x768	22	11								1100070	541231	
71	max000171.xls	00	813	480	1	17	30	1024x768	22	11								1100071	541232	
72	max000172.xls	00	813	480	1	17	30	1024x768	22	11								1100072	541233	
73	max000173.xls	00	813	480	1	17	30	1024x768	22	11								1100073	541234	
74	max000174.xls	00	813	480	1	17	30	1024x768	22	11								1100074	541235	
75	max000																			

Non-business uses for spreadsheets applications include the following:

- Personal Expenses
- Household Budgeting Recipe library
- Fitness Tracking
- Calorie Counting & Weight Monitoring
- Sports Leagues such as Fantasy Football
- Cataloging Music Libraries
- Contact Lists
- Shopping Lists
- Christmas Card Lists

### How a Data Analyst Uses Spreadsheets?

- Personal Expenses
- Household Budgeting Recipe library
- Fitness Tracking
- Calorie Counting & Weight Monitoring
- Sports Leagues such as Fantasy Football
- Cataloging Music Libraries
- Contact Lists
- Shopping Lists
- Christmas Card Lists

# Spreadsheet Basics - Part 1

Now that we have a basic understanding of what spreadsheet software is available, and why spreadsheets might be a useful tool for a Data Analyst, let's get started on looking at some of the basics of using a spreadsheet application.

In these videos we will be using the full 'desktop' version of Excel, but the majority of the tasks that we will perform can also be done using Excel 'on the web', also known as Excel Online, and other spreadsheet applications such as Google Sheets.

Let's first cover some basic spreadsheet terminology.

When you open **Excel**, you have the option of creating a new blank workbook or opening an existing workbook. We're going to choose New, and then Blank workbook.

Workbooks are the

- highest-level component in Excel and are represented as a .XLSX file. So, when you open an existing workbook or create a new workbook you are in fact working with a .XLSX file.
- The workbook contains all your data, calculations, and functions, and contains several other underlying elements that make up a workbook.
- A workbook consists of one or more worksheets, each of which is represented by a tab in Excel. Each worksheet is given a name which is displayed on the corresponding tab for the worksheet. By default, each tab is named Sheet1, then Sheet2, and so on. To make these worksheet tabs more meaningful it is usual to rename them, so they make more sense in relation to the worksheet's purpose.

For example, you might call a worksheet January Sales, or perhaps the name of a region or store, or even an office or department. To do this, right-click the tab and choose Rename. Instead of right-clicking to rename, you can also just double-click the name of a worksheet tab to rename it.

Essentially, worksheet tabs can be named anything you want to fit your particular needs to make it easier to understand what that worksheet represents. Note that a worksheet that is highlighted, as the Tire Sales worksheet tab is here, is referred to as the active worksheet. If you want to order your worksheets in a different way, that is very simple to do. Either drag a worksheet tab to the left or right and drop it in the place you want, which is represented by the little black arrow, or if you are not comfortable with dragging and dropping, then the longer way of doing that is to right-click the worksheet tab, select Move or Copy, and then in the list titled Before sheet, select where you want your worksheet tab to be placed, and click OK.

Every worksheet is made up of a lot of rectangular boxes called **cells**. These cells will contain your data, which may be text, numbers, formulas, or calculation results. Cells are organized in columns, which run vertically down the screen and use a letter system; this is column B for instance. And rows, which run horizontally across the screen and use a numeric system; this is row 7 for example. Each cell is represented by a cell reference which is essentially just its column letter and row number.

For example, if we click somewhere near the center of this worksheet, we now have the cell M20 selected.

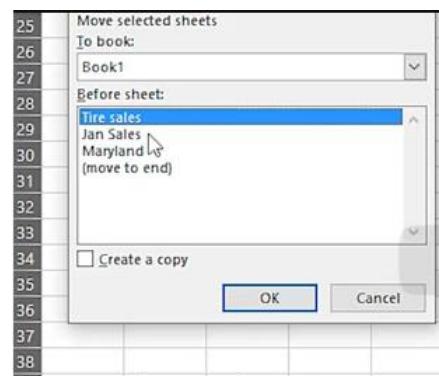
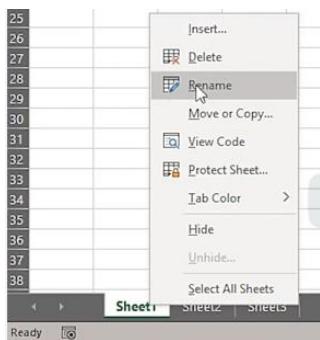
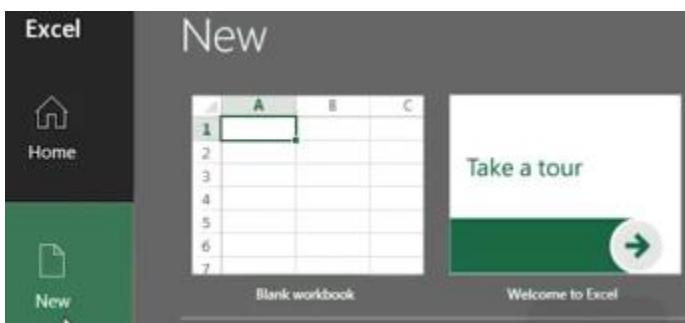
This is usually referred to as the ‘active cell’. This is not only indicated by the highlighted edges of the cell but also if you look in the top left corner of the worksheet, you will see its cell reference is noted in the little box.

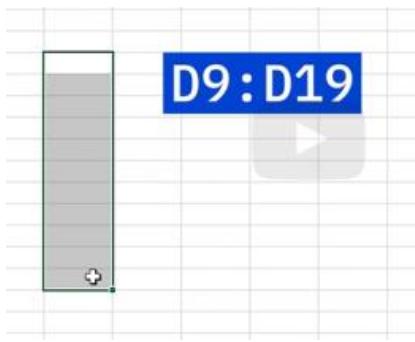
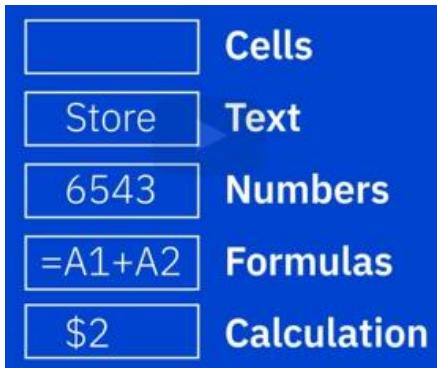
Here you can see it says M20. One important thing to note here is that cells are always referenced by their column letter first then their row number; so, column M, and row 20. The last element of a workbook I want to mention is a cell range. This identifies a collection of several cells selected together; that could mean a few cells in the same row or the same column, or it could mean several rows and columns together. This can either be done using the mouse by selecting the first cell then ‘dragging’ down or across to include other cells; or you can use SHIFT+ arrow keys.

This range of cells is often referred to as an array, and it’s most commonly used as a reference in calculations and formulas. For example, if you wanted to add up all the values in a column between cells D9 and D19 you would specify this cell range within a formula. Note that cell ranges are notated using a full colon (:) between the cell references; so, in this example it would be D9:D19, or to specify a few cells in the same row it might be D9:H9, or to select several rows and columns it might be D9:H19. We will see this notation in use later in this course when we start looking at calculations and formulas. These cell ranges could even be a reference point to cells contained on another worksheet; this is usually referred to as a 3D reference. We can now close this workbook and we don’t need to save it. In this video, we learned about some of the basic terminology of spreadsheet elements. In the next video, we will discuss how to navigate around a spreadsheet, how to use the ribbon and menus, and how to select data.

## Spreadsheet Basics - Part 1

### Basics of using spreadsheet application





## Spreadsheet Basics - Part 2

Now that we have a basic understanding of the main elements that make up a worksheet, let's see how to move around a spreadsheet, get familiar with the ribbon and menus, and learn how to select data in a worksheet.

To open a sample file, we click File. This opens Backstage View. Here you can create a new workbook, or open, save or print a workbook. You can also access Excel Options. Now, we want to open our sample file. So, we click Open, and either select it from my Recent list, or click Browse to find the data file we want.

The first thing we should do is get acquainted with the ribbon and menus. Notice that on the ribbon at the top we have several tabs. Some of these tabs may be familiar to you from other Office products, such as the Home, Insert, and View tabs, while others might be new to you, such as Formulas, Data, and Power Pivot.

To make a little more workspace for ourselves we can hide this ribbon by double-clicking any tab, and to unhide it, we do the same. The other option is to use the shortcut key CTRL+F1. The ribbon is organized into groups of buttons to make them easier to find.

So, on the Home tab we have groups for Font, Alignment, Number, Styles, and so on. Some of these groups contain all the available buttons on the ribbon when viewing in full screen, such as Styles and Cells, but other ribbon groups have more options, which we access by clicking the little arrow icon in the bottom right corner of the group, as can be seen here on the Number group for example.

The next item I want to point out is the Quick Access Toolbar at the top of the screen above the ribbon. As the name suggests this is where you can quickly access the tools you use most often. You can see we already have some tools in this toolbar such as Save, Undo, Redo, New, and Open.

But we can add other tools to the toolbar if we wish. So if we click the drop-down arrow in the toolbar and then select a tool we will use a lot, such as Sort Ascending, that will be added, and we will also add the Sort Descending button too.

Now we need to be comfortable with moving around a worksheet. You can simply use the arrow keys to move left, right, up, and down 1 cell at a time. But you can also use Page Down and Page Up to move around a bit faster, which is especially useful if you have lots of rows of data. And to move even quicker up or down a large datasheet use the vertical scroll bar, and to move left or right use the horizontal scroll bar. Again, these can be very useful when you have a large data set.

There are also some useful shortcuts you can use.

CTRL+Home key for example takes you back to the start of the worksheet (i.e. cell A1).

CTRL+End takes you to the cell at the end of your data in the worksheet.

**CTRL+Down** arrow takes you to the end of the column you're in, while **CTRL+Up** arrow takes you back to the top of that column.

So a quick way to find out how many rows of data you have in your worksheet is to go to the first cell in your data and press **CTRL+Down** arrow to see the last row of data.

So here you can see we have 160 rows. Now how do we go back to the top again? **CTRL+Home** will do it. So far, we have seen how to navigate around our worksheet and its data, now we need to look at how we select data. This is very important because you often need to select data to move it, copy it, or select it in a formula. The simplest selection is a single cell, usually done with a mouse or maybe a directional arrow key.

The next step up is to select multiple cells together, and this can be done either with a mouse by dragging from one cell to additional adjoining cells, or you can use the **SHIFT** key with directional arrow keys.

Next up is selecting a single column or row which is done simply by selecting the letter at the top of a column, or the number on the left of a row. Then we can progress to selecting multiple columns and rows, by clicking the mouse button, holding it down and dragging across more columns. Or if you are not comfortable with dragging you can also select the column first, then hold **SHIFT+Arrow** keys to select multiple columns. The same applies to rows too.

However, if you have data in non-contiguous rows or columns (i.e. not next to each other) you can select the first column, then use the **CTRL** key to select another unconnected column, such as columns C and F here. The largest thing you might want to select is the whole worksheet which you can do by clicking in the top left corner of the cells. However, this selects the entire worksheet including all the empty rows and columns; so if you only want the data in your worksheet, you can use the shortcut **CTRL+A**. A word of warning when selecting data in cells, rows, and columns; there are 3 types of cross symbols that you might see when working with selected cells.

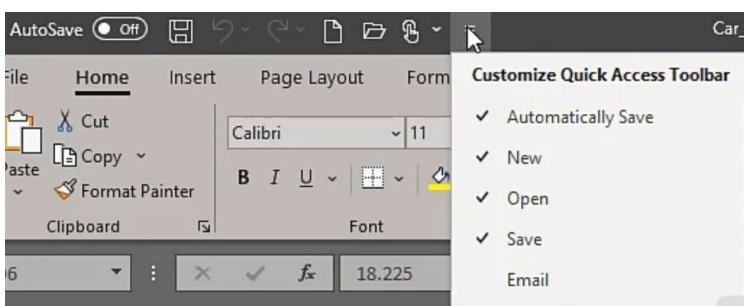
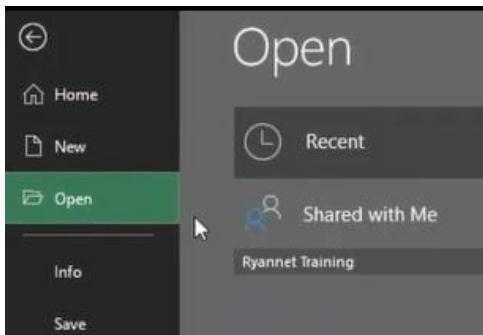
The first one is the large white cross that you see when you select a cell as can be seen here in cell A4, this is the Select cross that we have been using already in this video to select cells.

The second type you might see is when you hover over the bottom edge of a cell and see a thin black CROSS-type symbol with arrows on each point.... this is the Move symbol and would move the cell data to another location.

The last type is the small thin black cross that is seen when you hover over the bottom right corner of a cell; this is the Fill Handle or Copy symbol and it fills (or copies) the cell data to another location.

In this video, we learned how to move around a spreadsheet, became familiar with the ribbon and menus, and learned how to select data in a worksheet. In the next video, we will discuss how to enter data, how to copy and paste data, and how to format data in a spreadsheet.

# Spreadsheet Basics - Part 2



Category	Car Model	11.105	18.425	Passenger
Cadillac	Escalade	14.785		Car
Chevrolet	Cavalier	145.519	9.25	Passenger

Move to  
first cell  
**Ctrl + Home**

CTRL+Home key for example takes you back to the start of the worksheet (i.e. cell A1).

Cadillac	Catera	11.185	18.225	Pass
Cadillac	Escalade	14.785		Car
Chevrolet	Cavalier	145.519	9.25	Pass
Move to last cell				
<b>Ctrl + End</b>				
Chrysler	Sebring Conv.	32.775	14.18	Pass
Chrysler	Concorde	31.148	13.725	Pass

**CTRL+End** takes you to the cell at the end of your data in the worksheet.

Cadillac	Escalade	14.785	Car
Chevrolet	Cavalier	145.519	9.25 Passenger
Move to end column <b>Ctrl + ↓</b>			
Chrysler	Sebring Conv.	32.775	14.18 Passenger
Chrysler	Concorde	21.148	12.775 Sedan

CTRL+Down arrow takes you to the end of the column

Chevrolet	Malibu	15.170	11.775
Chevrolet	Lumina	24.629	10.31
Move to end column <b>Ctrl + ↑</b>			
Chrysler	Cirrus	32.306	12.64

CTRL+Up arrow takes you back to the top of that column.

## Select data

Manufacturer	Model	Sales (\$ in thousands)	Year Resale Value	Vehicle_type	Price (\$ in thousands)	Engine Size
Acura	Integra	16.919	16.36	Passenger	21.5	1.8

With a mouse by dragging from one cell to additional adjoining cells, or you can use the SHIFT key with directional arrow keys.

Model
Integra
TL
CL
RL
A4
A6

selecting a single column or row which is done simply by selecting the letter at the top of a column, or the number on the left of a row.

	C4	A	B	C	D
1				16.919	
2					
3	Manufacturer	Model	Sales (\$ in thousands)	Year Resale	
4	Acura	Integra	16.919		
5	Acura	TL	39.384		
6	Acura	CL	14.114		
7	Acura	RL	8.588		
8	Audi	A4	20.397		
9	Audi	A6	18.78		
10	Audi	A8	1.38		

Select is the whole worksheet which you can do by clicking in the top left corner of the cells.

Chevrolet	Malibu	125,329
Chevrolet	Lumina	24,629
Select all data Ctrl + A		
Chrysler	Concorde	31,348

## 3 Types of Cross Symbols

1. large white cross

Manufacturer	Model
Acura	+ Integra

2. thin black cross-type Move symbol

Manufacturer	Model
Acura	Integra
Acura	TL

Move symbol and would move the cell data to another location.

3. small thin black cross

Manufacturer	Model
Acura	Integra
	+

seen when you hover over the bottom right corner of a cell; this is the Fill Handle or Copy symbol and it fills (or copies) the cell data to another location.



IBM Developer  
SKILLS NETWORK

# Reading: Excel Keyboard Shortcuts

**Estimated time needed:** n/a

The table below lists keyboard shortcuts for some of the most common Excel tasks.

Task	Shortcut
Close a workbook	Ctrl+W
Open a workbook	Ctrl+O
Save a workbook	Ctrl+S
Copy	Ctrl+C
Cut	Ctrl+X
Paste	Ctrl+V
Undo	Ctrl+Z
Remove cell contents	Delete
Bold	Ctrl+B
Open context menu	Shift+F10
Expand or collapse the ribbon	Ctrl+F1
Move up one cell in the worksheet	Up arrow key

Move down one cell in the worksheet	Down arrow key
Move one cell left in the worksheet	Left arrow key
Move one cell right in the worksheet	Right arrow key
Move to the edge of the current data region in the worksheet (e.g. end of column)	Ctrl+Arrow key (e.g. Ctrl+Down arrow)
Move to the last cell on a worksheet	Ctrl+End
Move to the beginning of a worksheet	Ctrl+Home
Extend the selection of cells to the last used cell on a worksheet (lower right corner)	Ctrl+Shift+End
Move to the cell in the upper-left corner of the window (when Scroll Lock is On)	Home+Scroll Lock
Move one screen down in a worksheet	Page Down
Move one screen up in a worksheet	Page Up
Move one screen to the right in a worksheet	Alt+Page Down
Move one screen to the right in a worksheet	Alt+Page Down
Move one screen to the left in a worksheet	Alt+Page Up
Move to the next sheet in a workbook	Ctrl+Page Down
Move to the previous sheet in a workbook	Ctrl+Page Up
Edit the active cell and put the cursor at the end of the cell's contents	F2
Enter the current time	Ctrl+Shift+colon (:)
Enter the current date	Ctrl+semi-colon (;)



# Hands-on Lab 1: Introduction to Excel for the web

**Estimated time needed:** 20 minutes

Microsoft Excel is the most widely used spreadsheet software even three decades after its initial release. For all these years it has been available as a standard application that needed to be installed on your desktop; but it is not just a desktop app anymore. Now, you can even use Excel when you're online by using 'Excel for the web' - and run it right in your web browser without installing anything on your desktop!

'Excel for the web' (sometimes referred to as Excel Online) can be used at no charge as part of a free Microsoft account. Although it does not have all of the capabilities of the desktop and paid online versions, the free web version provides many of the key features.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The dataset used in this lab comes from the following source: <https://www.kaggle.com/sudalairajkumar/indian-startup-funding> under a [CC0: Public Domain license](#). Acknowledgement and thanks also goes to <https://trak.in> who were generous enough to share the data publicly for free.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

# Objectives

After completing this lab, you will be able to:

- Sign up for a Microsoft Account to use Excel for the web
- Sign in, upload and open a workbook in Excel for the web

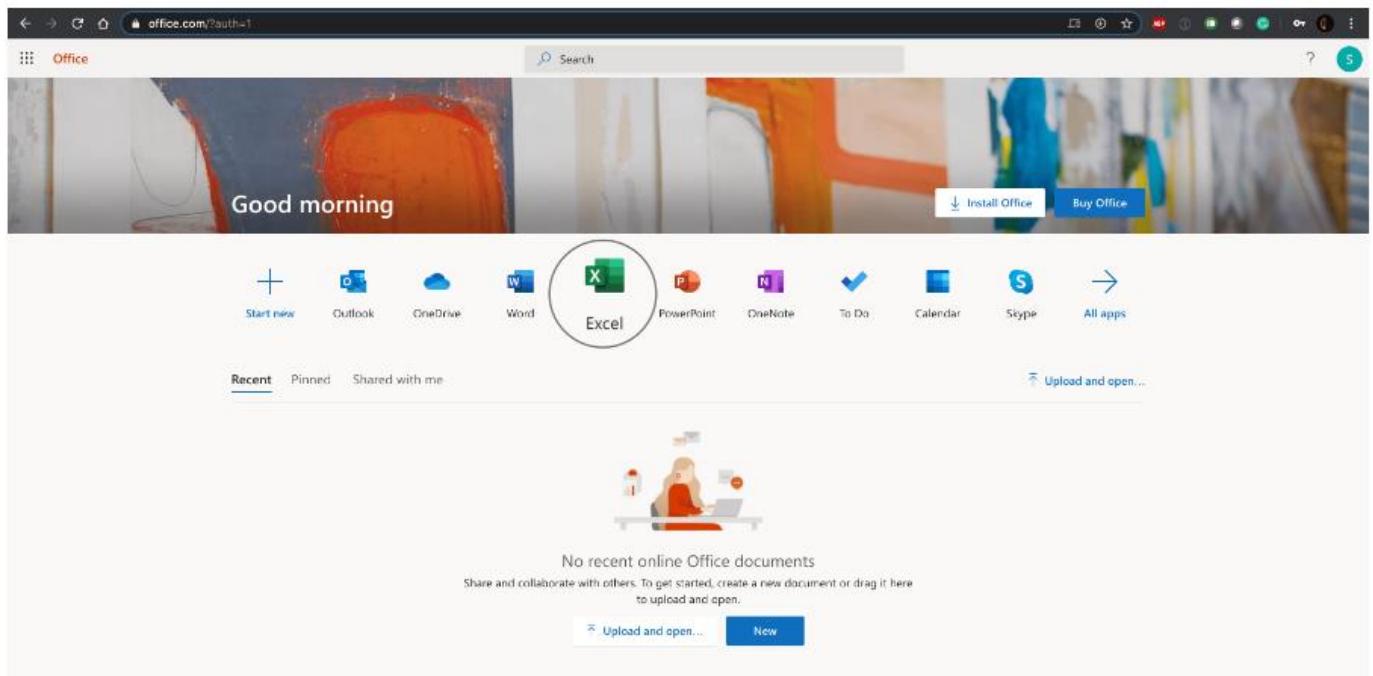
## Exercise 1: Sign-up for a Microsoft Account to Use Excel for the web

In this exercise, you will sign up for a Microsoft Account to use Excel for the web.

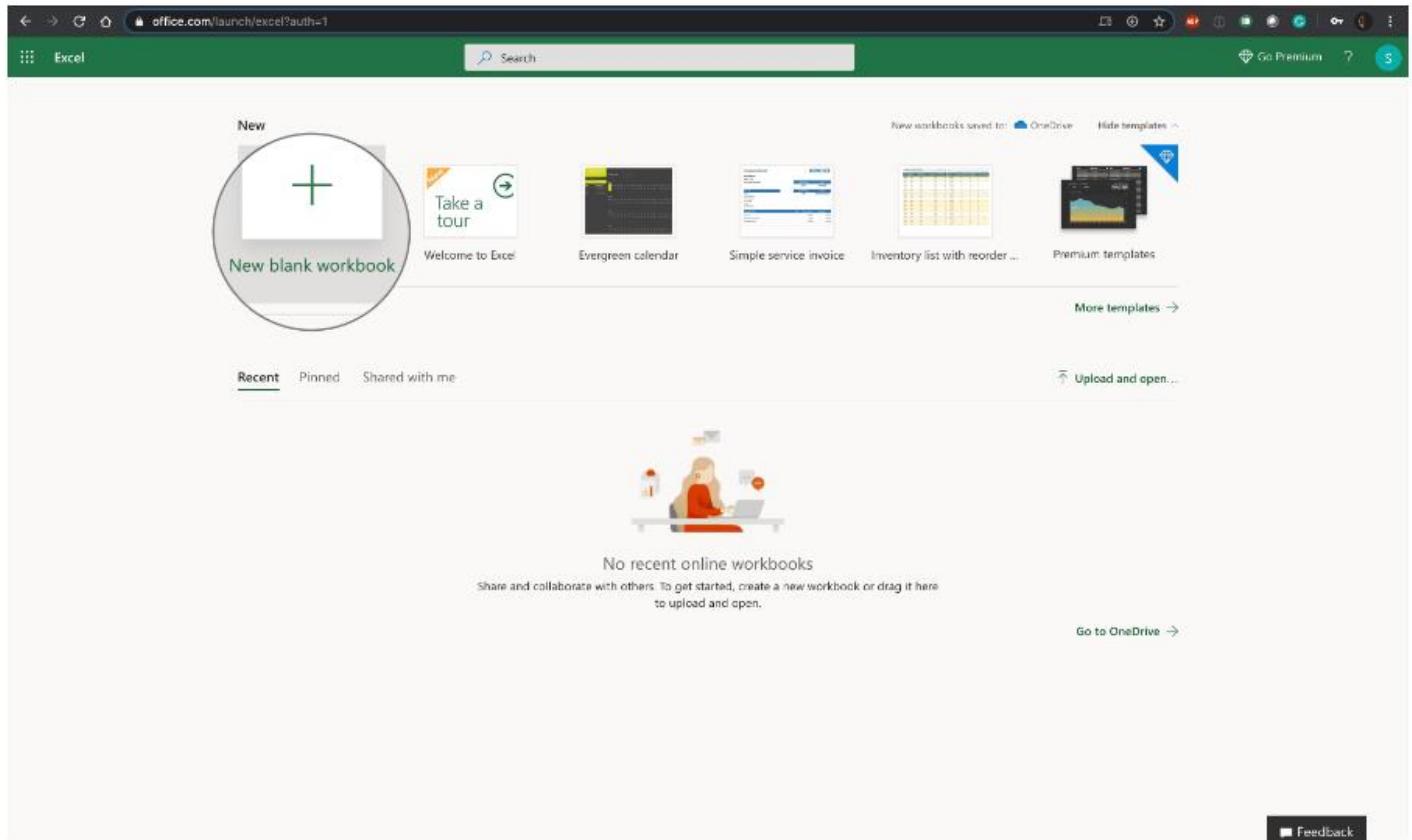
- If you already have a Microsoft account, you can skip Exercise 1 and proceed to Exercise 2 directly.
1. Go to [www.office.com](http://www.office.com). Click **Sign in**

# Task B: Open a new workbook in Excel for the web

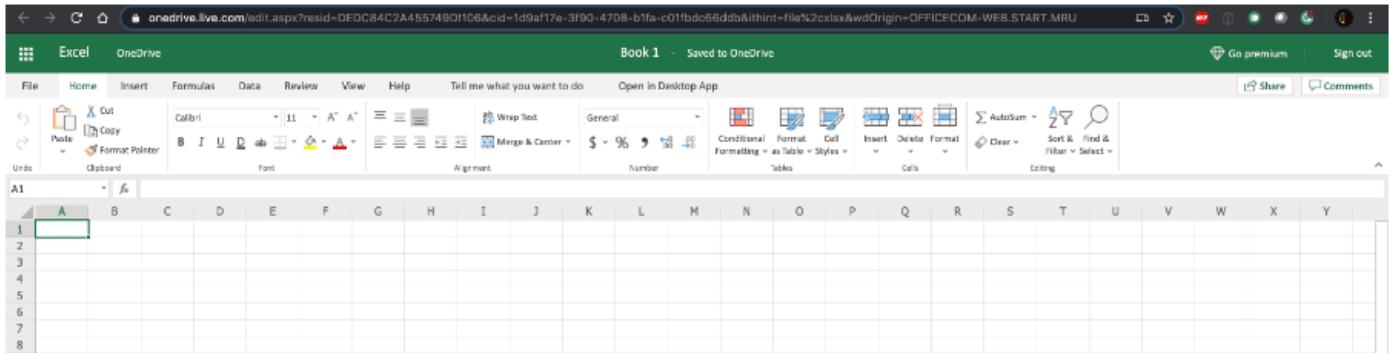
1. Click on the **Excel** icon.



2. Click **New blank workbook**.

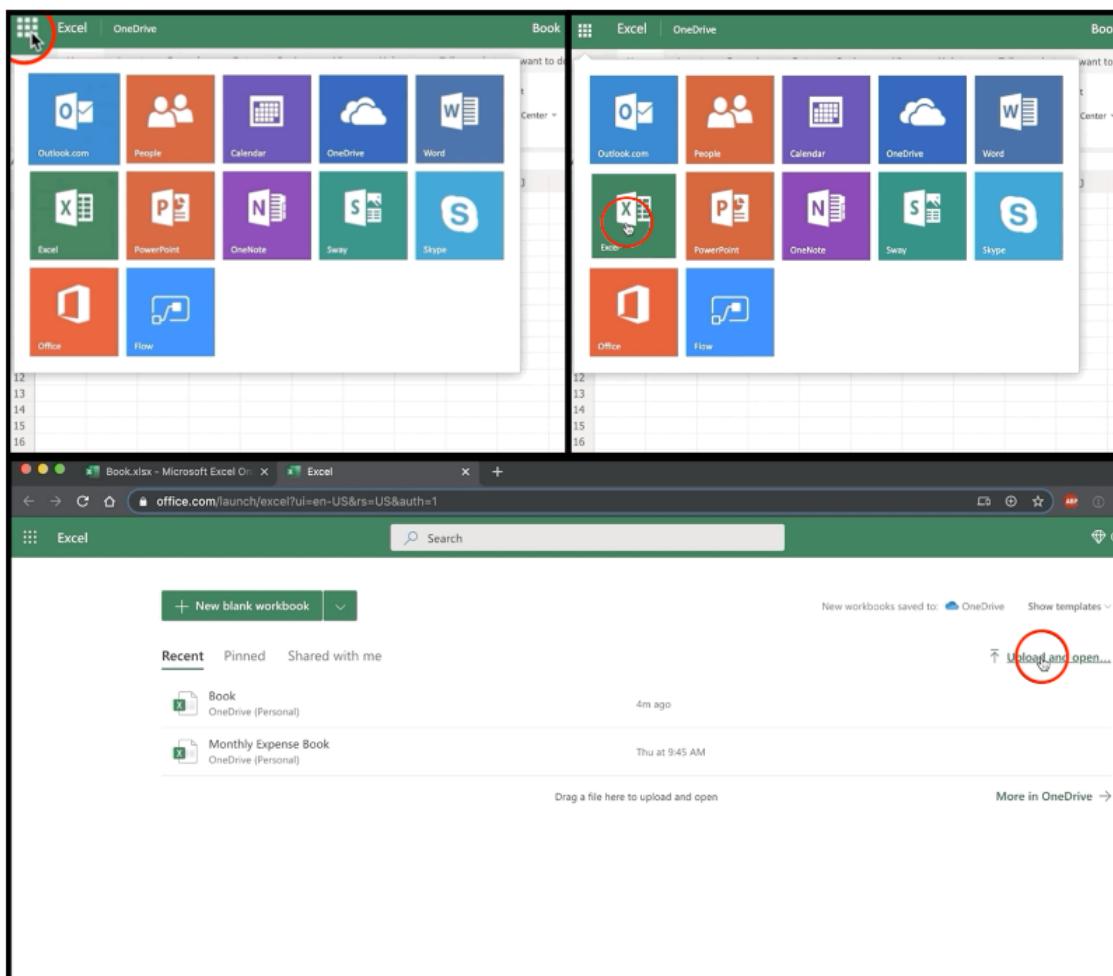


3. You have successfully opened a new workbook in Excel for the web.



## Task C: Upload and Open a workbook in Excel for the web

1. Download the file [\*\*indian\\_startup\\_funding\\_Lab1.xlsx\*\*](#).
2. To upload and open a workbook in Excel for the web, click the **App Launcher** (cube of dots) in the top left corner. Click **Excel** icon. Then click **Upload and open...** and select the [\*\*indian\\_startup\\_funding\\_Lab1.xlsx\*\*](#) file.



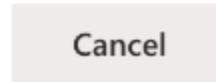
3. The file will be uploaded to your OneDrive of the Microsoft Account you signed up and used to open Excel for the web.

## Uploading to OneDrive

Your file will be opened automatically after upload is complete.

21 KB/21 KB

Tip: You can now drag and drop files to upload them.



4. You have successfully uploaded and opened a workbook in Excel for the web.

C1	Startup Name	Industry Vertical	SubVertical	City / Location	Investors Name	Investment Type	Amount in USD	Remarks
1	Leiskart.com	E-commerce	Online Eyewear Shopping Portal	Pune	SoftBank Vision Fund	Series G	231,000,000.00	
2	Healthians	B2B-focused foodtech startup	Food Solutions For Corporate	Bengaluru	Pavitr, NIFTX, Sabeer Partners and Neoplus	Series C	11,000,000.00	
3	Ustion	E-commerce	Online Meal And Snack Ordering Startups	Bengaluru	Ventex Growth Fund	Series E	1,000,000.00	
4	IndiCred	Finance	Buy Now Pay Later Company	Mumbai		Debt Funding	5,000,000.00	
5	Troll	Media	Experiential Discovery Platform	Bengaluru		Seed Round	2,000,000.00	
6	Itango	Technology	Logistics Services and Solutions	Gurgaon	Salt Partners, Spring Center Investment Ltd.	Series F	70,000,000.00	
7	Doveen	Technology	AgTech	Pune	Sathguru Catalyst Advisors	Series A	6,000,000.00	
8	Rein Games	Gaming	Real money-based gaming startup	Noida	Maxpol Education and Medical Group (MEMG)	Seed Round	50,000,000.00	
9	Carbitario	E-commerce	Automotive	Gurgaon	Ping An Global Voyager Fund	Series D	70,000,000.00	
10	Ornate Space	Aerospace	Satellite Communication	Bengaluru	Munroe Angels, Revalinch Ready	Seed	50,000,000.00	
11	Rystm	FinTech	Mobile Wallet	Noida	Vivek Shekhar Sharma	Runding Round	1,000,000.00	
12	Ave Finance	FinTech	Financial Services To MSMEs	Gurgaon	FinTech	Debt Funding	17,411,265.00	
13	Churnin	Seeds	Recovery software	San Jose	Altimeter Capital, Cuttermill Ventures	Series C	135,000,000.00	
14	Digital Mall Asia	E-commerce	Virtual e-commerce platform	Delhi	Amour Infrastructure	Seed Funding	220,000,000.00	
15	Fundalos School of Music	Education	Music Education	Tulsi Nagar	IAN Fund and 556 Consumer Partners	Seed Funding	200,000,000.00	
16	Healthians	Health and Wellness	Healthcare services	Gurgaon	DC Dawas Ventures, DG Investa	Series B	12,000,000.00	
17	Medibazaar	Healthcare	i2B platform for medical supplies	Mumbai	Ackermans & van Haaren, HealthQzard, Rebright Partners, Tropica	Series B	15,890,000.00	
18	Burger Singh	Food and Beverage	Indian Burger Brand	Gurgaon	RB Investments	Venture	unclosed	
19	Ninjascart	B2B Marketing	AgTech	Bengaluru	Trefecta Capital Advisors	Debt Funding	26,000,000.00	
20	Vago Automobile	Last Mile Transportation	Ride sharing app	Kolkata	Matrix Partners, Stellaris Venture Partners, Kaisen Capital	Series B	201,000,000.00	
21	Mites	Health and Wellness	Men's Health and Wellness brand	Gurgaon	Sauces, Raifores Ventures	Series B	486,000.00	
22	Friseworks	Software	Business and customer engagement tool	San Francisco	Sequoia, CapitalA, Accel	Series H	150,000,000.00	
23	Sunstone Education Pvt. Ltd	Education	E-learning	Gurgaon	Prime Venture Partners, LatentView, PSL Venture and GlobalLogic	Seed	1,100,000.00	
24	SuperGaming	Video Games	Social gaming platform	Pune	Dream Innovate	Seed Funding	1,700,000.00	
25	cBikeGo	Last Mile Transportation	Electric bike rental	Amritsar	Starus Buddy	Seed	300,000.00	
26	The Man Company	Consumer Goods	Beauty and Grooming	Gurgaon	Ayushman Khurana	Corporate Round	unknown	
27	Dunam	Customer Service	Delivery Service	Bengaluru	Lightbox	Series D	65,000,000.00	
28	Uzaan	B2B	Business development	Bengaluru	Altimeter Capital, DST Global	Series D	585,000,000.00	
29	PVI Technologies	FinTech	Financial Services	Pune	Matrix Partners India, Sequoia India	Maidan Round	4,500,000.00	
30	Cxentif	FinTech	Invoice discounting platform and SFA	Mumbai	Staff Partners	Series A	3,100,000.00	
31	Digital PS	Advertising, Marketing	Digital marketing firm	Mumbai	TIW Private Equity	Private Equity Round	6,000,000.00	
32	3rdDix	SaaS	Education Technology	Hyderabad	Infinity Ventures Pacific	Series A	5,000,000.00	
33	751	iOT	Building automation system	Bengaluru	Breathrough Energy Ventures	Series A	18,000,000.00	
34	Abeylin Foundry	Information Technology	Data Engineering	Bengaluru	India's Partners	Seed	1,000,000.00	
35	Alamdar Technology	Consumer Technology	Consumer Electronics, Home Appliances	Mumbai	AGL Partners	Series A	10,000,000.00	
36	GOOFIT	Health and Wellness	Wearable Fitness Bands	Mumbai Park	Bear Grylls and Company Ltd (BGCL)	Series C	450,000,000.00	
37	Vesper App	Assurance	Vehicle-based Assurance Software	Bengaluru	India Catalyst, Altera Ventures	Series A	5,000,000.00	
38	Reptile Bike Taxi	Transportation	Bike Taxi	Bengaluru	Westbridge Capital	Series B	3,000,000,000.00 min	
39	Zondrive	Automotives	Road Safety Analytics	San Francisco	X1 Innovate	Series B	37,000,000,000.00 min	
40	Lol Fizz	Consumer Goods	Low carb food for Diabetics	Bengaluru	Rashmi Chaps [founder, FreshMoms], Raveen Sastri (co-founder, M2Seed Round)	Series B	500,000.00 min	
41	Tels	FinTech	Digital Lending Platform	Santa Monica	RPS Ventures	Series D	110,000,000.00 min	
42	AdmittiCard	EdTech	University Admissions	Noida	Growth DNA	Seed Round	1,000,000.00 min	
43	Indowellth	FinTech	Health Management	Gurgaon	Tiger Global Management	Venture Round	15,000,000.00 min	
44	11.11	Financial Planning	Financial Planning	Bengaluru	Prudential Financial	Series C	4,000,000.00 min	



## Hands-on Lab 2: Spreadsheet Basics

**Estimated time needed:** 20 minutes

To get started with a spreadsheet app, you need to know:

- Some of the common terminology around it
- What its key features are
- How to use some basic tools on the ribbon
- How to move around a worksheet
- How to select data in it.

In this lab, you will go through some basic spreadsheet elements, explore the ribbon, navigate around a worksheet and select data.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The dataset used in this lab comes from the following source: <https://www.kaggle.com/sudalairajkumar/indian-startup-funding> under a [CC0: Public Domain license](#). Acknowledgement and thanks also goes to <https://trak.in> who were generous enough to share the data publicly for free.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

## Objectives

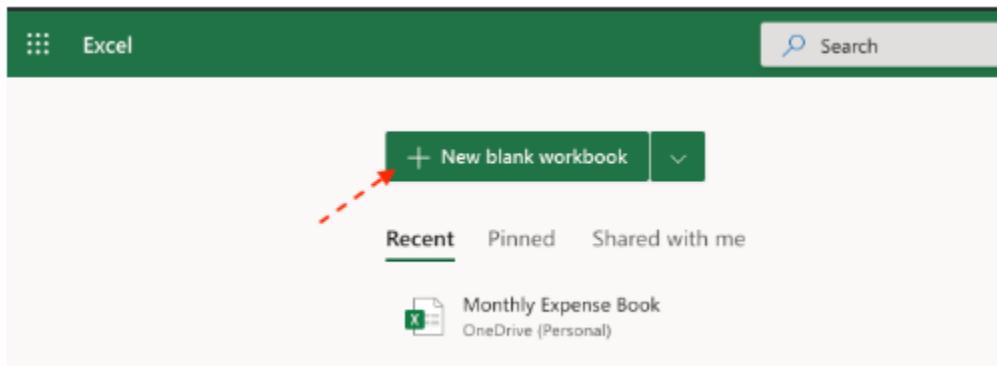
After completing this lab, you will be able to:

- Understand and use the basic elements of a spreadsheet.
- Explore the ribbon, navigate around a worksheet and select data.

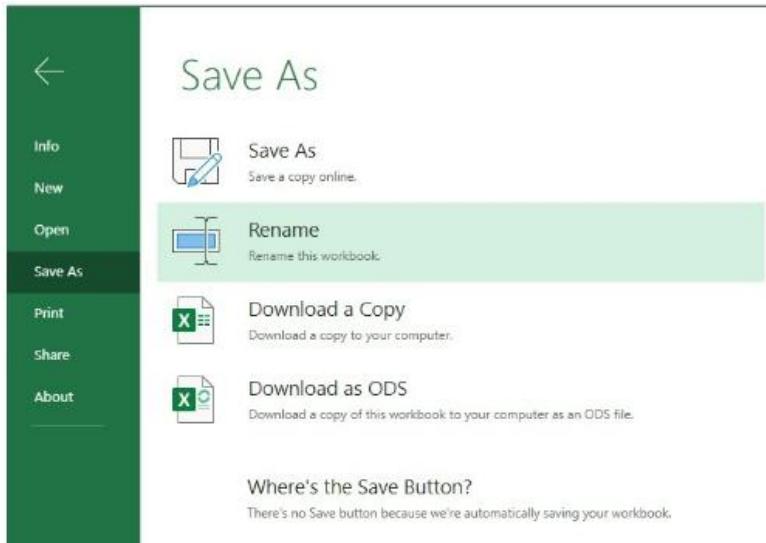
## Exercise 1: Introduction to Basic Spreadsheet Elements

In this exercise, you will learn about some common spreadsheet elements.

1. Open **Excel for the web**. Click on **New blank workbook**.



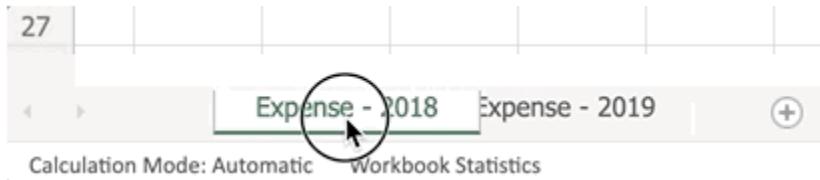
2. The new blank workbook will automatically be saved in Excel for the web as **Book**. To rename the workbook to something more meaningful, click **File**, **Save As**, then choose **Rename**.



3. In the file name box, type **Personal\_Monthly\_Expenditure\_Lab2** and click **OK**
4. In the saved workbook, you will have one worksheet opened, named *Sheet1*. Click + once to add another worksheet. Then, double-click the sheet name tab for **Sheet1** and rename it to **Expense - 2019**. Similarly, rename **Sheet2** as **Expense - 2018**.



5. To maintain an appropriate worksheet tab sequence, click on the worksheet tab **Expense - 2018**, then drag and drop it before the **Expense - 2019** tab.



6. Click on the **Expense - 2018** tab. Select an entire column by clicking on **B** in the top of the worksheet, then select an entire row by clicking on the number **5** in the left of the worksheet. Click cell **B5**, and a green outline will appear around the cell. Now check if you have clicked the correct cell by looking at the cell name box in the top left corner, circled in red below.

	B5	C	D
1			
2			
3			
4			
5			
6			
7			

7. Select several cells in the same row, such as A1:D1 by clicking cell **A1** and then drag the cursor across to **D1**. Similarly, select a cell range in the same column, such as A1:A5 by clicking **A1** and dragging the cursor down to **A5**.

A1	B	C	D
1	A1		
2			
3			
4			
5			
6			

8. Now select a cell range which includes several rows and columns together, such as A1:C5 by clicking **A1** and then dragging the cursor across and down to cell **C5**.

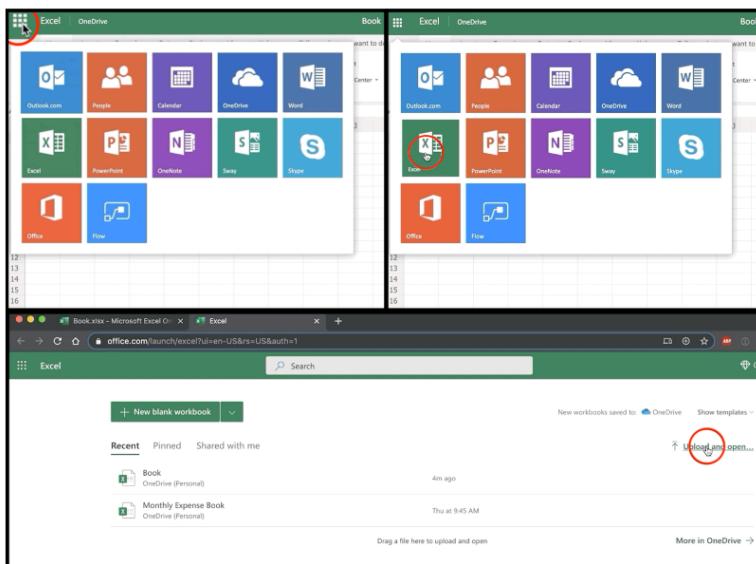
3R x 2C	B	C	D
1	A1	B1	
2			
3			
4			
5			
6			

## Exercise 2: Explore the Ribbon, Navigate around a Worksheet, and Select Data

In this exercise, you will explore the ribbon, then navigate around a worksheet, and select data.

### Task A: Explore the ribbon

1. Download the file [indian\\_startup\\_funding\\_Lab2.xlsx](#).
2. To open a sample file in Excel for the web, click the **App Launcher** (cube of dots) in the top left corner. Click **Excel**, and then click **Upload and open...** and select the **indian\_startup\_funding\_Lab2.xlsx** file.



- Click each of the following tabs in the ribbon; **File, Home, Insert, Formulas, Data, Review, View** to explore each of them and get acquainted with the ribbon. Double-click any of the tabs to hide the ribbon, then do the same again to unhide it.

## Task B: Navigate around a worksheet

- Click on **any cell** and move around the worksheet using the arrow keys; **Up, Down, Left, Right**.

	C7	A	B	C	D	E	F
1	Sr No	Date	Startup Name	Industry Vertical	SubVertical	City	Location
2	17	2019-12-20	Lenskart.com	E-Commerce	Online Eyewear Shopping Portal	Faridabad	
3	12	2019-12-17	Healthians	B2B-focused foodtech startu	Food Solutions For Corporate	Bengaluru	
4	13	2019-12-16	Licious	E-Commerce	Online Meat And Seafood Ordering Start	Bengaluru	
5	14	2019-12-16	InCred	Finance	Non-Banking Financial Company	Mumbai	
6	15	2019-12-14	Trell	Video	Experience Discovery Platform	Bengaluru	
7	11	2019-12-13	Rivigo	Technology	Logistics Services and Solutions	Gurgaon	
8	8	2019-12-12	Ecozen	Technology	Agritech	Pune	
9	16	2019-12-11	Rein Games	Gaming	Real money based gaming startup	Noida	
10	9	2019-12-06	CarDekho	E-Commerce	Automobile	Gurgaon	
11	10	2019-12-03	Dhruba Space	Aerospace	Satellite Communication	Bengaluru	
12	32	2019-11-25	Paytm	FinTech	Mobile Wallet	Noida	
13	24	2019-11-20	Aye Finance	FinTech	Financial Services To MSMEs	Gurgaon	
14	26	2019-11-20	Clumio	SaaS	Recovery software	San Jose	

- Click **Page Down** twice, and then **Page Up** twice to move around a bit faster, which is useful if you have lots of rows of data.
- Click and drag the **horizontal scroll bar** and then the **vertical scroll bar** to move even quicker up, down, and across a large datasheet.

- Try out these useful shortcuts in your worksheet:
  - Press **CTRL+End** to take you to the cell at the end of your data in the worksheet.
  - Press **CTRL+Home** to take you back to the start of the data in the worksheet (i.e. cell A2).
  - Press **CTRL+Down Arrow** to take you to the end of the column you're in
  - Press **CTRL+Up Arrow** to take you back to the top of the column.

## Task C: Select data

Perform the following steps to learn how to select different parts of your data (you can use the mouse to select cells if you prefer):

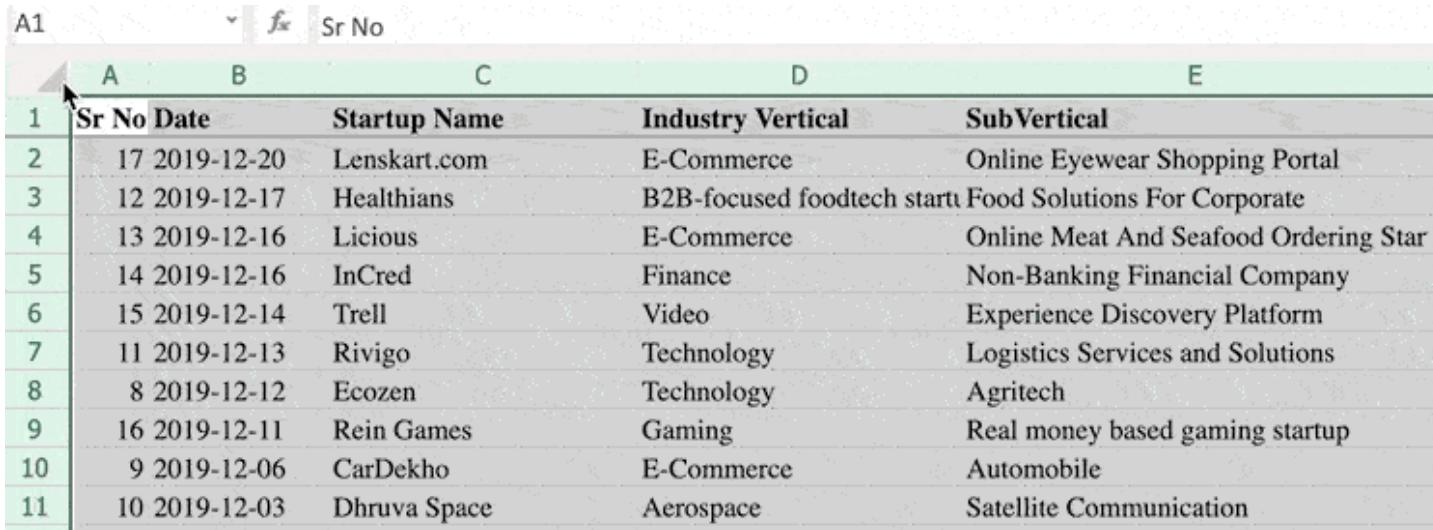
- [To select cells in a single row:](#) Select cell **A1**, then select cells **A1 to D1** by using **SHIFT+right arrow**.
  - [To select cells in a single column:](#) Select cell **A1**, then select cells **A1 to A10** by using **SHIFT+down arrow**.

The screenshot shows a Microsoft Excel spreadsheet with data from rows 1 to 11. Column A contains the index (1 to 11), Column B contains the date, Column C contains the startup name, Column D contains the industry vertical, and Column E contains the subvertical. The data includes various startups like Lenskart.com, Healthians, Licious, InCred, Trell, Rivigo, Ecozen, Rein Games, CarDekho, and Dhruva Space, categorized under E-commerce, B2B-focused foodtech, Finance, Video, Technology, Gaming, and Aerospace.

	A	B	C	D	E
1	Sr No	Date	Startup Name	Industry Vertical	SubVertical
2	17	2019-12-20	Lenskart.com	E-Commerce	Online Eyewear Shopping Portal
3	12	2019-12-17	Healthians	B2B-focused foodtech startu	Food Solutions For Corporate
4	13	2019-12-16	Licious	E-Commerce	Online Meat And Seafood Ordering Start
5	14	2019-12-16	InCred	Finance	Non-Banking Financial Company
6	15	2019-12-14	Trell	Video	Experience Discovery Platform
7	11	2019-12-13	Rivigo	Technology	Logistics Services and Solutions
8	8	2019-12-12	Ecozen	Technology	Agritech
9	16	2019-12-11	Rein Games	Gaming	Real money based gaming startup
10	9	2019-12-06	CarDekho	E-Commerce	Automobile
11	10	2019-12-03	Dhruva Space	Aerospace	Satellite Communication

- [To select multiple contiguous cols/rows:](#) Select column **A**, and use **SHIFT+ right arrow** to reach column **E**.
- [To select multiple non-contiguous cols/rows:](#) Select column **A**, then hold **CTRL** and select column **E**.
- [To select the entire worksheet:](#) Click the **corner button** (small grey triangle in top left corner of the worksheet).
  - [To select all your data:](#) Select any cell in the data, then press **CTRL+A**.

**Note:** The first time you press CTRL+A, it selects the current region if the worksheet contains data, the second time it selects the current data region and its header row, and the third time it selects the entire worksheet.



A screenshot of a Microsoft Excel spreadsheet titled "Sr No". The table has columns labeled A, B, C, D, and E. The data includes 11 rows of startup information, each with a unique ID (1-11), date, startup name, industry vertical, and subvertical. Row 1 is the header.

Sr No	Date	Startup Name	Industry Vertical	SubVertical
1	17 2019-12-20	Lenskart.com	E-Commerce	Online Eyewear Shopping Portal
2	12 2019-12-17	Healthians	B2B-focused foodtech startu	Food Solutions For Corporate
3	13 2019-12-16	Licious	E-Commerce	Online Meat And Seafood Ordering Star
4	14 2019-12-16	InCred	Finance	Non-Banking Financial Company
5	15 2019-12-14	Trell	Video	Experience Discovery Platform
6	11 2019-12-13	Rivigo	Technology	Logistics Services and Solutions
7	8 2019-12-12	Ecozen	Technology	AgriTech
8	16 2019-12-11	Rein Games	Gaming	Real money based gaming startup
9	9 2019-12-06	CarDekho	E-Commerce	Automobile
10	10 2019-12-03	Dhruva Space	Aerospace	Satellite Communication
11				

## Viewpoints: Using Spreadsheets as a Data Analysis Tool

What are the benefits and advantages of using spreadsheets as a tool for data analysis?

- ✓ Spreadsheets can be really useful in the right context
- ✓ You can see the data clearly laid out in a table
- ✓ It's easy to determine what the data is and the format
  
- ✓ I use Excel on a daily basis and have done throughout my career
- ✓ Using the functionality, pivot, tables, charts and formulas
- ✓ Excel is a one-stop shop to perform calculations and analyze financial ratios
- ✓ Export reports out of an ERP
  
- ✓ Spreadsheets are great for simple analysis
  
- ✓ Over the years spreadsheets have improved as systems and technology improve
- ✓ Excel is a good way to go if you have up to 20,000 lines of data
- ✓ Excel can take what seems like unmanageable data, but I can take it, massage it, filter it, sort it, and put it in a pivot table to get what I need
- ✓ It's all about making it more manageable by breaking it down into bite-sized chunks
  
- ✓ Spreadsheets are the easiest way to analyze data and to present data
- ✓ We don't need any fancy tools or additional software language to communicate

What are the drawbacks and limitations of using spreadsheets as a tool for data analysis?

- ✓ It's really hard to reproduce state
- ✓ There's no way to tell what steps have been taken on a dataset
- ✓ Because of the plethora of functions it's nearly impossible to know them all
- ✓ You can find yourself in 'analysis paralysis'
- ✓ You can spend a lot time and effort trying to figure one thing out

- ✓ If you have complex formulas, using VLOOKUPs and IF statements for example. At times they just stop working and you have to rebuild them
- ✓ Better to use Excel just for simple analysis and for a download of information
- ✓ If we start to get over 20,000 lines of data, it can get a little tricky
- ✓ Sometimes the spreadsheets will crash
- ✓ We might have to use Access or some of the other tools we use
- ✓ It's very difficult to handle extremely large datasets
- ✓ Spreadsheets have less flexibility for complicated analysis and presentation

## Summary and Highlights

### In this lesson, you have learned:

There are several spreadsheet applications available in the marketplace; the most commonly used and fully-featured spreadsheet application is Microsoft Excel.

Spreadsheets provide several advantages over manual calculation methods and they help you keep data organized and easily accessible.

As a Data Analyst, you can use spreadsheets as a tool for your data analysis tasks.

There are several elements that make up a workbook in a spreadsheet application.

The ribbon provides access to all the features and tools required to view, enter, edit, manipulate, clean, and analyze data in Excel.

There are several ways to navigate around a worksheet and workbook in Excel.

## Module 1: Practice Quiz



### Question 1

1/1 point (ungraded)

Of the following, which one would be considered a core capability of spreadsheets?

- Storage
- Statistical analysis
- Selling stocks
- Converting documents



### Question 2

1/1 point (ungraded)

What is the Excel keyboard shortcut to move to the last cell on a worksheet?

- Ctrl+Page Down
- Ctrl+End
- Ctrl+Down arrow
- Ctrl+Shift+End



### Question 3

1/1 point (ungraded)

According to the video, on the Excel Home tab what groups are listed?

- Font, Paragraph, Styles, Editing
- Font, Alignment, Number, Styles
- Tools, Pens, Convert, Replay
- Table, Illustrations, Add-ins, Charts



## Module 1: Graded Quiz

Bookmark

Graded Quiz due Jul 19, 2022 06:03 +08

### Question 1

1/1 point (graded)

Which of the following key Data Analyst tasks is typically done last?

Analyzing data

Visualizing data

Collecting data

Cleaning data



### Question 2

1/1 point (graded)

What is the Excel keyboard shortcut to move to the beginning of a worksheet?

Ctrl+Page Up

Ctrl+Home

Alt+Page Up

Up arrow key



### Question 3

1/1 point (graded)

Which of the following Excel keyboard shortcuts could be used to find how many rows of data you have in a worksheet, assuming you have no empty rows in your data? Select all that apply.

CTRL+Home

CTRL+Up

CTRL+Down

CTRL+End



## **Module Introduction**

In this module you will learn how to perform basic spreadsheet tasks, such as viewing, entering and editing data, and moving, copying and filling data. In addition, you will learn about the fundamentals of formulas, and learn about the most common functions used by a data analyst. Finally, you will learn how to reference data in formulas.

## **Learning Objectives**

**After completing this module, you will be able to:**

- View, enter, and edit data in a worksheet.
- Move, copy, and fill data in a worksheet. Navigate around an Excel worksheet and workbook.
- Describe the fundamentals of formulas.
- List some of the common functions used by a data analyst.
- Reference data in formulas.

# Viewing, Entering, and Editing Data

Now that you have learned basic spreadsheet terminology and learned how to navigate your way around worksheets and select data in Excel, it's now time to start entering some data.

First, we will look at some of the handy viewing features provided in Excel, and then we'll enter some data, and then edit that data.

When you have a lot of data in your worksheet it can be useful to zoom in closer to a specific area of the data. The Zoom Slider at the bottom right corner of the worksheet allows you to do just that. You can either click on the plus and minus buttons or drag the slider to select your preferred zoom value.

You also have some zoom controls in the ribbon on the View tab. Zoom lets you pick a predefined zoom level or a custom one, the 100% button zooms the worksheet back to its original size, and Zoom to Selection enables you to select an area of data and then zoom into that specific selection only.

If you want to see several areas of your data at the same time while zoomed in, you can use the Split button. This splits the screen into multiple sections; and you can scroll each section separately. If you only want two sections, you can remove either the horizontal or the vertical split by double-clicking on it.

If you have headings in your columns like a header row, then you might want those to remain on screen while you move down the sheet. To do that you need to use Freeze Panes.

You can freeze only the top row if you wish, or if that doesn't suit, as is the case here, then you can select the row (or even just a cell in the row) below the row or rows you want to freeze, and then select Freeze Panes. You can do a similar thing for columns you want to freeze too. And you can even freeze both rows and columns at the same time.

The trick here is to first select the cell that is both one row below where you want to freeze, and one column to the right of where you want to freeze. In this case, that is cell C4. Now we can scroll down the worksheet and across the worksheet and we can still see the header row and the Manufacturer and Model columns. Now, if you have multiple workbooks open (notice I said workbooks and not worksheets) then you can switch between them by using View, Switch Windows, or the faster method is to use the CTRL+F6 shortcut.

Now let's enter some data into a blank worksheet. The easiest way to open a new worksheet from within Excel is to click the New button in the Quick Access Toolbar (or CTRL+N if you prefer keyboard shortcuts). So let's enter some headings across the top of the worksheet; this is typically referred to as a 'header row'. Note, that if you press Enter after typing data into a cell the next active cell is the one directly below, which is not what we want in this case. But, if we press Tab after we enter data in a cell, it selects the next cell along in the row as the active cell.

Now we'll enter some headings and press Tab after each entry. Notice that the text is slightly longer in some of the cells and it either gets partly hidden by the next cell or overlaps it. If you click and hold the divider line between two columns, you can drag it left and right to resize it manually. If you want to do that automatically,

you can double-click the divider line between two columns. As these are going to be headings for our columns, let's make them bold.

Now let's add another column between the parts and accessories columns. Simply select the right-hand of those two columns, then right-click and choose Insert to put another column to the left of the selected column. Let's call it Servicing Sales.

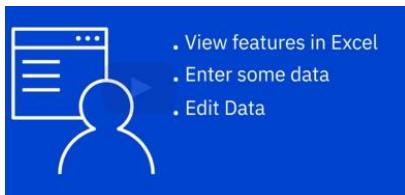
To tidy up all our column widths simultaneously, we select all the columns from A to E, then double-click any of the divider lines between columns; this automatically reduces or increases each column's width to fit the data in that column. OK, now we have some headings, let's enter some month data in column A. So, if we type Jan in cell A2 and press Enter then it takes us to the cell below, which is what we want in this case and we can type Feb in cell A3 and so on until we get to Dec in A13.

Now, let's suppose you need to change a couple of your headings. You have several ways of editing existing data in a cell; You can either select the cell and then just start over typing. Or you can select the cell and press F2 on your keyboard to put the cursor at the end of the cell and make your changes.

Or you can simply double-click somewhere on the cell to put the cursor at that position in the cell and make your changes. And you can even select the cell and then click in the formula bar to edit your cell data. Now let's do the same for the parts and accessories column headings.

In this video, we learned about some of the viewing options in Excel, and we learned how to enter and edit data in cells.

## Viewing, Entering, and Editing Data



# Copying, Filling, and Formatting Cells and Data

Now that we have learned about some of the handy viewing features provided in Excel, and entered and edited some data, let's discuss how to move, copy, and fill data, and how to format cells and data to suit our needs.

The first thing we are going to discuss is how to move data, so if you select a range of cells, in this case the headings in A1 to E1, and then hover over the top or bottom edge of a selected cell, and you will see the Move pointer, then you can drag the selection to another place on the worksheet. Alternatively, if you want to copy the data instead, you do the same thing but this time you also hold CTRL key as you select and drag the selection to another location and you will see the Copy pointer.

If you are not comfortable with dragging, you can also use Copy and Paste menu commands or keyboard shortcuts. So if you select some data in column A and copy it to the clipboard. Then you simply select the new location and paste the copied data. You can also move or copy between worksheets, so let's create a new worksheet. Then select some data from Sheet1, and this time let's use the CTRL+C keyboard shortcut to copy it to the clipboard. Then choose the other worksheet and use the CTRL+V shortcut to paste the data.

However, notice that the column widths are not the same as the original source data, so let's undo that and try another paste option. By default, when you paste the copied data, it uses the column width settings of the destination cells. So, to paste it and retain the column widths of the source data, you chose the special option under the Paste command, called Keep Source Column Widths.

As an alternative to having to enter data manually in a worksheet, you can use an Excel feature that automatically fill cells with data when it follows a sequential series or pattern. The feature is called AutoFill, and it can be especially useful when you need to enter lots of repetitive data into Excel, such as date information.

For example, if you enter a month in a cell, even using a shortened version of the name, you can use what's called the Fill Handle to select down to the end of the series, and AutoFill will work out what the series is, based on the selected data.

Let's try the same thing with days of the week. If you enter Mon in a cell, then drag the fill handle to use AutoFill, it will determine that you want to enter the days of the week sequentially. However, if you also enter Wed (for Wednesday) in the next cell down, and select both cells in the series, i.e. A16 and A17, and then drag the fill handle down, AutoFill determines that the sequence has changed to every other day, and fills in the data series for you.

It's important to select all cells that define the pattern when using AutoFill so that it can best determine what the pattern is, in this case cells A16 and A17. A similar thing applies to numerical patterns; if you enter 5 in a cell, and then use the fill handle to fill the data down the column. Because the data is not the name of a day or month for example, AutoFill can't determine what the pattern is yet. So, In this case, it just copies the value 5 into every selected cell.

However, if you enter the value 10 in B3, and then use the fill handle to fill the data down the column, AutoFill determines that the pattern is incrementing by 5 each time and it fills in the remainder of the data pattern for

you. We are now going to look at formatting our data, and there are essentially two distinct parts to this. First, there's formatting of the cells themselves (with a fill color and a bold border for example and bold text within it). And then there's formatting the data in the cells (for example, making it text format, number format, or a specific currency or accounting format).

Let's open the car sales worksheet we used previously. Then select the headings in cells A3 to P3 either using the mouse, or you could use the shortcut keys CTRL+SHIFT+Right Arrow. On the Home tab, click the Styles drop-down arrow, and select a style color for your cells. Then you can make the selected cells bold. Then you select the data in the Manufacturer column either using the mouse, or the shortcut keys CTRL+SHIFT+Down Arrow.

In the Styles drop-down arrow, select another style color for the selected cells. Again you can make the cells bold. Then you select the data in the Model column again either using the mouse, or the shortcut keys CTRL+SHIFT+Down Arrow.

In the Styles drop-down arrow, select another style color for the selected cells. This time you could make the selected cells italic. And you can also change the font size and style. Lastly, you can select all the other cells in the data by using the mouse or the CTRL+SHIFT+Right Arrow then Down Arrow, and apply borders to the data cells.

Now it's time to format the cell data. The sales figures in columns C and D can be formatted to display only two decimal places; just select the data and click the Decrease Decimal button. We also have an issue with a couple of the car models.

If you look in cells B129 and B130, where the model name is supposed to be displayed, you can see there are actually two dates listed instead. And if you look in the Number Format box, the format type is Custom. This has happened because the model numbers are supposed to be the Saab 9-5 and the Saab 9-3 but when the files were imported from CSV files these two cells must have been incorrectly determined to be date values and not just numbers.

You can fix this by formatting these two cells as Text, and then enter the correct values of 9-5 and 9-3. The last thing we shall do is format some data as currency. If you look at the heading in column F it says it is Price in thousands of dollars, and cell F4 is using the General format. So, let's change the format of this column to American currency format.

We select the column, F in this case, then select More Number Formats from the drop-down list, then we choose the Currency option, and the correct currency symbol and format. And we're done.

In this video, we learned how to move, copy, and fill data, and how to format cells and cell data to suit our needs.

In the next video, we will look at the basics of formulas, learn how to perform simple calculations, and learn how to select ranges and copy formulas.

# Copying, Filling, and Formatting Cells and Data

## Autofill

Month	Car Sales	Parts	Servicing	Accessories
Jan				

Month	Car Sales	Parts	Servicing	Accessories
1	Jan			
2				
3				

Month	Car Sales	Parts	Servicing	Accessories
1	Jan			
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				

Month	Car Sales	Parts
1	Jan	
2	Feb	
3	Mar	
4	Apr	
5	May	
6	Jun	
7	Jul	
8	Aug	
9	Sep	
10	Sep	
11	Oct	
12	Nov	
13	Dec	
14		

Month	Car Sales	Parts
1	Jan	
2	Feb	
3	Mar	
4	Apr	
5	May	
6	Jun	
7	Jul	
8	Aug	
9	Sep	
10	Sep	
11	Oct	
12	Nov	
13	Dec	
14		
15		
16	Mon	
17	Tue	
18	Wed	
19	Thu	
20	Fri	
21	Sat	
22	Sun	

Month	Car Sales	Parts
1	Jan	
2	Feb	
3	Mar	
4	Apr	
5	May	
6	Jun	
7	Jul	
8	Aug	
9	Sep	
10	Sep	
11	Oct	
12	Nov	
13	Dec	
14		
15		
16	Mon	
17	Tue	
18	Wed	
19	Thu	
20	Fri	
21	Sat	
22	Sun	

Month	Car Sales	Parts
1	Jan	
2	Feb	
3	Mar	
4	Apr	
5	May	
6	Jun	
7	Jul	
8	Aug	
9	Sep	
10	Sep	
11	Oct	
12	Nov	
13	Dec	
14		
15		
16	Mon	
17	Tue	
18	Wed	
19	Thu	
20	Fri	
21	Sat	
22	Sun	

Mon
Wed
Fri
Sun
Tue
Thu
Sat
Mon
Wed
Fri
Sun
Tue
Thu
Sat

Month	Car Sales	P
Jan	5	
Feb	10	
Mar		
Apr		
May		
Jun		
Jul		
Aug		

Month	Car Sales	Parts
Jan	5	
Feb	10	
Mar	15	
Apr	20	
May	25	
Jun	30	
Jul	35	
Aug	40	
Sep	45	
Oct	50	
Nov	55	
Dec	60	

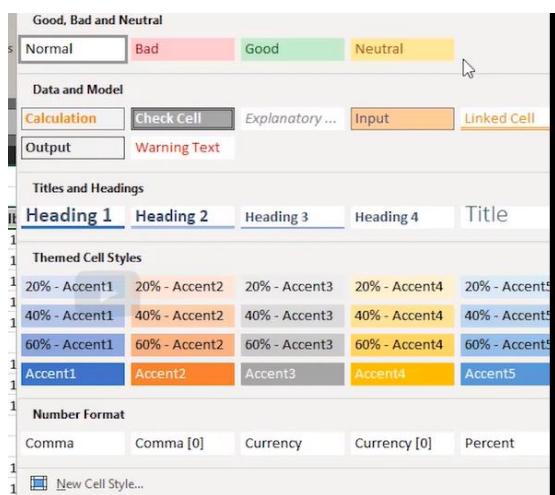
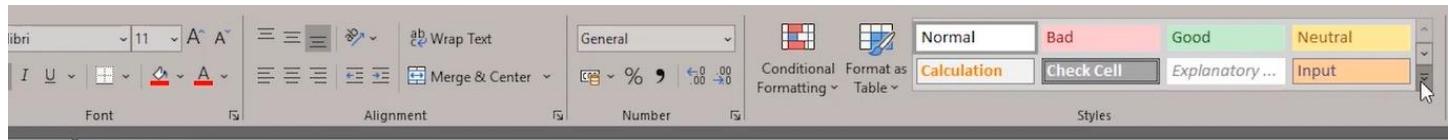
## Formatting Data

Formatting our data, and there are essentially two distinct parts

Formatting cells themselves



Formatting the data in the cells



Sales (\$ in thousands)		Year Resale Value	Vehicle_type
16.92		16.36	Passenger

125	Pontiac	Montana
126	Porsche	Boxter
127	Porsche	Carrera Coupe
128	Porsche	Carrera Cabrio
129	Saab	09-May
130	Saab	09-Mar
131	Saturn	SL
132	Saturn	SC

3	Manufacturer	Model	Sales (\$ in thousands)	Y
4	Acura	Integra	16.92	
5	Acura	TL	39.38	
6	Acura	RL	8.59	
7	Audi	A4	20.40	
8	Audi	A6	18.78	
9	Audi	A8	1.38	
0	BMW	323i	19.75	
1	BMW	328i	9.23	
2	BMW	528i	17.53	
3	Buick	Century	91.56	
4	Buick	Regal	39.35	
5	Buick	Park Avenue	27.85	
6	Buick	LeSabre	83.26	
7	Cadillac	DeVille	63.73	
8	Cadillac	Seville	15.94	
9	Cadillac	Eldorado		

CTRL+SHIFT+ → + ↓

23 Chevrolet Malibu

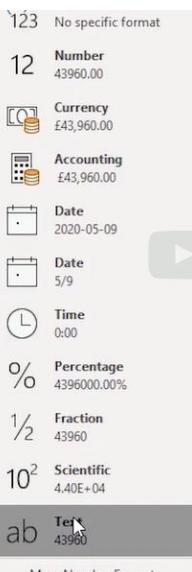
135.13

Manufacturer	Model	Sales (\$ in thousands)	Year	Resale Value	Vehicle_type
Acura	Integra	16.92		16.36	Passenger
Acura	TL	39.38		19.875	Passenger
Acura	RL	8.59		29.725	Passenger

128	Porsche	Carrera Cabrio
129	Saab	09-May
130	Saab	09-Mar
131	Saturn	SL



	B	C	D	E
Silhouette		24.4	15.2	Car
Neon		32.7	7.8	Passenger
Breeze		5.2	9.8	Passenger
Voyager		24.2	12.0	Car
Prowler		1.9	30.4	Passenger
Sunfire		51.6	13.8	Passenger
Grand Am		131.1	10.3	Passenger
Firebird		19.9	17.8	Passenger
Grand Prix		92.4	14.0	Passenger
Bonneville		35.9	13.2	Passenger
Montana		39.6	14.2	Car
Boxter		9.0	41.3	Passenger
Carrera Coupe		1.3	60.6	Passenger
Carrera Cabrio		1.9	67.6	Passenger
09-May		9.2	22.8	Passenger
09-Mar		12.1	18.3	Passenger
SL		80.6	9.2	Passenger
SC		24.5	10.6	Passenger
SW		5.2	10.8	Passenger
LW		8.5	13.7	Passenger
LS		50.0	9.9	Passenger



Vehicle_type	Price (\$ in thousands)
Passenger	21.5
Passenger	28.4
Passenger	42

Power Pivot

ext & Center Number

Vehicle Type	Price (\$ in thousands)	Engine Size
Passenger	21.5	1.8
Passenger	28.4	3.2
Passenger	42	3.5
Passenger	23.99	1.8
Passenger	33.95	2.8
Passenger	62	4.2
Passenger	26.99	2.5
Passenger	33.4	2.8
Passenger	38.9	2.8

Help Power Pivot

Wrap Text Merge & Center Conditional Formatting

Number

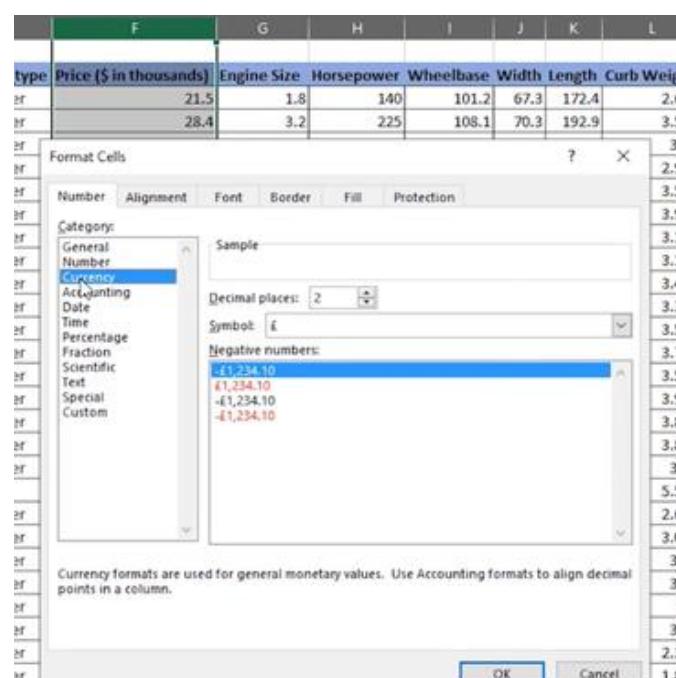
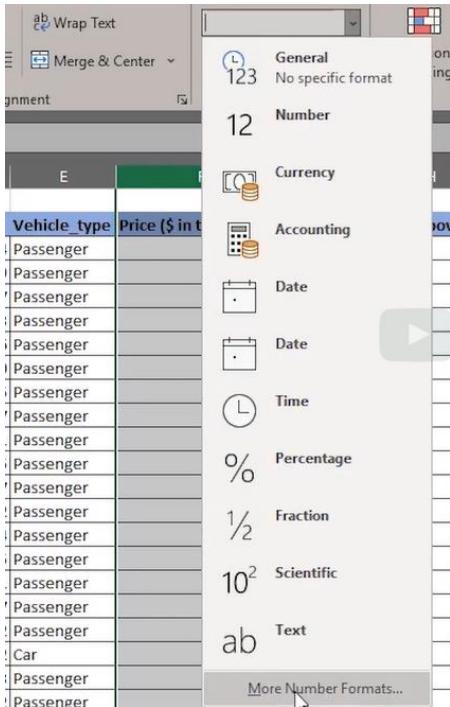
Number Format

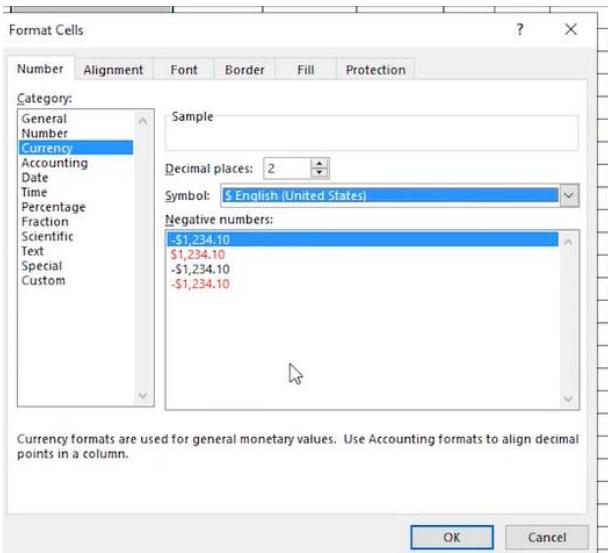
Choose the format for your cells, such as percentage, currency, date or time.

Vehicle Type Price (\$ in thousands) Engine Size

Passenger 21.5 1.8

Passenger 28.4 3.2





E	F	G
Vehicle_type	Price (\$ in thousands)	Engine Size
Passenger	\$21.50	1
Passenger	\$28.40	1
Passenger	\$42.00	1
Passenger	\$23.99	1
Passenger	\$33.95	1
Passenger	\$62.00	2
Passenger	\$26.99	2
Passenger	\$33.40	2
Passenger	\$38.90	2
Passenger	\$21.98	2
Passenger	\$25.30	2
Passenger	\$31.97	2
Passenger	\$27.89	2
Passenger	\$39.90	2
Passenger	\$44.48	2
Passenger	\$39.67	2
Passenger	\$31.01	2
Car	\$46.23	5

# Hands-on Lab 3: Entering and Formatting Data

**Estimated time needed:** 30 minutes

In this lab, first you will learn some of the viewing options in Excel, and then learn how to enter and edit data in cells. Then, you will learn how to move, copy, paste, and fill data, and how to format cells and cell data in a worksheet.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The first dataset used in this lab comes from the following source: <https://www.kaggle.com/sudalairajkumar/indian-startup-funding> under a [CC0: Public Domain license](#). Acknowledgement and thanks also goes to <https://trak.in> who were generous enough to share the data publicly for free.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

The second dataset used in this lab is an internal dataset.

## Objectives

After completing this lab, you will be able to:

- Use viewing options, and enter and edit data

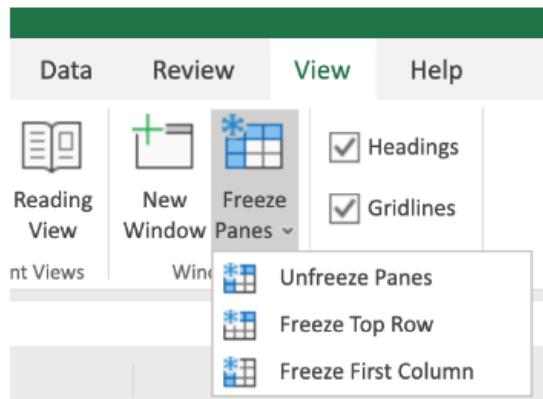
- Copy and fill data, and format cells and data

# Exercise 1: Viewing, Entering and Editing Data

In this exercise, you will learn some of the viewing options in Excel, how to enter and edit data in cells.

## Task A: Viewing Data

1. Download the file [indian startup funding Lab3.xlsx](#). Upload and open it using Excel for the web.
2. Select **F20:H26** (if required, use the vertical and horizontal scroll bars to bring the selected cell range area to the center of the screen). Hold **CTRL and +** to zoom in closer to the specific area of the data. Then hold **CTRL and -** to zoom the worksheet back out to its original size.  
**(Note: Zoom to Selection** which is found under the **View** tab of Excel Desktop, is not available for Excel for the web)
3. On the ribbon, click **View, Freeze Panes, Freeze Top Row**. Now you have headings in your columns like a header row, which will remain static on screen while you move down the worksheet. Next, click **Unfreeze Panes**, and click **Freeze First Column**. The **Sr No** column will remain static on the screen while you move right across the worksheet. Lastly, click **Unfreeze Panes** to end this step.



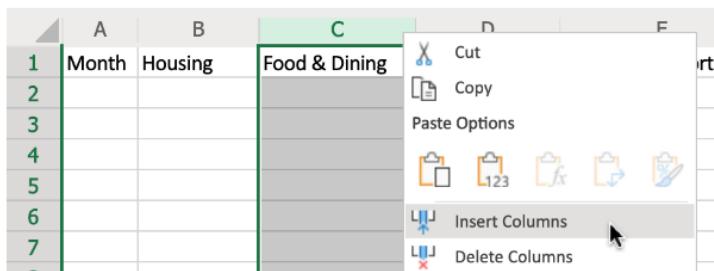
4. To freeze both the top row and the first column at the same time, select cell **B2** and click **View, Freeze Panes, Freeze Panes**.

Sr No	City	Location	Investors Name
17	Faridabad		SoftBank Vision Fund
12	Bengaluru		Paytm, NPTK, Sabre Partners and Neoplux
13	Bengaluru		Vertex Growth Fund
14	Mumbai		
15	Bengaluru		Ruizheng Investment
11	Gurgaon		SAIF Partners, Spring Canter Investment Ltd
8	Pune		Sathguru Catalyzer Advisors

5. You can open multiple workbooks in multiple browser tabs in Excel for the web, and to switch between them, you just click each browser tab. (In Excel Desktop you have to click the **View** tab, then click **Switch Windows**)

## Task B: Entering Data

1. Download the file [Personal Monthly Expenditure Lab3.xlsx](#). Upload and open it using Excel for the web. Go to the **Expense - 2018** worksheet.
2. In cell **A1**, type **Month** and press **Tab**. Then type **Housing** and press **Tab**, type **Food & Dining**, and press **Tab**, type **Personal**, and press **Tab**, type **Auto & Transport**, then press **Tab**, type **Health & Fitness**, then press **Tab**. You are now done with the header row.
3. To enter some data as rows in column **A**, in **A2**, type **Jan** and press **Enter**. Then type **Feb**, and press **Enter**, type **Mar**, and press **Enter**, type **Apr**, and press **Enter**.
4. To add another column between the **Housing** and **Food & Dining**, select column **C**, then right-click column **C**, and choose **Insert Columns**. In the top row header cell **C1**, type **Bills & Utilities**.

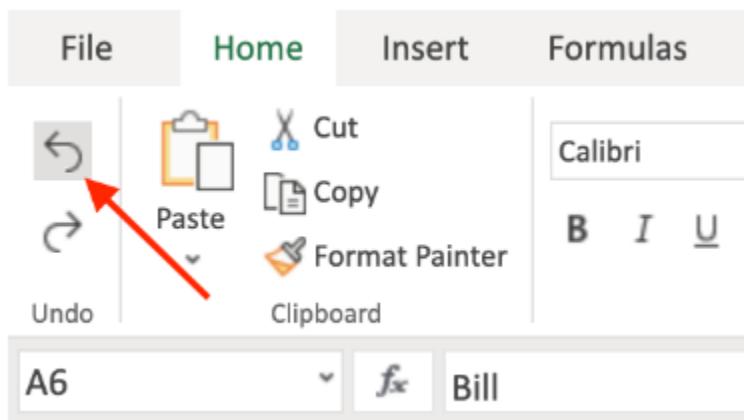


5. Select columns **A to G**, then double-click the divider between **A** and **B** to adjust the column widths.

Month	Housing	Bills & Utilities	Food & Dining	Personal	Auto & Transport	Health & Fitness
1						
2						
3						

### Task C: Editing Data

1. Select cell **C1** and press **Backspace** to clear the contents. Then type **Bills**.
2. Click **Undo** to undo the change.



## Exercise 2: Copying, Filling, and Formatting Cells and Data

In this exercise, you will learn how to move, copy, paste, and fill data, and how to format cells and cell data in a worksheet.

### Task A: Copying and Filling Data

1. Select **A2:A5**. Hover over the edge of the selected cells to get the **Move** pointer and then drag the selection to move the selected cells to **B6**. Click **Undo**.

	A	B	C
1	Month	Housing	Bills & Utilities
2	Jan		
3	Feb		
4	Mar		
5	Apr		
6			
7			
8			
9			
10			
11			

2. Select cell **A5**. Hover over the bottom right corner of cell **A5** to get the **+** (**Fill Handle**) symbol, then drag to **A13**.

	A	B
1	Month	Housing
2	Jan	
3	Feb	
4	Mar	
5	Apr	
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

3. On the **Expense - 2018** sheet, select **A1:G13** and press **CTRL+C**. Then on the **Expense - 2019** sheet, select cell **A1** and press **CTRL+V**.
4. Select cell **A1** and press **CTRL+A** to select the whole datasheet. On the **Home** tab, in the **Cells** group, click the drop-down arrow under **Format**, and click **Auto-Fit Column Width**.

	A	B	C	D	E	F	G
1	Month	Housing	Bills & Utilities	Food & Dining	Personal	Auto & Transport	Health & Fitness
2	Jan						
3	Feb						
4	Mar						
5	Apr						
6	May						
7	Jun						
8	Jul						
9	Aug						
10	Sep						
11	Oct						
12	Nov						
13	Dec						

## Task B: Formatting Cells and Data

1. **Formatting Cells:**
  - a. Select **A1:G13**. On the **Home** tab, in the **Tables** group, click **Format as Table**, and choose a table style from the list. In the pop-up dialog box, ensure that the option **My table has headers**, is checked and then click **OK**.

Month	Housing	Bills & Utilities	Food & Dining	Personal	Auto & Transport	Health & Fitness
Jan	800	210	400	100	100	60
Feb	800	180	350	100	125	70
Mar	800	170	420	100	120	60
Apr	800	160	400	120	100	60
May	800	150	420	100	100	80
Jun	800	150	380	100	130	60
Jul	800	150	420	120	100	60
Aug	800	150	420	100	100	80
Sep	800	150	400	120	110	60
Oct	800	170	420	100	100	60
Nov	800	200	390	120	100	50
Dec	800	220	400	100	115	60

1.

- o b. Select **A2:A13**. In the **Font** group click **Italic**. In the font size box, select **10**. In the font style drop-down box, select **Arial**.

## 2. Formatting Cell Data:

- o a. Select column **B**, and use **SHIFT+right arrow** to select across to include column **G**. On the **Home** tab, in the **Number** group, click the **Number Format** drop-down list and choose **Currency**.
- o b. Select columns **B** to **G** again. On the **Home** tab, in the **Number** group, click **Decrease Decimal** once.
- o c. Select columns **B** to **G** again. On the **Home** tab, in the **Number** group, click the **Accounting Number Format (\$)** drop-down list, and select **£ English (United Kingdom)**.

A	B	C	D	E	F	G
Month	Housing	Bills & Utilities	Food & Dining	Personal	Auto & Transport	Health & Fitness
Jan	£ 800.00	£ 210.00	£ 400.00	£ 100.00	£ 100.00	£ 60.00
Feb	£ 800.00	£ 180.00	£ 350.00	£ 100.00	£ 125.00	£ 70.00
Mar	£ 800.00	£ 170.00	£ 420.00	£ 100.00	£ 120.00	£ 60.00
Apr	£ 800.00	£ 160.00	£ 400.00	£ 120.00	£ 100.00	£ 60.00
May	£ 800.00	£ 150.00	£ 420.00	£ 100.00	£ 100.00	£ 80.00
Jun	£ 800.00	£ 150.00	£ 380.00	£ 100.00	£ 130.00	£ 60.00
Jul	£ 800.00	£ 150.00	£ 420.00	£ 120.00	£ 100.00	£ 60.00
Aug	£ 800.00	£ 150.00	£ 420.00	£ 100.00	£ 100.00	£ 80.00
Sep	£ 800.00	£ 150.00	£ 400.00	£ 120.00	£ 110.00	£ 60.00
Oct	£ 800.00	£ 170.00	£ 420.00	£ 100.00	£ 100.00	£ 60.00
Nov	£ 800.00	£ 200.00	£ 390.00	£ 120.00	£ 100.00	£ 50.00
Dec	£ 800.00	£ 220.00	£ 400.00	£ 100.00	£ 115.00	£ 60.00

# The Basics of Formulas

Now that we have learned how to move, copy, and fill data, and how to format cells and data, next we will take a look at the basics of formulas, including some basic calculations, selecting ranges in formulas, and how to copy formulas.

A typical **formula** is made of several key components. The equal sign starts the formula off and lets Excel know you are creating a formula in this cell.

The next part is the function, which performs the calculation. For example, the SUM function adds up the values in referenced cells or cell ranges. Then comes the reference, which is the cell or range of cells you want to include in your calculation, and these need to be enclosed in parentheses.

You also have operators, which specify what type of calculation to perform. Common arithmetic operators include: addition, subtraction, multiplication and division. And these are represented by symbols.

The plus symbol for addition, the minus symbol for subtraction, the asterisk for multiplication, and the forward slash for division. There are other types of operators too. Namely comparison, text concatenation, and reference. You may also use constants in your formulas, which as the name suggests are numbers or values which you can enter directly into a formula, and which don't change.

This might be a whole number such as 5, it might be a percentage such as 10%, or it might even be a date. So, a typical formula might be =SUM(B5\*20), which would take the value in cell B5 and multiply it by 20.

Let's start with a few basic calculations. Suppose you want to add up January and February sales of accessories. You would start by typing an equal sign, which lets Excel know you are entering a formula. Then you type in the function you wish to use, in this case the SUM function.

Note the description. Next you type an open parenthesis, then you select your cell range, which in this case would be E2 to E3, so you could enter that as 'E2,E3' then a close parenthesis and press Enter. And if you wanted to add March sales as well, then you would have to extend the cell range to include E4.

So you could type E2,E3,E4 as your range and it will work. Remember, to edit a cell, you select the cell, and either edit it directly in the formula bar, or press F2, or double-click the cell. However, it's very cumbersome and not very flexible to do it this way, because if you wanted to add up the entire column then you'd have to type every cell reference, one after the other.

So thankfully, there's a better way. Instead of typing each cell to include in the reference, you just put a colon between the first and last values in our range, so E2:E4, in this case. And if you wanted the whole column, then you would enter E2:E13 in your formula. But there's another way of doing it, and that's by using your mouse to select the range, so you still type =sum then open parenthesis, but select the range with your mouse (or SHIFT + arrow keys) and just press Enter.

Excel will add the close parenthesis for you. To total these columns up, and add some tax, you'd add some headings first for Subtotals, and Tax at 20%. Then your formula will need to multiply the value in Subtotals by 20%. If you want to add up all the column subtotals and calculate the taxes, then you could repeat

the previous process for each column, but that's very time consuming, and you don't need to, because Excel has some neat tricks to do this for you. Just select the fill handle in the bottom right corner of the cell, and drag across to the other cells to copy the formula; this is called AutoFill. Notice how the formula is copied, but the row references change in relation to the cells' position on the worksheet.

So what was E2:E13 has become B2:B13. These are known as relative references, but more on that later in the course. And you can do the same thing for the tax values in row 16. Now, you need a row for showing the totals. The calculation here is simply the subtotal value in cell B15, added to the tax in B16. And again, you can use the fill handle to copy the formula across.

If you want to total the sales of all products by month, you'd add a column heading; notice how the cell style is copied to the new heading automatically. Remember, to widen a column, either drag the divider manually, or double-click the divider. Then enter the formula in cell F2 as you've done before. However, Excel has another trick up its sleeve.

It's called **AutoSum** and is found on the Home tab, in the Editing group. This is a great little shortcut for some simple common functions like Sum, Average, Count, Max, and Min, but you can choose other functions too. You want 'Sum' for this particular calculation.

Notice that it also has a keyboard shortcut of 'Alt plus equals', and then press Enter, and it's done. Now you can use the fill handle to copy down the remaining values. But hold on, there is one more Excel trick to show and it's a good one!

Suppose your column of data was very long; you might have to drag the fill handle down over several pages, which isn't easy to do and can easily lead to errors when selecting large lists of data values. Rather than needing to drag down to the rest of the column, you can just double-click the fill handle, and it will automatically copy the formula to all the remaining cells in that column.

This one is a real time-saver. Finally, let's format all these values to use the US dollar currency format. In this video, we learned about the basics of formulas, how to perform simple calculations, how to select ranges in formulas, and how to copy formulas.

In the next video, we will look at how to use some of the common functions used by Data

Analysts and discover some more advanced functions.

# The Basics of Formulas

Basic calculations   Selecting ranges in formula   How to copy formulas



- = starts the formula
- SUM is the function
- (B5) is the cell reference
- \* is the operator for multiplication
- 20 is the constant



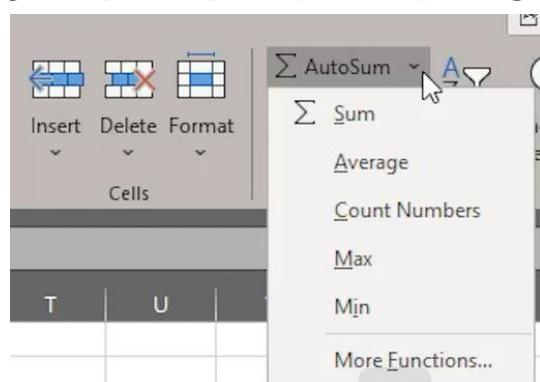
=SUM(B5\*20)

+      -  
Addition      Subtraction  
\*      /  
Multiplication      Division

## Autofill

Month	Car Sales	Parts	Servicing	Accessories
Jan	193670	11240	5980	1710
Feb	221445	13657	6124	1844
Mar	233378	13889	6578	2180
Apr	260135	14112	7210	2258
May	289940	12499	7324	2612
Jun	312692	12102	7688	2407
Jul	323446	13190	7200	2356
Aug	303575	12778	7734	2170
Sep	260336	14232	8114	2004
Oct	212291	15110	8396	1822
Nov	198440	14125	7855	2135
Dec	177783	12667	6320	2337
Subtotals	2987131	159601	86523	25835
Tax(20%)	597426.2	31920.2	17304.6	5167
Totals	3584557.2	191521.2	103827.6	31002

If you want to total the sales of all products by month, you'd add a column heading; notice how the cell style is copied to the new heading automatically. Remember, to widen a column, either drag the divider manually, or double-click the divider. Then enter the formula in cell F2 as you've done before.



Month	Car Sales	Parts	Servicing	Accessories	Monthly Total
Jan	193670	11240	5980	1710	=SUM(B2:E2)
Feb	221445	13657	6124	1844	SUM(number1, [number2], ...)
Mar	233378	13889	6578	2180	
Apr	260135	14112	7210	2258	
May	289940	12499	7324	2612	
Jun	312692	12102	7688	2407	
Jul	323446	13190	7200	2356	
Aug	303575	12778	7734	2170	
Sep	260336	14232	8114	2004	
Oct	212291	15110	8396	1822	
Nov	198440	14125	7855	2135	
Dec	177783	12667	6320	2337	
<b>SUM Function</b>					
Alt + =					

It's called AutoSum and is found on the Home tab, in the Editing group. This is a great little shortcut for some simple common functions like Sum, Average, Count, Max, and Min, but you can choose other functions too. You want 'Sum' for this particular calculation. Notice that it also has a keyboard shortcut of 'Alt plus equals', and then press Enter, and it's done.

Month	Car Sales	Parts	Servicing	Accessories	Monthly Total
Jan	193670	11240	5980	1710	212,600.00
Feb	221445	13657	6124	1844	243,070.00
Mar	233378	13889	6578	2180	256,025.00
Apr	260135	14112	7210	2258	283,715.00
May	289940	12499	7324	2612	312,375.00
Jun	312692	12102	7688	2407	334,889.00
Jul	323446	13190	7200	2356	346,192.00
Aug	303575	12778	7734	2170	326,257.00
Sep	260336	14232	8114	2004	284,686.00
Oct	212291	15110	8396	1822	237,619.00
Nov	198440	14125	7855	2135	222,555.00
Dec	177783	12667	6320	2337	199,107.00

Rather than needing to drag down to the rest of the column, you can just double-click the fill handle, and it will automatically copy the formula to all the remaining cells in that column.

Month	Car Sales	Parts	Servicing	Accessories	Monthly Total
Jan	\$ 193,670.00	\$ 11,240.00	\$ 5,980.00	\$ 1,710.00	\$ 212,600.00
Feb	\$ 221,445.00	\$ 13,657.00	\$ 6,124.00	\$ 1,844.00	\$ 243,070.00
Mar	\$ 233,378.00	\$ 13,889.00	\$ 6,578.00	\$ 2,180.00	\$ 256,025.00
Apr	\$ 260,135.00	\$ 14,112.00	\$ 7,210.00	\$ 2,258.00	\$ 283,715.00
May	\$ 289,940.00	\$ 12,499.00	\$ 7,324.00	\$ 2,612.00	\$ 312,375.00
Jun	\$ 312,692.00	\$ 12,102.00	\$ 7,688.00	\$ 2,407.00	\$ 334,889.00
Jul	\$ 323,446.00	\$ 13,190.00	\$ 7,200.00	\$ 2,356.00	\$ 346,192.00
Aug	\$ 303,575.00	\$ 12,778.00	\$ 7,734.00	\$ 2,170.00	\$ 326,257.00
Sep	\$ 260,336.00	\$ 14,232.00	\$ 8,114.00	\$ 2,004.00	\$ 284,686.00
Oct	\$ 212,291.00	\$ 15,110.00	\$ 8,396.00	\$ 1,822.00	\$ 237,619.00
Nov	\$ 198,440.00	\$ 14,125.00	\$ 7,855.00	\$ 2,135.00	\$ 222,555.00
Dec	\$ 177,783.00	\$ 12,667.00	\$ 6,320.00	\$ 2,337.00	\$ 199,107.00
<b>Subtotals</b>					
Tax(20%)					
<b>Totals</b>					

Finally, let's format all these values to use the US dollar currency format.

# Intro to Functions

Now that you have learned about the basics of formulas, learned how to perform some basic calculations, and how to select ranges and copy formulas, next we will have an introduction to functions, including using some common statistical functions. And then we will learn about some more advanced functions that a Data Analyst might also use.

First, let's look at some common functions used for statistical calculations. So, we'll add some row headings for average, minimum, maximum, count, and median. Then in cell B20, let's work out the average of the car sales for the year, from the table above.

On the Home tab, in the Editing group, we click the AutoSum drop-down list and choose Average. Now, because AutoSum tries to add up the values directly above it in the column, we need to modify the cell range here to B2 to B13. Then we can use the Fill Handle as we've seen before to copy the formula across to column E.

For the minimum calculation in B21, we select Min from the AutoSum list. And again, we need to modify the cell range. So this calculates the lowest value in our range. And fill across to column E. And for the maximum calculation, we select Max from the list. And then modify the range. And once again, copy the formula across. This calculates the highest value in our range.

In B23 we will calculate the Count, which basically just means the number of values that exist in the selected range. So, we select Count Numbers from the list. Then modify the range.

For the Median calculation, we can select 'More Functions' from the AutoSum list, then select 'Statistical' as the category, and scroll down to find the MEDIAN function. The 'median' returns the exact middle of a range of selected values.

Note that if you're selecting an odd number of values it will return the figure that is the middle value in your selected range, but if you have selected an even number of values in your range, it will return the middle figure between the two middle values in your range. Once again, we need to change the cell range to B2 to B13. And we can then copy this formula across to column E.

You've seen AutoSum and some of the common statistical functions in Excel, but there are another 400-plus other functions available, so let's explore just a few of those now.

On the Formulas tab, in the Function Library group, there are drop-down lists for several function categories. The first is a list of 'Recently Used' functions, which updates automatically as you use them. Then you have functions related to 'Financial' calculations.

If you hover over the name of a function, you see a short description for each one; so here we have the accrued interest function, and here is the interest rate function. The 'Logical' list has BOOLEAN operator functions such as AND, IF, and OR. There are several functions related to Text, such as CONCAT, which is an

updated version of a previous function called CONCATENATE (which is still supported by the way for backwards compatibility), FIND, and SEARCH.

There are also several functions related to dates and times, such as NETWORKDAYS, WEEKDAY, and WEEKNUM.

In the ‘Lookup & Reference’ list there are functions such as AREAS, HLOOKUP, SORTBY, and VLOOKUP. In the ‘Math & Trig’ list you’ll find lots of useful mathematical functions, such as POWER, SUMIF, and SUMPRODUCT, alongside many functions for trigonometric purposes, such as cosine, sine and tangent.

There is also a ‘More Functions’ list which provides several more function categories, such as Statistical, Engineering, and Information.

In the ‘Statistical’ list you’ll find functions such as Average, Count, Max, Median, and Min; we saw some of these used earlier in this video. If you’re struggling to find the function you want in these lists, you can also search for a function; just click the ‘Insert Function’ button on the Formulas tab, and then either browse the category lists available, or choose ‘All’ and look down the alphabetical list for the function you want. Alternatively, type the name of a function you want to find, and click ‘Go’ to search for it, then select the one you want from the returned search.

In this video, we learned about the basics of functions, how to use some of the more common functions that a Data Analyst might employ, and looked at some of the more advanced functions available in Excel.

In the next video, we will look at referencing data in formulas; specifically differentiating between relative and absolute references, and error handling in formulas.

## Intro to Functions

Month	Car Sales	Parts	Servicing	Accessories	Monthly Total
Jan	\$ 193,670.00	\$ 11,240.00	\$ 5,980.00	\$ 1,710.00	\$ 212,600.00
Feb	\$ 221,445.00	\$ 13,657.00	\$ 6,124.00	\$ 1,844.00	\$ 243,070.00
Mar	\$ 233,378.00	\$ 13,889.00	\$ 6,578.00	\$ 2,180.00	\$ 256,025.00
Apr	\$ 260,135.00	\$ 14,112.00	\$ 7,210.00	\$ 2,258.00	\$ 283,715.00
May	\$ 289,940.00	\$ 12,499.00	\$ 7,324.00	\$ 2,612.00	\$ 312,375.00
Jun	\$ 312,692.00	\$ 12,102.00	\$ 7,688.00	\$ 2,407.00	\$ 334,889.00
Jul	\$ 323,446.00	\$ 13,190.00	\$ 7,200.00	\$ 2,356.00	\$ 346,192.00
Aug	\$ 303,575.00	\$ 12,778.00	\$ 7,734.00	\$ 2,170.00	\$ 326,257.00
Sep	\$ 260,336.00	\$ 14,232.00	\$ 8,114.00	\$ 2,004.00	\$ 284,686.00
Oct	\$ 212,291.00	\$ 15,110.00	\$ 8,396.00	\$ 1,822.00	\$ 237,619.00
Nov	\$ 198,440.00	\$ 14,125.00	\$ 7,855.00	\$ 2,135.00	\$ 222,555.00
Dec	\$ 177,783.00	\$ 12,667.00	\$ 6,320.00	\$ 2,337.00	\$ 199,107.00
Subtotals	\$ 2,987,131.00	\$ 159,601.00	\$ 86,523.00	\$ 25,835.00	
Tax (20%)	\$ 597,426.20	\$ 31,920.20	\$ 17,304.60	\$ 5,167.00	
Totals	\$ 3,584,557.20	\$ 191,521.20	\$ 103,827.60	\$ 31,002.00	
Avg	\$ 248,927.58	\$ 13,300.08	\$ 7,210.25	\$ 2,152.92	
Min	\$ 177,783.00	\$ 11,240.00	\$ 5,980.00	\$ 1,710.00	
Max	\$ 323,446.00	\$ 15,110.00	\$ 8,396.00	\$ 2,612.00	
Count	12				
Median	=				

Insert Function

Search for a function:

Type a brief description of what you want to do and then click Go

Or select a category: Statistical

Select a function:

- LOGEST
- LOGNORM.DIST
- LOGNORM.INV
- MAX
- MAXA
- MAXIFS
- MEDIAN**

**MEDIAN**

**MEDIAN(number1,number2,...)**

Returns the median, or the number in the middle of the set of given numbers.

Help on this function

OK Cancel

## Formulas

AutoSave (Off) File Home Insert Page Layout Formulas Data Review View Help Power Pivot

**fx** Insert Function **Sum** AutoSum Recently Used **Median**

**WEEKNUM** **SUM** **AVERAGE**

Month	Servicing	Accessories	Monthly Total		
Jan	\$ 5,980.00	\$ 1,710.00	\$ 212,600.00		
Feb	\$ 6,124.00	\$ 1,844.00	\$ 243,070.00		
Mar	\$ 6,578.00	\$ 2,180.00	\$ 256,025.00		
Apr	\$ 7,210.00	\$ 2,258.00	\$ 283,715.00		
May	\$ 7,324.00	\$ 2,612.00	\$ 312,375.00		
Jun	\$ 7,688.00	\$ 2,407.00	\$ 334,889.00		
Jul	\$ 7,200.00	\$ 2,356.00	\$ 346,192.00		
Aug	\$ 7,734.00	\$ 2,170.00	\$ 326,257.00		
Sep	\$ 8,114.00	\$ 2,004.00	\$ 284,686.00		
Oct	\$ 8,396.00	\$ 1,822.00	\$ 237,619.00		
Nov	\$ 198,440.00	\$ 14,125.00	\$ 2,135.00	\$ 222,555.00	
Dec	\$ 177,783.00	\$ 12,667.00	\$ 6,320.00	\$ 2,337.00	\$ 199,107.00

AutoSave (Off) File Home Insert Page Layout Formulas Data Review

**fx** Insert Function **Sum** AutoSum Recently Used **Median**

**Text** **Date & Time** **Lookup & Reference** **Math & Trig** **More Functions**

**AND** **FALSE** **IF** **E** **Accessories**

**IFERROR** **IFNA** **IFS** **NOT** **OR** **SWITCH** **TRUE** **XOR**

**B24** A B E

Month	Car Sales	Accessories
Jan	\$ 193,670.00	\$ 1,710.00
Feb	\$ 221,445.00	\$ 1,844.00
Mar	\$ 233,378.00	\$ 2,180.00
Apr	\$ 260,135.00	\$ 2,258.00
May	\$ 289,940.00	\$ 2,612.00
Jun	\$ 312,692.00	\$ 2,407.00
Jul	\$ 323,446.00	\$ 2,356.00
Aug	\$ 303,575.00	\$ 2,170.00
Sep	\$ 260,336.00	\$ 2,004.00
Oct	\$ 212,291.00	\$ 1,822.00
Nov	\$ 198,440.00	\$ 2,135.00
Dec	\$ 177,783.00	\$ 2,337.00

File Home Insert Page Layout Formulas Data Review View Help Power Pivot

**fx** Insert Function... **More Functions** **Statistical** **AVEDEV**

**Lookup & Reference** **Math & Trig** **Name Manager** **Create from Selection**

Servicing	Accessories
\$ 5,980.00	\$ 1,710.00
\$ 124.00	\$ 1,844.00
\$ 578.00	\$ 2,180.00
\$ 2,210.00	\$ 2,258.00
\$ 324.00	\$ 312,375.00
\$ 688.00	\$ 334,889.00
\$ 2,000.00	\$ 346,192.00
\$ 7,734.00	\$ 326,257.00
\$ 114.00	\$ 2,004.00
\$ 396.00	\$ 284,686.00
\$ 1,822.00	\$ 237,619.00
\$ 855.00	\$ 222,555.00
\$ 320.00	\$ 2,337.00
\$ 523.00	\$ 199,107.00
\$ 2,152.92	
\$ 1,710.00	
\$ 2,612.00	
\$ 267.00	\$ 2,175.00

File Home Insert Page Layout Formulas Data Review

**fx** Insert Function... **More Functions** **Statistical** **MEDIAN**

**Text** **Date & Time** **Lookup & Reference** **Math & Trig** **Name Manager**

**Engineering** **Cube** **Information** **Compatibility** **Web**

Servicing	Accessories
\$ 5,980.00	\$ 1,710.00
\$ 6,124.00	\$ 1,844.00
\$ 6,578.00	\$ 2,180.00
\$ 7,210.00	\$ 2,258.00
\$ 7,324.00	\$ 2,612.00
\$ 7,688.00	\$ 2,407.00
\$ 7,200.00	\$ 2,356.00
\$ 7,734.00	\$ 326,257.00

# Referencing Data in Formulas

Now that you've had an introduction to functions, seeing the use of some common statistical functions and learned about some of the more advanced functions that a data analyst might use, in this video will look at the difference between relative, absolute, and mixed references in formulas as well as how to use them. And we'll learn about formula errors in Excel.

It's important to understand the difference between relative and absolute references when creating your formulas. By default, in Excel, cell references are always relative references. The term relative is the key here, because it means that when you reference a cell, you are in fact referencing the cells position in relation to the cell that the formula is in.

That is why when we have been copying formulas from one cell to another so far in this course, using either copy and paste or the fill handle, we haven't needed to modify the cell references because Excel assumes you are using relative references. When the formulas are copied, the cell references are changed to match the relative positions of the cells that are being copied to.

So now we know that relative references are the default in Excel, but how do we make it so that the cell references don't change when we copy them? For that you need to use absolute references in contrast to relative references. **Absolute references to cells stayed the same.** When you copy a formula containing such references.

Lastly, there may also be some instances where you only want one of the cell reference identifiers to be absolute and the other one to be relative. For example, you might want the row identifier to be absolute, but the column Identifier to be relative, or vice versa. These are called Mixed references and.

An example of this would be equal sign a dollar sign one plus A3 where a dollar one. Has a relative column and an absolute row or dollar 8. Three has an absolute column. Ando relative RO. In contrast to relative and absolute references, when you copy a formula containing mixed cell references, any relative cell references will change, whereas any absolute cell references will stay the same in the copied formula.

First, let's look at an example of using relative references in a formula. For example, if we enter the formula equals A1 plus a 3IN cell, four note the blue an red highlighted cells in a one, and a three. These denote the cells being relatively referenced in the formula. If we copy the formula to the cell directly below using the fill handle, we can see that the result changes, and if we look at the copied formula.

You can see that the blue and red cell references have changed relative to their position on the worksheet. The formula has been changed to equals A2 plus a four in the copied formula. That is, each cell reference has moved one cell down and if we copy and paste the formula to see seven, you can see that the results also changes and again we can see that the blue and red cell references in the copied formula have changed now.

Let's look at an example of how to use absolute references in a formula. All you need to do to make a cell reference absolute is put a dollar sign in front of the column and or row identifiers in the formula. For example, if we enter the formula equals dollar sign a one plus sign a dollar 3IN cell E4. Note the blue and red highlighted cells in a one and a three. These denote the cells being. Absolutely referenced in the formula.

When we copy the formula using the fill handle, you can see that the result stays the same this time and if we look at the copied formula you can see that the blue and red cell references haven't changed. The formula is still equal sign dollar a dollar one plus a dollar three in the copied formula. That is, the cell references haven't changed.

Similarly, if we then copy and paste the formula to E7, you can again see that the result stays the same this time and we can see that the blue and red cell references haven't changed. The formula is still equal sign dollar a dollar one plus dollar a dollar three in the copied formula. That is, the cell references haven't changed.

Lastly, will look at an example of how to use mixed references in a formula so. If we enter the formula equals a dollar one plus dollar 8, three in cell G4. Note the blue and red highlighted cells in A1A three. These denote the cells being referenced in the formula. If we copy the formula to the cell below using the fill handle, you can see that the result changes, but it's a different result from the previous examples. And if we look at the copied formula, you can see that the first blue cell reference has stayed the same.

But the second red cell reference has changed. If we copy and paste the formula to G7, you can see that the same thing happens. The result changes and again we can see that the first blue cell reference has stayed the same in the copied formula, while only the red cell reference has changed.

Now we'll have a quick introduction to dealing with **formula errors** in Excel.

Because of the complexity of writing formulas, especially the more complicated ones, there are bound to be occasions when you make a mistake in the syntax or in the data selection which will lead to a formula error. Errors are typically denoted by displaying in the cell that is supposed to be displaying the result. One of the error codes in this list when you see multiple hash symbols in a cell, it's not really an error, it just means the column either isn't wide enough to display the whole word or value. Or it contains a negative date or time value?

So if we type control plus semi colon, then space then control plus shift plus semi colon, it enters today's date and the current time. But the cell is too narrow to display it. So what we see is multiple hash symbols. If we adjust the column width we can now see the cell contents. So as I said, this really shouldn't be considered as an error. However if we enter the formula seen in Cell I7.

When we press enter, we see a hash name error. This error was caused by trying to use an X as a multiplication operator when in fact it should be an asterisk. Note the small green triangle in the top left corner of the cell. Also note that when you select the cell and exclamation mark appears, providing you with a hint about what caused the error. In this case it says the formula contains unrecognized text.

When you click the dropdown error next to the exclamation mark for an error, you see several options.

The first line also gives you a clue on the nature of the error. This one says invalid name error, so it was probably a mistyped cell reference value or function name. If you click help on this error, uh, help pane opens with specific information related to this error. If you click show calculation steps, a dialog box opens displaying the current syntax with the error underlined. And you can try to evaluate the error if you are certain the error is incorrect, you can choose ignore error, and if you want to edit the formula, click edit in Formula Bar and the cursor will be focused in the formula bar so that you can try and correct the formula error.

If you click error checking options, the Excel Options Dialog Box is opened at the section related to error checking rules and you can modify these options to suit your needs. Each of the errors you make which generate one of the error codes listed at the start of this video will have a different reason and a different solution. For more information on each of these errors and typical solutions visit the link provided. In this video we learned about referencing data in formulas, specifically differentiating between relative, absolute, and mixed references, and how to use them. And we learned about formula errors in Excel.

## Referencing Data in Formulas

- The difference between relative, absolute, and mixed references in formulas
- How to use them
- Formula errors in Excel

### Relative vs Absolute References

=A3+A5  
=\$A3+\$A5

### Relative References

	A	B	C	D	E	F	G
1	1						
2	2						
3	3						
4	4			4			
5			=A2+A4				
6	6						
7	7						
8	8						
9	9						

- By default, Excel uses relative references
- When you reference a cell in a formula, you are actually referencing its relative position
- The cell references are automatically changed in the copied formula

By default, in Excel, cell references are always relative references. The term relative is the key here, because it means that when you reference a cell, you are in fact referencing the cell's position in relation to the cell that the formula is in. That is why when we have been copying formulas from one cell to another so far in this course, using either copy and paste or the fill handle, we haven't needed to modify the cell references because Excel assumes you are using relative references. When the formulas are copied, the cell references are changed to match the relative positions of the cells that are being copied to. So now we know that relative references are the default in Excel.

# Absolute References

	A	B	C	D	E	F	G
1	1						
2	2						
3	3						
4	4			4	=A\$2+\$A\$3		
5	5			6			
6	6						
7	7				10		
8	8						
9	9						

- When you copy a formula containing absolute references, the cell references stay the same in the copied formula

How do we make it so that the cell references don't change when we copy them? For that you need to use absolute references in contrast to relative references. Absolute references to cells stayed the same. When you copy a formula containing such references.

# Mixed References

	A	B	C	D	E	F	G
1	1						
2	2						
3	3						
4	4			4			=A\$2+\$A3
5	5			6			
6	6				4		
7	7			10			
8	8				4		
9	9						

- When you copy a formula containing mixed references, the relative cell reference changes, while the absolute cell reference stays the same in the copied formula

There may also be some instances where you only want one of the cell reference identifiers to be absolute and the other one to be relative. For example, you might want the row identifier to be absolute, but the column Identifier to be relative, or vice versa. These are called mixed references and. An example of this would be equal sign a dollar sign one plus A3 where a dollar one. Has a relative column and an absolute row or dollar 8. Three has an absolute column. Ando relative RO. In contrast to relative and absolute references, when you copy a formula containing mixed cell references, any relative cell references will change, whereas any absolute cell references will stay the same in the copied formula.

## Relative reference

	A	B	C
1	1		
2	2		
3	3		
4	4		=A1+A3
5	5		

First, let's look at an example of using relative references in a formula. For example, if we enter the formula equals A1 plus a 3 in cell 4, note the blue and red highlighted cells in a one, and a three. These denote the cells being relatively referenced in the formula.

	A	B	C
1	1		
2	2		
3	3		
4	4		+ 4
5	5		=A2+A4
6	6		

If we copy the formula to the cell directly below using the fill handle, we can see that the result changes, and if we look at the copied formula. You can see that the blue and red cell references have changed relative to their position on the worksheet.

	A	B	C	D
1	1			
2	2			
3	3			
4	4			4
5	5			6
6	6			
7	7			=A4+A6
8	8			
9	9			

The formula has been changed to equals A2 plus a four in the copied formula. That is, each cell reference has moved one cell down and if we copy and paste the formula to see seven, you can see that the results also changes and again we can see that the blue and red cell references in the copied formula have changed now.

## Absolute Reference

C	D	E
4	+	= <span style="color: blue;">\$A\$1</span> + <span style="color: red;">\$A\$3</span>
6		
10		

an example of how to use absolute references in a formula. All you need to do to make a cell reference absolute is put a dollar sign in front of the column and or row identifiers in the formula. For example, if we enter the formula equals dollar sign a one plus sign a dollar 3IN cell E4. Note the blue and red highlighted cells in a one and a three. These denote the cells being Absolutely referenced in the formula.

C	D	E
4	+	4
6		= <span style="color: blue;">\$A\$1</span> + <span style="color: red;">\$A\$3</span>
10		

C	D	E
4		4
6	+	4
10		4

When we copy the formula using the fill handle, you can see that the result stays the same this time and if we look at the copied formula you can see that the blue and red cell references haven't changed. The formula is still equal sign dollar a dollar one plus a dollar three in the copied formula. That is, the cell references haven't changed. The formula is still equal sign dollar a dollar one plus a dollar three in the copied formula. That is, the cell references haven't changed. Similarly, if we then copy and paste the formula to E7, you can again see that the result stays the same this time and we can see that the blue and red cell references haven't changed. The formula is still equal sign dollar a dollar one plus dollar a dollar three in the copied formula. That is, the cell references haven't changed.

	A	B	C	D	E	F	G
1	1						
2	2						
3	3						
4	4		4		4		=A\$1+\$A3
5	5		6		4		
6	6						
7	7		10		4		
8	8						
9	9						

an example of how to use mixed references in a formula so. If we enter the formula equals a dollar one plus dollar 8, three in cell G4. Note the blue and red highlighted cells in A1A three. These denote the cells being referenced in the formula. If we copy the formula to the cell below using the fill handle, you can see that the result changes, but it's a different result from the previous examples.

	E	F	G
4			4
4			+ 5
4			

If we copy the formula to the cell below using the fill handle, you can see that the result changes, but it's a different result from the previous examples.

	E	F	G
4			4
4			=A\$1+\$A4
4			

And if we look at the copied formula, you can see that the first blue cell reference has stayed the same. But the second red cell reference has changed. If we copy and paste the formula to G7, you can see that the same thing happens. The result changes and again we can see that the first blue cell reference has stayed the same in the copied formula, while only the red cell reference has changed.

# List of Error Codes in Excel

#N/A

#NAME?

#NULL!

#REF!

#VALUE!

#DIV/0

One of the error codes in this list when you see multiple hash symbols in a cell, it's not really an error, it just means the column either isn't wide enough to display the whole word or value. Or it contains a negative date or time value?

G	H	I	J
4	10/07/2020 22:09		
5			
7			

G	H	I	J
4			
5			
7			

G	H	I	J
4			
5			
7			#####

G	H	I	J
4	10/07/2020 22:09		
5			
7			

So if we type control plus semi colon, then space then control plus shift plus semi colon, it enters today's date and the current time. But the cell is too narrow to display it. So what we see is multiple hash symbols. If we adjust the column width we can now see the cell contents. So as I said, this really shouldn't be considered as an error.

G	H	I	J
4	10/07/2020 22:09		
5			
7	=A1xC7		

G	H	I	J
4	10/07/2020 22:09		
5			
7		#NAME?	

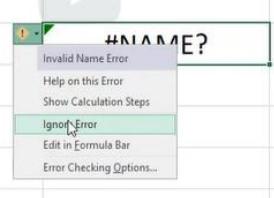
if we enter the formula seen in Cell I7. When we press enter, we see a hash name error. This error was caused by trying to use an X as a multiplication operator when in fact it should be an asterisk. Note the small green triangle in the top left corner of the cell.

G	H	I
4	10/07/2020 22:09	
5		
7		#NAME?



Also note that when you select the cell and exclamation mark appears, providing you with a hint about what caused the error. In this case it says the formula contains unrecognized text. When you click the dropdown error next to the exclamation mark for an error, you see several options. The first line also gives you a clue on the nature of the error. This one says invalid name error, so it was probably a mistyped cell reference value or function name. If you click help on this error, uh, help pane opens with specific information related to this error.

G	H	I
4	10/07/2020 22:09	
5		
7		#NAME?



If you click show calculation steps, a dialog box opens displaying the current syntax with the error underlined. And you can try to evaluate the error if you are certain the error is incorrect, you can choose ignore error, and if you want to edit the formula, click edit in Formula Bar and the cursor will be focused in the formula bar so that you can try and correct the formula error.

## Detect errors in formulas

<https://support.office.com/en-us/article/detect-errors-in-formulas-3a8acca5-1d61-4702-80e0-99a36a2822c1>



# Hands-on Lab 4: Simple Use of Functions

**Estimated time needed:** 30 minutes

In this lab, first you will learn the basics of formulas, how to perform simple calculations, how to select ranges in formulas, and how to copy formulas. Next, you will learn the basics of functions, how to use some of the more common functions that a Data Analyst might employ, and look at some of the more advanced functions available in Excel. Finally, you will learn about referencing data in formulas; specifically how to differentiate between relative and absolute references, and you will also learn about error handling in formulas.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, [it is recommended that you use Excel for the web for the hands-on labs](#) as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The dataset used in this lab is an internal dataset.

## Objectives

After completing this lab, you will be able to:

- Understand the basics of formulas
- Perform simple calculations
- Select ranges in formulas and copy formulas
- Understand the basics of functions
- Use common functions

- Understand the more advanced functions available
- Reference data in formulas
- Differentiate between relative and absolute references
- Understand how to handle formula errors

## Exercise 1: Basics of Formulas

In this exercise, you will learn the basics of formulas, how to perform simple calculations, how to select ranges in formulas, and how to copy formulas.

1. Download the file [Personal Monthly Expenditure Lab4.xlsx](#). Upload and open it using Excel for the web. Go to the **Expense - 2018** worksheet.

	A	B	C	D	E	F	G
1	Month	Housing	Bills & Utilities	Food & Dining	Personal	Auto & Transport	Health & Fitness
2	Jan	£ 800.00	£ 210.00	£ 400.00	£ 100.00	£ 100.00	£ 60.00
3	Feb	£ 800.00	£ 180.00	£ 350.00	£ 100.00	£ 125.00	£ 70.00
4	Mar	£ 800.00	£ 170.00	£ 420.00	£ 100.00	£ 120.00	£ 60.00
5	Apr	£ 800.00	£ 160.00	£ 400.00	£ 120.00	£ 100.00	£ 60.00
6	May	£ 800.00	£ 150.00	£ 420.00	£ 100.00	£ 100.00	£ 80.00
7	Jun	£ 800.00	£ 150.00	£ 380.00	£ 100.00	£ 130.00	£ 60.00
8	Jul	£ 800.00	£ 150.00	£ 420.00	£ 120.00	£ 100.00	£ 60.00
9	Aug	£ 800.00	£ 150.00	£ 420.00	£ 100.00	£ 100.00	£ 80.00
10	Sep	£ 800.00	£ 150.00	£ 400.00	£ 120.00	£ 110.00	£ 60.00
11	Oct	£ 800.00	£ 170.00	£ 420.00	£ 100.00	£ 100.00	£ 60.00
12	Nov	£ 800.00	£ 200.00	£ 390.00	£ 120.00	£ 100.00	£ 50.00
13	Dec	£ 800.00	£ 220.00	£ 400.00	£ 100.00	£ 115.00	£ 60.00
14							

2. In **A14**, type **Totals** and in **B14**, type **=SUM(** then select cells **B2 to B13** with the mouse, and press **Enter**.
3. Select the **fill handle** on cell **B14** and drag to **G14** to copy the formula.

13	Dec	£ 800.00	£ 220.00	£ 400.00	£ 100.00	£ 115.00	£ 60.00
14	<b>Totals</b>	£ 9,600.00	£ 2,060.00	£ 4,820.00	£ 1,280.00	£ 1,300.00	£ 760.00
15							

3. In cell **H1**, type **Monthly Total** and double-click the divider between **H** and **I**.
4. In **H2**, type **=SUM(** then select cells **B2 to G2** with the mouse, and press **Enter**. If necessary, select the **fill handle** on cell **H2** and drag to **H14** to copy the formula.
5. Select columns **B to H**. On the **Home** tab, in the **Number** group, click the **Accounting Number Format (\$)** drop-down list, and select **\$ English (United States)**.

# Exercise 2: Basics of Functions

In this exercise, you will have an introduction to functions, including using some common statistical functions, and then you will learn about some more advanced functions that a Data Analyst might also use.

1. In cells **A16-A20**, type the following:
  - o **Avg**
  - o **Min**
  - o **Max**
  - o **Count**
  - o **Median**
2. In **B16**, type **=AVERAGE(** then select cells **B2 to B13** with the mouse, and press **Enter**. Select the **fill handle** on cell **B16** and drag to **G16** to copy the formula.
3. In **B17**, type **=MIN(** then select cells **B2 to B13** with the mouse, and press **Enter**. Select the **fill handle** on cell **B17** and drag to **G17** to copy the formula.
4. In **B18**, type **=MAX(** then select cells **B2 to B13** with the mouse, and press **Enter**. Select the **fill handle** on cell **B18** and drag to **G18** to copy the formula.
5. In **B19**, type **=COUNT(** then select cells **B2 to B13** with the mouse, and press **Enter**. Select the **fill handle** on cell **B19** and drag to **G19** to copy the formula. Select row **19**. On the **Home** tab, click the **Number Format** drop-down list, and select **Number**.
6. In **B20**, type **=MEDIAN(** then select cells **B2 to B13** with the mouse, and press **Enter**. Select the **fill handle** on cell **B20** and drag to **G20** to copy the formula.

	A	B	C	D	E	F	G
16	Avg	\$ 800.00	\$ 171.67	\$ 401.67	\$ 106.67	\$ 108.33	\$ 63.33
17	Min	\$ 800.00	\$ 150.00	\$ 350.00	\$ 100.00	\$ 100.00	\$ 50.00
18	Max	\$ 800.00	\$ 220.00	\$ 420.00	\$ 120.00	\$ 130.00	\$ 80.00
19	Count	12.00	12.00	12.00	12.00	12.00	12.00
20	Median	\$ 800.00	\$ 165.00	\$ 400.00	\$ 100.00	\$ 100.00	\$ 60.00

7. Explore some more commonly used functions of a data analyst by clicking the arrow under **AutoSum**, then select **More Functions** and look at some of the functions in various categories to see what actions they perform:
  - o Financial: **ACCRINT, INTRATE**
  - o Logical: **AND, IF, OR, NOT**
  - o Text: **CONCAT, FIND, SEARCH**
  - o Date & Time: **NETWORKDAYS, WEEKDAY**
  - o Lookup & Reference: **AREAS, SORTBY, VLOOKUP, HLOOKUP**
  - o Math & Trig: **POWER, SUMIF, SUMPRODUCT**
  - o Statistical: **AVERAGE, COUNTIF, MAX, MEDIAN, MIN**

# Exercise 3: Referencing Data in Formulas (relative vs absolute) & Formula Errors

In this exercise, you will learn how to reference data in formulas; specifically differentiating between relative and absolute references, and you will also learn about error handling in formulas.

1. In cells **A31-A40**, type **1-10**. Select row **31 to 40**. On the **Home** tab, click the **Number Format** drop-down list, and select **General**.
2. Relative References : In cell **B33**, type **=A31+A32** and press **Enter**. Select the **fill handle** on cell **B33** and drag to **B40** to copy the formula. Here, both first and second cell reference will move 1 cell down. For example, on cell **B34** formula will be changed to **=A32+A33**, on cell **B35** formula will be changed to **=A33+A34** and so on.
3. Absolute References : In cell **C33**, type **=\$A\$31+\$A\$32** and press **Enter**. Select the **fill handle** on cell **C33** and drag to **C40** to copy the formula. Here, both first and second cell references will not change. For example, on cell **C34** formula will remain **=\$A\$31+\$A\$32**, on cell **C35** formula will remain **=\$A\$31+\$A\$32** and so on.
4. Mixed References : In cell **D33**, type **=\$A\$31+\$A32** and press **Enter**. Select the **fill handle** on cell **D33** and drag to **D40** to copy the formula. Here, first cell reference will stay the same, but the second reference will change. For example, on cell **D34** formula will be changed to **=\$A\$31+\$A33**, on cell **D35** formula will be changed to **=\$A\$31+\$A34** and so on.

	A	B	C	D
30		Relative	Absolute	Mixed
31		1		
32		2		
33	3	3	3	3
34	4	5	3	4
35	5	7	3	5
36	6	9	3	6
37	7	11	3	7
38	8	13	3	8
39	9	15	3	9
40	10	17	3	10

5. In cell **B31**, type **=A16+A17**. Now this will lead to a formula error **#VALUE!** since cells **A16** and **A17** do not contain any number.

A	B	C	D	E
30	Relative	Absolute	Mixed	
31	1 #VALUE!	Error in Value		
32	2	A value used in the formula is of the wrong data type.		
33	3			
34	4			
35	5			
36	6			
37	7	11	3	/
38	8	13	3	8
39	9	15	3	9
40	10	17	3	10

6. Click the **question mark icon** in the error message box. This will open the **Help** for this topic. Read through this help file for more information about **#VALUE!** errors in formulas.

B31      f<sub>x</sub> =A16+A17

A	B	C	D	E	F
30	Relative	Absolute	Mixed		
31	1 #VALUE!	Error in Value			
32	2	A value used in the formula is of the wrong data type.			
33	3				
34	4				
35	5				
36	6				
37	7	11	3	/	
38	8	13	3	8	
39	9	15	3	9	
40	10	17	3	10	
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					

Expense - 2018    Expense - 2019    +

Calculation Mode: Automatic    Workbook Statistics

Help

Search help

How to correct a #VALUE! error

#VALUE! is Excel's way of saying, "There's something wrong with the way your formula is typed. Or, there's something wrong with the cells you are referencing." The error is very general, and it can be hard to find the exact cause of it. The information on this page shows common problems and solutions for the error. You may need to try one or more of the solutions to fix your particular error.

Fix the error for a specific function

Which function are you using? ▾

Problems with subtraction

How to do basic subtraction ▾  
#VALUE! with basic subtraction ▾  
How to subtract dates ▾  
#VALUE! error subtracting dates stored as text ▾

Problems with spaces and text

Remove spaces that cause #VALUE! ▾

Check for text as general numbers ▾

Help Improve Office

# **Reading: Summary and Highlights**

## **In this lesson, you have learned:**

There are several features to modify views in Excel, and it is very straightforward to enter and edit data in a spreadsheet.

You can move or copy data within a worksheet or between worksheets, and you can use AutoFill to automatically enter data that is in a series or that fits a pattern.

You can format both cells and data in Excel.

A formula is made up of several component parts, and formulas can perform calculations using numbers directly or by using references to data in the worksheet.

You can use the Fill Handle in Excel to quickly copy formulas to other cells.

There are several different categories of function you can use for different purposes, and you can search for a function by name, or by category.

You can reference cells in the worksheet in your formulas by using relative, absolute, or mixed references.

You can make a formula absolute by adding a dollar symbol (\$) to a cell reference.

If you get errors in your formulas, you can use the error-checking capabilities of Excel to resolve them.

## Module 2: Practice Quiz

 Bookmark this page

### Question 1

1/1 point (ungraded)

How can you zoom to a specific area of data in an Excel spreadsheet?

- Use the Zoom button
- Use the Zoom to Selection button
- Use the Freeze Frames option
- Use the 100% button



### Question 2

1/1 point (ungraded)

What is one of the key components of a typical formula?

- Calculation
- Reference
- Division
- Percentage



### Question 3

1/1 point (ungraded)

What happens when you use the median calculation but select an even number of values in a range?

- Divides the selected range in half
- Returns middle figure between the two middle values in the selected range
- Returns one of the middle value of the selected range
- Returns both of the middle values of the selected range



---

### Question 4

1/1 point (ungraded)

When creating formulas, what is a mixed reference?

- Both cell references are either absolute, or relative
- One cell reference is absolute, the other one is relative
- One cell reference is absolute, the other must be manually assigned
- Both cell references are manually assigned



## Module 2: Graded Quiz

 Bookmark this page

Graded Quiz due Jul 22, 2022 04:03 +08

### Question 1

1/1 point (graded)

Which of the following is a valid way of editing existing data in a cell? Select all that apply

- Press F2
- Select the cell you want to edit and press Enter
- Select the cell you want to edit and then click in the formula bar
- Press CTRL+N



### Question 2

1/1 point (graded)

In Excel for the web, how can you format data in cells to use a currency? Select all that apply

- Select the data and click the Decrease Decimal button
- Select "Format cells" from the Format drop-down list in the Cells group
- Select "More Number Formats" from the Number Format drop-down list in the Number group
- Right-click on a cell and select Number Format



### Question 3

1/1 point (graded)

What is one of the functions found in the AutoSum drop-down list?

- General
- Accounting
- Count Numbers
- Number



## Question 5

1/1 point (graded)

How do you make a cell reference absolute in a formula?

- Put a plus sign (+) between the column and row identifiers in the formula
- Put a percentage sign (%) behind the column or row identifiers in the formula
- Put a dollar sign (\$) in front of the column and/or row identifiers in the formula
- Put an asterisk (\*) in front of the column or behind the row identifiers in the formula



## Question 4

1/1 point (graded)

In Excel Desktop, what is one of the function categories on the Formulas tab, in the Function Library group?

- Medical
- Analytical
- Lookup and Reference
- Functional



## Question 5

1/1 point (graded)

How do you make a cell reference absolute in a formula?

- Put a plus sign (+) between the column and row identifiers in the formula
- Put a percentage sign (%) behind the column or row identifiers in the formula
- Put a dollar sign (\$) in front of the column and/or row identifiers in the formula
- Put an asterisk (\*) in front of the column or behind the row identifiers in the formula



<https://github.com/MicrosoftLearning/Analyzing-Visualizing-Data-Excel>

<https://github.com/kanuptiwari/Data-analysis-in-Excel-/blob/main/Complete%20Data%20Analysis.xlsx>

(refer to folder excel capstone) Data-analysis-in-Excel-

THIS IS A CASE STUDY ON REVENUE OF A BICYCLE COMPANY WHICH SALES PRODUCT IN MORE THAN ONE COUNTRY AND CASE STUDY HAD PERFORMED USING SALES DATA YEAR OVER YEAR. THE TABLE INVOLVES THE COUNTRY NAME, PRODUCT CATEGORY , SUB CATEGORY DETAILS & YEAR. IN THE SHEET YOU CAN FIND THE COMPLETE DATA INFORMATION.

<https://github.com/sonpn82/Data-analysis-using-Excel>

---

## **Module 3 - Basics of Data Quality and Privacy**

---

-  [Module Introduction and Learning Objectives](#)
-  [Video: Introduction to Data Quality](#)
-  [Video: Importing File Data](#)
-  [Video: Basics of Data Privacy](#)
-  [Video: Viewpoints: Data Quality and Privacy](#)
-  [Reading: Summary and Highlights](#)
-  [Module 3: Practice Quiz](#)
-  [Module 3: Graded Quiz \(3 Questions\)](#)

Graded Quiz due Jul 25, 2022, 2:03 AM GMT+8

## **Module Introduction**

In this module, you will learn about the importance of data quality and how to import file data in to Excel from external sources. You will also learn about the fundamentals of data privacy.

## **Learning Objectives**

**After completing this module, you will be able to:**

- Explain the importance of data quality.
- Import file data into Excel.
- Describe the fundamentals of data privacy.

# Introduction to Data Quality

Data analysis can play a pivotal role in business decisions and processes. In order to use the data to make confident decisions, we must have the right information for the project and the data must be free from errors.

In this video we will learn how to profile data to discover inconsistencies.

Whether we are working with small sets of data or analyzing a spreadsheet with thousands of rows, one of the most difficult parts of the data analysis is finding and keeping clean data. To help with this process and qualify the data, look for these five traits: Accuracy, Completeness, Reliability, Relevance and Timeliness.

## Accuracy

- Is the first and most significant aspect to data quality. A data analyst must clean the data set by removing duplicates, correcting formatting errors, and removing blank rows.

## Completeness

Another important aspect of data quality is determining if the information required to **complete** the data set is readily available.

- Why does this matter as a trait for quality data?
- Let's say we are given the task to calculate the revenues of all sales per region. After collecting the data, we discover that no regions were specified. This data would then be considered incomplete and other sources would have to be considered to obtain the data required.

## Reliability

- is another vital factor in determining the quality of the data.
- For instance, let's say we are given the task to determine the agent revenue by customer. When gathering the data, we find the agents keep their own records and do not always update the information in the shared company database.
- With those factors in mind, we would then determine that the data in the shared company database was unreliable and new processes would need to be established to ensure reliable data.

**Relevance** is another trait of quality data.

- When collecting information, a data analyst must consider if the data being assembled is really necessary for the project.
- For example, when reviewing the data related to the sales revenue per customer, information such as customer birthdays and other personal information is also included.

- By making the determination early to exclude the personal information from the data set, the analyst would save themselves from having to review unnecessary information.

The last factor in determining the quality of the data is **timeliness**.

- This trait refers to the availability and accessibility of the selected data.
- Let's say our sales report is going to be used for weekly employee reviews, but our report is only refreshed once a month.
- This error in refreshing the data would cause our report to become outdated, and would have serious consequences for employee reviews.

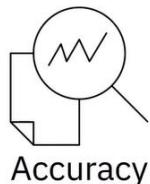
In this video we learned the important role of a data analyst in qualifying data. By considering the five traits of good quality data, an analyst can save time, avoid serious issues, and have data that is free from errors. In the next video we will take the collected data and learn how to import it to our spreadsheet.

## Introduction to Data Quality

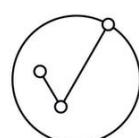
Data analysis can play a pivotal role in business decisions and processes.

In order to use the data to make confident decisions, we must have the right information for the project and the data must be free from errors.

### 5 Traits of Good Data



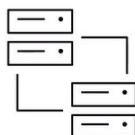
Accuracy



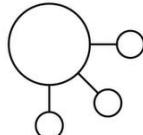
Completeness



Timeliness



Reliability



Relevance

## 1. Accuracy

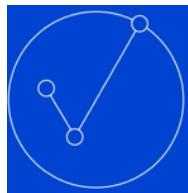


Accuracy is the first and most significant aspect to data quality.

ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONT
10103	26	100		11 5404.62	1/29/2003 0:00	Shipped	1	
10112	29	100		1 7209.11	3/24/2003 0:00	Shipped	1	
10112	29	100		1 7209.11	3/24/2003 0:01	Shipped	1	
10183	23	100		8 5372.57	11/13/2003 0:00	Shipped	4	
10194	42	100		11 7290.36	11/25/2003 0:00	Shipped	4	

- is the first and most significant aspect to data quality. A data analyst must clean the data set by removing duplicates, correcting formatting errors, and removing blank rows.

## 2. Completeness



Another important aspect is a complete and readily available dataset.

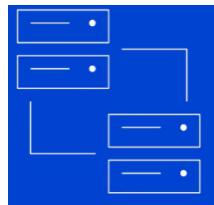
CUSTOMERNAME	COUNTRY	TERRITORY	CONTACTLASTNAME	CONTACTFIRSTNAME	DEALSIZE
Baane Mini Imports			Bergulsen	Jonas	Medium
Volvo Model Replicas, Co			Berglund	Christina	Large
Volvo Model Replicas, Co			Berglund	Christina	Large
Classic Gift Ideas, Inc			Cervantes	Francisca	Medium
Saveley & Henriot, Co.			Saveley	Mary	Large

Another important aspect of data quality is determining if the information required to complete the data set is readily available.

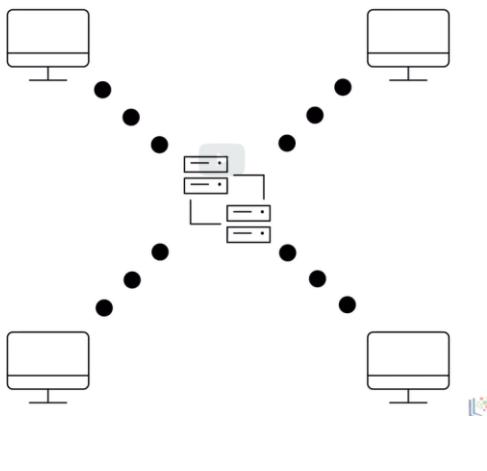
- Why does this matter as a trait for quality data?

Let's say we are given the task to calculate the revenues of all sales per region. After collecting the data, we discover that no regions were specified. This data would then be considered incomplete and other sources would have to be considered to obtain the data required.

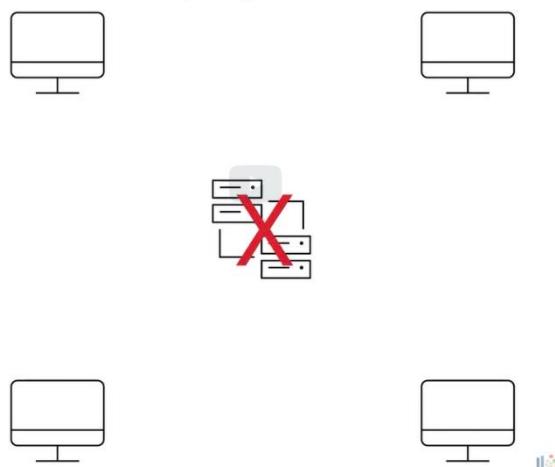
### 3. Reliability



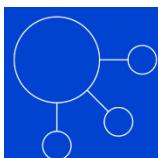
Reliability is another vital factor in determining the quality of the data.



Reliability is another vital factor in determining the quality of the data.



### 4. Relevance



Consider if the data being assembled is really necessary for the project.

Customer ID	Customer Name	Customer Birthday	Customer Phone
1235-4	Pam B.	10/11/75	555-0191
1235-5	Jim H.	03/04/78	555-0101

Customer ID	Customer Name	Customer Phone
1235-4	Pam B.	555-0191
1235-5	Jim H.	555-0101

Relevance is another trait of quality data.

- When collecting information, a data analyst must consider if the data being assembled is really necessary for the project.
- For example, when reviewing the data related to the sales revenue per customer, information such as customer birthdays and other personal information is also included.
- By making the determination early to exclude the personal information from the data set, the analyst would save themselves from having to review unnecessary information.

5. Timeliness



The last factor in determining the quality of the data is timeliness.



- This trait refers to the availability and accessibility of the selected data.
- Let's say our sales report is going to be used for weekly employee reviews, but our report is only refreshed once a month.
- This error in refreshing the data would cause our report to become outdated, and would have serious consequences for employee reviews.

# Importing File Data

Learn to:

- How to import data
- How to adjust column widths
- How to add and remove columns and rows

Now that you have learned about the importance of data quality, in this video you will learn how to import data from a text file using the Text Import Wizard, learn how to adjust column widths, and learn how to add and remove columns and rows.

As you know, by default Excel works with .XLSX or .XLS files and opens them as workbooks. But Excel can also use data that is in other formats, such as plain text, or data that has been comma-separated and tab-separated. Sometimes, these source files will be saved with a .TXT extension and referred to as 'text' files, but others might be saved with a .CSV file extension, and are typically referred to as CSV files.

Here in Notepad, I have opened a text file that contains data about car sales, and it uses comma separated values (or CSVs) to separate each bit of data in a record. Notice that the top line holds headings, such as Manufacturer, Model, Engine\_size, and so on, and each one is separated by a comma.

We want these to become our headers when we import the file into Excel. The line below these headings is the first line of real data, and again you can see that each piece of data is also separated by a comma. There are 16 headings and there are also 16 pieces of data on each of the lines below the headings. If we scroll to the bottom, we can see that last data record is for the Volvo S80.

Now, to open the file in Excel, we choose File, Open, and then either select the file from the recently used list, or click Browse to find the file we want to import. When we open the file, the Text Import Wizard launches automatically, and it will start to try and determine what your file is. Note that it has been detected as being a delimited file; that is, one that has its data fields separated by a character such as a comma or a tab. As we want the headings to become headers in Excel, we need to ensure that we select the option 'My data has headers'.

We can see a mini preview of the data in the preview box below. Then we click Next to proceed in the wizard. In step 2 of the wizard, we need to select our delimiter; that is, which character is separating our pieces of data; so we select Comma, and deselect any others. Note the data preview now starts to show us what the imported data will look like. You can scroll down and across this preview window to ensure that the data is going to look as you want and expect. It all looks OK, so we'll continue with the wizard.

In step 3 of the wizard, we can set the data format for each column. For example, you might want to change a column to Text or Date format. In this case we can just accept the default General format, and finish the import wizard.

In Excel we can see that the headings in the text file have been imported as a header row. But also notice that some of the columns are not showing all the data; some of the headings are not showing in full and some of the data is not shown either; all you can see are a number of hashes in the cells. This is because the column widths are too narrow in some cases. If you remember, we can manually adjust a column's width by dragging the divider across. But to change them all in one go, we select all the columns first, then double-click one of the selected column dividers.

We can do a similar thing with rows by dragging to make them bigger or smaller, or double-clicking a row divider to autosize it. There are some columns that we have decided we don't really need; namely Vehicle\_type and Latest\_Launch, so let's remove those. This can either be done using the Delete drop-down menu in the Cells group on the Home tab, and select Delete Sheet Columns, or by selecting and right-clicking a column and deleting it that way. To add another column, you simply select the column to right of where you want your new column to be, then right-click the column and choose Insert.

And let's give the header a name, such as Year. To delete a row you don't need, select the row, right-click it, and choose Delete. And to add a row, select the row below the place you want to add your new row, right-click the row, and choose Insert. If you want to save the file as an Excel file, you can either choose File, Save As, or you can click Save As in the yellow tooltip that appeared at the top of the worksheet when we imported the file, and then you would choose 'Excel Workbook (\*.xlsx)' in the 'Save as type' box.

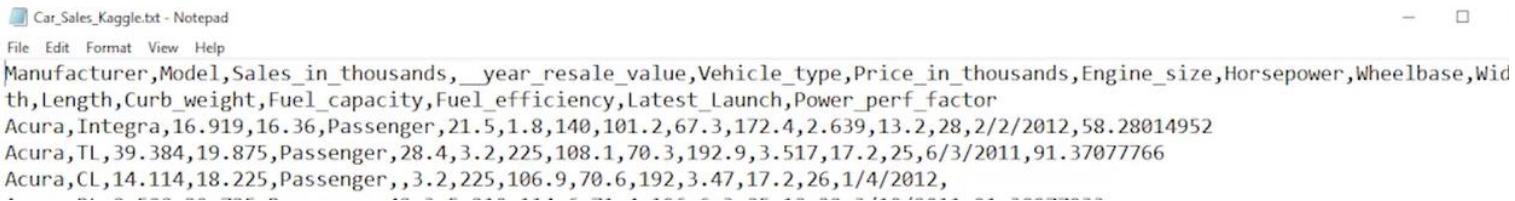
In this video, we learned how to import data using the Text Import Wizard, we learned how to adjust column widths, and we learned how to add and remove columns and rows.

In the next video, we will discuss the importance of data privacy, including sensitive information, and personally identifiable data.

## Importing File Data

Learn to:

- How to import data
- How to adjust column widths
- How to add and remove columns and rows



Sometimes, these source files will be saved with a .txt extension and referred to as 'text' files, but others might be saved with a .CSV file extension, and are typically referred to as CSV files. Here in Notepad, I have opened a text file that contains data about car sales, and it uses comma separated values (or CSVs) to separate each bit of data in a record.

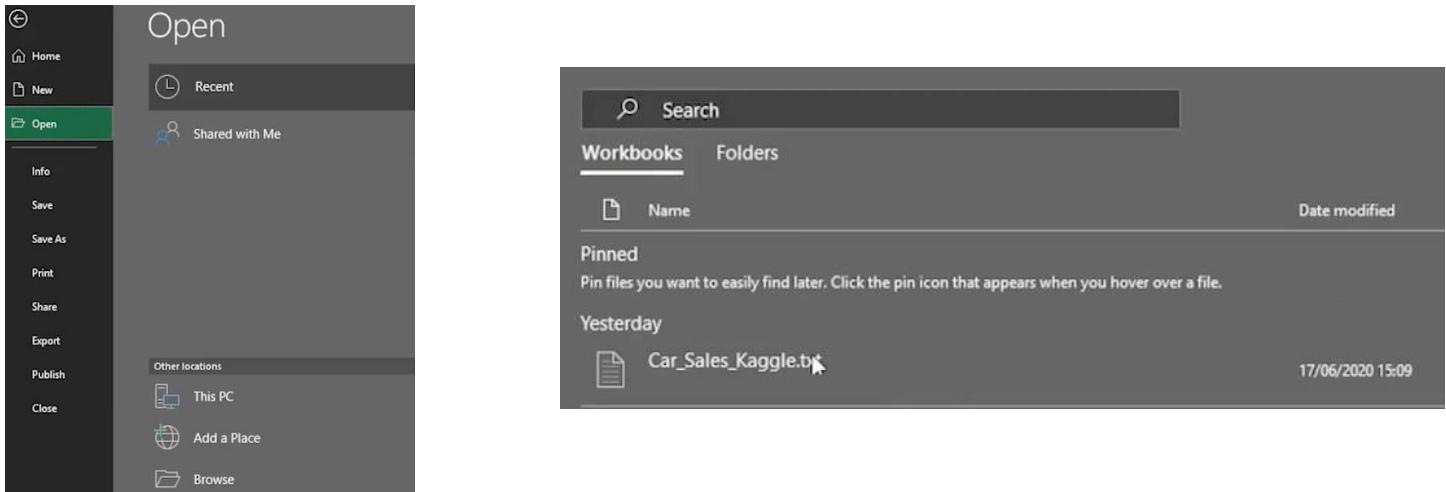
---

```
Car_Sales_Kaggle.txt - Notepad
File Edit Format View Help
Manufacturer,Model,Sales_in_thousands,__year_resale_value,Vehicle_type,Price_in_thousands,
th,Length,Curb_weight,Fuel_capacity,Fuel_efficiency,Latest_Launch,Power_perf_factor
Acura,Integra,16.919,16.36,Passenger,21.5,1.8,140,101.2,67.3,172.4,2.639,13.2,28,2/2/2012,
Acura,TL,39.384,19.875,Passenger,28.4,3.2,225,108.1,70.3,192.9,3.517,17.2,25,6/3/2011,91.3
Acura,CL,14.114,18.225,Passenger,,3.2,225,106.9,70.6,192,3.47,17.2,26,1/4/2012,
```

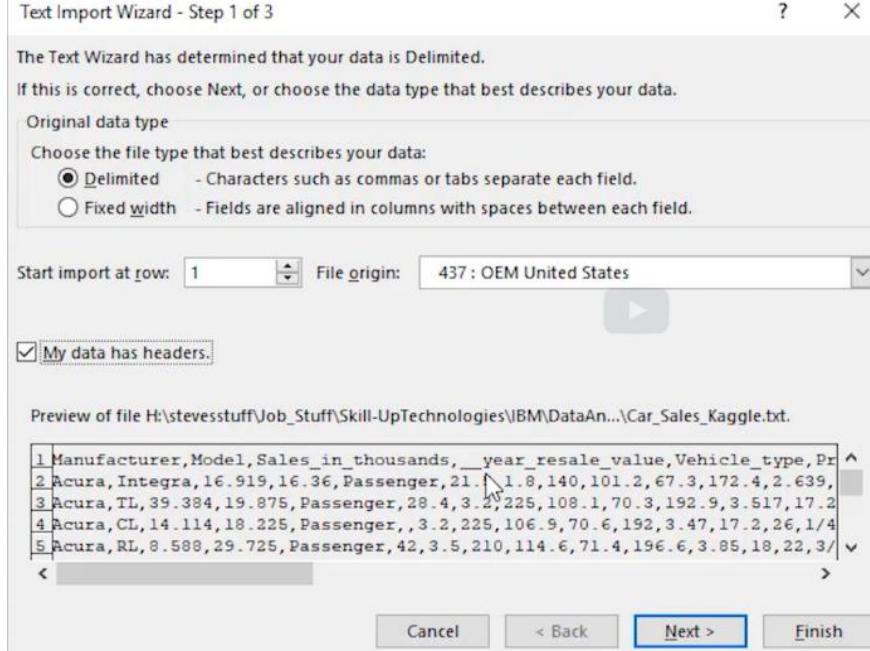
We want these to become our headers when we import the file into Excel. The line below these headings is the first line of real data, and again you can see that each piece of data is also separated by a comma. There are 16 headings and there are also 16 pieces of data on each of the lines below the headings.

## To transfer and open file in excel

1. Now, to open the file in Excel, we choose File, Open, and then either select the file from the recently used list, or click Browse to find the file we want to import.

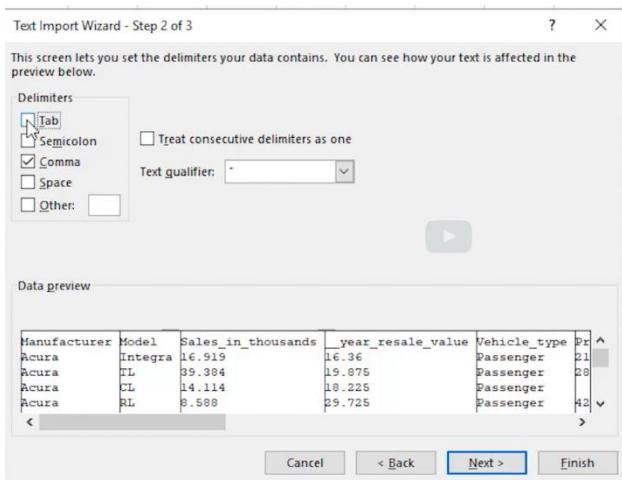


2.



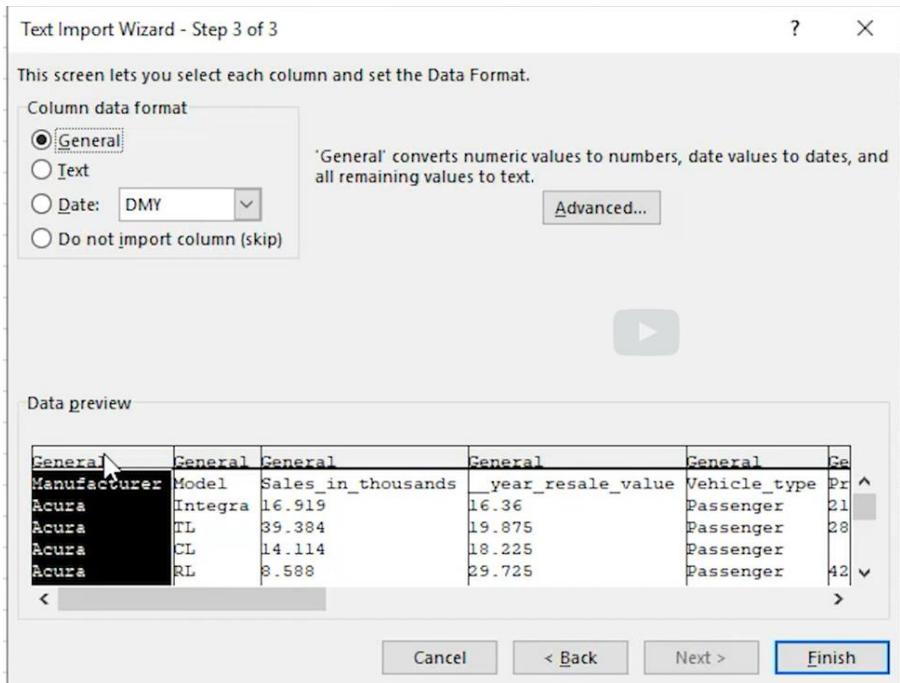
When we open the file, the Text Import Wizard launches automatically, and it will start to try and determine what your file is. Note that it has been detected as being a delimited file; that is, one that has its data fields separated by a character such as a comma or a tab. As we want the headings to become headers in Excel, we need to ensure that we select the option 'My data has headers'. We can see a mini preview of the data in the preview box below.

3.



In step 2 of the wizard, we need to select our delimiter; that is, which character is separating our pieces of data; so we select Comma, and deselect any others. Note the data preview now starts to show us what the imported data will look like. You can scroll down and across this preview window to ensure that the data is going to look as you want and expect. It all looks OK, so we'll continue with the wizard.

4.



In step 3 of the wizard, we can set the data format for each column. For example, you might want to change a column to Text or Date format. In this case we can just accept the default General format, and finish the import wizard.

## 5.Edit

Manufacturer	Model	Sales_in_thousands	year_resale_value	Vehicle_type	Price_in_thousands	Engine_size	Horsepower	Wheelbase	Width	Length	Curb_weight	Fuel_capacity	Fuel_efficiency	Latest_Lat	Power_perf_factor
Acura	Integra	16.919	16.36	Passenger	21.5	1.8	140	101.2	67.3	172.4	2.639	13.2	28	#####	58.28015
Acura	TL	39.384	19.875	Passenger	28.4	3.2	225	108.1	70.3	192.9	3.517	17.2	25	#####	91.37078
Acura	CL	14.114	18.225	Passenger		3.2	225	106.9	70.6	192	3.47	17.2	26	#####	
Acura	RL	8.588	29.725	Passenger	42	3.5	210	114.6	71.4	196.6	3.85	18	22	#####	91.38978

If you remember, we can manually adjust a column's width by dragging the divider across. But to change them all in one go, we select all the columns first, then double-click one of the selected column dividers. We can do a similar thing with rows by dragging to make them bigger or smaller, or double-clicking a row divider to autosize it.

5.1

	Acura	CL		14.114	18.225	3.2	225	106.9	70.6	192	3.47	17.2	26		
1	Acura	RL		8.588	29.725	42	3.5	210	114.6	71.4	196.6	3.85	18	22	91.3897
2	Audi	A4		20.397	22.255	23.99	1.8	150	102.6	68.2	178	2.998	16.4	27	62.7771
3	Audi	A6		18.78	23.555	33.95	2.8	200	108.7	76.1	192	3.561	18.5	22	84.5651
4	Audi	A8		1.38	39	62	4.2	310	113	74	198.2	3.902	23.7	21	134.6561
5	BMW	323i		19.747		26.99	2.5	170	107.3	68.4	176	3.179	16.6	26	71.1912
6	BMW	328i		9.231	28.675	33.4	2.8	193	107.3	68.5	176	3.197	16.6	24	81.8770
7	BMW	528i		17.527	36.125	38.9	2.8	193	111.4	70.9	188	3.472	18.5	25	83.998
8	Buick	Century		91.561	12.475	21.975	3.1	175	109	72.7	194.6	3.368	17.5	25	71.1814
9	Buick	Regal		39.35	13.74	25.3	3.8	240	109	72.7	196.2	3.543	17.5	23	95.6367
10	Buick	Park Avenue		27.851	20.19	31.965	3.8	205	113.8	74.7	206.8	3.778	18.5	24	85.8284
11	Buick	LeSabre		83.257	13.36	27.885	3.8	205	112.2	73.5	200	3.591	17.5	25	84.2545
12	Cadillac	DeVille		63.729	22.525	39.895	4.6	275	115.3	74.5	207.2	3.978	18.5	22	113.854
13	Cadillac	Seville		15.943	27.1	44.475	4.6	275	112.2	75	201		18.5	22	115.621

In Excel we can see that the headings in the text file have been imported as a header row. But also notice that some of the columns are not showing all the data; some of the headings are not showing in full and some of the data is not shown either; all you can see are a number of hashes in the cells. This is because the column widths are too narrow in some cases. If you remember, we can manually adjust a column's width by dragging the divider across. But to change them all in one go, we select all the columns first, then double-click one of the selected column dividers.

We can do a similar thing with rows by dragging to make them bigger or smaller, or double-clicking a row divider to autosize it. There are some columns that we have decided we don't really need; namely `Vehicle_type` and `Latest_Launch`, so let's remove those. This can either be done using the Delete drop-down menu in the Cells group on the Home tab, and select Delete Sheet Columns, or by selecting and right-clicking a column and deleting it that way. To add another column, you simply select the column to right of where you want your new column to be, then right-click the column and choose Insert.

## Basics of Data Privacy

In this video, we will learn about data privacy and the regulations that govern the collected data. When collecting customer data, specific regulations apply to how that data can be used. By understanding data privacy regulations and getting familiar with the following three fundamentals, you can eliminate the risk of financial penalties and keep the trust of your customers. Confidentiality, Collection and Use, and Compliance.

**Confidentiality** is an important element in data privacy and it acknowledges that the customer's personal information belongs to them. The types of information that can be accessed by a data analyst can range from sales forecasts, to employee information, or even patient records. When accessing these types of records the analyst must be able to recognize the different types of personal data.

**Personal Information** or PI is any type of information that can be traced back to a specific individual. This type of information can include anything from emails to images.

**Personally Identifiable Information** or PII is specific information that could be used to identify an individual. This type of information could include a social security number or a driver's license number.

**Sensitive Personal Information** or SPI, may not necessarily identify a specific individual, but contains private information that needs to be protected because if made public it could possibly be used to harm the individual. The type of information can include data about race, sexual orientation, biometric or genetic information. By understanding personal data and the associated regulations, we can efficiently anonymize our data by removing unnecessary information. This type of action can help build consumer confidence and continue to develop the free flow of information.

When searching through data, the analyst must know the location of the company collecting the data and the location of the respondent. Knowing where the data was collected is an essential element of data privacy and what regulations must be applied.

The General Data Protection Regulation or GDPR is a regulation specific to the European Union, and only applies to the jurisdiction of the individual. A new law created in Brazil, the LGPD, will take effect in August 2020. These new data policy regulations apply to individuals within Brazil, and ignores the location of the data processor.

While the United States does not have one country-wide principle law for data privacy. Because of this individual states began to make their own regulations. For instance, California created the California Consumer Privacy Act (CCPA) to better protect customer data. There are also industry specific regulations that govern the collection and use of sensitive and personal data. For example, in Healthcare,

HIPAA privacy rules govern the collection and disclosure of protected health information.

In retail, the PCI standards govern credit card data, and failure to safeguard cardholder information can result in hefty fines. With a basic understanding of these policies, we are able to remain compliant when handling any sensitive information.

Unfortunately, breeches in customer data is an all too common occurrence and understanding how to remain compliant is essential. Understanding the data privacy regulations of the European Union, the United States, and other countries as well as industries is key to keeping data safe.

Companies must comply with these privacy regulations at all times and also make sure policies are readily accessible to employees. For example, let's say a data analyst downloads a spreadsheet of sensitive information. In order to complete the report by Monday morning, the analyst decided to take their work laptop home for the weekend. After driving home, the analyst accidentally left the laptop in their car. The next morning, they found their car had been stolen along with the laptop. Because it is the responsibility of the company to keep customer data safe, this was a breach of privacy when the data left company property.

This type of action could not only cost the company large amounts of money in fines and penalties, but could also reduce consumer confidence causing a significant impact to revenue.

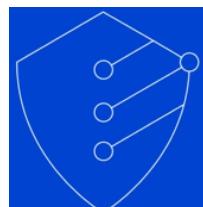
While data privacy applies to most data that is collected, there are some instances where these regulations do not apply. In order for these laws and regulations not to apply, the particular collection of data must be completely anonymous. To make data anonymous means to exclude all data which ties it back to a particular individual. While this approach might not be practical in all circumstances, collecting data with privacy in mind could remove privacy limitations and make data collections more accessible.

In this video we learned about the importance of data privacy and the challenges that a data analyst can face when collecting and sorting through data. In the videos in the next lesson, we will learn about different methods for cleaning data in a spreadsheet.

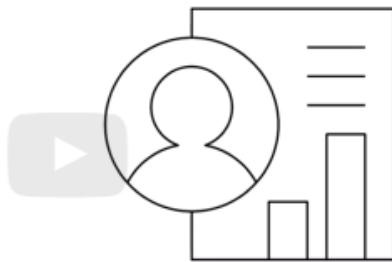
## Basics of Data Privacy

Learn about data privacy and the regulations that govern the collected data

## Fundamentals of Data Privacy



When collecting customer data, specific regulations apply to how that data can be used.

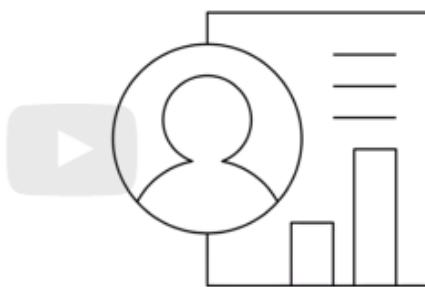


**Confidentiality   Collection and Use   Compliance**



## 1. Confidentiality

Confidentiality is an important element in data privacy and it acknowledges that the customer's personal information belongs to them.



Confidentiality is an important element in data privacy and it acknowledges that the customer's personal information belongs to them.



Sales Forecasts



Employee Information



Patient Records

# Different types of personal data

1.

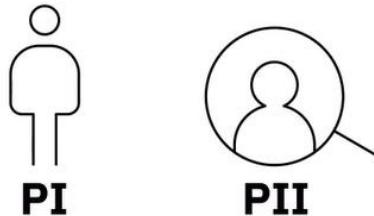


Customer ID	Customer Name	Customer Photos	Customer Emails
1235-4	Pam B.	 Link	Pamela.H@...
1235-5	Jim H.	 Link	Jimmy@...

## **Personal Information or PI**

is any type of information that can be traced back to a specific individual. This type of information can include anything from emails to images.

2.



Customer ID	Customer Name	Driver's License	Social Security
1235-4	Pam B.	 24558574	874-58-4758
1235-5	Jim H.	 45157647	874-98-8745

## **Personally Identifiable Information or PII**

is specific information that could be used to identify an individual. This type of information could include a social security number or a driver's license number.

3.

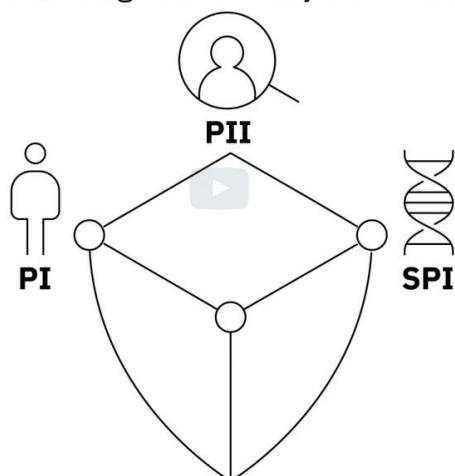


Customer ID	Customer Name	Fingerprint	Genetic Test
1235-4	Pam B.	Link	Link
1235-5	Jim H.	Link	Link

### Sensitive Personal Information or SPI

may not necessarily identify a specific individual, but contains private information that needs to be protected because if made public it could possibly be used to harm the individual. This type of information can include data about race, sexual orientation, biometric or genetic information.

By understanding personal data and the associated regulations, we can efficiently anonymize our data by removing unnecessary information.

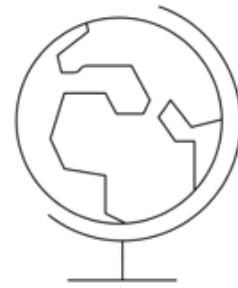


## 2. Collecting the Data

When searching through data, the analyst must know the location of the company collecting the data and the location of the respondent.



**The General Data Protection Regulation or GDPR** is a regulation specific to the European Union



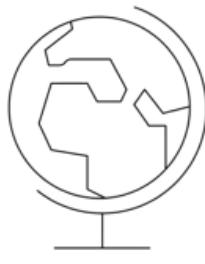
### **Lei Geral de Proteção de Dados Pessoais or LGPD**

These new data policy regulations apply to individuals within Brazil, and ignores the location of the data processor



**United States does not have one country-wide principle law for data privacy**

California created the California Consumer Privacy Act (CCPA) to better protect customer data



### **There are also industry specific regulations**

HIPAA privacy rules govern the collection and disclosure of protected health information. In retail, the PCI standards govern credit card data, and failure to safeguard cardholder information can result in hefty fines.

## 3. Compliance



Unfortunately, breaches in customer data is an all too common occurrence and understanding how to remain compliant is essential.

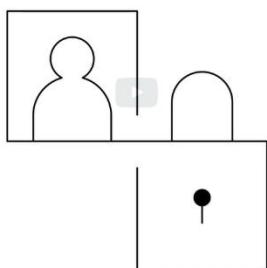


Understanding the data privacy regulations of the European Union, the United States, and other countries as well as industries is key to keeping data safe.



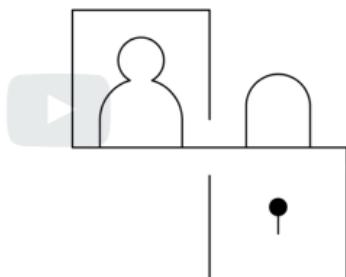
For example, let's say a data analyst downloads a spreadsheet of sensitive information. In order to complete the report by Monday morning, the analyst decided to take their work laptop home for the weekend. After driving home, the analyst accidentally left the laptop in their car. The next morning, they found their car had been stolen along with the laptop. Because it is the responsibility of the company to keep customer data safe, this was a breach of privacy when the data left company property. This type of action could not only cost the company large amounts of money in fines and penalties, but could also reduce consumer confidence causing a significant impact to revenue.

While data privacy applies to most data that is collected, there are some instances where these regulations do not apply.



While data privacy applies to most data that is collected, there are some instances where these regulations do not apply. In order for these laws and regulations not to apply, the particular collection of data must be completely anonymous.

To make data anonymous means to exclude all data which ties it back to a particular individual.



To make data anonymous means to exclude all data which ties it back to a particular individual. While this approach might not be practical in all circumstances, collecting data with privacy in mind could remove privacy limitations and make data collections more accessible.

# Viewpoints: Data Quality and Privacy

In this video, we will listen to several data professionals discuss the importance of data quality and data privacy as they relate to data analysis. Let us start with, "What is the importance of data quality as it relates to data analysis?"

Data quality is of the utmost importance in terms of data and analytics, but the reason behind this is because as soon as what you're presenting does not align with what someone expects, that's the first thing that they tend to go after. Where did you get the data? What's happened to the data? How's it been transformed? Because people like to think that they know and understand their, their business. And when you start to challenge that if you don't have the ground to stand on of the data that it's quality that it's clean and then it is from a trusted source, that's when you start to get into a lot of discussions. A lot of debate. And ultimately, the plot of what you're trying to present gets lost.

The backbone of any successful data analysis project is good quality data. There is a common term in computer science called garbage in garbage out, which is essentially if you read in bad quality data, you can expect to get bad quality results. So, there's really nothing more important when doing a data analysis than making sure that you're working with good quality data, and it's really important to sense-check the data yourself and really feel comfortable that the data you're using is of a really high quality.

Data accuracy is above all: garbage in garbage out. It's a waste of time to analyze data of poor quality, and it might mislead the business direction.

The integrity of the data that you're using or providing for someone else to use is of the utmost importance. Data is used determine, when or where to launch a product, if a division is profitable or not and it's easy to get things confused if you're not paying attention to the details. Using inventory as an example, if you're looking at inventory at a SKU level and you accidentally pick the wrong SKU to analyze and then you draw these conclusions that this particular item isn't profitable when in fact it is.

So, that's a major, major decision for a company to make obviously, so the expectation is that there will be lots of due diligence, but in the beginning if you start off with that data and then you build on that only to later realized that it wasn't a good idea, you've lost time, energy, effort, and in some cases, trust. Thank you for those viewpoints.

What about the importance of data privacy as it relates to data analysis?

Data privacy is incredibly important, especially when you're working in industries like pharmaceuticals or healthcare, but that's not where it stops. We have to have the ability to make sure that the users are getting the appropriate level of data based on their roles and their permissions.

Now we can do this through a number of cuts of the data specific to each geography or each function, or in some tools such as Cognos Analytics, we can start to build out that as part of our model. Within there you can say who has access to what, whether it's at a granular level of this person can see data in Canada or the US or whether it's simply this person can see this report in its entirely or not.

There's lots of different ways to handle this, but data privacy is of the utmost important across all industries.

In today's world, data privacy is a huge thing on the tax side, especially of our business we have what we have what we call PII: personal, identifiable information. We have to protect that and so we can't just send things through email. We don't send tax returns or even actually in our business. We don't send things through email. They have sensitive PII data in it. We encrypt it. We make sure the email is encrypted or we use software. Some certain softwares that will allow us to not show the social security numbers or the names or the date of birth and what will happen is it has a certain sequence, and we share that with the client by calling them. We don't put that in an email and we certainly don't put that in the same email with the encrypted information because we want to make sure that you are always safe. So, we have to make sure we're protecting it. At all costs.

## Importance of Data Quality & Privacy : Experts Viewpoints

What is the importance of data quality as it relates to data analysis?"

- ✓ Data Quality is of the utmost importance to data analytics
- ✓ People like to think they know and understand their business
  
- ✓ Good quality data is the backbone of any successful data analysis project
- ✓ Common data science term “garbage in > garbage out”
- ✓ If you input bad quality data, you can expect bad quality results
- ✓ Nothing more important when analyzing data than making sure you have good quality data
- ✓ Important to sense check the data yourself
  
- ✓ Data accuracy is above all else
- ✓ Garbage in>garbage out
- ✓ Waste of time to analyze data of poor quality
- ✓ Might mislead the business direction
  
- ✓ Integrity of the data you're using or providing is of the utmost importance
- ✓ Easy to get things confused if you're not paying attention to the details
- ✓ Expectation is that there will be lots of due diligence
- ✓ If you start off with bad data and then build on that, you can waste time, energy, and effort and sometimes lose trust

What about the importance of data privacy as it relates to data analysis?

- ✓ Data privacy is incredibly important especially when working in industries such as pharmaceuticals or healthcare
  - ✓ Ensure that users are getting appropriate levels of data based on their roles and permissions
  - ✓ Who has access to what?
- 
- ✓ Data privacy is a huge thing right now
  - ✓ What we do to protect PII is to encrypt the data when we are sending it back and forth
  - ✓ We call customers to share a specific sequence with them to protect data
  - ✓ We want to make sure this kind of data is always safe and protected

## Reading: Summary and Highlights

In this lesson, you have learned the following information:

The Five Traits of Data Quality:

Accuracy

Completeness

Reliability

Relevance

Timeliness

Importing Text:

You can use the ‘Text Import Wizard’ to import data from other formats, such as plain text, or comma-separated value files.

The Three Fundamentals of Data Privacy:

Confidentiality      Collection and Use      Compliance

## Module 3: Practice Quiz

 [Bookmark this page](#)

### Question 1

1/1 point (ungraded)

What is the first and most significant part of data quality?

- Relevance
- Accuracy
- Completeness
- Reliability



### Question 2

1/1 point (ungraded)

What is a delimited file?

- file with data fields separated by fixed-widths
- file with data fields separated by characters like commas or tabs
- file with text qualifiers between data fields
- file with no spaces between data fields



## Question 3

1/1 point (ungraded)

What feature do you use to import data from a text file?

- Function Wizard
- Text Import Wizard
- Formula Wizard
- Chart Wizard



## Question 4

1/1 point (ungraded)

What are the fundamentals of data privacy? Select all that apply.

- Confidentiality
- Collection and Use
- Compliance
- Completeness



## Module 3: Graded Quiz

 Bookmark this page

Graded Quiz due Jul 25, 2022 02:03 +08

### Question 1

1/1 point (graded)

Which data quality trait refers to the availability and accessibility of the data?

Timeliness

Relevance

Completeness

Accuracy



### Question 2

1/1 point (graded)

After importing a text file into Excel, you find some columns aren't showing all data. How can you fix all column widths at the same time?

drag a divider across

click the Format button

select all columns and double-click one of the selected column dividers

shorten the text so it fits



### Question 3

1/1 point (graded)

What are different types of personal data? Select all that apply.

PI - Personal Information

PII - Personally Identifiable Information

SPI - Sensitive Personal Information

SSN - Social Security Number



## **Module 4 - Cleaning Data**

---

-  [Module Introduction and Learning Objectives](#)
-  [Video: Removing Duplicated or Inaccurate Data and Empty Rows](#)
-  [Video: Dealing with Inconsistencies in Data](#)
-  [Video: More Excel Features for Cleaning Data](#)
-  [Video: Viewpoints: Issues with Data Quality](#)
-  [Hands-on Lab 5: Cleaning Data](#)
-  [Reading: Summary and Highlights](#)
-  [Module 4: Practice Quiz](#)
-  [Module 4: Graded Quiz \(5 Questions\)](#)  
Graded Quiz due Jul 28, 2022, 12:03 AM GMT+8

## **Module Introduction**

In this module, you will learn how to remove duplicate and inaccurate data, and how to remove empty rows in your data. You will then learn how to deal with inconsistencies in your data and how to use the Flash Fill and Text to Columns features to help you manipulate and standardize your data.

## **Learning Objectives**

After completing this module, you will be able to:

- Remove duplicate and inaccurate data, and empty rows.
- Handle inconsistencies in data.
- Use the Flash Fill and Text to Columns features in Excel.

# Removing Duplicated or Inaccurate Data and Empty Rows

Now that we have learned about the importance of data quality and data privacy, in this video we will learn how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data.

It's very common when collecting or importing data - whether through manual or automated processes - to get errors and inconsistencies in your data. This can be as simple as spelling mistakes, extra white space, or the wrong case used in text, to empty rows or missing values in your data, to inaccurate or duplicated data. Having these errors and inconsistencies in your data can lead to issues with formulas not working, with unsuccessful sorting and filtering operations and therefore inadequately visualized and presented data findings. These data errors and inconsistencies require you to carry out some form of data-cleaning routine to improve the quality and usability of the data.

## Inconsistencies

Let's start off with one of the easier of those tasks, which is **1. spell checking**.

In Excel, this works in pretty much the same way as you may have already encountered in applications such as Microsoft Word or other common word processing applications. I have some data here relating to the sales of toy vehicles, and the first thing we need to do is select what data we wish to check for spelling; in this case we will try column K which contains the product line data. Then we click Spelling which is on the Review tab.

Well that seems to be OK, so let's try the Country information in column T. So, we do have an error here, where a country name has been misspelt, or more likely, mistyped. We just click Change if we are happy with the spelling suggestion, or we could choose another suggestion from the list, or even ignore this error if we know the data is correct, but in this case we will change it. Here's another typo for a country name and here's one more. So, that seems to be all the errors in this column, let's try the final column now which is the deal size in column X. Here is a misspelling of the word small and another for medium. And that seems to be all for this column.

The next inconsistency we will look for is **2. empty rows**.

Empty rows in your data can cause lots of issues relating to moving around your data, working with formulas, and sorting and filtering. Therefore, it's very important to remove them from your data.

If you remember from an earlier lesson, when we click CTRL+DOWN ARROW, it should take us to the end of that column of data, but notice if we do that in this dataset, the cursor keeps stopping when it gets to an empty row meaning that the dataset is essentially being split into multiple sections, separated by these empty rows. That's not good, so let's resolve that now.

We have a couple of options; one option is 2.1 to just manually scroll down the sheet looking for empty rows and deleting each one, which is easy enough, and fine to do if you only have a small amount of data, but imagine if you were dealing with hundreds, or thousands, or even tens of thousands of rows? That would be a very laborious and time-consuming process.

There is a much better way - which involves **2.2** selecting all our data first, either using the mouse, or the CTRL+SHIFT+END keyboard shortcut. Then we select the Filter icon on the Data tab. We can now see that each column has a filter icon next to the column header. If we then select the Customer Name column-filter in column M then uncheck Select All then scroll down to the bottom of the list, we can check the item called Blanks, and then click OK.

This will now show only the empty rows at the top of our sheet; this can be quite hard to see, but if you look in the row numbers, you can see that rows 28,29,65,73,74,75 and 117 are listed at the top and are highlighted in blue text. We can now select these rows, either using the mouse or going to the first cell in the first data row, which is A28, and then using the CTRL+SHIFT+END keyboard shortcut then delete the offending empty rows. We then need to clear the filter and turn it off, so we can view our data again. Now, if we go back to the first row in the top of the datasheet and try the CTRL+DOWN shortcut again, to go to the end of the data column, it will work.

The next inconsistency we'll look for is **3. duplicated rows of data**; it's quite common for duplicate data rows to exist in your imported data, caused either by human input error, or an error in the import process.

There are two ways of doing this in Excel; the **3.1** first way includes reviewing the data you plan to remove first, before deleting it, to ensure you are deleting the right data. This is our preferred method as it provides an additional level of data security.

The **3.2 second method**, which we will also show you, is simpler, as you don't review the data to be removed first, but it lacks the security of the first method.

It's important to select a column of data that you would NOT expect to have duplicate values in. For example, if we consider the Price Each column, which is C, we would expect lots of these values to be repeated, because the unit price of some products is the same, so this is a bad example of a column to use to find duplicates.

Instead, let's use the Sales column in column E, because it is far less likely that these values will be duplicated in the normal process of things, as they are the total sales for each order. So, we select the column...and choose Conditional Formatting, then Highlight Cells Rules, and then Duplicate Values.

When we click OK, and scroll down the sheet, we can see that only a few values have been identified as being duplicates. There seem to be duplicate values in rows 36 to 40 and in rows 74 to 78

**3.1** Let's zoom out so we can see both duplicate sections together. It seems like these are in fact exact duplicate entries, and are likely to be an input error.

Let's delete the second section of duplicate rows as they are out of sequence; as they relate to Motorcycles sales and are in the Ships section of the sheet. So that was the first, and recommended method of removing duplicate rows of data, which previews the data to be removed first.

**3.2** Now, let's try the second, simpler, but less secure method. Let's go back to 100% zoom, and go back to the top of the worksheet. This time, we select the whole datasheet, and on the Data tab, we use the Remove

Duplicates button. We then unselect all the columns, then only select the Sales column. And the duplicate rows are deleted.

The last cleaning process we'll look at in this video is using the **4. Find & Replace** feature to repair some misspelt surnames in the customer contacts column. Find and Replace tools are under Find & Select on the Home tab in Excel, and if you have used other Office products such as Word, it should be familiar to you already.

We've had an email from a Swedish customer, informing us that we have his surname spelt incorrectly on his order sheets. So, we type the misspelt surname in to the 'Find what' box and click Find Next, then click it again to see there are multiple incorrect entries. If we click Find All, all instances are listed, and we can open the Replace tab to enter a name to replace the incorrect spellings. His surname should be Larsson with a double 's', so we'll replace all instances with that corrected spelling. That looks better, and we are finished.

In this video, we learned how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data.

In the next video, we will look at changing the case of text, fixing date formatting errors, and trimming whitespace from data.

# Removing Duplicated or Inaccurate Data and Empty Rows

## 1. Spell Checking

The screenshot shows a Microsoft Excel spreadsheet titled "PRODUCTLINE" with data from rows 1 to 4. The "Review" tab is selected in the ribbon. A spell check dialog box is open, showing the message "Spell check complete. You're good to go!" with an "OK" button. The status bar at the bottom right of the Excel window also displays "OK".

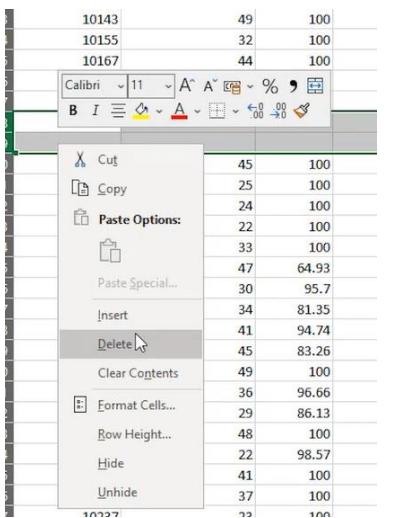
K1	A	B	C	D	E	F	G	H	I	J	K
1	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	PRODUCTLINE
2	10103	26	100	11	5404.62	29/01/2003	Shipped	1	1	2003	Classic Cars
3	10112	29	100	1	7209.11	24/03/2003	Shipped	1	3	2003	Classic Cars
4	10183	23	100	8	5372.57	13/11/2003	Shipped	4	11	2003	Classic Cars

Review -> Spelling -> (or) Change



## 2. Removing Empty Rows

### 2.1 for small data



Manually scroll down and look for empty rows and delete

### 2.2 for big data

#### 2.2.1

A screenshot of Microsoft Excel showing a large dataset. A blue box highlights the entire sheet area with the text "Select Whole Datasheet" and "CTRL+SHIFT+END". The table contains several rows of data, such as "2003 Vintage Cars", "S18\_1342", "Vitachrome Inc.", etc.

	Customer Name	Order ID	Shipper
1	2003 Vintage Cars	S18_1342	Vitachrome Inc.
5	2003 Vintage Cars	S18_1342	Australian Collectors, Co.
4	2004 Vintage Cars	S18_1342	Euro Shopping Channel
6	2004 Vintage Cars	S18_1342	Danish Wholesale Imports

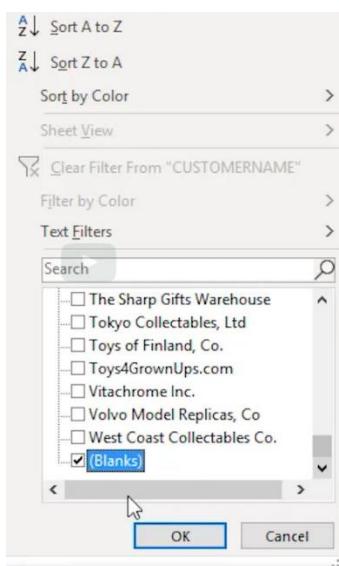
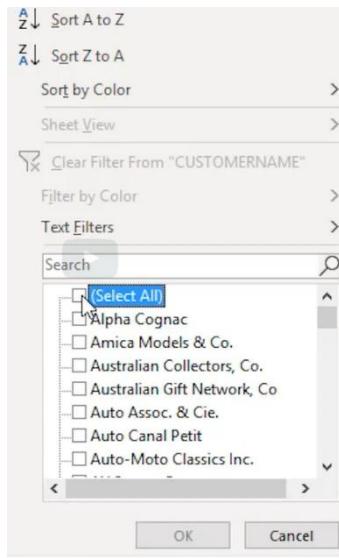
2.2.2 Data->Filter->Click down arrow-> Uncheck select all -> check “blanks” -> check ok -> select rows in blue color by selecting whole datasheet (ctrl+shft+end)-> ok -> Filter and turned it off -> then check

A screenshot of Microsoft Excel showing the ribbon with the 'Data' tab selected. Below the ribbon is a table with columns: ORDERLINENUMBER, SALES, ORDERDATE, STATUS, QTR, MONTH, and YEAR. To the right of the table are filter icons for each column.

ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR	MONTH	YEAR
11	5404.62	29/01/2003	Shipped	1	1	2003
1	7209.11	24/03/2003	Shipped	1	3	2003
8	5372.57	13/11/2003	Shipped	4	11	2003

A screenshot of Microsoft Excel showing a list of customer names. The first item is "CUSTOMERNAME" followed by "Baane Mini Imports", "Volvo Model Replicas, Co", "Classic Gift Ideas, Inc", and "Saveley & Henriot, Co.". A dropdown arrow is shown next to the last item.

CUSTOMERNAME
Baane Mini Imports
Volvo Model Replicas, Co
Classic Gift Ideas, Inc
Saveley & Henriot, Co.



A	ORDERNUMB
28	
29	
65	
73	
74	
75	
117	

A	B	C	D	E	F	G	H	I	J	K	L	
1	ORDERNUMB	QUANTITYORDERD	PRICEEA	ORDERLINENUMB	SALES	ORDERDATE	STATUS	QTR	MONTH	YEAR	PRODUCTLINE	PRODUCTCOI
28												
29												
65												
73												
74												
75												
117												
17												
28												
29												
31												
32												
33												
34												
35												
36												
37												
38												
39												
40												
41	Select Whole Datasheet CTRL+SHIFT+END											
42												
43												
44												
45												



H	I	J	K
R	MONTH	YEAR	PRODUCTL
1	1	2003	Classic Car
1	3	2003	Classic Car

### 3. Duplicate Data Rows

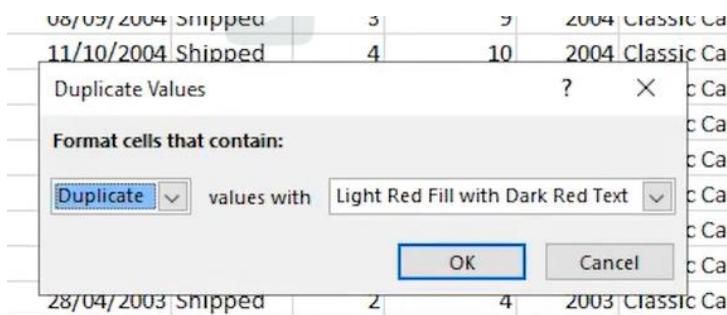
Select column which you would NOT expect to have duplicate

3.1 Choose column -> conditional formatting-> Highlight cells rules-> Duplicate values-> click ok-> check then delete

The screenshot shows the Microsoft Excel ribbon with the 'Conditional Formatting' tab selected. A context menu is open over a table, with the 'Highlight Cells Rules' option being the active choice. Within this menu, the 'Duplicate Values...' option is highlighted.

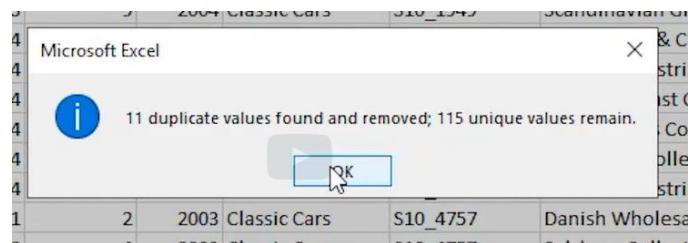
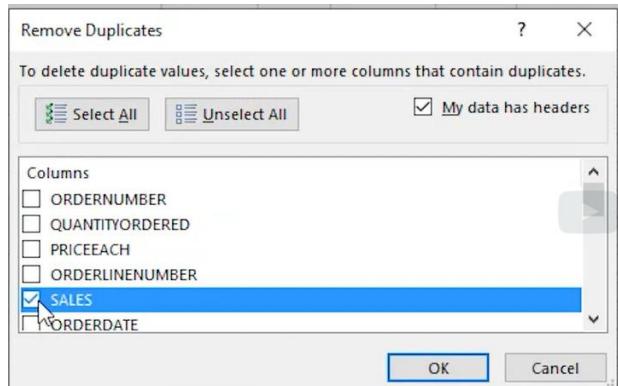
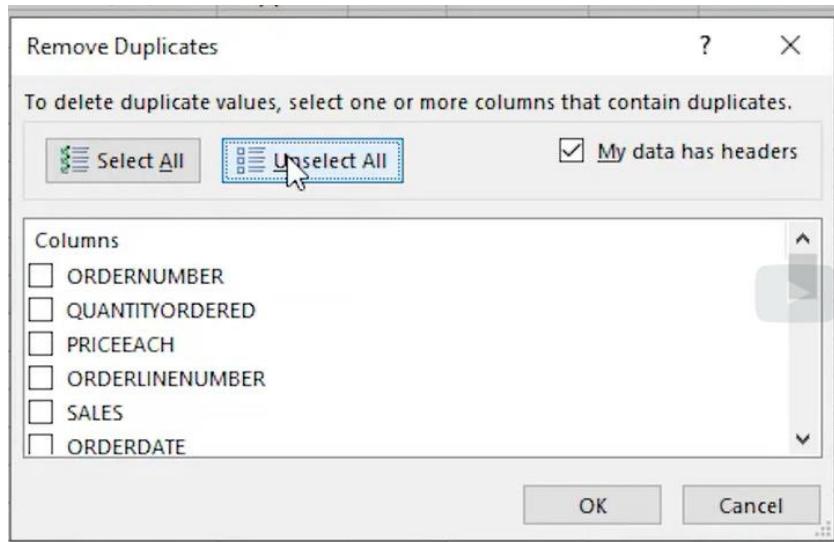
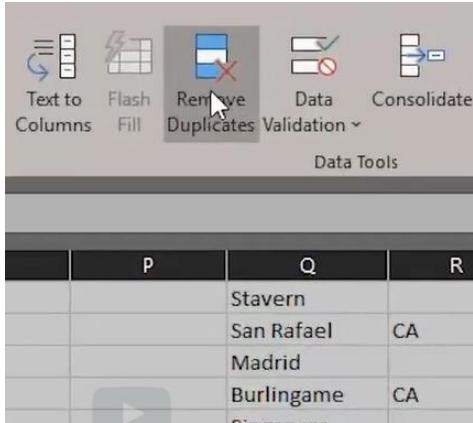
NUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID
11	5404.62	29/01/2003	Shipped	1	
1	7209.11	24/03/2003	Shipped	1	
8	5372.57	13/11/2003	Shipped	4	
11	7290.36	25/11/2003	Shipped	4	
6	9064.89	05/12/2003	Shipped	4	
3	6075.3	29/01/2004	Shipped	1	
2	6463.23	10/03/2004	Shipped	1	
9	6120.34	04/05/2004	Shipped	2	
6	7680.64	15/06/2004	Shipped	2	
9	4905.39	19/07/2004	Shipped	3	
2	8014.82	17/08/2004	Shipped	3	
11	7136.19	08/09/2004	Shipped	3	

So, we select the column...and choose Conditional Formatting, then Highlight Cells Rules, and then Duplicate Values.



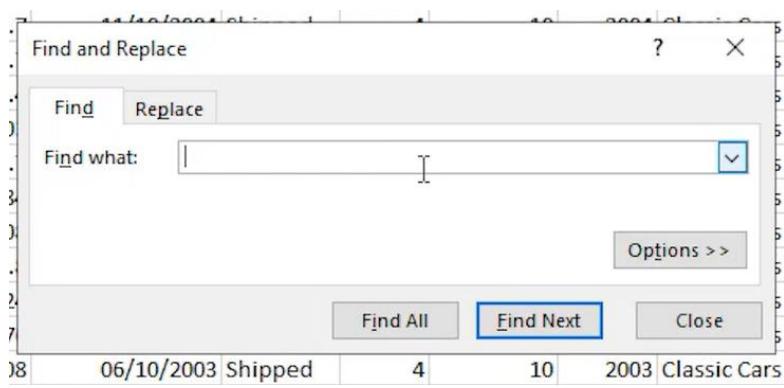
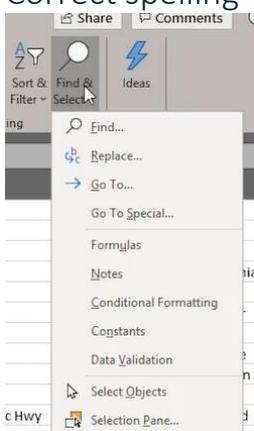
D	E	F	G
2	2871	24/02/2003	Shipped
5	2765.9	07/05/2003	Shipped
2	3884.34	01/07/2003	Shipped
6	3746.7	25/08/2003	Shipped
14	5205.27	10/10/2003	Shipped
1	3479.76	28/10/2003	Shipped
9	2497.77	11/11/2003	Shipped
1	5512.32	18/11/2003	Shipped
2	2168.54	01/12/2003	Shipped
14	4708.44	15/01/2004	Shipped

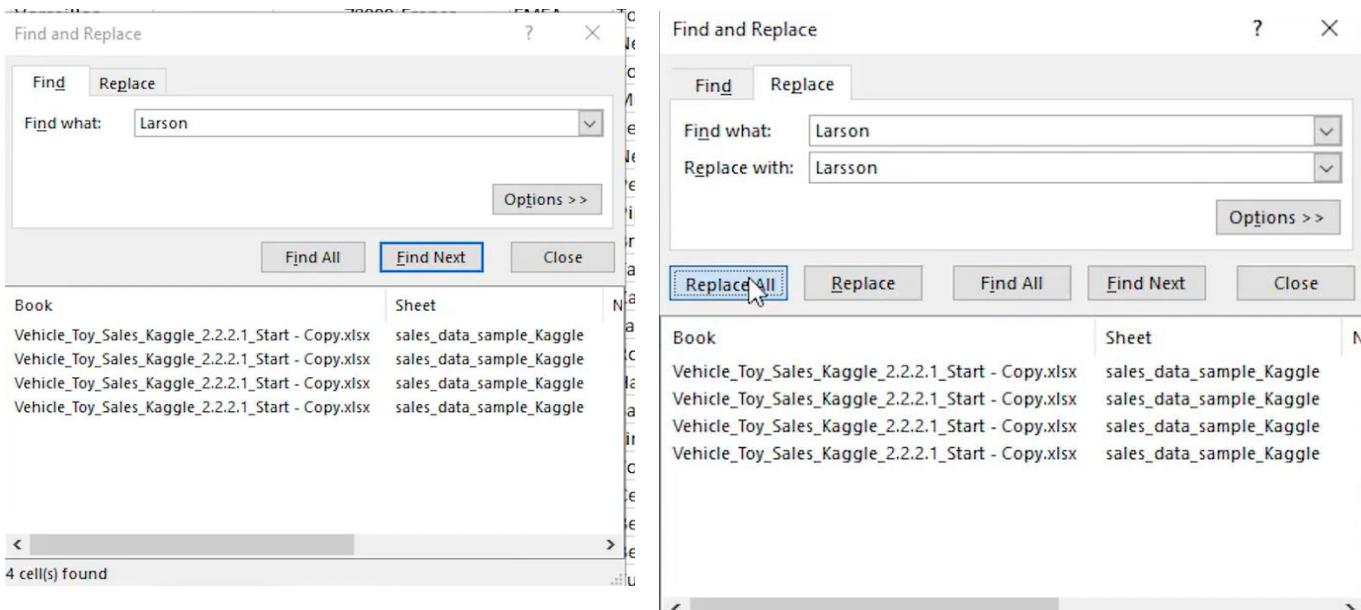
Select the whole worksheet then...



## 4. Find and Replace (extreme upper right)

Select Find & Replace-> type name to find -> click find next (2x) -> Find all -> Replace->type  
Correct spelling -> replace all





We've had an email from a Swedish customer, informing us that we have his surname spelt incorrectly on his order sheets. So, we type the misspelt surname in to the 'Find what' box and click Find Next, then click it again to see there are multiple incorrect entries.

If we click Find All, all instances are listed, and we can open the Replace tab to enter a name to replace the incorrect spellings. His surname should be Larsson with a double 's', so we'll replace all instances with that corrected spelling.

# Dealing with Inconsistencies in Data

Now that we've learned how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data, in this video we'll look at changing the case of text, fixing date formatting errors, and trimming whitespace from data.

When you collect or receive data from varying sources, it's quite common to find that your data contains text in mixed case; that is, some in uppercase, some in lowercase and some in capitalized proper case (also known as sentence case). Some of this may be intentional; but often it's not.

Excel doesn't have a Change Case button like there is in Microsoft Word, so you need to use other methods to perform this data cleaning task.

Those methods are functions; namely the UPPER, LOWER, and PROPER functions. You can use these functions to help you change the case of text in your data. You can see that the header row here is using all uppercase characters,

## Proper Case

so if you want to change that to use proper case then you need to

1. add another row to put the function in; this is referred to as a 'helper' row. The PROPER function is simple to use;
2. just type equals, then PROPER, then open parenthesis, then the cell reference - In this case A1 - then close parenthesis, and press Enter. Here you can see that the result in A2 is in proper case.
3. Now you can try and drag the formula right across to column X by using the Fill Handle on A2... but this can be very tricky when you have a lot of columns, so let's try another way.
4. Instead of dragging, you can use SHIFT+RIGHT ARROW to select the columns across to X first then press F2 to bring the cursor into focus in cell A2 then you hold down the CTRL key while you press Enter, and it fills across for you.

You might think that you could now remove the original row; but look what happens when you do; you get a REF error because the formula is referencing an invalid reference, and the header row cells now contain just the failed formula rather than the actual header text. So, you need to undo that, and instead, you

5. copy the contents of the helper row to row 1 but when you paste you need to choose the Paste Values option.

Now the header row cells just contain header text, and you can

6. remove the helper row in row 2.

A	B	C	
1	ORDERNUMBER	QUANTITYORDERED	PRICEEACH
2			
3	10103	26	100

ORDERNUMBER	QUANTITYORDERED	PRICEEACH
=PROPER(A1)		
10103	26	100
10112	29	100
10183	23	100

ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES
Ordernumber				
10103	26	100	11	5404.62
10112	29	100	1	7209.11
10119				
SHIFT + →				
10105	50	100		
10119	46	100		
	33	100		
	49	100		
	32	100		
	44	100		
	24	100		
10186	26	100		
10147	25	100		

ORDERNUMBER	QUANTITYORDERED	PRICEEACH
=PROPER(A1)		
10103	26	100
10112	29	100
10183	23	100

10105	50	100
10119	46	100
	33	100
	49	100
	32	100
	44	100
	24	100
10186	26	100
10147	25	100

ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES
Ordernumber	Quantityordered	Priceeach	Orderlinenumber	Sales
10103	26	100	11	5404.62
10119	46	100		
10186	26	100		

Ordernumber	Quantityordered	Priceeach
10103	26	100
101	Paste Options:	100
101		100
101		100
10206	47	100
10206		100

## Upper Case

Let's now use the UPPER function to change text from proper case to upper case.

1. Insert a column to the right of the column you want to change. This will be a 'helper' column.
2. Then you type the formula containing the UPPER function in the first data cell in this new helper column.
3. Again, it's a simple formula; you type equals, then UPPER, then open parenthesis, then the cell reference – in this case T2 – and then close parenthesis and press Enter
4. You can see the result is the country name in upper case, and you can then copy that formula down the rest of the column by double-clicking the Fill Handle cross symbol.
5. Copy and paste the contents of the helper column to the original column, but use the Paste Values option.
6. Now you can delete the helper column

The screenshot shows a step-by-step process in Microsoft Excel:

- Step 1:** A table with columns T and U. Column T contains the country names: Norway, Sweden, USA, France, Canada, USA. Column U is empty.
- Step 2:** The formula `=UPPER(T2)` is entered into cell U2. The cell reference T2 is highlighted in blue, indicating it is being used as a relative reference for the formula.
- Step 3:** The formula is copied down the column, resulting in the following table:

Country	Country
Norway	NORWAY
Sweden	SWEDEN
USA	USA
France	FRANCE
Canada	CANADA
USA	USA
- Step 4:** The user right-clicks on the cell containing "NORWAY" in column U. A context menu appears with options: Cut, Copy, Paste Options, Paste Special..., and Delete. The "Paste Options" button is highlighted with a cursor.
- Step 5:** The "Paste Options" menu is open, showing various options for pasting. The "Values" icon (a clipboard with a checkmark) is selected.
- Step 6:** The "Values" icon is highlighted in the "Paste Options" menu, indicating it is the chosen option for pasting the converted values back into column T.

## Lower Case

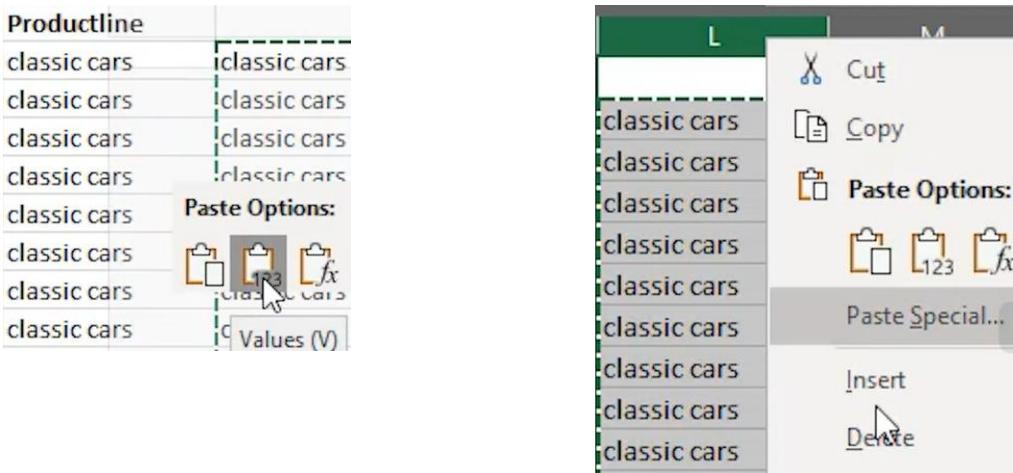
Next, we'll use the LOWER function to change text from proper case to lower case.

1. Insert a column to the right of the column you want to change. This will be another 'helper' column.
2. Type the formula containing the LOWER function in the first data cell in the helper column.
3. Type equals, then LOWER, then open parenthesis, then the cell reference – in this case K2 – and then close parenthesis, and press Enter.
4. See the result is the product line data in lower case,
5. Copy the formula down to the rest of the column by double-clicking the Fill Handle once more.
6. Copy and paste the contents of the helper column to the original column, but ensuring you use the Paste Values option.
7. Delete the helper column

It's quite common to receive data that has a mixture of date formats, or that uses a date format that isn't suitable to your region.

The screenshot shows a Microsoft Excel interface with three main parts:

- Left Panel:** A table with columns K and L. Column K contains the text "Productline", "Classic Cars", "Classic Cars", "Classic Cars", "Classic Cars", and "Classic Cars". Column L is empty.
- Middle Panel:** A close-up view of the first two rows. Cell L2 contains the formula `=LOWER(K2)`. The formula bar above also shows `=LOWER(K2)`. The fill handle is visible at the bottom-right corner of cell L2, indicating it can be copied down.
- Bottom Panel:** A screenshot of the entire Excel window showing the results. Column K now contains "classic cars" and column L contains "classic cars" with a small green plus sign icon indicating the data is pasted as values.



## Change Format of Dates

Now let's look at how to change the format of some dates. You can see that this date format is currently using a 2-digit day, a 2-digit month, and a 4-digit year value.

1. Open the Number format dialog box, you can see in the Locale box, that this is an English (United Kingdom) date format.

2. To use a US date format, so you first change the locale to English (United States). In this list, you can see there are several date options to choose from; let's choose one which uses the full month name, then a 2-digit day, and a 4-digit year value.

3. Copy this format to the rest of the date cells.

However, if you want to format these dates using your own custom format, you can do that too.

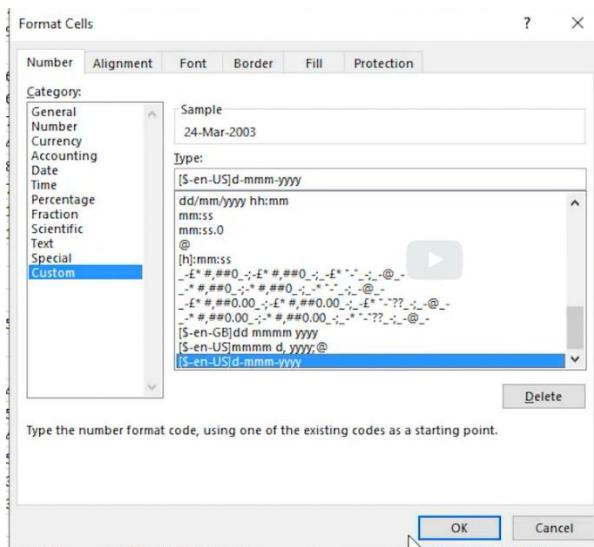
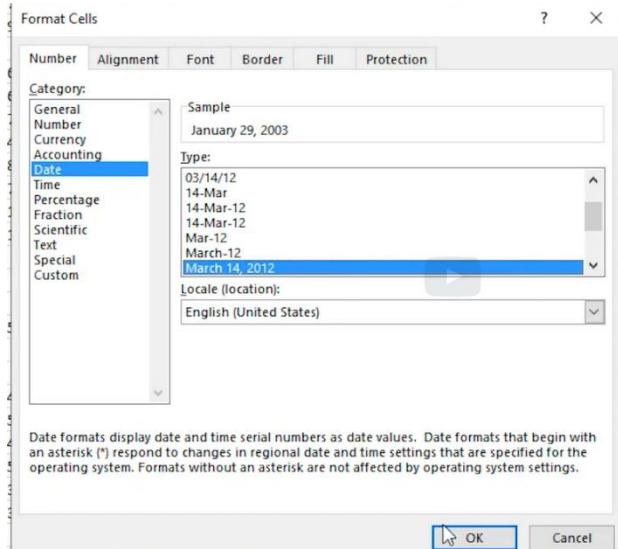
4. In the Number format list, select Custom

5. Choose an existing format that is similar to what you want

6. Modify it to create a new custom format; here we'll have the day, then 3-letter month, then 4-digit year.

7. To apply that new custom date format to the rest of the column you could either use the Format Painter tool, or you can select the rest of the column

8. Choose the new custom format from the Custom list in the Number Format dialog box.



## Whitespace or Unwanted Spaces

You might find that your data has some whitespace; that is, unwanted spaces in your data. Here you can see that we have some spaces at the start some spaces at the end and some unwanted double spaces in the middle of our data.

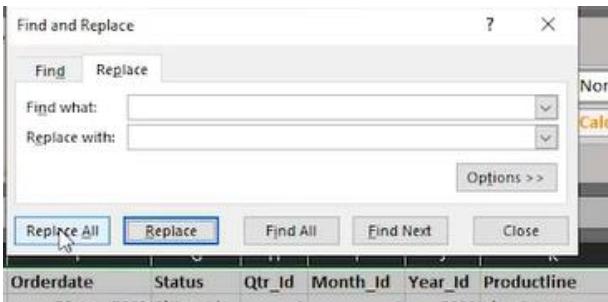
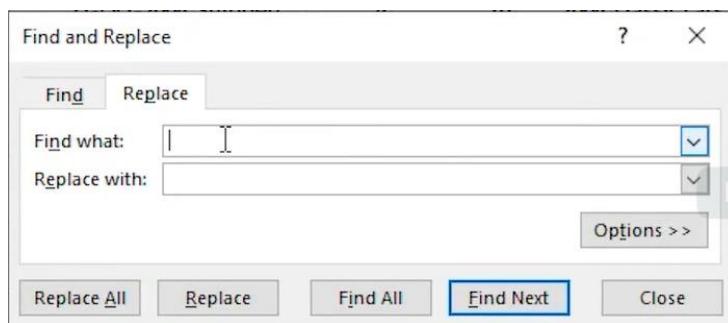
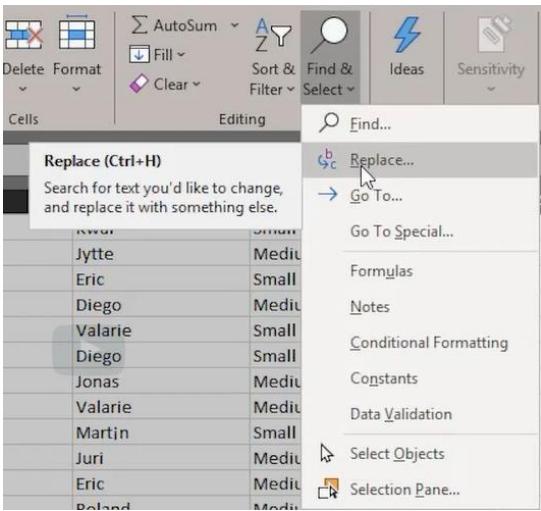
We'll first have a look at what you can do to clean up these unwanted spaces in your data

1. by using the Find & Replace feature in Excel.
2. First select all the data then on the Home tab,
3. Click Find & Select, then Replace.
4. To get rid of double spaces, you enter a double space in the 'Find what' box, and a single space in the 'Replace with' box.
5. Click Find Next. And choose Replace for each item you want to change. You could click Replace All to do all the fixes in one go but unless you are absolutely sure of the changes, it's better practice to check and replace each one in sequence in case there are some valid reasons for these extra spaces.
6. If you have a very large dataset you might also choose Replace All to save you a lot of time.

So using the Find & Replace feature got rid of most of those unwanted whitespaces, but not all of them; we removed double spaces using that feature, but we also have some single spaces left at the start and end of some of the cells. You can't use Find & Replace to remove single spaces otherwise you would lose ALL spaces in your data - including standard spaces between words - which you don't want to remove. But, there is another tool you can use to clear spaces from cells, and that's the

Australian Collectors, Co.
Mini Gifts Distributors Ltd.
Danish Wholesale Imports
Salzburg Collectables
Stylish Desk Decors, Co.
Mini Creations Ltd.
Toys of Finland, Co.
Scandinavian Gift Ideas
Alpha Cognac
Double Decker Gift Stores. Ltd

Enaco Distributors
Royal Canadian Collectables, Ltd.
Gifts4AllAges.com
Classic Gift Ideas, Inc
giftsbymail.co.uk



## Remove Single Spaces at the Start and End of Some Cells

TRIM function.

To use the TRIM function, you once again 1. Insert a helper column.

2. Type equals, then TRIM, then open parenthesis, then the cell reference – in this case M2 – then close parenthesis, and press Enter.
3. Double-click the Fill Handle symbol to copy this formula down to the remainder of the column.
4. Copy the contents of the new column N to column M, and remember once again to paste using the Paste Values option.

You can now see that those erroneous spaces have been removed, or more accurately speaking, have been trimmed. And lastly, you can

## 5. Remove the helper column.

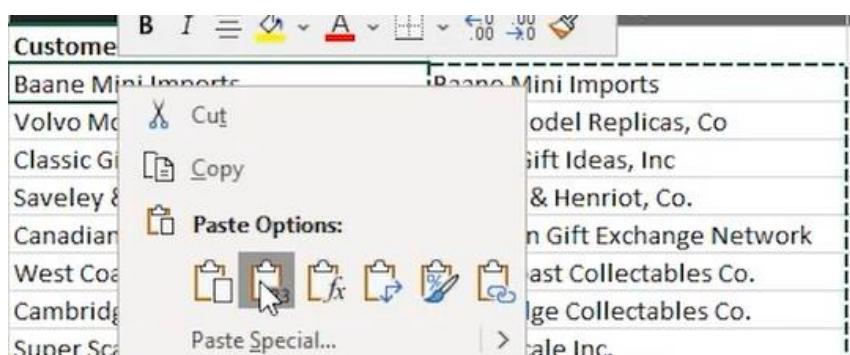
In this video, we learned how to change the case of text, how to change date formatting, and how to trim whitespace from data. In the next video, we will discuss how to use the Flash Fill and Text to Columns features in Excel to help clean data.

M	N
Customername	
Baane Mini Imports	
Volvo Model Replicas, Co	
Classic Gift Ideas, Inc	

Customername	
Baane Mini Imports	=TRIM(M)
Volvo Model Replicas, Co	TRIM(text)

Customername	
Baane Mini Imports	Baane Mini Imports

M	N
Customername	
Baane Mini Imports	Baane Mini Imports
Volvo Model Replicas, Co	Volvo Model Replicas, Co
Classic Gift Ideas, Inc	Classic Gift Ideas, Inc
Saveley & Henriot, Co.	Saveley & Henriot, Co.
Canadian Gift Exchange Network	Canadian Gift Exchange Network



# More Excel Features for Cleaning Data

Now that we've learned how to change the case of text, how to change date formatting, and how to trim whitespace from data, in this video we'll discuss how to use the Flash

Fill and Text to Columns features in Excel to help clean data. We used Flash Fill briefly earlier in the course as a quick method of entering data that fits a specific pattern, such as the names of months or days of the week, but it can also be useful as a data cleaning tool.

## Flash Fill

It can split a column of full names into two separate columns for the forename and surname, and it can also help to modify the naming convention used in a column of names. For example, in the vehicle toy sales worksheet there is a column containing the last names of contacts, and another containing their first names.

If you want to use the Flash Fill feature to combine these names into one name column,

you first 1.

1. insert a helper column; let's call it 'Contactname'.
2. in the first row in the new column you enter the full name of the first contact in the format of your choice; for example you might want surname, then a comma, then the forename, or you might want surname, and just an initial, and so on; in this case let's just enter the name in the standard format of forename then surname with a space between them, and then we press Enter.
3. start typing the second contact's name in, and you'll see that Flash Fill displays a preview of the remaining names for you. If you're happy with what's in the preview, all you have to do is press Enter, and it fills in the remaining names for you right down the column.

It even works when there are two names in one of the columns such as Wing C here ... and Da Cunha here.

4. Now you can remove the original columns if you no longer need them. So, in the previous task we saw how to combine two columns of data into one column using Flash Fill; now let's see how to use it to modify the naming convention in a column. Let's switch to the customer contacts worksheet.

Then in the first data row of the next column, that is B2, 5. we type the name of the first contact using whatever naming convention we want. We'll use surname, then comma, then a space, then the forename, and press Enter. Again, 6. when we start typing the second contact's name in the next row down, that is B3, Flash Fill detects the pattern and fills in the remaining names in column B when we press Enter.

7. You could then copy and paste the column header, and delete the original column A.

What we couldn't do with Flash Fill was take a single column with two names in and split that into two separate columns.

## Insert a Helper Column

V	W	X
Contactlastname	Contactfirstname	
Bergulfsen	Jonas	
Berglund	Christina	
Cervantes	Francisca	
Saveley	Mary	
Tannamuri	Yoshi	
Thompson	Steve	

V	W	X
Contactlastname	Contactfirstname	Contactname
Bergulfsen	Jonas	Jonas Bergulfsen
Berglund	Christina	Christina Berglund
Cervantes	Francisca	Francisca Cervantes
Saveley	Mary	Mary Saveley
Tannamuri	Yoshi	Yoshi Tannamuri

V	W	X
Contactlastname	Contact	
Bergulfsen	Jonas	
Berglund	Christin	
Cervantes	Franciso	
Saveley	Mary	
Tannamuri	Yoshi	
Thompson	Steve	
Tseng	Kyung	
Murphy	Leslie	
Shimamura	Akiko	
Huxley	Adrian	
Accorti	Paolo	
Larsson	Maria	
Tonini	Daniel	
Nelson	Valarie	
Young	Valarie	
Murphy	Julie	
Ferguson	Peter	Peter Ferguson
Nelson	Valarie	Valarie Nelson

The screenshot shows a context menu open over the 'Contact' column of the second table. The 'Delete' option is highlighted with a gray background and white text. Other options visible include Cut, Copy, Paste Options, Insert, Delete, Clear Contents, Format Cells, Column Width, Hide, and Unhide.

## Modify Naming Convention

The screenshot shows the Excel ribbon with the 'Window' tab selected. A dropdown menu is open, listing two windows: '1 Vehicle\_Toy\_Sales\_Kaggle\_2.2.2.3\_Start.xlsx' and '2 Customer\_Contacts\_2.2.2.3\_Start - Copy.xlsx'. The second window is currently active. Below the ribbon, a portion of a table is visible with columns Q, R, S, T, U, and V. The columns are labeled State, Postalcode, Country, Territory, and Contactname.

Q	R	S	T	U	V
	State	Postalcode	Country	Territory	Contactname
I		4110	NORWAY	EMEA	Jonas Bergulfse
		S-958 22	SWEDEN	EMEA	Christina Berglu

A	B	C
Customer Contact		
Mary Saveley	Saveley, Mary	
Yoshi Tannamuri	Tannamuri, Yoshi	
Daniel Tonini	Tonini, Da	
Valarie Nelson	Nelson, V	

A	B	C
Customer Contact		
Mary Saveley	Saveley, Mary	
Yoshi Tannamuri	Tannamuri, Yoshi	
Daniel Tonini	Tonini, Da	
Valarie Nelson	Nelson, Valarie	
Leslie Murphy	Murphy, Leslie	

A	B	C
Customer Contact		
Saveley, Mary		
Tannamuri, Yoshi		
Tonini, Daniel		
Nelson, Valarie		

We need to use the ‘Text to Columns’ feature to do that. So, we’ll close this worksheet and we won’t save the changes. Now, let’s see how the ‘Text to Columns’ feature can help with data cleaning too.

As the name suggests, and unlike Flash Fill, the ‘Text to Columns’ feature can take a column containing multi-part text and split that text into one or more other columns. This can be useful for splitting any multi-part text, such as names or addresses, into separate component parts.

Let’s open the customer contacts worksheet again.

Then we’ll 1. add column headings for the next two columns, and

2. copy the cell format used in the first column header. Then we’ll widen the columns.

If we then 3. select the data in column A from A2 to A23, and on the Data tab, click Text to Columns, a wizard is launched.

On the first page of the wizard, ensure that 4. ‘Delimited’ is selected.

On the second page, 5. ensure that only ‘Space’ is selected as the delimiter.

6. On the third page of the wizard, click the little arrow next to ‘Destination’...

7. Select cell B2 on the worksheet, then click the little arrow again to return to the wizard.

We’re now finished with this wizard. You can see that the full customer contact names in column A have now been successfully split into two new columns in B and C, and you could now remove column A if you no longer

need it. We'll close this worksheet, and again we won't save the changes. You can also achieve the same result using functions.

**Convert Text to Columns Wizard - Step 1 of 3**

The Text Wizard has determined that your data is Delimited.  
If this is correct, choose Next, or choose the data type that best describes your data.

Original data type  
Choose the file type that best describes your data:  
 Delimited - Characters such as commas or tabs separate each field.  
 Fixed width - Fields are aligned in columns with spaces between each field.

Preview of selected data:

2 Mary Saveley	
3 Yoshi Tannamuri	
4 Daniel Tonini	
5 Valarie Nelson	
6 Leslie Murphy	

Cancel < Back Next > Finish

**Convert Text to Columns Wizard - Step 2 of 3**

This screen lets you set the delimiters your data contains. You can see how your text is affected in the preview below.

Delimiters  
 Tab  
 Semicolon  
 Comma  
 Space  
 Other:   
 Treat consecutive delimiters as one

Data preview

Mary	Saveley
Yoshi	Tannamuri
Daniel	Tonini
Valarie	Nelson
Leslie	Murphy

Cancel < Back Next > Finish

**Convert Text to Columns Wizard - Step 3 of 3**

This screen lets you select each column and set the Data Format.

Column data format  
 General  
 Text  
 Date: DMY  
 Do not import column (skip)  
 Advanced...

Destination: \$A\$2

Data preview

Mary	Saveley
Yoshi	Tannamuri
Daniel	Tonini
Valarie	Nelson
Leslie	Murphy

Cancel < Back Next > Finish

**Convert Text to Columns Wizard - Step 3 of 3**

=\$B\$2

	A	B	C
1	Customer Contact	Forename	Surname
2	Mary Saveley	Mary	Saveley
3	Yoshi Tannamuri	Yoshi	Tannamuri
4	Daniel Tonini	Daniel	Tonini
5	Valarie Nelson	Valarie	Nelson

## In Excel for the Web

This would be required if you were using 'Excel for the web', the online version of Excel, as this doesn't have the 'Text to Columns' feature.

There's also a bit more flexibility with functions, which can be especially useful if you have names that are complex and mixed, such as having hyphenated names or some names with a middle initial, some with two middle initials, and some with no middle initial.

So, we 1. Open the customer contacts worksheet again.

2.add column headings for the next two columns, and

3. copy the cell format used

in the first column header. Then we'll widen the columns.

4.enter the formula in B2 to extract the forename part of the name.

This formula extracts five characters from cell A2, starting from the left and including the space.

5. in cell C2 we enter the formula to extract the surname part of the name.

This formula extracts seven characters from cell A2, starting from the right.

6.Double-click the Fill Handle in cell B2 to use AutoFill to complete the column.

And we do the same to the Fill Handle in cell C2 to use AutoFill to complete that column also. In this video, we learned how to use the Flash Fill and Text to Columns features in Excel to help clean data.

	A	B	C	D	E
1	Customer Contact	Forename	Surname		
2	Mary Saveley	=LEFT(A2,SEARCH(" ",A2,1))			
3	Yoshi Tannamuri				
4	Daniel Tonini				
5	Valarie Nelson				

	A	B	C	D	E	F
1	Customer Contact	Forename	Surname			
2	Mary Saveley	Mary	=RIGHT(A2,LEN(A2)-SEARCH(" ",A2,1))			
3	Yoshi Tannamuri					

Double Click Fill Handle

	A	B	C	D	E	F
1	Customer Contact	Forename	Surname			
2	Mary Saveley	Mary	Saveley			
3	Yoshi Tannamuri	Yoshi	Tannamuri			
4	Daniel Tonini	Daniel	Tonini			

## Viewpoints: Issues with Data Quality

Can you tell us your experience with poor quality data and the cleaning of that data?

- ✓ A large portion of my time is spent cleaning, verifying, checking data before I run an analysis.
  - ✓ Most information captured is based on what someone has put in
  - ✓ Two people can have a similar situation and look at things slightly differently.
  - ✓ To ensure your results are accurate, always check the integrity of the information before you do your analysis
- 
- ✓ No data is going to be perfect
  - ✓ Data often has things missing or it's not in the correct format
  - ✓ Lots of data cleansing activities that can be undertaken
- 
- ✓ Is what I'm looking at directionally correct?
  - ✓ Am I looking at the right thing?
  - ✓ Are all of those costs relevant to the period I'm analyzing
- 
- ✓ Scrub the information to validate it is correct
  - ✓ If the data is incorrect or our period, adjustment need to be made
- 
- ✓ Can cause you not to be able to present your data reliably to ensure it's being collected correctly
  - ✓ Outline what transformations have been done to the data
- 
- ✓ Time spent fixing things could have been spent working on other deliverables
  - ✓ Data integrity can be called into question
  - ✓ Constantly redoing certain parts of data can be frustrating
  - ✓ Pay attention the details and minute



# Hands-on Lab 5: Cleaning Data

**Estimated time needed:** 45 minutes

In this lab, first you will learn how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data. Next, you will learn how to change the case of text, how to change date formatting, and how to trim whitespace from data. Finally, you will learn how to use the Flash Fill feature and functions in Excel to help clean data.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The dataset used in this lab comes from the following source: <https://dataplatform.cloud.ibm.com/exchange/public/entry/view/f8ccaf607372882403a37d9019b3abf4>. This dataset is published by **IBM**, and includes fictitious customer demographics and sales data.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

## Objectives

After completing this lab, you will be able to:

- Understand how to deal with irrelevant or inaccurate data

- Remove empty rows and duplicated data
- Change text case and date formatting
- Trim whitespaces from data
- Use Flash Fill and functions to clean data

# Exercise 1: Removing Duplicated, Irrelevant or Inaccurate Data

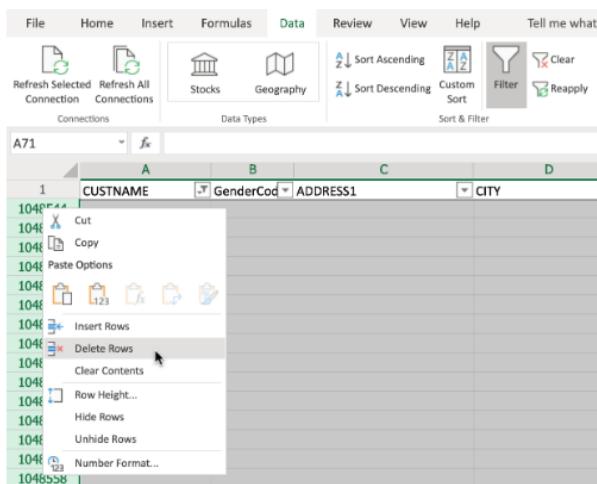
In this exercise, you will learn how to deal with inaccurate data, how to remove empty rows, and how to remove duplicated data.

## Task A: Check spelling

1. Download the file [\*\*Customer demographics and sales Lab5.xlsx\*\*](#). Upload and open it using Excel for the web.
2. Select column **L (CREDITCARD\_TYPE)**, then click **Review** tab, and select **Spelling**.
3. Click the correct suggestion to change the spelling.
  - o **Note:** Don't change 'jcb' spelling when doing the spell check. We will need 'jcb' for the Exercise 1 Task D.
4. Close the **Spelling** pane.

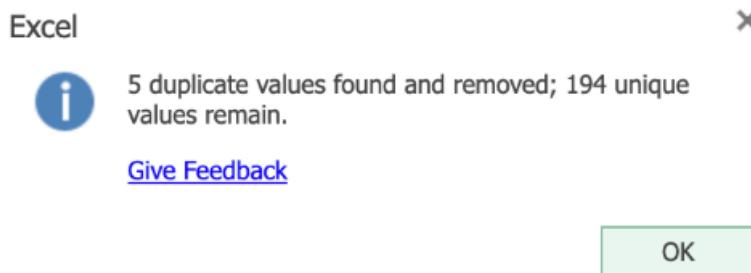
## Task B: Remove empty rows

1. Press **CTRL+HOME**, then press **CTRL+SHIFT+END** to select the whole datasheet.
2. On the **Data** tab, click **Filter**.
3. Press **CTRL+HOME**, click the **filter arrow** in the **CUST\_NAME** column, and then click **Filter**.
4. Click the **Select All** checkbox to deselect all of them. Then select just **Blanks**, then **OK**.
5. Select **first row**, then press **CTRL+SHIFT+END** to select all rows.
6. Right-click the selected rows and then click **Delete Rows**.
7. Finally, on the **Data** tab, click **Clear**, then click **Filter**.



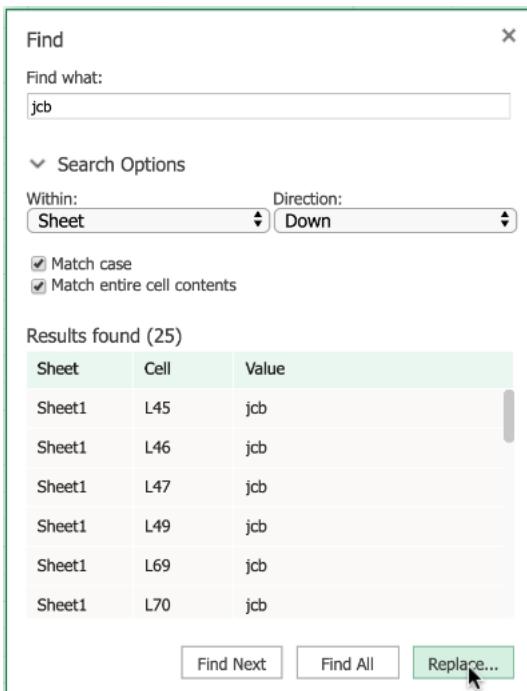
## Task C: Remove duplicate rows

1. Select Column **T (ORDER\_ID)** since ORDER\_ID values are unique.
2. On the **Home** tab, click **Conditional Formatting > Highlight Cells Rules > Duplicate Values**, and then click **OK**.
3. Select the whole datasheet (**CTRL+SHIFT+END**)
4. On the **Data** tab, click **Remove Duplicates**.
5. In the Remove Duplicates dialog box, ensure that **Select all columns** is checked and that **My data has headers** is also checked, then click **OK**.
6. In the pop-up box informing you how many duplicate values were found and removed, click **OK**.



## Task D: Use Find & Replace to correct misspelling

1. On the **Home** tab, click **Find & Select**.
2. Click **Find**. In Find what, type **jcb**, and click **Find All**.
3. Click **Replace**.
4. In Replace with, type **JCB**, click **Replace All**, and then click the **Close** icon.
5. On the **Home** tab, click **Conditional Formatting > Clear Rules > Clear Rules from Entire Sheet**.



# Exercise 2: Dealing with Inconsistencies in Data

In this exercise, you will learn how to change the case of text, how to change date formatting, and how to trim whitespace from data.

## Task A: Use the PROPER function to change text from upper case to proper case

1. Select row **2**, then right-click it and choose **Insert Rows**.
2. In cell **A2**, type **=PROPER(A1)** and press **Enter**.
3. Hover over the bottom-right corner of cell **A2**, and drag the **Fill Handle** across to the last column.
  - o If dragging across is too difficult with the mouse, then select the cells in the row **2** using **SHIFT+RIGHT ARROW**, then press **F2** to put the cursor focus back in cell **A2**, then hold **CTRL** while you press **Enter**.
4. Select row **2**, then press **CTRL+C**.
5. Select row **1**, Right-click and choose **Paste Options>Values**.
6. Select row **2**, right-click it and choose **Delete Rows**.

## Task B: Use the UPPER function to change text from proper case to upper case

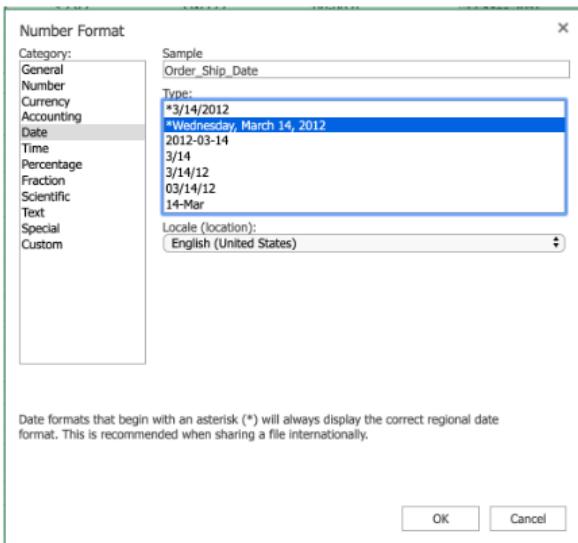
1. Select column **AG (Generation)**. Then right-click and choose **Insert Columns**. In cell **AG1**, type **Generation**.
2. In cell **AG2**, type **=UPPER(AH2)** and press **Enter**.
3. Hover over the bottom-right corner of cell **AG2** and double-click the **Fill Handle**.
4. Select column **AG**, then press **CTRL+C**.
5. Select column **AH**, right-click and choose **Paste Options>Values**.
6. Select column **AG**, right-click it and choose **Delete Columns**.

## Task C: Use the LOWER function to change text from proper case to lower case

1. Select column **AC (T\_Type)**. Then right-click and choose **Insert Columns**. In cell **AC1**, type **T\_Type**.
2. In cell **AC2**, type **=LOWER(AD2)** and press **Enter**.
3. Hover over the bottom-right corner of cell **AC2** and double-click the **Fill Handle**.
4. Select column **AC**, then press **CTRL+C**.
5. Select column **AD**, right-click and choose **Paste Options>Values**.
6. Select column **AC**, right-click it and choose **Delete Columns**.

## Task D: Change date formatting

1. Select column **Z (Order\_Ship\_Date)**.
2. On the **Home** tab, in the **Number** group click **Number Format> More Number Formats**.
3. In the Category list, select **Date**.
4. In the **Format Cells** box, under **Locale**, select **English (United States)**.
5. Under **Type**, select **Wednesday, March 14, 2012** and click **OK**.



## Task E: Use Find & Replace to trim whitespace

1. Click **CTRL+HOME**.
2. Select all the data using **CTRL+SHIFT+END**.
3. On the **Home** tab, click **Find & Select**, then **Replace**.
4. In **Find what**, type **2 spaces**. In **Replace with**, type **1 space**.
5. Click **Find All**, then click **Replace All**.
6. Click the **Close** icon.

# Exercise 3: More Excel Features for Cleaning Data

In this exercise, you will learn how to use the Flash Fill feature and functions in Excel to help clean data.

## Task A: Use the Flash Fill feature to clean data:

1. Select column **A (Cust\_Name)**, right-click and choose **Insert Columns**.
2. In cell **A1** type **Customer\_Name** and press **Enter**.
3. In cell **A2**, type **Mr. Allen Perl** and press **Enter**.
4. Select column **A (Customer\_Name)**, on the **Data** tab, click **Flash Fill**.
5. Click **Undo** to undo this step.

**If you are using the desktop version of Excel, you could use the 'Text to Columns' feature to perform this next task (see the corresponding topic video for instructions).**

**If you are using 'Excel for the web' (the online version of Excel), the 'Text to Columns' feature is not available, but you can achieve the same results using functions, as shown in the steps below.**

## **Task B: Use LEFT, RIGHT, LEN, and SEARCH functions to clean data:**

1. Select column **A (Cust\_Name)**, right-click and choose **Insert Columns**.
2. Select column **A** again, right-click and choose **Insert Columns**.
3. In cell **A1**, type **Customer\_Firstname** and in cell **B1**, type **Customer\_Lastname**.
4. Click **C1**, then on the **Home** tab, click **Format Painter**, then drag **across to A1 and B1**.
5. Double-click the **divider between columns A and B**.
6. In cell **A2** type **=LEFT(C2, SEARCH(" ",C2,1))** and press **Enter**.
7. In cell **B2** type **=RIGHT(C2,LEN(C2)-SEARCH(" ",C2,1))** and press **Enter**.
8. Double-click the **Fill Handle** on cell **A2**.
9. Double-click the **Fill Handle** on cell **B2**.

**Congratulations! You have completed Lab 5, and you are ready for the next topic.**

## Reading: Summary and Highlights

In this lesson, you have learned the following information:

It's important to remove any duplicated or inaccurate data, and it's important to remove any empty rows in your dataset.

There are several other types of data inconsistency that you may need to resolve, in order to properly clean your data:

Change the case of text

Fix date formatting errors

Trim whitespace from your data

You can use the Flash Fill and Text to Columns features in Excel to manipulate and standardize your data, and functions can also be used to help manipulate and standardize your data.

## Question 1

1/1 point (ungraded)

According to the video, what's one of the easiest common errors or inconsistencies to fix when importing data?

Duplicated data

Spelling mistakes

Extra white space

Empty rows



## Question 2

1/1 point (ungraded)

What is the first thing you should do when checking spelling errors in Excel?

Click the Spelling button

Select the data you want to check for spelling

Find the misspelled data

Use a dictionary



### Question 3

1/1 point (ungraded)

What is one of the functions you can use to change the text case in your data?

STRUCTURE

LOWER

CASE

CAP



### Question 4

1/1 point (ungraded)

Which of the following tasks can Flash Fill do? Select all that apply.

Set column width

Remove empty rows

Combine two columns of data into one

Split one column of data into two



## Module 4: Graded Quiz

 Bookmark this page

Graded Quiz due Jul 28, 2022 00:03 +08

### Question 1

1/1 point (graded)

What is one of the issues that empty rows cause in your spreadsheet?

Spell-check will fail

Formula errors

Data will be hidden

Indexing function error



### Question 2

1/1 point (graded)

There are two methods to locate and remove duplicated rows in Excel, what is the easiest way?

Select a column and choose Conditional Formatting > Highlight Cells Rules > Duplicate Values

Search for the first character in rows

Select all data and click the Remove Duplicates button

Use the HLOOKUP function



### Question 3

1/1 point (graded)

What does the PROPER function do?

Converts upper case text to lower case

Converts lower case text to upper case

Changes text to sentence case

Repairs a REF error



## Question 4

1/1 point (graded)

Why do you need to use the Paste Values option when you paste contents from a helper row to the original row?

- Ensures consistency
- Validates the copied function
- Makes sure formulas aren't also copied

- Keeps cell formatting the same



## Question 5

1/1 point (graded)

What is one of the ways to apply new data formats to the rest of a column?

- Format button
- Paste Special tool
- Format Painter tool

- Text Import wizard



## **Module 5 - Data Analysis Basics, Filtering and Sorting Data**

---

[Module Introduction and Learning Objectives](#)

---

[Video: Intro to Analyzing Data Using Spreadsheets](#)

---

[Video: Filtering and Sorting Data in Excel](#)

---

[Video: Viewpoints: Filtering and Sorting](#)

---

[Video: Useful Functions for Data Analysis](#)

---

[Video: Using VLOOKUP and HLOOKUP Functions](#)

---

[Hands-on Lab 6: Filtering and Sorting Data](#)

---

[Reading: Summary and Highlights](#)

---

[Module 5: Practice Quiz](#)

---

[Module 5: Graded Quiz \(5 Questions\)](#)

Graded Quiz due Jul 30, 2022, 10:03 PM GMT+8

## **Module Introduction**

In this module, you will learn about the fundamentals of analyzing data using an Excel spreadsheet and how to filter and sort your data. You will also learn how to use some of the most useful functions for a data analyst, and how to use the VLOOKUP and HLOOKUP reference functions

## **Learning Objectives**

**After completing this module, you will be able to:**

- Describe the fundamentals of analyzing data using a spreadsheet.
- Filter and sort data in a worksheet.
- Use some of the most useful functions for a data analyst.
- Use the VLOOKUP and HLOOKUP functions.

# Intro to Analyzing Data Using Spreadsheets

Now that we have learned how to collect and clean our data, it is time to decide the best method for analysis. In this video, we will discuss the importance of filtering, sorting, performing calculations, and shaping our data to provide meaningful information.

Deciding how to manipulate our data can sometimes be difficult. Before we make any changes or adjustments, we will need to visualize the final output. Below are some questions to ask before beginning the task.

How big is the dataset?

What type of filtering is required to find the necessary information?

How should the data be sorted?

What type of calculations are needed?

Now that we have visualized the final output, we must decide the best approach to shape our data.

The most basic step would be to **filter and sort the data**. By sorting the data, we are able to organize it based on conditions such as alphabetically or numerically.

For example, if we wanted to check for duplicate order numbers, we could sort the data and quickly see any duplicates. After sorting and removing the duplicate row, we find that the view needs to be more specific to meet our requirements.

We now decide that we only want to see the data for the month of November. By adding a filter, we can now choose to only see items with a ‘MONTH\_ID’ that is equal to “11”. By filtering our data, we are now able to only see the rows that meet the filter criteria and it allows us to better analyze our information.

Becoming familiar with all of the tools to analyze data can seem daunting, but one key benefit of using a spreadsheet is the ability to use functions. Functions in Excel are organized by several categories, including mathematical, statistical, logical, financial, and date and time-based.

Let’s say we wanted to get an average of company revenue for the month of June. We realize there are over 100 items that would need to be calculated. In normal circumstances, to get an average, we would have to create a formula to add each row and divide by the total number of rows.

This type of calculation would not only be very long, but can expose the analyst to possibly making a mistake.

=B1+B2+B3...../160

With the use of a function, we would be able to simplify our calculation in one easy step.

=AVERAGE(B1:B160)

While sorting and filtering data on our spreadsheet can be useful on its own, first converting your data to a table has many benefits. When we convert our data into a table, we are able to filter and calculate the data

more efficiently. One example is the ability to easily calculate columns. For the column 'MSRP', we choose 'Sum' and we're able to quickly calculate the sum of the column. If we then look at the data, and only want to calculate the 'MSRP' total based on Japan, we would filter the 'Country' column to only display Japan, and the column would then only add the values in the rows that were associated with Japan.

While all data may not work in a table, there are quite a few advantages to formatting your data as a table:

Automatic calculations even when filtering

Column headings never disappear

Banded rows to make reading easier

Tables will automatically expand when adding new rows. Sometimes data needs to be more organized than what a basic tabular format can give us, and creating pivot tables with charts can be a better way to analyze and display the required information.

In Excel we have the option of creating a pivot table to display and analyze our data, and optionally, an associated pivot chart. For example, let's say we want to know what company ordered products in the month of October.

From the original table of data, we create a pivot table to organize and analyze the required data, along with a pivot chart to display the information. By then adding the month filter to the newly created pivot table, we can see the results for the month of October not only in the table, but the changes are automatically updated in the pivot chart.

When trying to single out specific information in a large dataset, a pivot table is a nice way to show only the information that is required. This allows us to quickly and easily scan the essential information.

Pivot charts are a nice accessory to pivot tables, as they allow us to visually process data, and in most cases, will let the audience grasp the information quicker.

The advantages of selecting a pivot table and chart are:

Manipulate data without using formulas

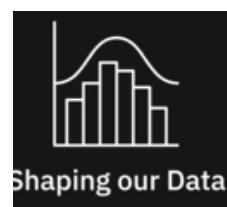
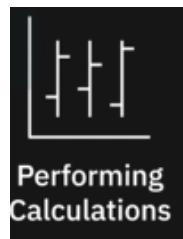
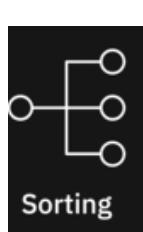
Quickly summarize large data sets

Ability to display engaging charts and graphs

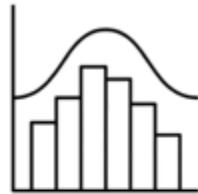
In this video, we learned about the importance of filtering, sorting, performing calculations, and shaping our data to provide meaningful information, and we learned about some of the tools to begin analyzing our data. In the next video, we will learn more about filtering and sorting our data.

# Intro to Analyzing Data Using Spreadsheets

Time to decide the best method for analysis"



How to Shape your Data



- How big is the dataset?
- What type of filtering is required to find the necessary information?
- How should the data be sorted?
- What type of calculations are needed?

## Filter and Sort the data

By sorting the data, we are able to organize it based on conditions such as alphabetically or numerically

ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	PRODUCTLINE
10103	26	100	11	5404.6	1/29/2003 0:00	Shipped	1	1	2003	Classic Cars
10112	29	100	1	7209.1	3/24/2003 0:00	Shipped	1	3	2003	Classic Cars
10112	29	100	1	7209.1	3/24/2003 0:01	Shipped	1	3	2003	Classic Cars
10183	23	100	8	5372.6	11/13/2003 0:00	Shipped	4	11	2003	Classic Cars
10194	42	100	11	7290.4	11/25/2003 0:00	Shipped	4	11	2003	Classic Cars

ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	PRODUCTLINE
10103	26	100	11	5404.6	1/29/2003 0:00	Shipped	1	1	2003	Classic Cars
10112	29	100	1	7209.1	3/24/2003 0:01	Shipped	1	3	2003	Classic Cars
10183	23	100	8	5372.6	11/13/2003 0:00	Shipped	4	11	2003	Classic Cars
10194	42	100	11	7290.4	11/25/2003 0:00	Shipped	4	11	2003	Classic Cars

QTR_ID	MONTH_ID	YEAR_I	PRODUCTLINE
1	1	2003	Classic Cars
1	3	2003	Classic Cars
4	11	2003	Classic Cars
4	11	2003	Classic Cars

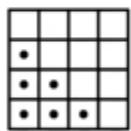
  

QTR_ID	MONTH_ID	YEAR_I	PRODUCTLINE
4	11	2003	Classic Cars
4	11	2003	Classic Cars

## Using Functions in Data Analysis

Becoming familiar with all of the tools to analyze data can seem daunting, but one key benefit of using a spreadsheet is the ability to use functions

Functions in Excel are organized by several categories



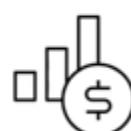
Mathematical



Statistical



Logical



Financial



Date and Time

=B1+B2+B3..../160

With the use of a function, we would be able to simplify our calculation in one easy step.

=AVERAGE(B1:B160)

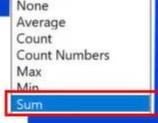
## Analyzing Data in a Table

While sorting and filtering data on our spreadsheet can be useful on its own, first converting your data to a table has many benefits

MONTH_ID	YEAR_ID	PRODUCTLINE	MSRP	PRODUCTCODE	CUSTOMERNAME	COUNTRY	TERRITORY	CONTACTLA
1	2004	Classic Cars	214	S10_1949	West Coast Collectables Co.	USA	NA	Thompson
3	2004	Classic Cars	214	S10_1949	Cambridge Collectables Co.	USA	NA	Tseng
5	2004	Classic Cars	214	S10_1949	Super Scale Inc.	USA	NA	Murphy
6	2004	Classic Cars	214	S10_1949	Tokyo Collectables, Ltd	Japan	Japan	Shimamura
7	2004	Classic Cars	214	S10_1949	Souveniers And Things Co.	Australia	APAC	Huxley
8	2004	Classic Cars	214	S10_1949	Amica Models & Co.	Italy	EMEA	Accorti
9	2004	Classic Cars	214	S10_1949	Scandinavian Gift Ideas	Sweden	EMEA	Larsson
6	2004	Trucks and Buses	136	S12_1666	Tokyo Collectables, Ltd	Japan	Japan	Shimamura
10	2004	Classic Cars	214	S10_1949	Auto Assoc. & Cie.	France	EMEA	Tonini
10	2004	Classic Cars	214	S10_1949	Mini Gifts Distributors Ltd.	USA	NA	Nelson

136

MONTH_ID	YEAR_ID	PRODUCTLINE	MSRP	PRODUCTCODE	CUSTOMERNAME
1	2004	Classic Cars	214	S10_1949	West Coast Collectables Co.
3	2004	Classic Cars	214	S10_1949	Cambridge Collectables Co.
5	2004	Classic Cars	214	S10_1949	Super Scale Inc.
6	2004	Classic Cars	214	S10_1949	Tokyo Collectables, Ltd
7	2004	Classic Cars	214	S10_1949	Souveniers And Things Co.
8	2004	Classic Cars	214	S10_1949	Amica Models & Co.
9	2004	Classic Cars	214	S10_1949	Scandinavian Gift Ideas
6	2004	Trucks and Buses	136	S12_1666	Tokyo Collectables, Ltd
10	2004	Classic Cars	214	S10_1949	Auto Assoc. & Cie.
10	2004	Classic Cars	214	S10_1949	Mini Gifts Distributors Ltd.



MONTH_ID	YEAR_ID	PRODUCTLINE	MSRP	PRODUCTCODE
1	2004	Classic Cars	214	S10_1949
3	2004	Classic Cars	214	S10_1949
5	2004	Classic Cars	214	S10_1949
6	2004	Classic Cars	214	S10_1949
7	2004	Classic Cars	214	S10_1949
8	2004	Classic Cars	214	S10_1949
9	2004	Classic Cars	214	S10_1949
6	2004	Trucks and Buses	136	S12_1666
10	2004	Classic Cars	214	S10_1949
10	2004	Classic Cars	214	S10_1949

2062

COUNTRY	TERRITORY	CONTACTLASTNAME
USA	NA	Thompson
USA	NA	Tseng
USA	NA	Murphy
Japan	Japan	Shimamura
Australia	APAC	Huxley
Italy	EMEA	Accorti
Sweden	EMEA	Larsson
Japan	Japan	Shimamura
France	EMEA	Tonini
USA	NA	Nelson

COUNTRY	TERRITORY	CONTACTLASTNAME
Japan	Japan	Shimamura
Japan	Japan	Shimamura

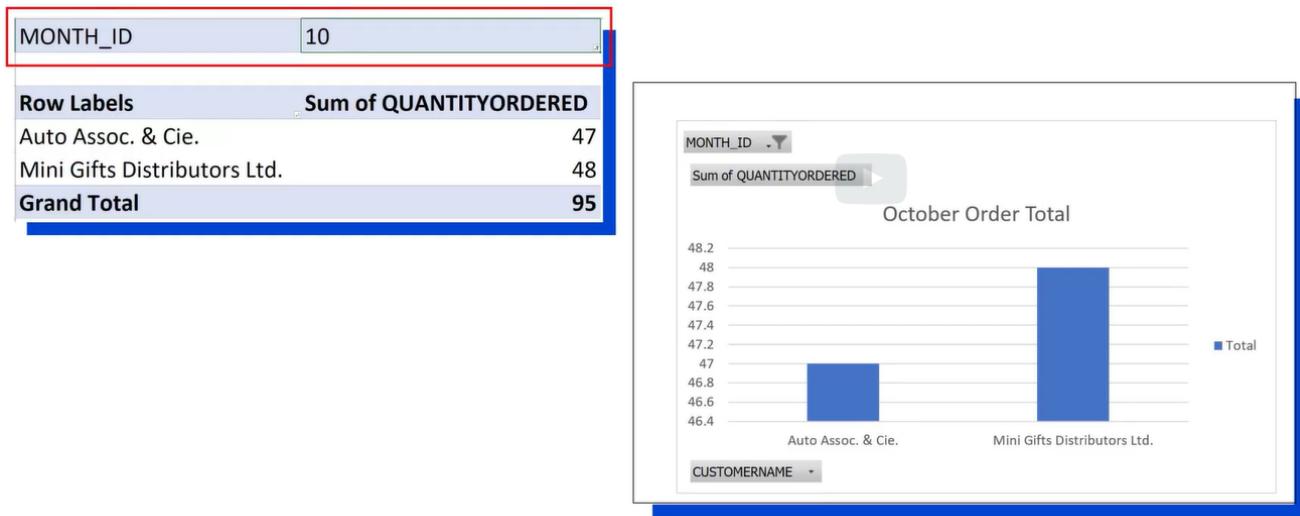
MONTH_ID	YEAR_ID	PRODUCTLINE	MSRP	PRODUCTCODE	CUSTOMERNAME	COUNTRY	TERRITORY	CONTACTLASTNAME
6	2004	Classic Cars	214	S10_1949	Tokyo Collectables, Ltd	Japan	Japan	Shimamura
6	2004	Trucks and Buses	136	S12_1666	Tokyo Collectables, Ltd	Japan	Japan	Shimamura

350

## Advantages of formatting Data to a Table:

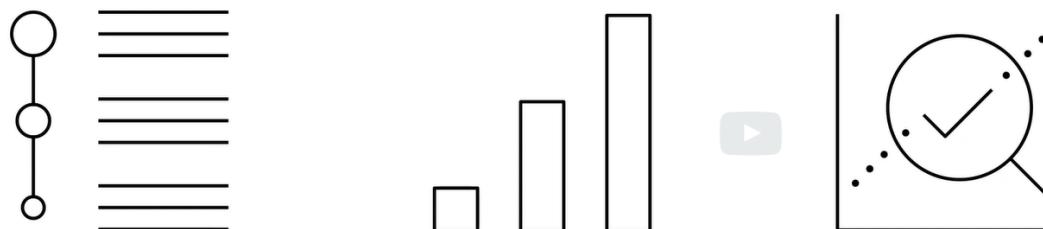
- Automatic calculations even when filtering
- Column headings never disappear
- Banded rows to make reading easier
- Tables will automatically expand when adding new rows

# Analyzing and Displaying Data in Pivot Tables and Pivot Charts



By then adding the month filter to the newly created pivot table, we can see the results for the month of October not only in the table, but the changes are automatically updated in the pivot chart.

## Advantages of Pivot Tables and Pivot Charts



- Manipulate data without using formulas
- Quickly summarize large data sets
- Ability to display engaging charts and graphs

# Filtering and Sorting Data in Excel

In the previous video we learned how to use the flash fill and text-to-columns features in Excel to help clean data. In this video we will discuss how to filter and sort our data to enable us to control what information is displayed and how it's displayed in our worksheets.

Filtering your data enables you to gain more control over which parts of your data are displayed at any given time in excel. This can help with the visibility of data by narrowing down the data to within specified criteria and parameters and it can also help when searching for specific pieces of data.

To filter your data the first thing you need to do is

1. turn filtering on, which is very simple.
2. On the data tab click "filter" and that's it, you will now see a small filter icon next to each of the column headers. As a side note if you want to only filter on one or more columns, select those columns first, then click filter. As another side note, if you format your data as a table, the columns automatically have filter controls added to them. So now each column has a filter that can be applied to the data in that column. In the order date column you can filter on the years.

In product line you can filter on the different product types. And in customer name you can filter on each customer by name. Let's first filter on the year. 1. We'll select orders from 2004. only by deselecting the other year, and if you wanted to we could expand the year and filter by months also, but we won't do that for now. 2. If you look at the status bar at the bottom of the worksheet you can see that there are only 50 out of 114 records now displayed. If you want to clear a filter you can either click the "clear filter" from option or click the "select all item" in the filter list.

Now let's filter on the product line column to display only the rows that hold data for sales of classic cars. And again we'll clear the filter. Lastly, we'll filter on the customer name column and only displays sales to "mini gifts distributors limited" and then clear that filter. So far we've only applied one filter at a time, but suppose you want to **filter down to a greater degree**. 1. We can do that too by just enabling all those filters together.

2. And now we are only displaying sales of classic cars to mini gifts distributors limited in 2004.
3. Remember if you only want to clear one filter then click its filter button in the column header, and
4. click the "clear filter from option," but if you want to
5. quickly clear all filters you can use the clear button in the sort and filter group on the data tab.

So far we've used what are commonly referred to as auto filters, but you can also use **Custom filters** to specify other criteria to apply to a filter to text or numbers. For example, if you wanted to see sales orders that are over or under a certain value you can do that with custom filters. For the

sales column let's add a number filter that only displays sales that are over two thousand dollars. If you look in the status bar you can see that we are now showing 111 out of 114 records. Then let's clear that filter and filter it the other way to display the sales orders that are below two thousand dollars.

We can see that there are only three orders that are below two thousand dollars. It's important to note that the data rows that we don't see have not been removed. They are still there, they have just been hidden from view by the filters and this is indicated by the row numbers you see on the left in blue. The row numbers start at 69 and jump in large increments indicating that there are many more rows of data in our data set than are currently being displayed. Let's clear those filters.

If we look at a column filter for a column that contains text, you will see that the menu item changes to text filters instead of number filters and you can see that there are several text filter options. And if you want to turn off filtering altogether for a worksheet, just click the filter button on the data tab.

Now let's take a look at the basic sorting capabilities in Excel. Sorting is a very important part of the role of a typical data analyst. You might need to organize your text-based data alphabetically, your number-based data numerically, or your date-based data chronologically. When you sort data using these logical parameters it makes it easier for you to conceptualize and visualize your data in a more meaningful way.

When **Sorting data** the first thing you need to do is 1. select which data to sort. For example, if you want to sort your customers alphabetically,

1. select a cell in the customer name column first and then either sort by a to z or by z to a. And if you want to sort your sales figures numerically, select a cell in the sales column first and then either sort from smallest to largest or from largest to smallest. And lastly,
2. if you want to sort your customers' order dates chronologically, select a cell in the order date column first, then sort from oldest to newest or from newest to oldest.

But you can also **Sort your data by more than one column at a time.**

1. Simply select a cell in your data then on the data tab click "sort," then either use the sort-by column suggested or use the drop-down list to select a different column.
2. In this case we'll choose the order date column as our first sorting criteria and we'll choose oldest to newest in the order drop down list.
3. To add a further sorting level you click "add level" then you choose another sort column in the "then by" drop-down list. In our case we'll choose sales, and for this sort level we'll choose largest to smallest in the order list. If you have a header row in your data as we do here, then ensure you select the "my data has headers" check box, then click "ok" to sort.

So, the data is now sorted to list the oldest orders by order date first, then within each order date if there are multiple instances with the same order date, then the net sorting level lists data by the largest order values first, down to the smallest order values.

In this video we learned how to use the filter and sort tools in Excel to filter and sort our data to enable us to control what information is displayed and how it is displayed in our worksheets.

## Filtering and Sorting Data in Excel

To control what information is displayed and how it's displayed in our worksheets.

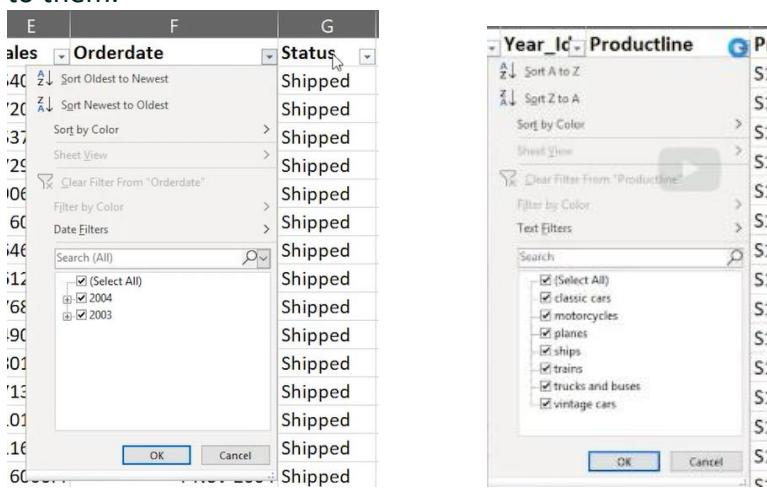
### Filtering Data

1. To filter your data the first thing you need to do is turn filtering on, which is very simple.
2. On the data tab click "filter" and that's it, you will now see a small filter icon next to each of the column headers.



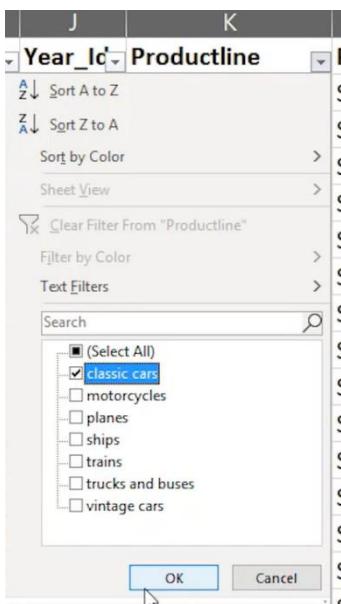
F	G	H	I	J	K	
	Status	Qtr_Id	Month_Id	Year_Id	Productline	
29-Jan-2003	Shipped		1	1	2003	classic cars
24-Mar-2003	Shipped		1	3	2003	classic cars
13-Nov-2003	Shipped		4	11	2003	classic cars
25-Nov-2003	Shipped		4	11	2003	classic cars
5-Dec-2003	Shipped		4	12	2003	classic cars

3. As a side note if you want to only filter on one or more columns, select those columns first, then click filter. As another side note, if you format your data as a table, the columns automatically have filter controls added to them.



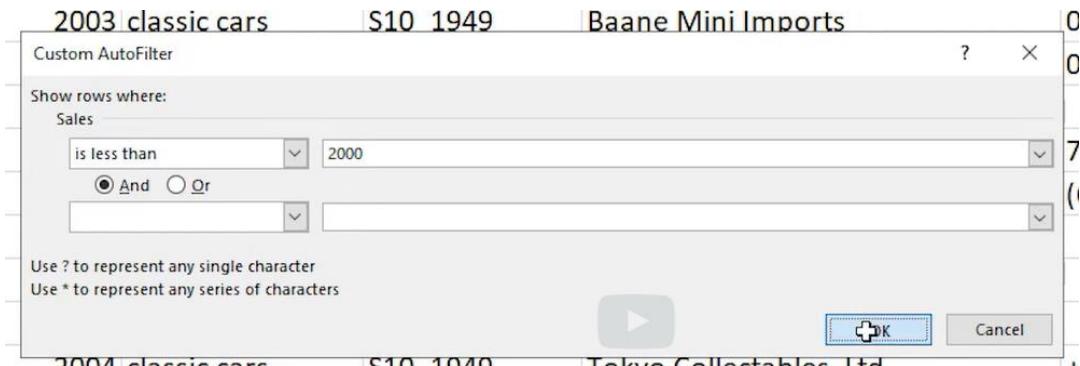
	Priceeach	Orderlinenumber	Sales	Orderdate
7	100	3	6075.3	29-Jan-2004
8	100	2	6463.23	10-Mar-2004
9	100	9	6120.34	4-May-2004
10	100	6	7680.64	15-Jun-2004
11	100	9	4905.39	19-Jul-2004
12	100	2	8014.82	17-Aug-2004
13	100	11	7136.19	8-Sep-2004
14	100	6	10172.7	11-Oct-2004
15	100	3	11623.7	21-Oct-2004
16	100	1	6000.4	4-Nov-2004
17	100	3	3003	18-Nov-2004

If you look at the status bar at the bottom of the worksheet you can see that there are only 50 out of 114 records now displayed. If you want to clear a filter you can either click the "clear filter" from option or click the "select all item" in the filter list.



filter on the product line column to display only the rows that hold data for sales of classic cars. And again we'll clear the filter.

Remember if you only want to clear one filter then click its filter button in the column header, and click the "clear filter from option," but if you want to quickly clear all filters you can use the clear button in the sort and filter group on the data tab.



filter it the other way to display the sales orders that are below two thousand dollars.

C	D	E	F	G	H	I	J	
Priceeach	Orderlinenumber	Sales	Orderdate	Status	Qtr_Id	Month_Id	Year_Id	Prod
89.46	9	1878.66	28-Apr-2003	Shipped	2	4	2003	ships
73.98	1	1553.58	7-May-2004	Cancelled	2	5	2004	ships
36.11	3	1516.62	29-Nov-2004	Shipped	4	11	2004	vinta

We can see that there are only three orders that are below two thousand dollars.

\*\*\*It's important to note that the data rows that we don't see have not been removed. They are still there, they have just been hidden from view by the filters and this is indicated by the row numbers you see on the left in blue.

**Sorting** is a very important part of the role of a typical data analyst.

- ✓ to organize your text-based data alphabetically
- ✓ number-based data numerically,
- ✓ date-based data chronologically.

When you sort data using these logical parameters it makes it easier for you to conceptualize and visualize your data in a more meaningful way.

The screenshot shows the Excel ribbon with the Data tab selected. The Sort & Filter group is highlighted, displaying various sorting and filtering tools. Below the ribbon, a table is shown with columns: Qtr\_Id, Month\_Id, Year\_Id, Productline, Productcode, and Customername. The last two rows of the table are highlighted in green, representing sorted data.

Qtr_Id	Month_Id	Year_Id	Productline	Productcode	Customername
1	1	2003	classic cars	S10_1949	Baane Mini Imports
1	3	2003	classic cars	S10_1949	Volvo Model Replicas, Co

if you want to sort your customers alphabetically,

1. select a cell in the customer name column first and then either sort by a to z or by z to a.

The screenshot shows the Excel ribbon with the Data tab selected. The Sort & Filter group is highlighted, displaying various sorting and filtering tools. A context menu is open over the 'Customername' column, showing sorting options: 'Sort Smallest to Largest' and 'Sort Largest to Smallest'. Below the ribbon, a table is shown with columns: Sales, Orderdate, Status, Qtr\_Id, and Month\_Id. The last row of the table is highlighted in green, representing sorted data.

Sales	Orderdate	Status	Qtr_Id	Month_Id
6075.3	29-Jan-2004	Shipped	1	1
2413.26	29-Jan-2004	Shipped	1	1
7209.11	24-Mar-2003	Shipped	1	3

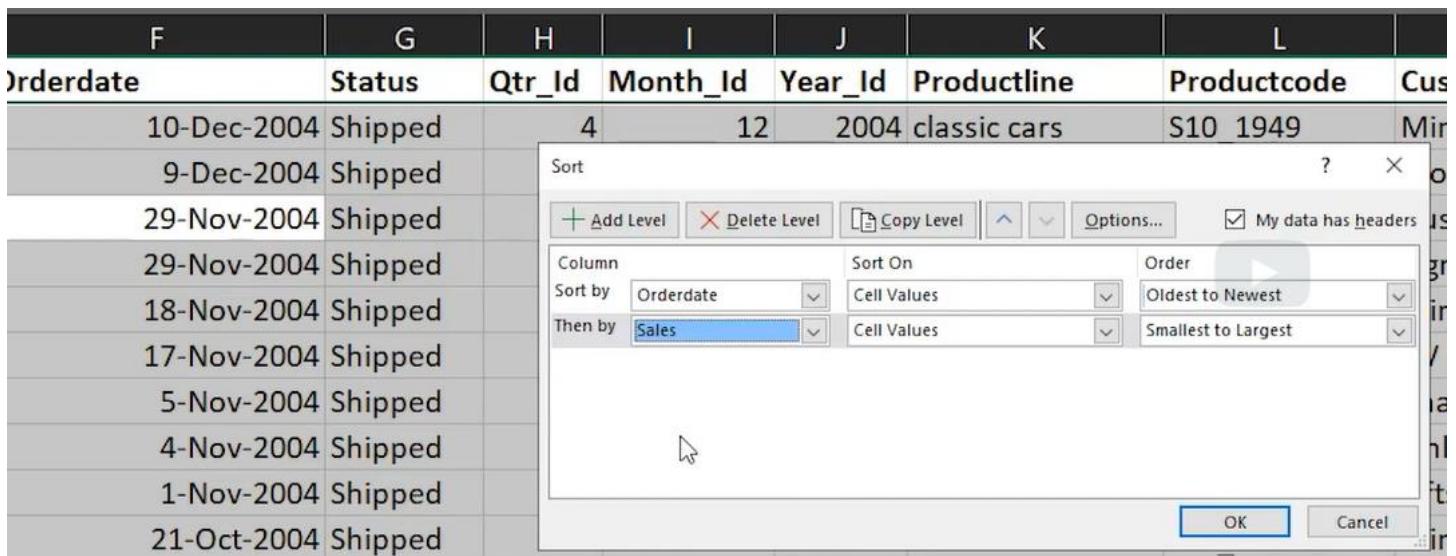
- select a cell in the sales column first and then either sort from smallest to largest or from largest to smallest.



A screenshot of the Microsoft Power Pivot ribbon. The 'Data Types' section is open, showing icons for 'Stocks' and 'Geography'. Below the ribbon is a small preview of a table with columns F and G, containing data for Orderdate and Status.

	F	G	H
	Orderdate	Status	
13.7	21-Oct-2004	Shipped	
12.7	11-Oct-2004	Shipped	
1.89	5-Dec-2003	Shipped	

- Select a cell in the order date column first, then sort from oldest to newest or from newest to oldest.



A screenshot of a Power Pivot table with the 'Sort' dialog box open. The table has columns F through L. The 'Sort' dialog shows two levels of sorting: 'Sort by' Orderdate (Oldest to Newest) and 'Then by' Sales (Smallest to Largest). The 'My data has headers' checkbox is checked. The 'OK' button is highlighted.

F	G	H	I	J	K	L	M
Orderdate	Status	Qtr_Id	Month_Id	Year_Id	Productline	Productcode	Cus
10-Dec-2004	Shipped	4	12	2004	classic cars	S10 1949	Min
9-Dec-2004	Shipped						
29-Nov-2004	Shipped						
29-Nov-2004	Shipped						
18-Nov-2004	Shipped						
17-Nov-2004	Shipped						
5-Nov-2004	Shipped						
4-Nov-2004	Shipped						
1-Nov-2004	Shipped						
21-Oct-2004	Shipped						

- Can also sort your data by more than one column at a time. Simply select a cell in your data then on the data tab click "sort," then either use the sort-by column suggested or use the drop-down list to select a different column.
- In this case we'll choose the order date column as our first sorting criteria and we'll choose oldest to newest in the order drop down list.
- To add a further sorting level you click "add level" then you choose another sort column in the "then by" drop-down list. In our case we'll choose sales, and for this sort level we'll choose largest to smallest in the order list.
- If you have a header row in your data as we do here, then ensure you select the "my data has headers" check box, then click "ok" to sort.

# Filtering and Sorting Data

Why is it important to filter and sort your data?

- ✓ Filtering and sorting are very important as part of your analysis
- ✓ Allows you to create one single view of the data
- ✓ Provides a function for people to do their own analysis on the data
  
- ✓ Sorting tends to be highest to lowest or alphabetical
- ✓ You may want to do some custom sorting
  
- ✓ Filtering and sorting is a great part of Excel
- ✓ Allows you to get into the heart of your data
- ✓ Can drill down and see how much revenue a client has for a specific timeframe
  
- ✓ Filtering means we have a value by which we want to see the data
- ✓ If we have a bar chart showing product sales over months, filtering allows us to filter down, so we only see one geography or one product line

# Useful Functions for Data Analysis

## IF Function

- IF function is one of the most used logical functions in Excel.
- IF compares a value against criteria then returns a result based on TRUE or FALSE comparison
- IF says; “if something is TRUE, return ‘X’, but if it is FALSE, return ‘Y’”

Example:

Vehicle toy sales worksheet, if we wanted to have a column that recorded whether the order had been shipped or not

1. add a new column to the right of the existing column – let’s call it Shipped?
2. enter the formula seen in cell H2 This formula is saying – if the text in G2 says ‘shipped’ then return ‘Yes’, and if it doesn’t then return ‘No’.
- 3.

F	G	H	I	J
Orderdate	Status	Shipped?	Qtr_Id	Month_Id
10-Jan-2003	Shipped	=IF(G2="Shipp		1
10-Jan-2003	Shipped	IF(logical_test, [value_if_true], [value_if_false])		
29-Jan-2003	Shipped		1	1
29-Jan-2003	Shipped		1	1

This formula is saying – if the text in G2 says ‘shipped’ then return ‘Yes’, and if it doesn’t then return ‘No’

4. You can then use the Fill Handle to copy this formula down the column. You can see that most of the cells do say ‘Yes’, but some don’t, as the order hasn’t been shipped for one reason or another.

Sales	Orderdate	Status	Shipped?
4808.3	10-Jan-2003	Shipped	Yes
2055.7	10-Jan-2003	Shipped	Yes
5404.6	29-Jan-2003	Shipped	Yes
3395	29-Jan-2003	Shipped	Yes
7208	11-Feb-2003	Shipped	Yes

Example: To emphasize the size of an order

1. Add a new column to the right of 'Sales', and name it '3K plus or minus'
2. Then enter the formula seen in cell F2 This formula is saying – if the order is over three thousand, then return the text "Over 3k", but if it isn't, then return the text "Under 3k".
3. We can copy the formula down the column.

E	F	G	H
Sales	3K + / -	Orderdate	Status
4808.3	=IF(F2>3000,"Over 3k","Under 3k")		
2055.7		10-Jan-2003	Shipped
5404.6		29-Jan-2003	Shipped

Sales	3K + / -
4808.3	Over 3k
2055.7	Under 3k
5404.6	Over 3k
3395	Over 3k
7208	Over 3k

Use multiple 'nested' Ifs when several conditions exist

1. First Method

F	G	H	I	J
Order Size (IF)	3K + / -	Orderdate	Status	Shipped?
=IF(E2>7000,"Large",IF(E2>4000,"Medium",IF(E2>0,"Small")))		-2003	Shipped	Yes
IF(logical_test, [value_if_true], [value_if_false])	Over 3k	29-Jan-2003	Shipped	Yes

If we add another column here for the order size. And then enter the formula seen in cell F2. You can see that this formula, contains multiple IF functions; one is needed for each condition one for Large, one for Medium, and one for Small and it requires three sets of parentheses. So, it's a relatively long and complex formula, but it does work. Again, we can copy the formula down the column.

#### IFS Function

- Excel theoretically supports up to 64 nested IFS – but it's not best practice
- Having large numbers of 'IFS' in a formula is hard to manage and understand
- Newer function called 'IFS' eliminates the need for multiple nested 'IFS'
- IFS only supported on EXCEL 2019, Excel for M365 and Excel for the web

## 2. IFS Function

E	F	G	H	I	J	K	L
Sales	Order Size (IF)	Order Size (IFS)	3K + / -	Orderdate	Status	Shipped?	Qtr_Id
4808.3	Medium	=IFS(E2>7000,"Large",E2>4000,"Medium",E2>0,"Small")					
2055.7	Small		IFS(logical_test1,value_if_true1,[logical_test2,value_if_true2],[logical_test3,value_if_true3],[logical_t				

As you can see in cell G2, this formula only has one set of parentheses instead of three, and only uses one function instead of three. Let's copy that formula down the column too.

Order Size (IF)	Order Size (IFS) 3K + / -	Orderdate	Status
Medium	Medium	Over 3k	10-Jan-2003 Shipped
Small	Small	Under 3k	10-Jan-2003 Shipped
Medium	Medium	Over 3k	29-Jan-2003 Shipped
Small	Small	Over 3k	29-Jan-2003 Shipped

1. With Conditional Formatting. If we switch to the car sales worksheet and add a new column to the right of the Year Resale Value column and call it 'Retention %'.

Then, we enter the formula seen in cell G2, which will divide the 'Year Resale Value', by the original 'Retail Price'. We need to format this as a percentage.

And then we can copy it down the column.

E	F	G
Retail Price	Year Resale Value	Retention %
\$21,500	\$16,360	=SUM(F2/E2)
\$28,400	\$19,875	
\$42,000	\$29,725	

Year Resale Value	Retention %
\$16,360	76%
\$19,875	70%
\$29,725	71%
\$22,255	93%
\$23,555	69%

2. Add a column to highlight the retention value for each car. The formula we add here in cell H2 uses the IF function to state that if the percentage in the previous column is greater than 69%, then mark it as 'Good', but if it isn't, then mark it as 'Poor'.

G	H	I
Retention %	Retention Value	Engine Size
76%	=IF(G2>69%, "GOOD", "POOR")	
70%		3.2
71%		3.5
93%		1.8

3. Once again, we'll copy the formula down the column.

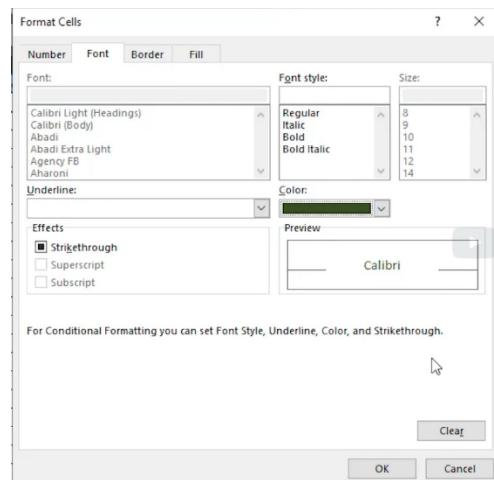
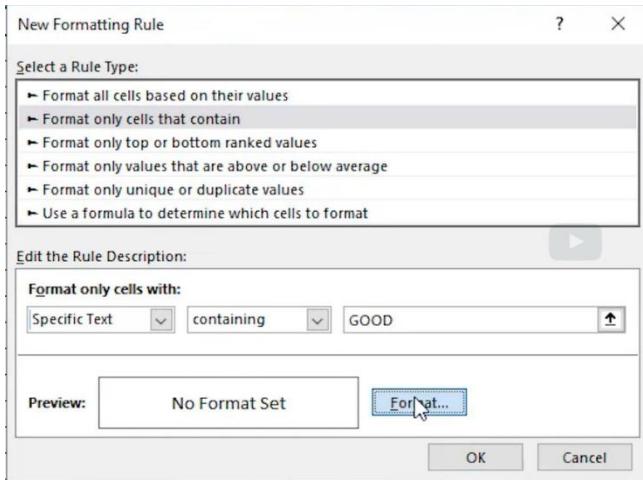
G	H
Retention %	Retention Value
76%	GOOD
70%	GOOD
71%	GOOD

Retention %	Retention Value
76%	GOOD
70%	GOOD
71%	GOOD
93%	POOR
54%	POOR

We could also use Conditional Formatting to highlight the retention value percentages even more.

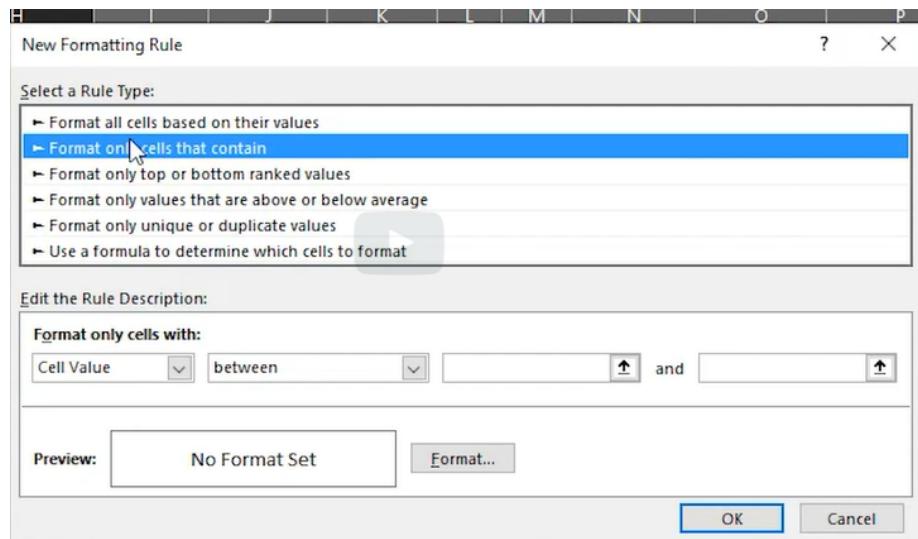
We select H2, and on the Home tab, click Conditional Formatting, and make a new rule.

The condition in our rule will only format cells that contain a specific text value... and that value is the word 'GOOD'.

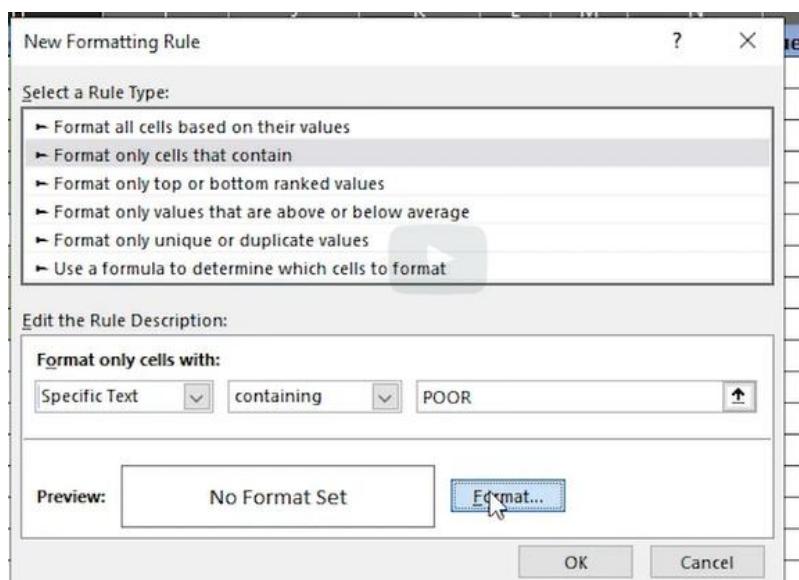


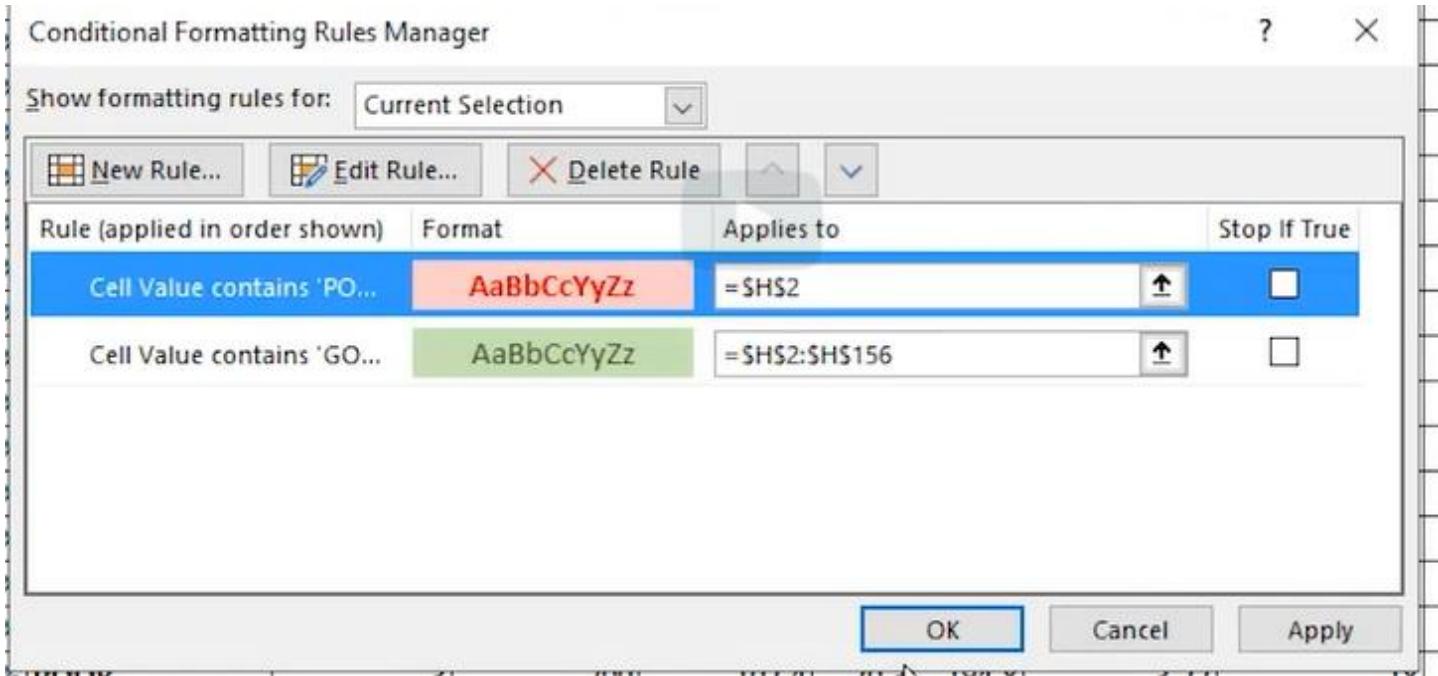
The condition in our rule will only format cells that contain a specific text value... and that value is the word 'GOOD'. And if it does match that condition, then format it with a dark green font and fill the cell in pale green.

Let's add another conditional format rule. This time, we'll select Manage Rules, because we are going to add another rule to our existing rule. The new rule will be the same as the previous one, with the exception of looking for a match with the word 'poor' instead, and formatting those matching cells with red text and a pink background fill.



Let's add another conditional format rule. This time, we'll select Manage Rules, because we are going to add another rule to our existing rule. The new rule will be the same as the previous one, with the exception of looking for a match with the word 'poor' instead, and formatting those matching cells with red text and a pink background fill.





This time, we'll select Manage Rules, because we are going to add another rule to our existing rule. The new rule will be the same as the previous one, with the exception of looking for a match with the word 'poor' instead, and formatting those matching cells with red text and a pink background fill.

H	I	J	K
Retention Value	Engine Size	Horsepower	Wheelbase
GOOD	1.8	140	101.2
GOOD	3.2	225	108.1
GOOD	3.5	210	114.6
GOOD	1.8	150	102.6
GOOD	2.8	200	108.7
POOR	4.2	310	113
GOOD +	2.5	170	107.3
GOOD	2.8	193	107.3
GOOD	2.8	193	111.4
POOR	3.1	175	109
POOR	3.8	240	109
POOR	3.8	205	113.8
POOR	3.8	205	112.2
POOR	4.6	275	115.3
POOR	4.6	275	112.2

## COUNTIF Function

- COUNTIF is a common statistical Excel Function
- COUNTIF counts the number of cell values that meet specified criteria

### Examples:

- How many times does “Employee X’s name appear in sales orders?
- How many times has ‘part number Y’ been ordered this month/year?

Suppose you want to find out how many of the sales orders in the list went to customers based in the United Kingdom. We enter the formula you see in cell AD7. Note that when we are using text as a criterion, we have to enclose the text in quotation marks. So there were 6 sales orders in the UK.

COUNTIF	=COUNTIF(X2:X115,"UK")	COUNTIF(range, criteria)	6
SUMIF			
SUMIF			
COUNTIF	=COUNTIF(X2:X115,"UK")	COUNTIF(range, criteria)	14
SUMIF			
SUMIF			

## COUNTIFS Function

- Newer ‘COUNTIFS’ function eliminates need for multiple nested ‘COUNTIF’s
- COUNTIFS only supported on Excel 2019, Excel for M365, and Excel for the web

## SUMIF Function

- SUMIF is a very common mathematical Excel function
- SUMIF sums the values in a given range that match specified criteria

### Example:

- What is the total of all employee salaries that are above 30k per annum?
- What is the total of all sales of ‘product x’ in 2016 in the UK?

SUMIF	=SUMIF(E2:E115,>3000")
-------	------------------------

This formula will add up each of the sales orders that have a total of more than 3,000 dollars. Again, notice that because we have used an arithmetic operator, that is the ‘greater than’ operator, we must enclose the criterion in quotes. If we specify a criterion that is only a number, we don’t enclose it in quotes. So, the total sum of all orders that were over 3,000 dollars is almost 470,000 dollars.

SUMIF	469782
SUMIF	=SUMIF(O2:O115,"*cars",E2:E115) SUMIF(range, criteria, [sum_range])

Can also use wildcards such as ‘question mark’ (?) and ‘asterisk’ (\*) when searching for partial matches, and you can also specify to extract values from a different column than the column where you have specified the criteria.

For example, if we enter the formula you can see in cell AD13, it will sum all the car sales in column E, for only those products in the ‘productline’ column that end in ‘cars’.

## SUMIFS Function

- Newer ‘SUMIFS’ functions eliminates need for multiple nested ‘SUMIFS’
- SUMIFS only supported on Excel 2019, Excel for M365 and Excel for the web

## Using VLOOKUP and HLOOKUP Functions

### VLOOKUP Function

- Enables you find data referenced in a lookup table
- Vertical Lookup – find by row
- Horizontal Lookup – find by column
- Works by using a common shared key

=VLOOKUP(B3,A2:B12,2, FALSE)



B3 Lookup Value

A2:B12 Lookup Range

2 Lookup Column

FALSE Optional  
(exact or approx)

=VLOOKUP(

VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup])

lookup value (B3) = cell or spot to write what you want to find

table\_array (A2:B12) = tables or columns to be used for VLOOKUP searching

col\_index\_num (2) = number of columns that is included in your VLOOKUP searching

True or False = True is approximate match; False is exact match

If not specified default is FALSE; that is, an exact match is required. Can also use: 0 for FALSE, and 1 for TRUE.

Steps:

1. The first thing we need to do, is put the column containing the value you want to search for, in the leftmost column, as VLOOKUP requires this.

2. Construct VLOOKUP cell. Suppose we wanted a quick price list of our favorite cars.

3. The first thing we need to do, is put the column containing the value we want to search for, in the leftmost or the first column, as VLOOKUP requires this.

Model	Manufacturer	Model	Value
Integra	Acura	Integra	F
TL	Acura	TL	F
RL	Acura	RL	F
A4	Audi	A4	F
A6	Audi	A6	F
A8	Audi	A8	F
323i	BMW	323i	F
328i	BMW	328i	F
528i	BMW	528i	F
Century	Buick	Century	F
Regal	Buick	Regal	F
Park Avenue	Buick	Park Avenue	F
LeSabre	Buick	LeSabre	F
DeVille	Cadillac	DeVille	F
Seville	Cadillac	Seville	F

4. Delete the original column or the 3<sup>rd</sup> column in the example above.

5. Enter the formula in constructed VLOOKUP function cell.

VLOOKUP	Corvette			
	Camaro			
	Concorde			
	Viper			
	Mustang			
	Taurus			
	Cherokee			
	Firebird			
	Boxter			
		=VLOOKUP(A25,A2:G156,5,FALSE)		
		VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])		

VLOOKUP	Corvette		
	Camaro		
	Concorde		
	Viper		
	Mustang		
	Taurus		
	Cherokee		
	Firebird		
	Boxter		
		45705	

VLOOKUP	Corvet	!	#REF!
	Camaro		
	Concorde		
	Viper		
	Mustang		
	Taurus		
	Cherokee		
	Firebird		
	Boxter		
		\$ 45,705	

6. But there's a problem, because when we copied the formula, the cell references changed. This happened because as we learned earlier in this course, the default state of cell references is relative, and we want them to be absolute in this case. So, let's undo that copy operation. To make the cell references absolute, we need to add dollar symbols to all the cell references in the formula.

=VLOOKUP(\$V\$5,\$A\$2:G156,5, FALSE)
VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup])

7. Let's try and copy the formula again and this time it works. If we use the Fill Handle on cell W5 to copy it down to the rest of the cars, it doesn't work; in fact, we end up with the same result in every cell. Why? Because each one is referencing the same cells in the lookup value, because we used an absolute reference. All we need to do now, is modify the formula to remove the absolute reference for just the row parameter, in the lookup value part of the formula, by removing the dollar symbol. So in cell W5 we change \$V\$5 to \$V5, then when we drag the Fill Handle down it will copy the formula correctly, and all the prices will be changed to reflect their correct retail price.

VLOOKUP	Corvette	\$ 45,705
	Camaro	\$ 24,340
	Concorde	\$ 22,245
	Viper	\$ 69,725
	Mustang	\$ 21,560
	Taurus	\$ 17,885
	Cherokee	\$ 21,620
	Firebird	\$ 25,310
	Boxter	\$ 41,430
		\$ 45,705

7. To show that the two tables are now connected by this VLOOKUP function, if we change the retail price for the Chevrolet Corvette in the main data table in cell E25... the price will also change in the favorite cars price list.

135126	\$16,535	\$11,225
24629	\$18,890	\$10,310
42593	\$19,390	\$11,525
26402	\$24,340	\$13,025
17947	\$55,705	\$36,225

## HLOOKUP Function

- Enables you to find data referenced in a lookup table
  - Looks for data in columns, rather than rows
  - Looks for a word or value in the top row of a table
  - Use HLOOKUP if your comparison values were situated in a row along the top of a data table
  - Use VLOOKUP if your comparison values were located in a column to the left of the data
  - Syntax for HLOOKUP is identical to that of VLOOKUP except that you specify a row index number

1. Create a small lookup table on the right hand-side of our main data table; a few columns have been hidden in this worksheet to make viewing a little easier.

HLOOKUP	Low HP	Medium HP	High HP
	:(	:)	:)

2. This function will look for the value in cell K2, which in this case is ‘Medium HP’, and it will look for it in the cell range from Y21 to AA22, which is our little lookup table, and it will return the answer it finds in row 2 of the table under medium HP, and use an exact value. Note that we’ve used some absolute references in this formula too.

=HLOOKUP(\$K2,\$Y\$21:\$AA\$22,2,FALSE)

**HLOOKUP(lookup\_value, table\_array, row\_index\_num, [range\_lookup])**

3. Notice that what is returned is the text 'KK', so we need to format the cell using the Wingdings font.

K	L	M
HP Level	HP Rating	Wheelbase
Medium HP	KK	101.2
Medium HP		108.1

4. When we double-click the Fill Handle, the whole column shows the HP Rating symbols relevant to each row's HP Level value.

HP Level	HP Rating	Wheelbase
Medium HP	☺☺	101.2
Medium HP	☺☺	108.1
Medium HP	☺☺	114.6
Medium HP	☺☺	102.6
Medium HP	☺☺	108.7
High HP	☺☺☺	113

## XLOOKUP Function

- Only supported on Excel desktop versions from Excel for Microsoft 365, and on Excel for the web, as well as on Excel for iPad and iPhone, and Excel for Android tablets and phones.
- XLOOKUP is an improved and combined version of VLOOKUP and HLOOKUP together
- Works in any direction
- Uses separate lookup array and return array values



IBM Developer  
SKILLS NETWORK

# Hands-on Lab 6: Filtering and Sorting Data using Functions for Data Analysis

**Estimated time needed:** 30 minutes

In this lab, first you will learn how to use the Filter and Sort tools in Excel to filter and sort our data to enable us to control what information is displayed, and how it is displayed in our worksheets. Next, you will learn how to use some of the most common functions a Data Analyst might use; namely IF, IFS, COUNTIF, and SUMIF. Finally, you will learn how to use the VLOOKUP and HLOOKUP functions in Excel to reference data contained in both vertical and horizontal lookup tables.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Datasets Used in this Lab

The first dataset used in this lab comes from the following source: <https://dataplatform.cloud.ibm.com/exchange/public/entry/view/f8ccaf607372882403a37d9019b3abf4>. This dataset is published by **IBM**, and includes fictitious customer demographics and sales data.

The second dataset used in this lab comes from the following source: <https://www.kaggle.com/sudalairajkumar/indian-startup-funding> under a **CC0: Public Domain license**. Acknowledgement and thanks also goes to <https://trak.in> who were generous enough to share the data publicly for free.

We are using modified subsets of these datasets for the lab, so to follow the lab instructions successfully please use the datasets provided with the lab, rather than the datasets from their original sources.

The third dataset used in this lab is an internal dataset.

## Objectives

After completing this lab, you will be able to:

- Use the Filter and Sort tools
- Use IF, IFS, COUNTIF, and SUMIF functions for data analysis
- Use the VLOOKUP and HLOOKUP reference functions

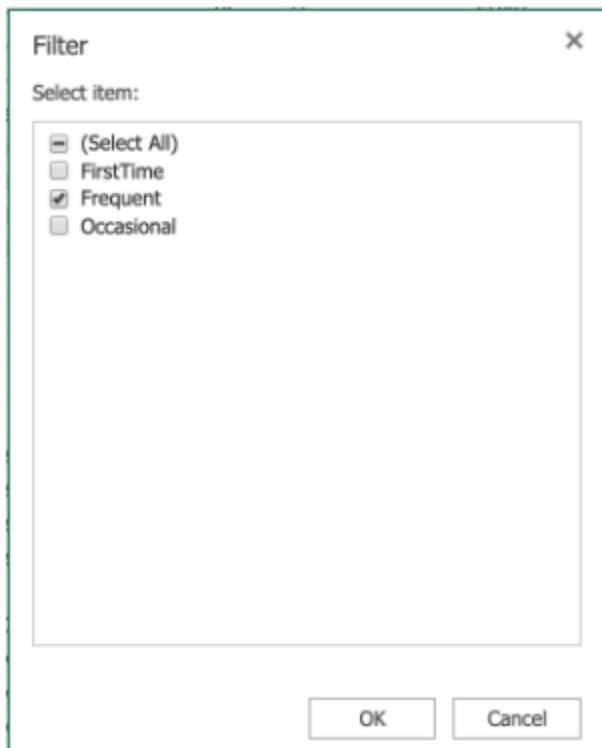
## Exercise 1: Filtering and Sorting Data

In this exercise, you will learn how to use the Filter and Sort tools in Excel to filter and sort our data to enable us to control what information is displayed, and how it is displayed in our worksheets.

### Task A: Filtering data

#### To use Auto Filters to filter data:

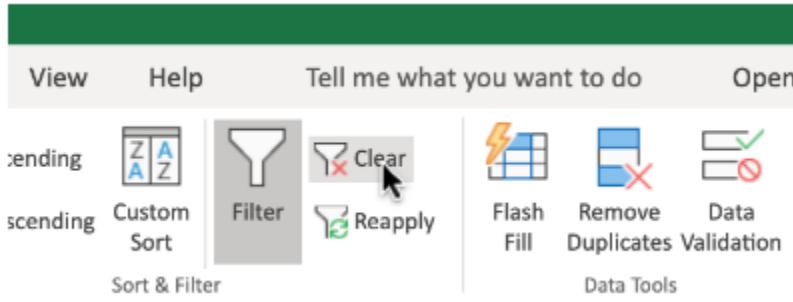
1. Download the file [\*\*Customer\\_demographics\\_and\\_sales\\_Lab6.xlsx\*\*](#). Upload and open it using Excel for the web.
2. Select **any cell** in the data, and click the **Data** tab, then click **Filter**.
3. Click the **filter drop-down** in column **AG (Purchase\_Status)**, and select **Filter....**
4. In the list, only select **Frequent** and click **OK**.



5. Click the filter drop-down in the column AG, and click Clear Filter From "Purchase\_Status".

AG	AH	AI
Purchase_Status	Order_Type	Generation
Frequent	<span>Sort Ascending</span>	
Frequent	<span>Sort Descending</span>	
Frequent	<span>Custom Sort</span>	
Frequent	<span>Sheet View</span>	
Frequent	<span>Clear Filter from 'Purchase_Status'</span>	
Frequent	<span>Text Filters</span>	
Frequent	<span>Filter...</span>	

6. Click the filter drop-down in column AE (T\_Type), and select Filter....
7. In the list, only select Cancelled and click OK.
8. Click the filter drop-down in column AF (Purchase\_Touchpoint), and select Filter....
9. In the list, only select Desktop and click OK.
10. On the Data tab, click Clear.

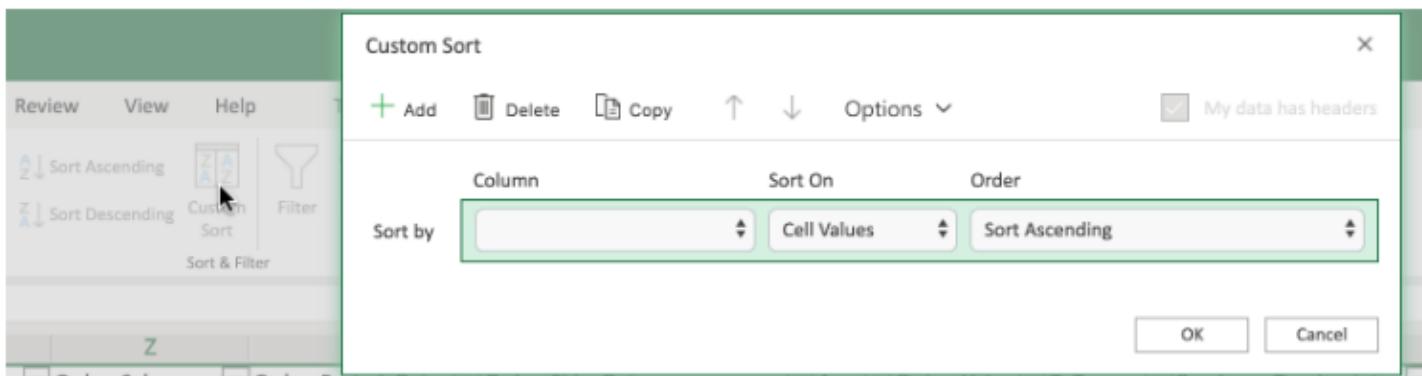


### To use Custom Filters to filter data:

1. Click the **filter drop-down** in column **AD (Order\_Value)**, then **Number Filters>Top 10...**
2. Change the value from **10 to 50** and Click **OK**.
3. Click the **filter drop-down** in the column **AD**, and click **Clear Filter From “Order\_Value”**.

## Task B: Sorting data

1. On the **Data** tab, click Custom Sort to open a dialog box like below.



2. Click the **Column drop-down** of row **Sort By**, select **Order\_Ship\_Date**.
3. Click the **Order drop-down** of row **Sort By**, select **Sort Ascending**.
4. Click **Add**.
5. Click the **Column drop-down** of row **Then By**, select **Order\_Value**.
6. Click the **Order drop-down** of row **Then By**, select **Sort Descending**.
7. Click **OK**.

## Exercise 2: Useful Functions for Data Analysis

In this exercise, you will learn how to use some of the most common functions a Data Analyst might use; namely IF, IFS, COUNTIF, and SUMIF.

### Task A: Use of IF to apply one condition

1. Select column **AF**, right-click, **Insert**.
2. In cell **AF1**, type **Complete?**.
3. In cell **AF2**, type **=IF(AE2="Complete","Yes","No")** and press **Enter**.
4. Double-click the **Fill Handle of AF2** to copy down the column.

## Task B: Use of Nested IF to apply multiple conditions

1. Select column **AE**, right-click, **Insert**.
2. In cell **AE1**, type **Order Size (IF)**.
3. In cell **AE2**, type **=IF(AD2>300,"Large",IF(AD2>100,"Medium",IF(AD2>0,"Small")))** and press **Enter**.
4. Double-click the **Fill Handle of AE2** to copy down the column.

## Task C: Use of IFS to apply multiple conditions (alternative of Nested IF)

1. Select column **AE**, right-click, **Insert**.
2. In cell **AE1**, type **Order Size (IFS)**.
3. In cell **AE2**, type **=IFS(AD2>300,"Large",AD2>100,"Medium",AD2>0,"Small")** and press **Enter**.
4. Double-click the **Fill Handle of AE2** to copy down the column.

## Task D: Use of COUNTIF to count the number of cells that meet a specified criterion

1. Select cell **BX2** and type **count VISA card**.
2. Select cell **BY2** and type **=COUNTIF(N2:N195,"VISA")** and press **Enter**.

## Task E: Use of SUMIF function to sum the values within a specified range that meet a specified criterion

1. Select cell **BX3** and type **sum Large order**.
2. Select cell **BY3** and type **=SUMIF(AE2:AE195,"Large", AD2:AD195)** and press **Enter**.
  - o Formula: **=SUMIF(range, criteria, [sum range])**.

## Task F: Use of SUMIFS function to sum the values within a specified range that meet multiple specified criteria

1. Select cell **BX4** and type **sum Large order with Baby Gen**.

- Select cell **BY4** and type **=SUMIFS(AD2:AD195, AE2:AE195,"Large", AL2:AL195,"\*BABY\_BOOMERS\*")** and press **Enter**.
  - Formula: **=SUMIFS ([sum range], range1, criteria1, range2, criteria2, ...)**.

## Exercise 3: Using the VLOOKUP and HLOOKUP Functions

In this exercise, you will learn how to use the VLOOKUP and HLOOKUP functions in Excel to reference data contained in both vertical and horizontal lookup tables.

### Task A: Use of VLOOKUP to look up data in a table organized vertically

- Download the file [indian startup funding Lab6.xlsx](#). Upload and open it using Excel for the web.
- In cell **K2,L2,M2**, type **VLOOKUP, Startup Name, Amount in USD** respectively.
- Select and copy cells from **C9 to C15** and paste in cell **L3**.
- In cell **M3**, type **=VLOOKUP(L3, C2:I113, 7, FALSE)** and press **Enter**.
  - Formula: **=VLOOKUP (value, table, col\_index, [range\_lookup])**.
- Hover over the bottom-right corner of cell **M3**, and drag the Fill Handle down to the cell **M9**.
- Select cells from **M3 to M9** and **select Number Format>Currency**.

K	L	M
VLOOKUP	Startup Name	Amount in USD
	Rein Games	=VLOOKUP(L3, C2:I113, 7, FALSE)
	CarDekho	\$70,000,000.00
	Dhruva Space	\$50,000,000.00
	Paytm	\$1,000,000,000.00
	Aye Finance	\$17,411,265.00
	Clumio	\$135,000,000.00
	Digital Mall Asia	\$220,000,000.00

### Task B: Use of HLOOKUP to look up data in a table organized horizontally

- Download the file [Personal Monthly Expenditure Lab6.xlsx](#). Upload and open it using Excel for the web.

2. In cell **J2,K2,L2,M2**, type **HLOOKUP, Month, Food & Dining, Health & Fitness** respectively.
3. Select and copy cells from **A10 to A12** and paste in cell **K3**.
4. In cell **L3**, type **=HLOOKUP(D1, A1:H14, 10, FALSE)** and press **Enter**.
  - o Formula: **=HLOOKUP (value, table, row\_index, [range\_lookup])**.
5. Hover over the bottom-right corner of cell **L3**, and drag the Fill Handle down to the cell **L5**.
6. Select cells from **L3 to L5** and **select Number Format>Currency**.
7. In cell **M3**, type **=HLOOKUP(G1, A1:H14, 10, FALSE)** and press **Enter**.
8. Hover over the bottom-right corner of cell **M3**, and drag the Fill Handle down to the cell **M5**.
9. Select cells from **M3 to M5** and **select Number Format>Currency**.

J	K	L	M	N	O
HLOOKUP	Month	Food & Dining	Health & Fitness		
Sep		\$400.00	=HLOOKUP(G1, A1:H14, 10, FALSE)		
Oct		\$420.00	60		
Nov		\$390.00	50		

Congratulations! You have completed Lab 6, and you are ready for the next topic.

## Author(s)

- [Sandip Saha Joy](#)

## Other Contributor(s)

- [Steve Ryan](#)

# Reading: Summary and Highlights

**In this lesson, you have learned the following information:**

Before shaping your data, you need to visualize the final output, and ask yourself the following questions:

How big is the dataset?

What type of filtering is required to find the necessary information?

How should the data be sorted?

What type of calculations are needed?

There are several advantages to formatting your data as a table:

Automatic calculations even when filtering

Column headings never disappear

Banded rows to make reading easier

Tables will automatically expand when adding new rows

The most basic way of shaping your data is to sort and filter it:

Sorting data helps you to organize it by a specified criteria, such as numerically, alphabetically, or chronologically.

Filtering our data makes it easier to control what data is displayed and what is hidden, based on filtered fields.

Excel Functions:

Functions in Excel are arranged into multiple categories; including mathematical, statistical, logical, financial, and date and time-based.

Common functions for a data analyst include IF, IFS, COUNTIF, SUMIF, VLOOKUP, HLOOKUP

## Module 5: Practice Quiz

Bookmarked

### Question 1

1/1 point (ungraded)

Which of the following are valid sorting orders you can choose when adding sorting levels to your data? Select all that apply.

Youngest to Oldest

Z to A

Oldest to Newest

Smallest to Largest

### Question 2

1/1 point (ungraded)

After filtering a column and getting the results, in which two ways can you return to showing all the data in a column? Select two answers.

refresh the worksheet

refresh the column

clear the filter

turn filtering off

✓

### Question 3

1/1 point (ungraded)

According to the video 'Useful Functions for Data Analysis', what is one of the most common functions a Data Analyst might use?

Average

Countif

Randbetween

Int



### Question 4

1/1 point (ungraded)

What does VLOOKUP stand for?

Volume Lookup

Vertical Lookup

Virtual Lookup

Variable Lookup



## Module 5: Graded Quiz

Bookmarked

Graded Quiz due Aug 18, 2022 22:48 +08    Completed

### Question 1

1/1 point (graded)

If you have multiple filters set, how can you clear all of them at once?

- You must clear each filter separately
- Data tab> Sort and Filter group> Clear
- Refresh the column
- Turn sorting off



### Question 2

1/1 point (graded)

After enabling Filtering, where can you see and access the filter controls?

- Each column header now has a filter control
- Only selected cells have filter controls
- Each row has filter controls
- You must manually select data for filtering



### Question 3

1/1 point (graded)

The IF function applies to one or two conditions, but what if you need to apply multiple conditions?

- Use separate IF functions
- Use the IF-Then function
- Use the nesting capabilities of the IF function

- Use the IFor function



### Question 4

1/1 point (graded)

When you use nested functions, what is required for each of the functions?

- =IF to begin each function
- Semi-colon between each function
- A set of parentheses for each function

- Each function must be alphabetically ordered



---

### Question 5

1/1 point (graded)

What do custom filters provide that AutoFilters don't?

- associate a formula with the filter control
- set multiple filters at the same time
- filter a column to see only above or below a certain value



PS: In anticipation of what I hope will be smooth sailing towards your being issued an official appointments, I have few things to confirm with you all:

- Where you would prefer to have a permanent desk (the Los Baños headquarters or the Diliman main office) if you start reporting physically to work?
- Do you have an online profile that contains your most complete and up-to-date listing of major research, teaching, and public service/public engagement highlights? (e.g., ORCID.org profile, Google Scholar Profile, LinkedIn profile, ResearchGate.net profile, Academia.edu profile, personal website, updated downloadable CV, etc.) Please send me the URL for this profile. Please send only one link.
- If you haven't signed up for an ORCID yet, please consider getting one. <https://orcid.org/>

I prefer the Diliman Main office since I'll be coming from Olongapo City. May I ask if I'll be required to be physically present on weekdays on a regular basis should I pass the final interview and be given an official appointment?

Considering that I'll be traveling from Subic, I prefer the Diliman Main Office. May I enquire as to whether, should I make it through the final interview and be given an official appointment, I will need to be physically present on a regular basis throughout the weekdays?

## Module 6 - Using Pivot Tables

---

-  [Module Introduction and Learning Objectives](#)
-  [Video: Introduction to Creating Pivot Tables in Excel](#)
-  [Video: Viewpoints: Pivot Tables](#)
-  [Video: Pivot Table Features](#)
-  [Hands-on Lab 7: Using Pivot Tables](#)
-  [Reading: Summary and Highlights](#)
-  [Module 6: Practice Quiz](#)
-  [Module 6: Graded Quiz \(5 Questions\)](#)

Graded Quiz due Aug 21, 2022, 8:48 PM GMT+8

## Module Introduction

In this module, you will learn how to create pivot tables in Excel and use several pivot table features to analyze data.

### After completing this module, you will be able to:

- Create pivot tables in Excel.
- Use pivot table features.

# Introduction to Creating Pivot Tables in Excel

Now that we've learned how to use the VLOOKUP and HLOOKUP functions, in this video we'll look at how to create and use Pivot Tables in Excel.

We'll first look at how to format our data as a table, then how to create Pivot Tables and use fields in a Pivot Table to analyze data, and lastly we'll see how to perform calculations in a Pivot Table. Having a worksheet full of informational data is all very well, but to really get some use out of it we need to analyze it from different perspectives to find answers to questions related to the data.

Now, we've already used features such as filters and formulas to draw mathematical and logical conclusions about our data but not all questions can be answered easily using filters and formulas alone. In order to obtain usable and presentable insights into your data you need something else... and that something else is Pivot Tables.

Pivot Tables provide a simple and quick way, in spreadsheets, to summarize and analyze data, to observe trends and patterns in your data and to make comparisons of your data.

A Pivot Table is dynamic, so as you change and add data to the original dataset on which the Pivot Table is based, so the analysis and summary information changes too.

A Data Analyst can use Pivot Tables to draw useful and relevant conclusions about, and create insights into, an organization's data in order to present those insights to interested parties within the company.

Before you start to create a Pivot Table in Excel, it can be very helpful to first format your data as a table. The reason for this is not only to make it more organized and defined and to add table styles to your data, but primarily it makes it a lot easier when adding records to the dataset.

In the car sales worksheet, let's first select any cell within the data, and then on the Home tab, in the Styles group, choose 'Format as Table'. Then choose a style from the gallery... note that Excel automatically knows the boundaries of our data range, but we can change this if we need to.

And ensure you select 'My table has headers', if indeed it does. After you click OK and the data has been formatted as a table, note the filter drop-downs at the top of each column – these are automatically added when you format as a table.

If we now scroll down to the bottom of the table... and start adding another row of data for another vehicle... when you click Tab or Enter, note that it is automatically formatted and included as part of our table.

OK, now let's see how to create a basic Pivot Table, and how to use fields to arrange data in a Pivot Table. Just before we do that, there are a few things you should use as a checklist to ensure your data is in a fit state to make a Pivot Table from, and these are:

Format your data as a table for best results

Ensure column headings are correct, and there is only one header row, as these column headings become the field names in a Pivot Table

Remove any blank rows and columns, and try to eliminate blank cells also

Ensure value fields are formatted as numbers, and not text

Ensure date fields are formatted as dates, and not text

In the worksheet, we can just select any cell in the table. Then, on the Insert tab, we click PivotTable. Note that in the 'Select a table or range' box, the table name – Table1 – is already entered for us. If we hadn't just formatted this data as a table, we would specify the cell range here instead.

Under that, we need to decide whether we want to create the Pivot Table on a separate new blank worksheet, or on this worksheet – a new worksheet is the default – and is the most commonly used option. So, a new blank worksheet opens, displaying some basic Pivot Table instructions in the graphic on the left of the worksheet, and a 'PivotTable Fields' pane on the right.

You can rename the worksheet for the Pivot Table if you wish. To build the Pivot Table report we need to add some fields from the top of the PivotTable Fields pane, to one or more of the sections in the bottom part of the pane. For example, if we want to find out the total sales for each model of car, let's drag the Manufacturer field to the Rows section of the report, ... and then we'll drag the Model field there too.

But this isn't really the way we want it to look, so we'll drag the Manufacturer field to appear at the top of the Rows section above the Model, which makes more sense with our data.

Next, we'll add the Price field to the Columns section, ... ... but again that really isn't the way we want to view the data, so we'll drag Price to the Values section instead, which makes a lot more sense and looks a lot better.

Next, we'll add the Unit Sales field to Values too, so now we can see both the individual price for each model and the number of unit sales of each model.

Let's add the Vehicle-type field to Columns, but that doesn't seem very useful, so let's remove that field, ... , which we can do in two ways. Either by using the drop-down menu, ... ( or, if we undo that, ... we can also do it by simply dragging the field out of the Columns section, either to the left over the worksheet, or to the top over the fields list above.

Let's now look at how to perform a simple calculation in a Pivot Table. If we look in the 'Sum of Price' column in our Pivot Table, we can see that the figures are formatted as General.

So first, let's change the format for these figures to US currency. This can be done by modifying the value field settings for the field in the relevant section of the PivotTable Fields pane.

We'll format the field as US dollars and show no decimal places. Next, we'll add a calculated field from the 'PivotTable Analyze' tab, using the 'Fields, Items & Sets' button.

We want this field to calculate the total sales for each model by multiplying the price by the number of unit sales. When we create and add this formula, it gets added to the PivotTable Fields pane, as a field called Total Model Sales.

And we can change the format to make it US dollars again. A new column called ‘Sum of Total Model Sales’ has now appeared in the Pivot Table in our worksheet. In row 5 we can see that there have been over 360 million dollars of sales of the Acura Integra model, ... and in row 7 we can see that there has been over a billion dollars in sales of the Acura TL model.

In this video, we learned how to format data as a table, how to create a Pivot Table and use fields to analyze data in a Pivot Table, and how to perform calculations using Pivot Table data.

In the next video, we'll look at some other features of Pivot Tables.

## Introduction to Creating Pivot Tables in Excel

Look at:

- ✓ Formatting data as a table
- ✓ Creating a Pivot Table
- ✓ Using fields in a Pivot Table
- ✓ Performing calculations in a Pivot Table

Why use Pivot Table?

- Filters and formulas can help draw mathematical and logical conclusions about data
- Not all questions can be answered using filters and formulas alone
- To obtain usable and presentable insights into your data you need Pivot Tables

What is a Pivot Table?

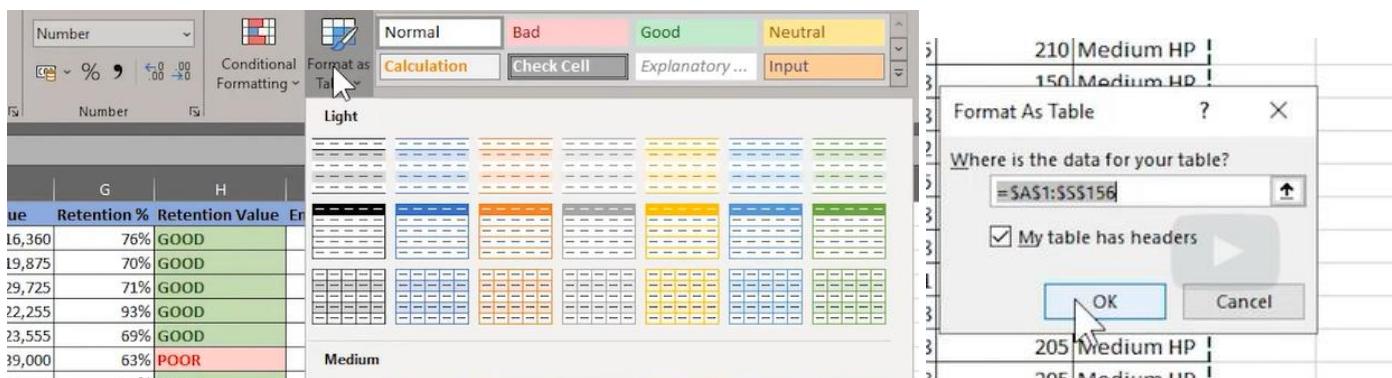
- Pivot Tables provide a simple and quick way, in spreadsheets, to summarize and analyze data, to observe trends and patterns in your data and to make data comparisons
- Pivot Tables are dynamic, so as data is changed so does the Pivot Table

- Data Analyst can use Pivot Tables to draw about, and create insights into, an organization's data
- 

## STEPS

### 1. Formatting data as a table to make it easier when adding records to the dataset.

1.1 In the car sales worksheet, let's first select any cell within the data, and then on the Home tab, in the Styles group, choose 'Format as Table'. Then choose a style from the gallery... note that Excel automatically knows the boundaries of our data range, but we can change this if we need to. And ensure you select 'My table has headers', if indeed it does.



1.2 After you click OK and the data has been formatted as a table, note the filter drop-downs at the top of each column – these are automatically added when you format as a table. If we now scroll down to the bottom of the table... and start adding another row of data for another vehicle... when you click Tab or Enter, note that it is automatically formatted and included as part of our table.

Model	Manufacturer	Vehicle_type	Unit Sales	Price
Integra	Acura	Passenger	16919	\$21,500
TL	Acura	Passenger	39384	\$28,400
RL	Acura	Passenger	8588	\$42,000

If we now scroll down to the bottom of the table... and start adding another row of data for another vehicle... when you click Tab or Enter, note that it is automatically formatted and included as part of our table.

S80	Volvo	Passenger	18969	\$36,000	\$26,403	73%	GOOD
E10 +	Zenos				#DIV/0!		#DIV/0!

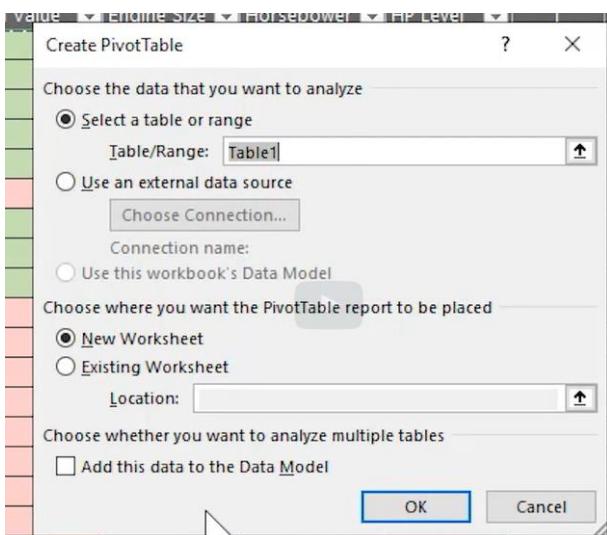
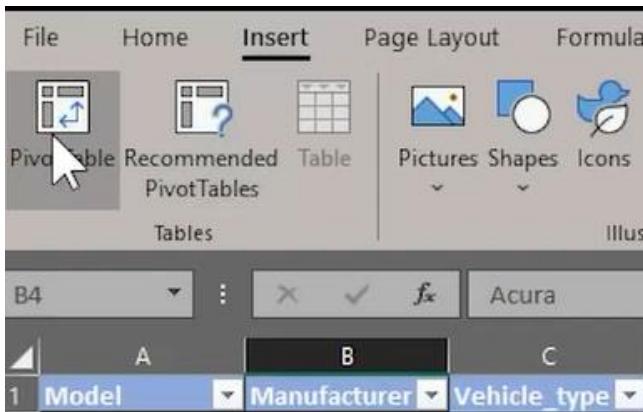
## Pivot Table Checklist

- Format your data as a table for best results
- Ensure column headings are correct, and there is only one header row, as these column headings become the field names in a Pivot Table
- Remove any blank rows and columns, and try to eliminate blank cells
- Ensure value fields are formatted as numbers, and not text
- Ensure date fields are formatted as dates, and not text

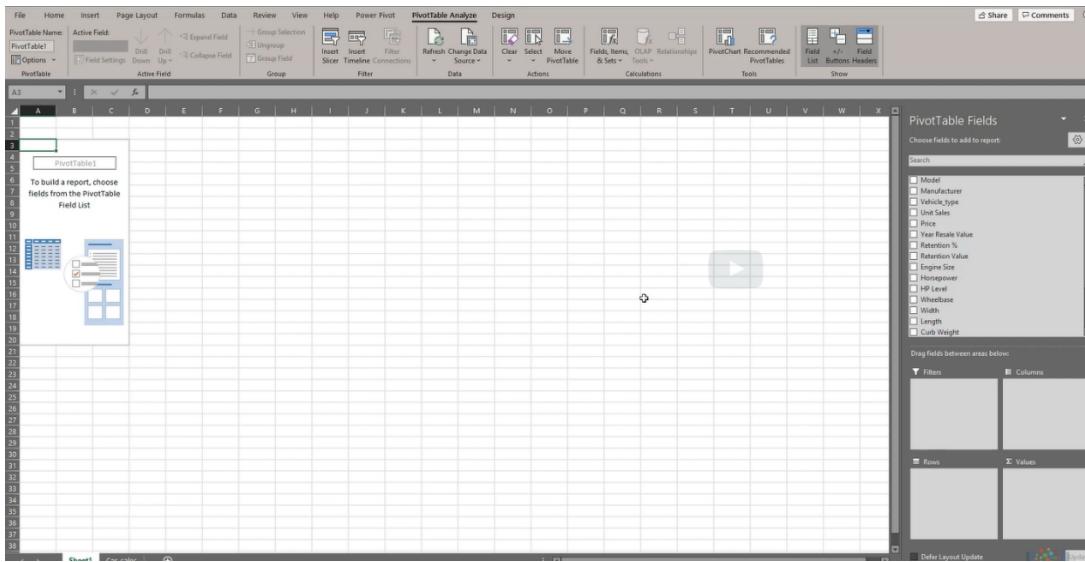
1.3 Just select any cell in the table then, on the Insert tab, we click PivotTable.

Then, on the Insert tab, we click PivotTable. Note that in the ‘Select a table or range’ box, the table name – Table1 – is already entered for us.

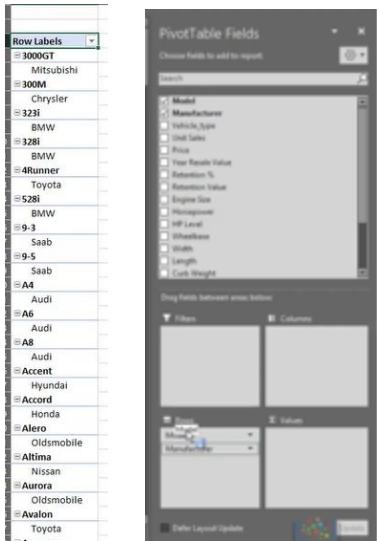
If we hadn’t just formatted this data as a table, we would specify the cell range here instead. Under that, we need to decide whether we want to create the Pivot Table on a separate new blank worksheet, or on this worksheet – a new worksheet is the default – and is the most commonly used option.



1.4 A new blank worksheet opens, displaying some basic Pivot Table instructions in the graphic on the left of the worksheet, and a ‘PivotTable Fields’ pane on the right. You can rename the worksheet for the Pivot Table if you wish.



To build the Pivot Table report we need to add some fields from the top of the PivotTable Fields pane, to one or more of the sections in the bottom part of the pane. For example, if we want to find out the total sales for each model of car, let’s drag the Manufacturer field to the Rows section of the report, ... and then we’ll drag the Model field there too.



But this isn’t really the way we want it to look, so we’ll drag the Manufacturer field to appear at the top of the Rows section above the Model, which makes more sense with our data.

Column Labels	\$9,235 \$9,699 \$10,685 \$11,528 \$11,799 \$12,050 \$12,070 \$12,315 \$12,535 \$12,640 \$12,885 \$13,108 \$13,260 \$13,499 \$13,960 \$13,987 \$14,290 \$14,460 \$14,900 \$14,999 \$15,010 \$15,350 \$15,900 \$16,080 \$16,240
Row Labels	\$9,235 \$9,699 \$10,685 \$11,528 \$11,799 \$12,050 \$12,070 \$12,315 \$12,535 \$12,640 \$12,885 \$13,108 \$13,260 \$13,499 \$13,960 \$13,987 \$14,290 \$14,460 \$14,900 \$14,999 \$15,010 \$15,350 \$15,900 \$16,080 \$16,240
= Acura Integra	

1.5 Add the Price field to the Columns section, ... ... but again that really isn't the way we want to view the data, so we'll drag Price to the Values section instead, which makes a lot more sense and looks a lot better.

Row Labels	Sum of Price
= Acura	91900
Integra	21500
RL	42000
TL	28400
= Audi	119940
A4	23990
A6	33950
A8	62000
= BMW	99290
323i	26990
328i	33400
528i	38900

1.6 Add the Unit Sales field to Values too, so now we can see both the individual price for each model and the number of unit sales of each model.

Drag fields between areas below:

Filters	Columns
	$\Sigma$ Values
Rows	$\Sigma$ Values
Manufacturer	Sum of Price
Model	Sum of Unit Sales
	Unit Sales

1.7 Let's add the Vehicle-type field to Columns, but that doesn't seem very useful, so let's remove that field, ... , which we can do in two ways. Either by using the drop-down menu, ... ( or, if we undo that, ... we can also do it by simply dragging the field out of the Columns section, either to the left over the worksheet, or to the top over the fields list above.

How to remove fields?

1. Either by using the drop-down menu, ... ( or, if we undo that, ...
2. Dragging the field out of the Columns section, either to the left over the worksheet, or to the top over the fields list above.

# How to perform a simple calculation in a Pivot Table

1. Change the format for these figures to US currency.

The 'Value Field Settings' dialog box shows the following settings:

- Source Name: Price
- Custom Name: Sum of Price
- Summarize Values By: Show Values As
- Summarize value field by:
  - Sum (selected)
  - Count
  - Average
  - Max
  - Min
  - Product

The 'Format Cells' dialog box shows the following settings:

- Category: Currency (selected)
- Sample: \$91,900
- Decimal places: 0
- Symbol: \$ English (United States)
- Negative numbers:
  - \$-1,234
  - \$1,234
  - \$1,234
  - \$1,234

2. Add a calculated field from the 'PivotTable Analyze' tab, using the 'Fields, Items & Sets' button.

The ribbon tabs are: Layout, Formulas, Data, Review, View, Help, Power Pivot, PivotTable Analyze (selected), Design.

The 'Fields, Items & Sets' button is highlighted in the PivotTable Analyze tab's ribbon group.

The dropdown menu for 'Fields, Items & Sets' includes:

- Calculated Item...
- Solve Order...
- List Formulas
- Create Set Based on Row Items...
- Create Set Based on Column Items...
- Manage Sets...

The Pivot Table data shows:

	Sum of Price	Sum of Unit Sales
\$91,900	64891	
\$21,500		16919
\$42,000		8588
\$28,400		39384

3. Add a calculated field from the 'PivotTable Analyze' tab, using the 'Fields, Items & Sets' button.

The ribbon tabs are: Layout, Formulas, Data, Review, View, Help, Power Pivot, PivotTable Analyze (selected), Design.

The 'Fields, Items & Sets' button is highlighted in the PivotTable Analyze tab's ribbon group.

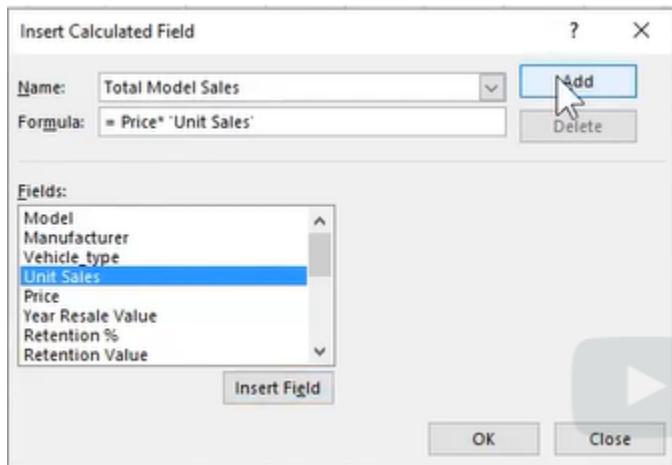
The dropdown menu for 'Fields, Items & Sets' includes:

- Calculated Item...
- Solve Order...
- List Formulas
- Create Set Based on Row Items...
- Create Set Based on Column Items...
- Manage Sets...

The Pivot Table data shows:

	Sum of Unit Sales
64891	
16919	

4. We want this field to calculate the total sales for each model by multiplying the price by the number of unit sales. When we create and add this formula, it gets added to the PivotTable Fields pane, as a field called Total Model Sales. And we can change the format to make it US dollars again.



5. A new column called 'Sum of Total Model Sales' has now appeared in the Pivot Table in our worksheet.

	Sum of Unit Sales	Sum of Total Model Sales
	64891	5963482900
	16919	363758500
	8588	360696000
	39384	1118505600
	40557	4864406580

6. In row 5 we can see that there have been over 360 million dollars of sales of the Acura Integra model, ... and in row 7 we can see that there has been over a billion dollars in sales of the Acura TL model.

	Sum of Unit Sales	Sum of Total Model Sales
	64891	\$5,963,482,900
	16919	\$363,758,500
	8588	\$360,696,000
	39384	\$1,118,505,600
	40557	\$4,864,406,580

## Viewpoints: Pivot Tables

What are your experiences using pivot tables to analyze data?

- My experience with pivot tables in Excel is extensive; I can use them all the time
- You can easily Sum, Average and Count
- Pivot tables are great if you've got a couple of thousand records up to Excel's maximum capability
- Pivot tables are a simply way of manipulation without having to do querying or use a development language
  
- I had a huge e-commerce sales dataset
- I needed to analyze the KPIs including the gross merchandise volume and take rate
- I could only generate limited insights if I stayed at a high level
- With pivot tables I was able to group the data in terms of countries, type of stores, type of products
- Enabled me to view the data and analyze the key KPI's at different levels of granularity.
  
- Pivot tables are especially useful when we look at large sets of data, for example, client audits
- Pivot tables help you take a large set of data and whittle it down to something that is meaningful
  
- We don't want to waste time looking at smaller invoice amounts which don't have as much impact on the audit and financial statements
- We use pivot tables to narrow it down to the invoices that are going to have the highest impact on the financial statement
- Pivot tables are a great way to understand your data quickly and effectively
- It's easy to open up a worksheet, put it into a pivot table, drag and drop things in to get a sense of what the numbers look like, what the values are

- Pivot tables can really help you get a good sense of the data
- Being able to understand the fields and what they mean can help you at the start of a project as you do your analysis
- Pivot tables are incredibly useful to get a quick view of your data and look at data at multiple levels of aggregation
- It's very easy to create a pivot table on a set of raw data aggregated by some level of interest, such as by country or date
- Pivot tables are really good for quickly seeing and understanding the more high-level summaries hidden within your data

# Pivot Table Features

Now that we've learned how to create and use Pivot Tables in Excel, in this video we'll look at some other features that we can use with Pivot Tables, including Recommended Pivot Tables, Filters, Slicers, and Timelines.

First, let's look at **Recommended Pivot Tables**, which isn't exactly a feature as such; it's really more of a list of suggested different combinations of data that could be used when creating a Pivot Table. These recommendations are based on the data we select in the worksheet, and they are a great way to get started creating Pivot Tables if you don't have much experience with them yet.

For example, in the vehicle toy sales worksheet, if we select column B, which contains data about the quantity of items ordered, ... when we choose Recommended Pivot Tables from the Insert tab, then we are presented with a list of potential data combinations related to the order quantity information.

However, if we select column F, which contains Order Size information, then the recommended pivot table list changes to reflect that data. And if we select column E, which contains sales information, then the pivot tables recommended are related to sales data.

Let's select the third one down, which is the sum of sales by territory; because that sounds like something we could get some useful insight from, by presenting it in a pivot table. Note that a new worksheet is opened containing the recommended pivot table, and a new pane opens on the right, called PivotTable Fields. Let's rename the worksheet to something more meaningful.

In the PivotTable Fields pane, you can see that some fields have already been added to the Rows and Values areas. Although it's a recommended pivot table, we can still make it our own, by adding more fields for example.

So, let's add the Productline item to the Columns area using drag and drop. Now we have columns for each of the product lines in our pivot table, such as motorcycles, ships, and trains.

In the pivot table, we can manually expand any field we want to view its contents. Here we can see that the order dates are located underneath the territory names in our pivot table.

Note that this matches the order of the fields in the Rows area of the PivotTable Fields pane. We can manually collapse each of the fields too. But we also have the option of expanding all the fields at once, ... and collapsing them all too.

The next feature we will delve into is **pivot table filtering**. Pivot table filters work in much the same way as the standard filters we used earlier in the course. Note that we already have some in-built filtering in this pivot table. For example, the Row Labels header is a filter, and we can filter on any of the listed territories, such as Japan. Just like standard filters, it's very simple to clear a filter in a pivot table.

We also have a Column Labels filter, allowing us to filter on any of the productline items in this pivot table; for example we could show data only for the trains product.

We also have the option of adding the Productline field as a standard filter instead of a column heading, by dragging it to the Filters area in the PivotTable Fields pane. And we can then use it as a standard filter, as we have done earlier in this course. The filter also allows us to select multiple filter items. But because it is now being used as a standard filter rather than a column header, we can't see the split of the information on these two product lines; we just see a combined total. When we had the filter as a column header, the information on each product line was presented separately in each column.

Let's display all the field totals again. And we'll drag the productline field back to the Columns area where it was previously, so we can see the split of our different product lines in the pivot table. The next pivot table feature we will look at are **Slicers**. **Slicers** are essentially on-screen graphical filter objects that enable you to filter your data using buttons.

Slicers make it easy to perform quick filtering of your pivot table data, and they also display the current filter state, making it easier for you to know, and see, what data is currently being shown, and which is being hidden, by the filter.

For example, if we remove the productline field from the pivot table by dragging it out of the PivotTable Fields pane, ... and then, from the PivotTable Analyze tab, we click Insert Slicer,... and then choose the Territory field as our slicer,... we can see that the slicer can be freely moved around anywhere on the worksheet, and it contains buttons for each of the territory items, such as EMEA, North America, and Japan.

We can also select the Multi-Select button to filter on multiple territories if we wish. We can click the Clear Filter button to clear all slicer filters. Let's add another slicer to our worksheet for the productline field. However, be sure to select a cell in the pivot table first, because if you don't, then the insert slicer button won't work.

Note that slicers can also be added from the Filters group on the Insert tab as well as from the PivotTable Analyze tab. We'll select the Productline field this time for our slicer, and drag it near the top of the worksheet.

As before, we can select only one slicer item, or we can turn on Multi-Select and choose several items to filter on in the slicer. Then let's clear the slicer filters, ... and now let's filter using both slicers. Note that when you use multi-select filtering, when you select an item, you are in fact filtering it out; that is, you are defining which items will NOT be displayed in the pivot table.

This is the opposite behavior to when you are selecting single items in a slicer. So now we are displaying only 'classic cars', 'trains', and 'trucks and buses' products for the EMEA and North America territories.

Now let's clear those slicer filters, and put the productline field back in the Columns area of the pivot table, so it's ready for the next feature we will explore. And let's move these slicers out of the way, further down the worksheet. The last useful feature for pivot tables we are going to look at, is Timelines.

A **Timeline** is another type of filter tool that enables you to filter specifically on date-related data in your pivot table. This is a much quicker and more effective way of dynamically filtering by date, rather than having to create and adjust filters on your date columns.

We can add a Timeline for our pivot table either from the PivotTable Analyze tab, or from the Insert tab. Again, ensure you select any cell in the pivot table first. We'll select the Orderdate field as our Timeline filter. Then we can drag it up the worksheet and enlarge it. The default for this timeline is to display data by month, but you can also filter by days, ... or by quarters.

You can select a single quarter; or you can select a range of quarters. In this case, we'll select twelve months between quarter 3 of 2003 and quarter 2 of 2004. You use the Clear Filter button to clear a timeline filter. You can also filter by years.

For example, here we have selected 2003 only. And you can combine slicers and timelines as filters in a pivot table. For example, here we can filter the slicers to display only data for trains, in the EMEA and North America territories, and only in the year 2003.

And if we filter on the year 2004 instead, you'll see that there is no data being displayed; meaning that there were no sales of train products in 2004 in either the EMEA or the North America territories.

Timelines and Slicers have their own tabs in the ribbon when you select them, and their properties can be modified to change how they look and how they work. For example, let's change this Timeline to a light green shade, ... and let's change this Slicer to a nice orange color.

And lastly, to remove a timeline or slicer, you can either select it and press the Delete key, ... or right-click it and choose Cut.

In this video, we learned about some of the other features in Excel that we can use with Pivot Tables, namely; Recommended Pivot Tables, Filters, Slicers, and Timelines.

# Pivot Table Features

Look at:

- Using Recommended Pivot Tables
- Using Filters in a Pivot Table
- Using Slicers in a Pivot Table
- Using Timelines in a Pivot Table

## Recommended Pivot Tables

- A list of suggested different combinations of data
- Based on the data we select in the worksheet
- A great way to get started creating Pivot Tables

The screenshot shows the Microsoft Excel ribbon with the 'Insert' tab selected. In the 'Tables' group, the 'Recommended PivotTables' button is highlighted. To the right, the 'Recommended PivotTables' dialog box is open, displaying several suggested pivot table configurations based on the selected data.

**Suggested Pivot Tables:**

- Sum of Sales by Order Size
- Sum of Sales by Status
- Sum of Sales by Territory
- Sum of Priceeach by Territory

**Sum of Sales by Order Size Data:**

Row Labels	Sum of Sales
Large	91044.69
Medium	291422.89
Small	145175.2
<b>Grand Total</b>	<b>527642.78</b>

**Sum of Sales by Status Data:**

Row Labels	Sum of Sales
Cancelled	23172.94
Resolved	6563.06
Shipped	497906.78
<b>Grand Total</b>	<b>527642.78</b>

**Sum of Sales by Territory Data:**

Row Labels	Sum of Sales
APAC	32280.72
EMEA	263940.95
Japan	33664.26
NA	197756.85
<b>Grand Total</b>	<b>527642.78</b>

**Sum of Priceeach by Territory Data:**

Row Labels	Sum of Priceeach
APAC	32280.72
EMEA	263940.95
Japan	33664.26
NA	197756.85
<b>Grand Total</b>	<b>527642.78</b>

**Sample Sales Dataset (Sales Table):**

OrderNumber	Quantityordered	Priceeach	OrderLineNum	Sales
10102	39	100	2	4808.31
10102	41	50.14	1	2055.74
10103	26	100	11	5404.62
10103	27	100	8	3394.98
10105	50	100	2	7208.00
10105	38	100	13	4330.10
10106	36	100	12	5279.40
10106	41	83.44	18	3421.04
10107	30	95.7	2	2871.00
10112	29	100	1	7209.11
10113	21	100	2	3415.44
10117	21	95.8	7	2011.80
10119	43	100	3	6916.12
10119	46	100	11	5004.80
10119	21	89.46	9	1878.66
10121	34	81.35	5	2765.90
10125	32	100	1	3254.72
10126	21	100	8	2439.57
10126	42	54.99	17	2309.58
10128	41	100	4	4837.18
10129	33	100	2	4398.24

For example, in the vehicle toy sales worksheet, if we select column B, which contains data about the quantity of items ordered, ... when we choose Recommended Pivot Tables from the Insert tab, then we are presented with a list of potential data combinations related to the order quantity information. However, if we select column F, which contains Order Size information, then the recommended pivot table list changes to reflect

that data. And if we select column E, which contains sales information, then the pivot tables recommended are related to sales data. Let's select the third one down, which is the sum of sales by territory; because that sounds like something we could get some useful insight from, by presenting it in a pivot table.

The screenshot shows a Microsoft Excel interface with a PivotTable on the left and a PivotTable Fields pane on the right.

**PivotTable Fields pane:**

- Choose fields to add to report: Search bar.
- Fields listed with checkboxes:
  - Priceeach
  - Orderlineumber
  - Sales
  - Order Size (IFS)
  - Orderdate
  - Status
  - Qtr\_Id
  - Month\_Id
  - Year\_Id
  - Productline
  - Productcode
  - Customername
  - Phone
  - Addressline1
  - Addressline2

**PivotTable grid:**

Sum of Sales	Column Labels	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
Row Labels		8850.09	3965.66				10911.94	8553.03	32280.72
APAC		92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
EMEA		7680.64		13203.48		2011.8	10768.34		33664.26
NA		74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
Grand Total		182796.63	65273.02	98050.33	38698.78	15904.36	72608.17	54311.49	527642.78

The screenshot shows a Microsoft Excel interface with a PivotTable containing detailed sales data for the EMEA region.

Sum of Sales	Column Labels	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
Row Labels		8850.09	3965.66				10911.94	8553.03	32280.72
APAC		92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
EMEA		7680.64		13203.48		2011.8	10768.34		33664.26
NA		74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
Grand Total		182796.63	65273.02	98050.33	38698.78	15904.36	72608.17	54311.49	527642.78
21-May-2003							3254.72		3254.72
16-Jul-2003							5298.31		5298.31
6-Nov-2003							6817.22		6817.22
20-Feb-2004			3965.66						3965.66
19-Jul-2004		4905.39					4094.72		9000.11
29-Nov-2004		3944.7							3944.7
EMEA		92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
29-Jan-2003		5404.62					3394.98		8799.6
11-Feb-2003		7208					4330.1		11538.1
17-Feb-2003				5279.4	3421.04				8700.44
24-Mar-2003		7209.11							7209.11
28-Apr-2003		5004.8		6916.12	1878.66				13799.58
7-May-2003			2765.9						2765.9
28-May-2003							2439.57		4749.15
6-Jun-2003									4837.18
12-Jun-2003		4398.24					4837.18		4398.24
1-Jul-2003			3884.34						3884.34
28-Sep-2003					2573.46				2573.46
6-Oct-2003		4526.08		6531.44	3482.16				14539.68
21-Oct-2003							6563.06		6563.06
23-Oct-2003		5924.16		5763.72	3363.52				15051.4
8-Nov-2003		3492.48		6490.68	3350.52				13333.68
11-Nov-2003			2497.77						2497.77
14-Nov-2003		3854.24		6004.8	2862.72				12721.76
18-Nov-2003			5512.32						5512.32

In the pivot table, we can manually expand any field we want to view its contents. Here we can see that the order dates are located underneath the territory names in our pivot table.

3 Sum of Sales Column Labels

4 Row Labels classic cars    motorcycles    planes

Select field:

Territory

A↓ Sort A to Z  
Z↓ Sort Z to A  
More Sort Options...  
Clear Filter From "Territory"

Label Filters >  
Value Filter >

Search

- (Select All)
- APAC
- EMEA
- Japan
- NA

Note that we already have some in-built filtering in this pivot table. For example, the Row Labels header is a filter, and we can filter on any of the listed territories, such as Japan. Just like standard filters, it's very simple to clear a filter in a pivot table. We also have a Column Labels filter, allowing us to filter on any of the productline items in this pivot table; for example we could show data only for the trains product.

Sum of Sales	Column Labels	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
Row Labels									
APAC		8850.09	3965.66				10911.94	8553.03	32280.72
EMEA		92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
Japan		7680.64		13203.48		2011.8	10768.34		33664.26
NA		74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
Grand Total		182796.63	65273.02	98050.33	38698.78	15904.36	72608.17	54311.49	527642.78

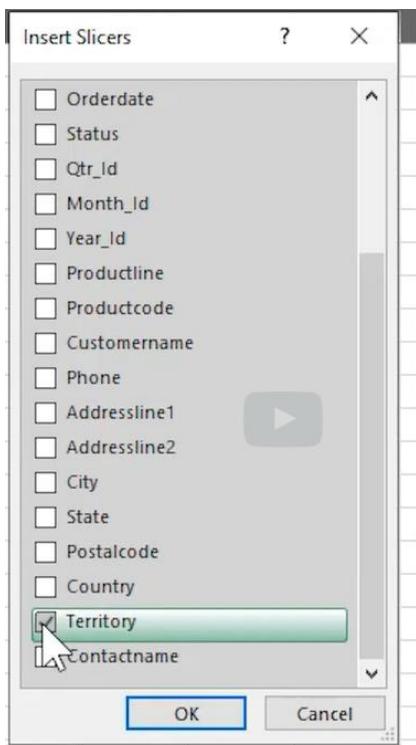
Let's display all the field totals again. And we'll drag the productline field back to the Columns area where it was previously, so we can see the split of our different product lines in the pivot table. The next pivot table feature we will look at are Slicers.

Slicers are essentially on-screen graphical filter objects that enable you to filter your data using buttons. Slicers make it easy to perform quick filtering of your pivot table data, and they also display the current filter state, making it easier for you to know, and see, what data is currently being shown, and which is being hidden, by the filter.

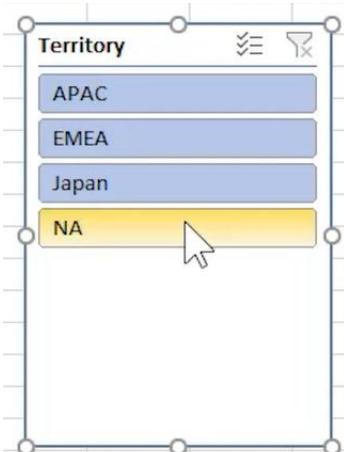
File	Home	Insert	Page Layout	Formulas	Data	Review	View	Help	Power Pivot	PivotTable Analyze
PivotTable Name:	PivotTable5	Active Field:	Sum of Sales	↑ Expand Field ↓ Drill Down Field Settings	↑ Drill Up Collapse Field	Group Selection Ungroup Group Field	Insert Slicer Insert Timeline Filter	Insert Connections Refresh Data Source	Change Data Source	Data
PivotTable				Active Field		Group				
B3										

Sum of Sales

	A	B	C	D	E	F	G	H	I	J
1										
2										
3	Row Labels	Sum of Sales								
4	APAC	32280.72								
5	EMEA	263940.95								
6	Japan	33664.26								
7	NA	197756.85								
8	Grand Total	527642.78								



For example, if we remove the productline field from the pivot table by dragging it out of the PivotTable Fields pane, ... and then, from the PivotTable Analyze tab, we click Insert Slicer,... and then choose the Territory field as our slicer,... we can see that the slicer can be freely moved around anywhere on the worksheet, and it contains buttons for each of the territory items, such as EMEA, North America, and Japan.



We can see that the slicer can be freely moved around anywhere on the worksheet, and it contains buttons for each of the territory items, such as EMEA, North America, and Japan.

\*\*Be sure to select a cell in the pivot table first, because if you don't, then the insert slicer button won't work.

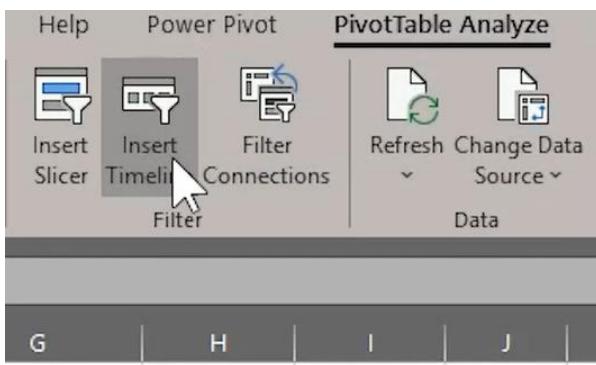
Row Labels	Sum of Sales
EMEA	156029.59
NA	120815.22
Grand Total	276844.81

Note that when you use multi-select filtering, when you select an item, you are in fact filtering it out; that is, you are defining which items will NOT be displayed in the pivot table. This is the opposite behavior to when you are selecting single items in a slicer. So now we are displaying only 'classic cars', 'trains', and 'trucks and buses' products for the EMEA and North America territories.

The screenshot shows a PivotTable with the following data:

	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
APAC	8850.09	3965.66				10911.94	8553.03	32280.72
EMEA	92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
Japan	7680.64		13203.48		2011.8	10768.34		33664.26
NA	74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
<b>Grand Total</b>	<b>182796.63</b>	<b>65273.02</b>	<b>98050.33</b>	<b>38698.78</b>	<b>15904.36</b>	<b>72608.17</b>	<b>54311.49</b>	<b>527642.78</b>

Now let's clear those slicer filters, and put the productline field back in the Columns area of the pivot table, so it's ready for the next feature we will explore. And let's move these slicers out of the way, further down the worksheet.



A Timeline is another type of filter tool that enables you to filter specifically on date-related data in your pivot table. This is a much quicker and more effective way of dynamically filtering by date, rather than having to create and adjust filters on your date columns. We can add a Timeline for our pivot table either from the PivotTable Analyze tab, or from the Insert tab. Again, ensure you select any cell in the pivot table first.

The screenshot shows the following data:

	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
APAC	3965.66				10911.94	8553.03	32280.72
EMEA	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
Japan		13203.48		2011.8	10768.34		33664.26
NA	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
<b>Grand Total</b>	<b>65273.02</b>	<b>98050.33</b>	<b>38698.78</b>	<b>15904.36</b>	<b>72608.17</b>	<b>54311.49</b>	<b>527642.78</b>

The screenshot shows a BI interface with a table of sales data and a modal dialog for inserting timelines.

	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
APAC	8850.09	3965.66				10911.94	8553.03	32280.72
EMEA	92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
Japan	7680.64		13203.48		2011.8	10768.34		33664.26
NA	74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
<b>Grand Total</b>	<b>182796.63</b>	<b>65273.02</b>	<b>98050.33</b>	<b>38698.78</b>	<b>15904.36</b>	<b>72608.17</b>	<b>54311.49</b>	<b>527642.78</b>

**Insert Timelines**

Orderdate

OK Cancel

Below the table are two filters: Territory (APAC) and Productline (Automobiles).

We'll select the Orderdate field as our Timeline filter. Then we can drag it up the worksheet and enlarge it. The default for this timeline is to display data by month, but you can also filter by days, ... or by quarters. You can select a single quarter; or you can select a range of quarters.

The screenshot shows a BI interface with a table of sales data and a timeline filter.

	classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
APAC	8850.09	3965.66				10911.94	8553.03	32280.72
EMEA	92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95
Japan	7680.64		13203.48		2011.8	10768.34		33664.26
NA	74125.03	30769.47	29868.04	16304.12	2151.82	27836.94	16701.43	197756.85
<b>Grand Total</b>	<b>182796.63</b>	<b>65273.02</b>	<b>98050.33</b>	<b>38698.78</b>	<b>15904.36</b>	<b>72608.17</b>	<b>54311.49</b>	<b>527642.78</b>

**Orderdate**

All Periods

2004

JN JUL AUG SEP OCT NOV DEC

MONTHS

We'll select the Orderdate field as our Timeline filter. Then we can drag it up the worksheet and enlarge it. The default for this timeline is to display data by month, but you can also filter by days, ... or by quarters. You can select a single quarter; or you can select a range of quarters.

The screenshot shows a timeline filter for the Orderdate field, displaying a range from Q3 2003 to Q2 2004.

**Orderdate**

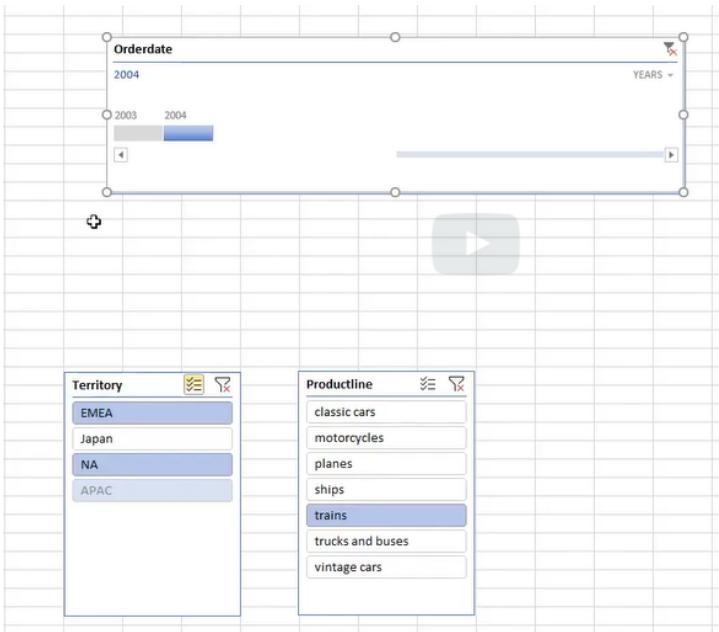
All Periods

2003 2004

Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4

QUARTERS

In this case, we'll select twelve months between quarter 3 of 2003 and quarter 2 of 2004. You use the Clear Filter button to clear a timeline filter. You can also filter by years.



For example, here we can filter the slicers to display only data for trains, in the EMEA and North America territories, and only in the year 2003. And if we filter on the year 2004 instead, you'll see that there is no data being displayed; meaning that there were no sales of train products in 2004 in either the EMEA or the North America territories.

	Sum of Sales	Column Labels							
Row Labels		classic cars	motorcycles	planes	ships	trains	trucks and buses	vintage cars	Grand Total
# APAC	8850.00	3065.66				10911.94	8553.03	32280.72	
# EMEA	92140.87	30537.89	54978.81	22394.66	11740.74	23090.95	29057.03	263940.95	
# Japan	7680.64		13203.48		2011.8	10768.34		33664.26	
# NA	74125.03	30769.47	29668.04	16304.12	2151.82	27836.94	16701.43	197756.85	
Grand Total	182796.63	65273.02	98050.33	38698.78	15904.36	72608.17	54311.49	527642.78	

Timelines and Slicers have their own tabs in the ribbon when you select them, and their properties can be modified to change how they look and how they work. For example, let's change this Timeline to a light green shade, ... and let's change this Slicer to a nice orange color. And lastly, to remove a timeline or slicer, you can either select it and press the Delete key, ... or right-click it and choose Cut.



# Hands-on Lab 7: Using Pivot Tables

**Estimated time needed:** 30 minutes

In this lab, first you will learn how to format data as a table, how to create a Pivot Table and use fields to arrange data in a Pivot Table, and how to perform calculations using Pivot Table data. Next, you will learn some other features that we can use with Pivot Tables, including Recommended Charts, Filters, Slicers, and Timelines.

## Software Used in this Lab

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

## Dataset Used in this Lab

The dataset used in this lab comes from the following source: <https://www.kaggle.com/sudalairajkumar/indian-startup-funding> under a **CC0: Public Domain license**. Acknowledgement and thanks also goes to <https://trak.in> who were generous enough to share the data publicly for free.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

# Objectives

After completing this lab, you will be able to:

- Format data as a table
- Create a Pivot Table and use fields to arrange data in a Pivot Table
- Perform calculations using Pivot Table data
- Use the Recommended Charts feature (does not work with the 'Basic' Office for the web plan.)
- Use the Filters feature
- Use the Slicers feature
- Use the Timelines feature

## Exercise 1: Introduction to Creating Pivot Tables in Excel

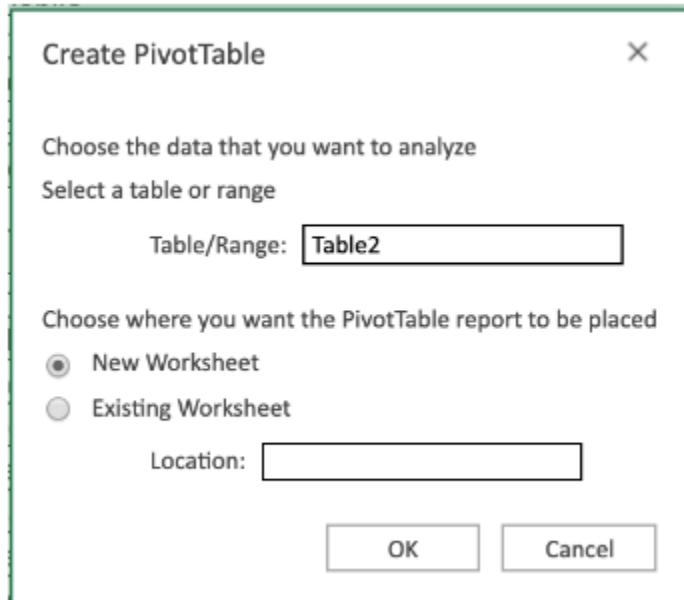
In this exercise, you will learn how to format data as a table, how to create a Pivot Table and use fields to arrange data in a Pivot Table, and how to perform calculations using Pivot Table data.

### Task A: Format data as a table

1. Download the file [\*\*indian\\_startup\\_funding\\_Lab7.xlsx\*\*](#). Upload and open it using Excel for the web.
2. Select cell **A2**.
3. On the **Home** tab, in the **Tables** group, click **Format as Table**.
4. Select **Light Gray, Table Style Medium 15**.

### Task B: Create a pivot table and use fields to arrange data in a pivot table

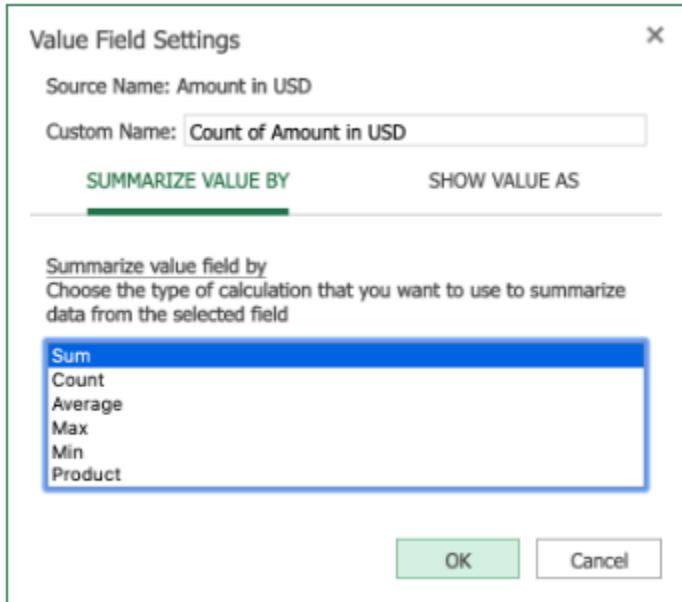
1. Select cell **D4**
2. On the **Insert** tab, click **PivotTable**.
3. Click **OK**.



4. Double-click **Sheet1**, type **Pivot1** and click **OK**.
5. In the fields list, drag **Industry Vertical** to **Rows**.
6. In the fields list, drag **City Location** to **Rows** below **Industry Vertical**.
7. In the fields list, drag **Startup Name** to **Rows** below **City Location**.
8. In the fields list, drag **Amount in USD** to **Values**.
9. In the ribbon, select the **PivotTable** tab, click **Settings**, then in the **PivotTable Settings** pane, under **Layout**, select **Single column**.

### Task C: Perform a simple calculation in a pivot table

1. In the **PivotTable Fields** pane, in the **Values** section, click the drop-down arrow next to **Count of Amount in USD**, and click **Value Field Settings**.
2. Select **Summarize value field by > Sum**.



3. Click **OK**.
4. Select the column called **Sum of Amount in USD** and then on the **Home** tab, select **Accounting Number Format > \$ English (United States)**.

## Exercise 2: Pivot Table Features

In this exercise, you will learn some other features that we can use with Pivot Tables, including Recommended Charts, Filters, Slicers, and Timelines.

**Note:** The 'Recommended Charts' feature only works with 'full' Office for the web plans (those plans that come with an Office 365 subscription). Recommended Charts do not work with the 'basic' plan that comes with a Microsoft Account.

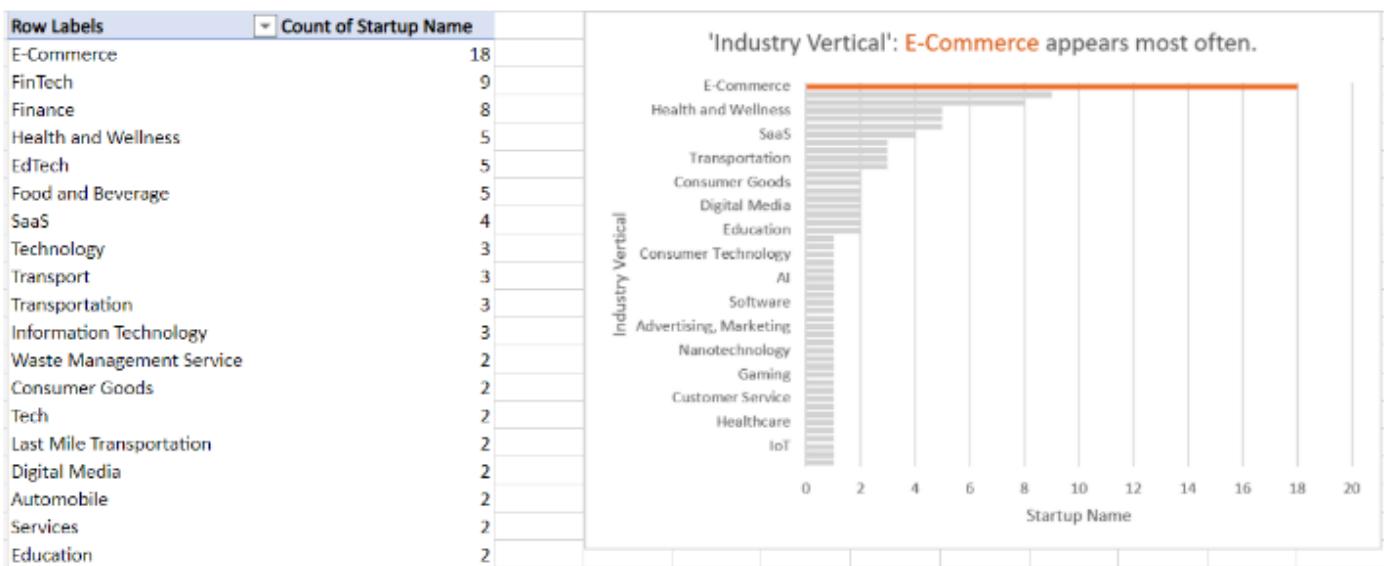
**Task A:** Use of the Recommended Charts feature (Optional: If you have a full Office for the web plan)

1. Switch to worksheet **indian-startup-funding**.
2. Select column **F (City Location)**.
3. On the **Insert** tab, select **Recommended Charts**.
4. Click + **Insert PivotChart**.

The screenshot shows a PivotTable on the left with columns 'City Location' and 'Investors Name'. A 'Recommended Charts' dialog box is open on the right, displaying a bar chart titled 'Count of City Location' where Bengaluru is the most frequent location. Below the chart are buttons for 'Insert PivotChart' and 'Is this helpful?'.

City Location	Investors Name
Faridabad	SoftBank Vision Fund
Bengaluru	Paytm, NPTK, Sabre Partners and Neoplux
Bengaluru	Vertex Growth Fund
Mumbai	
Bengaluru	RuiZheng Investment
Gurgaon	SAIF Partners, Spring Canter Investment L
Pune	Sathguru Catalyst Advisors
Noida	Manipal Education and Medical Group (M
Gurgaon	Ping An Global Voyager Fund
Bengaluru	Mumbai Angels, Ravikanth Reddy
Noida	Vijay Shekhar Sharma
Gurgaon	FinTech
San Jose,	Altimeter Capital, Sutter Hill Ventures
Delhi	Amour Infrastructure

5. Switch to worksheet **indian-startup-funding** again.
6. Select column **C, D, E**.
7. On the **Insert** tab, select **Recommended Charts**.
8. Choose the recommended chart, and click **+ Insert PivotChart**.



## Task B: Use of the Filters feature

1. Switch to worksheet **Pivot1**.
2. In the Pivot Table, click the **Row Labels** arrow.
3. Select **City Location**, then **Filter....**
4. Just select **Burnsville, Delhi, New York**, then click **OK** to display the amounts for startups in these three cities only.
5. In the Pivot Table, click the **Row Labels** arrow.

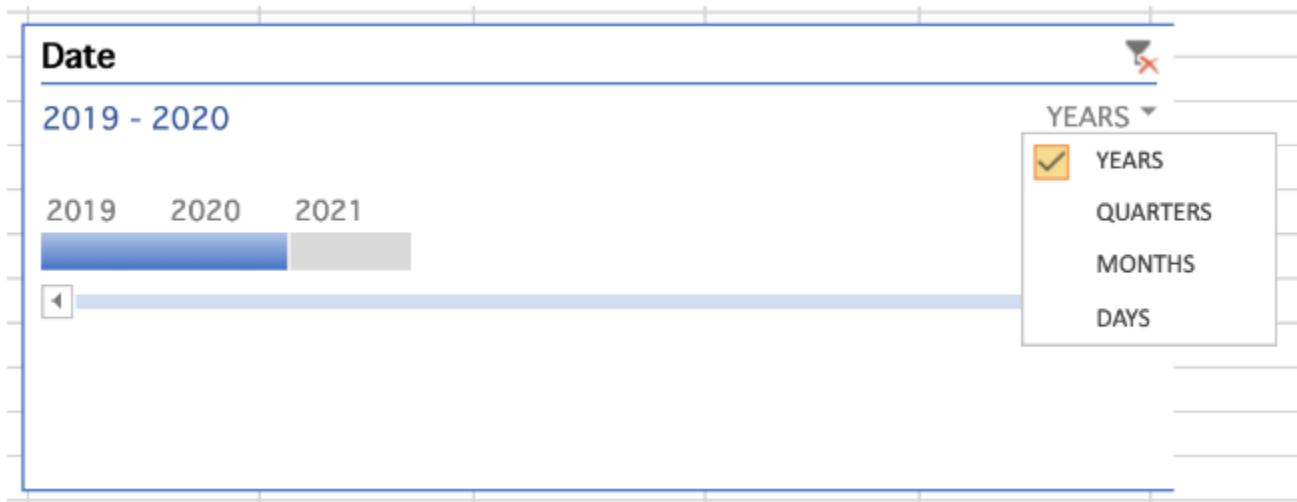
- Select **City Location**, then click **Clear Filter From 'City Location'** to display the startups in all cities again.

## Task C: Use of the Slicers feature

- Download the file [indian startup funding Lab7 with slicers timelines.xlsx](#). Upload and open it using Excel for the web.
- Switch to worksheet **Pivot1** if you are not there.
- In the **City Location** slicer, select **Burnsville**, then **Delhi**, then **New York**.
- To filter by multiple selection in the **City Location** slicer, with **New York** still selected, press **CTRL** and select **Burnsville**, and then **Delhi**.
- To filter using more than one slicer, in the **Investors Name** slicer, select **Amour Infrastructure**, then press **CTRL** and select **Westbridge Capital**, and then **Breakthrough Energy Ventures**.
- In the **City Location** slicer, click the **Clear Filter** button, then in the **Investors Name** slicer, click the **Clear Filter** button.

## Task D: Use of the Timelines feature

- In the Date timeline, click **top right drop-down** and select **DAYS**, then scroll **left and right**.



- In the Date timeline, click **top right drop-down** and select **QUARTERS**.
- In the Date Timeline, select **2019 Q1**, then drag **2019 Q1 to 2019 Q3**.
- In the Date timeline, click the **Clear Filter** icon.
- In the Date timeline, click **top right drop-down** and select **YEARS**, then select **2020** only.

# Reading: Summary and Highlights

In this lesson, you have learned the following information:

Pivot Tables:

- To obtain usable and presentable insights into your data you need to use Pivot Tables.
- Pivot tables provide a simple and quick way to summarize and analyze data, to observe trends and patterns in your data and to make comparisons of your data.
- Pivot tables are dynamic, so as you change and add data to the original dataset on which the pivot table is based, the analysis and summary information changes too.
- A Data Analyst can use pivot tables to draw useful and relevant conclusions about, and create insights into, an organization's data in order to present those insights to interested parties within the company.

Use this Pivot Table checklist to ensure your data is in a fit state to make a Pivot Table:

- Format your data as a table for best results.
- Ensure column headings are correct, and there is only one header row, as these column headings become the field names in a Pivot Table.
- Remove any blank rows and columns, and try to eliminate blank cells also.
- Ensure value fields are formatted as numbers, and not text, and ensure date fields are formatted as dates, and not text.

Arranging Pivot Tables with Filters and Recommended Tables:

- You use the Pivot Table Fields pane to add and arrange data fields in your pivot table.
- Recommended Pivot Tables are a list of suggested different combinations of data that could be used when creating a Pivot Table, based on the data selected in the worksheet.

Filters and Slicers:

- Slicers are on-screen graphical filter objects that enable you to filter your data using buttons, which makes it easier to perform quick filtering of your pivot table data.
- Timelines are another type of filter tool that enable you to filter specifically on date-related data in your pivot table. This is a much quicker and more effective way of dynamically filtering by date, rather than having to create and adjust filters on your date columns.

## Module 6: Practice Quiz

Bookmarked

### Question 1

1/1 point (ungraded)

Before creating a pivot table, how should you format your data?

- Cell size
- Sort and filter
- As a table
- Conditional formatting



### Question 2

1/1 point (ungraded)

How can you add more filters to the pivot table?

- use a formula
- drag a field to the Filters area of the PivotTable Fields pane
- use a function
- first add filters to the original table data



Submit

✓ Correct (1/1 point)

### Question 3

1/1 point (ungraded)

What are slicers?

- split pivot tables
- on-screen graphical filter objects
- standard filters
- header filters



Submit

✓ Correct (1/1 point)

## Module 6: Graded Quiz

Bookmarked

Graded Quiz due Aug 21, 2022 20:48 +08

### Question 1

1/1 point (graded)

According to the video checklist, what should you remove before making a Pivot Table?

Value fields

Blank rows, columns, and cells

Row labels

Date fields



### Question 2

1/1 point (graded)

What is automatically added after formatting data as a table?

Filter drop-downs at top of columns

Column headers

Alternate light/dark rows

Data bars



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

### Question 3

1/1 point (graded)

After creating a pivot table and selecting it, what pane appears to the right of the pivot table?

- The PivotTable Properties pane
- The PivotTable Metadata pane
- The PivotTable Fields pane
- Additional pivot table examples



Submit

You have used 1 of 2 attempts

Rese

✓ Correct (1/1 point)

### Question 4

1/1 point (graded)

What do timelines provide in pivot tables?

- provides details of all actions performed during a given period
- sets the pivot table to an earlier view
- Lets you filter specifically on date-related data
- previews views of different filters



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

### Question 5

1/1 point (graded)

What is one way to remove a slicer or timeline?

- Create a new one
- Press CTRL+Z
- Right-click it and select cut
- Remove all filters from pivot table



Submit

You have used 1 of 2 attempts

✓ Correct (1/1 point)

# Introduction to Final Project - Part 1

Now that you are equipped with the skills to clean and prepare data using Excel, you will have the opportunity to practice and apply your skills on a real world data set.

In this scenario, you are a Junior Data Analyst who has recently joined a growing company that markets classic and collector cars, motorcycles, and other vehicles to customers around the world. You have been tasked with cleaning data that has been imported from the sales department. The data is in comma-separated value (CSV) format and needs to be cleaned before you begin your analysis. In Part 2, you will be tasked with running an analysis of the sales data using pivot tables.

## Grading Criteria

After completing both Project - Part 1 and Project - Part 2, you will complete a Peer-graded Final Assignment. Your grade will be based on completing the following tasks:

In addition to the tasks below, you are required to submit the following URL:

- Excel file named **sales\_data\_sample\_PART1.XLSX** for Part 1 of the final assignment.

### Part 1:

Task 1: Save the CSV file saved as an XLSX file?

Task 2: Widen all columns so that all the data is visible

Task 3: Remove all the empty rows

Task 4: Remove duplicated records

Task 5: Sort column A

Task 6: Format the numbers

Task 7: Remove all double-spaces from the data

Task 8: Combine the two columns into a single column

Task 9: Save the Workbook

## Peer-graded Assignment Part 1 – Clean and Prepare the Data:

**Estimated time needed:** 45 minutes

In this hands-on lab you will open a CSV file in Excel for the web, convert it to an Excel format, and then clean and prepare the data.

### Software Used in this Assignment

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

### Dataset Used in this Assignment

The dataset used in this lab comes from the following source: <https://www.kaggle.com/kyanyoga/sample-sales-data> under a **Public Domain license**.

We are using a modified subset of the dataset for the lab. In order to follow the lab instructions successfully, please use the dataset provided with the lab, rather than the dataset from the original source.

### Guidelines for Submission

Download the file [sales\\_data\\_sample.CSV](#).

Upload and open the file with Excel for the web and convert it to an .XLSX file. Then, clean the data as detailed below.

Use the course videos from Module 3 and the Lab: Hands-on Lab 5: Cleaning Data to help you complete these tasks.

Tasks to perform:

1. **Save the CSV file as an XLSX file:** Click the 'Edit Workbook' button in the ToolTip to save the file as an XLSX file. The file is converted when you select 'OK' in the prompt.
2. **Column widths:** Adjust the widths of all columns so that the data is clearly visible in all cells.
3. **Empty rows:** Use the Filter feature to look for blanks and remove all empty rows from the data.
4. **Duplicate records:** Use either the Conditional Formatting or Remove Duplicates feature to look for and remove any duplicated records from the data based on the contents of Column A.
5. **Sorting:** Sort the data in Column A from smallest to largest.
6. **Format Cells:** In column C, set the type as Number and format it to remove the decimals. In column E, set the type as Number and format it to show up to two decimals.
7. **Whitespace:** Use the Find and Replace feature to remove all double-spaces from the data.
8. **Contact Name:** The contact name is divided into two columns as CONTACTLASTNAME (Column W) and CONTACTFIRSTNAME (Column X). Use Flash Fill to reduce the names to just one column, and then remove any unnecessary columns. Show the column values as FirstName LastName.
9. **Save your workbook:** Use 'Save As' to save your completed workbook as **sales\_data\_sample\_PART1.XLSX**.

# **Introduction to Final Project - Part 2**

Now that you are equipped with the skills to take clean and prepared data and create pivot tables, you will have the opportunity to practice and apply your skills on a data set.

## **Grading Criteria**

After completing both Projects, Part 1 and Part 2, you will complete a Peer-graded Final Assignment. Your grade will be based on completing the following tasks.

In addition to the tasks below, you are required to submit the following URL:

- Excel file named **sales\_data\_sample\_PART2.XLSX** for Part 2 of the final assignment.

### **Part 2:**

Task 10: Is the Format as Table option applied correctly?

Task 11: Are the five correct AutoSum values calculated correctly?

Task 12: Is the pivot table created correctly?

Task 13: Is the pivot table data sorted correctly?

Task 14: Has the pivot table been created two more times?

Task 15: Has the workbook been saved correctly?

## Peer-graded Assignment Part 2 – Analyze the Data:

**Estimated time needed:** 45 minutes

In this hands-on lab you will take the cleaned and prepared data and create pivot tables to help you analyze the data.

### Software Used in this Assignment

The instruction videos in this course use the full Excel Desktop version as this has all the available product features, but for the hands-on labs we will be using the free 'Excel for the web' version as this is available to everyone.

Although you can use the Excel Desktop software if you have access to this version, it is recommended that you use Excel for the web for the hands-on labs as the lab instructions specifically refer to this version, and there are some small differences in the interface and available features.

### Dataset Used in this Assignment

The dataset used in this lab comes from the following source: <https://www.kaggle.com/kyanyoga/sample-sales-data> under a **Public Domain license**.

We are using a modified subset of that dataset for the lab, so to follow the lab instructions successfully please use the dataset provided with the lab, rather than the dataset from the original source.

### Guidelines for Submission

Download and open the [sales data sample PART2 Start.xlsx](#) file in Excel for the web.

Use the course videos from Module 4 and the Lab; Hands-on Lab 7: Using Pivot Tables to help you complete these tasks.

Tasks to perform:

1. **Format the data as a table:** Use the Insert, Table command to format the data as a table.
2. **Use AutoSum to calculate values:** Use AutoSum to find the following values for column 'E' and record each of the values:
  - o SUM
  - o AVERAGE
  - o MIN
  - o MAX
  - o COUNT
3. **Create a Pivot Table:** Use the PivotTable feature to create a pivot table that displays the PRODUCTLINE field in the Rows section, and the QUANTITYORDERED in the Values section, so that the pivot table displays the total quantity count by product.
4. **Sort the pivot table data:** Use the Sort By Value setting on the pivot table to sort it in descending order by the sum of total quantity count.
5. **Make two more pivot tables that are the same as the one you created in task 3:** Follow the same steps you performed in Tasks 3 and 4 to create two more identical pivot tables so that you end up with 3 worksheets that contain identical pivot tables.
6. **Analyze data in the pivot table:** Use the PivotTable Fields pane to manipulate and analyze data in the two copied pivot table as follows:
  - o In pivot table 2, add the STATUS field below the PRODUCTLINE field so that the status of order shows up with their respective counts.
  - o Collapse all fields except the top one – **Classic Cars**.
  - o In pivot table 3, add the STATUS field above the PRODUCTLINE field so that the different status appears first, with the different product listed underneath each status with their respective counts.
  - o Collapse all fields except the top one – **Cancelled**.
7. **Save your workbook:** Use 'Save As' to save your workbook as **sales\_data\_sample\_PART\_2.XLSX**.

# Peer-graded Final Assignment for Part 1 and Part 2

## OVERVIEW

Please ensure you have successfully completed both Final Project - Part 1 and Final Project - Part 2 prior to taking the Peer-graded Final Assignment. This Final Assignment will be graded by your peers who are also completing this assignment within the same session.

## GRADING CRITERIA

There are a total of **20 points** possible for this final project. This Final Assignment is 20% of your total grade.

In addition to the tasks below, you are required to submit the following files:

- Excel file named **sales\_data\_sample\_PART1.XLSX** for Part 1 of the final assignment.
- Excel file named **sales\_data\_sample\_PART2.XLSX** for Part 2 of the final assignment.

**Your grade will be based on completing the following tasks:**

### Part 1:

Task 1: Is the CSV file saved as an XLSX file?

Task 1: Are all columns widened so that all the data is visible?

Task 3: Are all empty rows removed correctly?

Task 4: Are duplicated records removed correctly?

Task 5: Is column A sorted correctly?

Task 6: Are the numbers formatted correctly?

Task 7: Are all double-spaces removed from the data correctly?

Task 8: Are the two columns combined into a single column correctly?

Task 9: Has the workbook been saved correctly?

## **Part 2:**

Task 10: Is the Format as Table option applied correctly?

Task 11: Are the five correct AutoSum values calculated correctly?

Task 12: Is the pivot table created correctly?

Task 13: Is the pivot table data sorted correctly?

Task 14: Has the pivot table been created two more times?

Task 15: Has the workbook been saved correctly?

"C:\Users\mtina\OneDrive\Desktop\sales\_data\_sample\_PART1.XLSX"

"C:\Users\mtina\OneDrive\Desktop\sales\_data\_sample\_PART2.xlsx"