**Excel basics for data analysis**

**Module 1**

**Reading: excel keyboard shortcut**

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 |
| sBold | 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 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 (;) |

* 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 2**

**Getting started using spreadsheets**

* 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 3**

**3.1 Basics of data quality and privacy**

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

**3.2 Cleaning data**

* 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:

1. Change the case of text
2. Fix date formatting errors
3. Trim whitepace 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.
* Spelling
* Remove empty row
* Remove duplicate
* Replace text
* Fill flash
* =UPPER,LOWER,PROPER,LEFT,RIGHT,SEARCH,LEN

**Module 4**

**4.1 Data analysis basics, filtering and sorting data**

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

**More function:**

* COUNTIFS
* SUMIFS

| **Function** | **Purpose** | **Syntax** | **Example** | **Comparison** |
| --- | --- | --- | --- | --- |
| **COUNTIFS** | Counts cells based on multiple criteria | COUNTIFS(criteria\_range1, criteria1, [criteria\_range2, criteria2], ...) | =COUNTIFS(B2:B6, "North", C2:C6, ">500") | Handles multiple criteria |
| **COUNTIF** | Counts cells based on a single criterion | COUNTIF(criteria\_range, criteria) | =COUNTIF(B2:B6, "North") | Handles a single criterion |
| **SUMIFS** | Sums cells based on multiple criteria | SUMIFS(sum\_range, criteria\_range1, criteria1, [criteria\_range2, criteria2], ...) | =SUMIFS(C2:C6, B2:B6, "North", C2:C6, ">500") | Handles multiple criteria |
| **SUMIF** | Sums cells based on a single criterion | SUMIF(criteria\_range, criteria, sum\_range) | =SUMIF(B2:B6, "North", C2:C6) | Handles a single criterion |
| **XLOOKUP** | Searches a range and returns a match | XLOOKUP(lookup\_value, lookup\_array, return\_array, [if\_not\_found], [match\_mode], [search\_mode]) | =XLOOKUP("Cherry", B2:B5, C2:C5) | Flexible, handles both vertical and horizontal lookups, better error handling |
| **VLOOKUP** | Searches vertically and returns a match | VLOOKUP(lookup\_value, table\_array, col\_index\_num, [range\_lookup]) | =VLOOKUP("Cherry", B2:C5, 2, FALSE) | Searches vertically, left to right only |
| **HLOOKUP** | Searches horizontally and returns a match | HLOOKUP(lookup\_value, table\_array, row\_index\_num, [range\_lookup]) | =HLOOKUP("Cherry", B1:D2, 2, FALSE) | Searches horizontally, top to bottom only |

**Comparison Table**

**4.2 Using pivot tables**

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.