

HR Data Analysis in Excel - A Step-by-Step Guide

<https://www.aihr.com/blog/hr-data-analysis-excel/>

Posted by [Paul van der Laken](#)

Contents

HR data analysis process	2
Step 1: Data cleaning	2
1. Excel Tables.....	2
2. Remove duplicates	3
3. Sorting	4
4. Converting data	5
VALUE	5
TEXT.....	5
FIXED	5
5. Text manipulation	7
PROPER	7
LEFT & RIGHT	7
SUBSTITUTE	7
6. Find and Replace	8
7. Paste special.....	8
Step 2: Data analysis	9
1. Slicers and Filters	9
Step 3: Storytelling.....	9
Step 4: Dashboarding.....	10

Microsoft Excel is the market leader when it comes to data analysis, both in HR and other business functions. While it is no substitute for an HR Information System and does not offer the most advanced people analytics capabilities, it is the all-time favorite for quick analysis and data visualization.

Excel provides HR professionals with a dynamic, relatively easy-to-use analysis tool. This article showcases some lesser-known Excel tools and functions that will help you power up your HR data analysis capabilities.

HR data analysis process

This article goes through four common steps in any HR data analysis project.

- Cleaning your data
- Analyzing your data
- Storytelling
- Dashboarding

A subset of [Kaggle's IBM HR Analytics Attrition dataset](#) is used for the examples. If you want to put your own Excel skills to the test, the data is available to download.

Step 1: Data cleaning

Every HR data analysis project starts with data. Sometimes you receive data from a colleague or client. At other times you query the database yourself.

If you are lucky, the data is immediately ready for analysis, but more often than not, your data may be contaminated, **contains duplicates, inconsistencies, or other errors**. So, before you even start analyzing, you always want to check the quality and accuracy of your HR data.

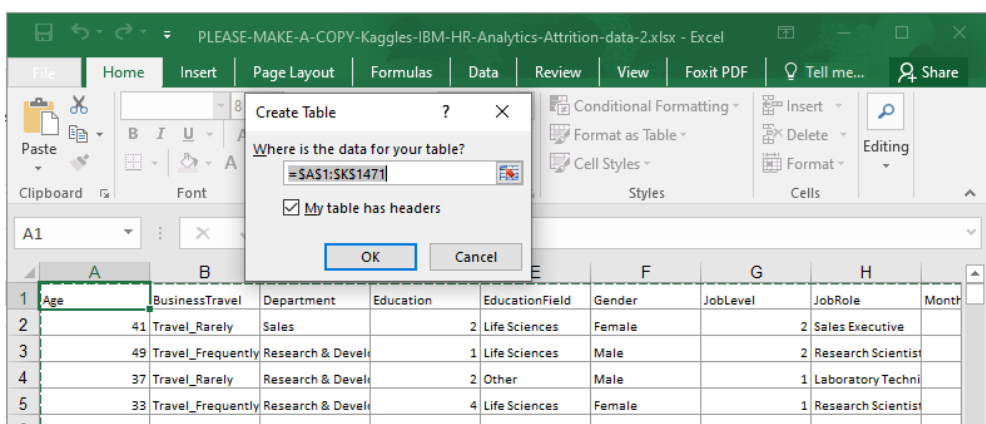
Data cleaning – also called data cleansing or data scrubbing – is often the most laborious phase of an HR data analysis project.

1. Excel Tables

You might be tempted to start working on your data from the moment you open Excel. However, this could mean missing out on utilizing the great functionality that Excel's built-in tables offer.

Excel Tables are handy containers for your data. They ensure that the values in a row stay together, which prevents you from accidentally scrambling the data when you sort and filter columns.

To make a table, click on any cell, then go to **Insert > Table** and select the range. Or use **CTRL + T** as a shortcut.



Excel adds out-of-the-box formatting to its tables, as well as advanced **filtering** and **slicing** functionality.

FileHomeInsertPage LayoutFormulasDataReviewViewFoxit PDFDesignTell me what you want to do...

Table Name:

Table1

Summarize with PivotTable

Remove Duplicates

Resize Table

Convert to Range

Properties

Tools

Insert Slicer

Export

Refresh

Open in Browser

Unlink

External Table Data

Header Row

First Column

Filter Button

Total Row

Last Column

Banded Rows

Banded Columns

Table Style Options

</

Your work will become much more organized when you use Excel Tables for your analyses.

2. Remove duplicates

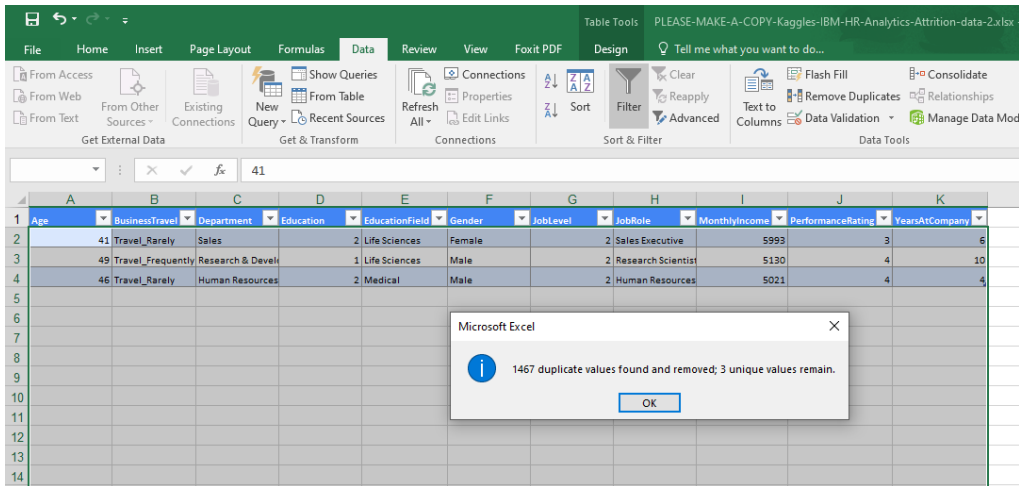
At times, your data tables may contain duplicate information. Maybe a specific employee **appears twice** in your database due to **an administration error**. Or, an employee has **a dual contract**, and therefore two records exist. If you are working with Excel Tables, the Table Design tab contains a **'Remove Duplicates'** button. It removes any rows that contain duplicate values on the selected columns.

The screenshot shows the Excel interface with the 'Data' tab selected. The 'Remove Duplicates' button is highlighted with a red box. The 'Remove Duplicates' dialog box is open, showing the 'Columns' list with 'Gender', 'JobLevel', 'JobRole', 'MonthlyIncome', 'PerformanceRating', and 'YearsAtCompany' selected. The 'My data has headers' checkbox is checked. The 'Remove Duplicates' button in the 'Data' tab is highlighted with a red box.

Fortunately, there are **no duplicate** rows in our dataset.

The screenshot shows a message box from Microsoft Excel that says 'No duplicate values found.' with an 'OK' button.

However, if you were to select the **'Department'** column only, Excel will remove all duplicates (such as Research and Development) and retain only the first rows in which it finds unique values.

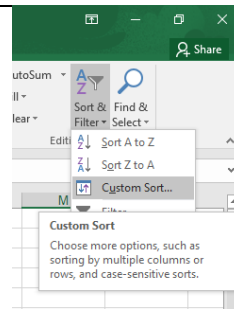
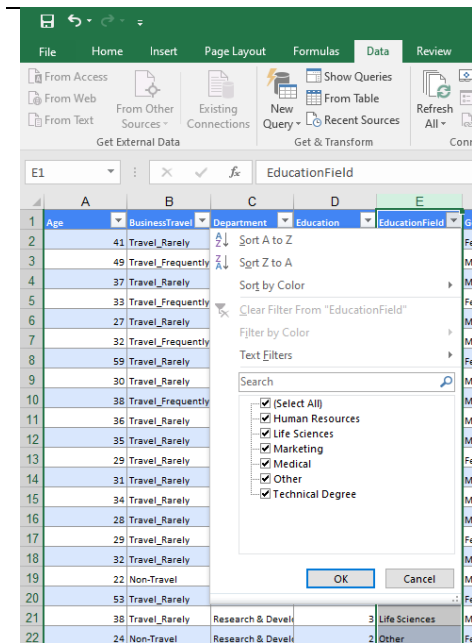


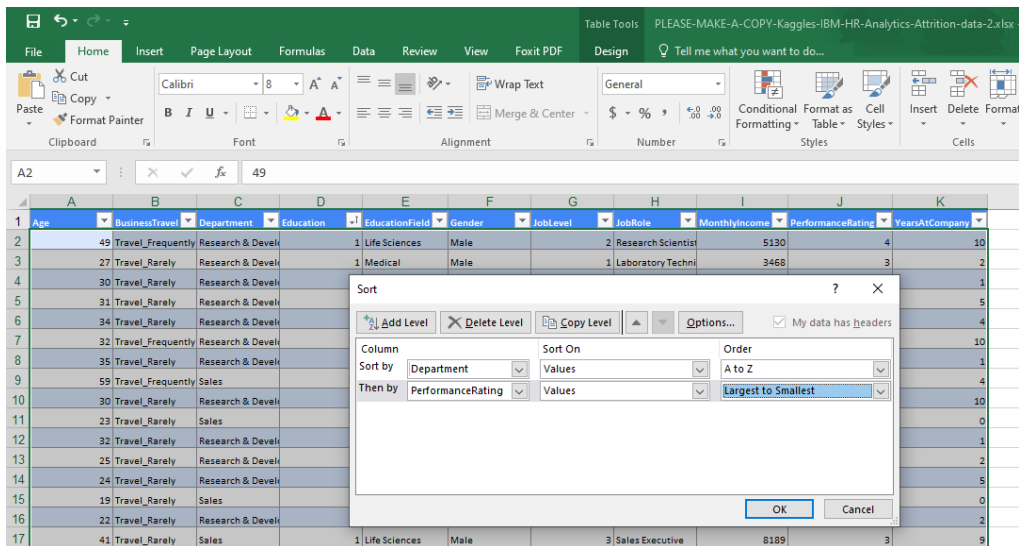
3. Sorting

Sorting is one of the most common tools of data management. You can sort your table by one or more columns in Excel, in ascending or descending order, or create a custom sort.

In Excel Tables, each column has a sortable filter included in its header.

Under **Data > Sort & Filter**, you find **Custom Sort**, allowing more advanced, multi-column sorting.





4. Converting data

Sometimes, numeric data is accidentally imported as text in Excel. This causes problems in your analysis, as Excel sorts text data differently.

Also, many advanced analyses, like correlation analysis, require that you provide numerical data. There are several handy functions to switch between data types in Excel:

VALUE

The VALUE function helps convert numbers stored as text to actual numeric data. Note that Excel, by default, aligns text strings to the left. Numeric data is default right-aligned. This gives you a visual cue as to how Excel interprets your data.

TEXT

The TEXT function helps turn numeric data into text strings in a specific number format. This [neat overview](#) shows the many ways in which you can format text numbers in Excel.

	A	B	C	D	E
1					
2					
3		-2.4		-2.4	=VALUE(B3)
4		-2.4		-\$2.40	=TEXT(B4,"\$#,##.00")
5					

FIXED

The FIXED function also returns numeric data as text, giving full control over formatting options like thousand separators and decimals.

Data	Formula Used	Result	Remarks
5129.631	=FIXED(5129.631)	5,129.63	Decimal argument is omitted so the default value of 2 is used.
529.631	=FIXED(529.631,1)	529.6	As the decimal argument is 1, we get the result with 1 decimal point.
5123.591	=FIXED(5123.591,-1)	5120	If decimals are negative, the number is rounded to the left of the decimal point.
5123.591	=FIXED(5123.591, -3, TRUE)	5000	

7	FIXED	5123.63		5,123.63	=FIXED(B7)
8		5123.63		5,123.6	=FIXED(B8,1)
9		5123.63		5,124	=FIXED(B9,0)
10		5123.63		5,120	=FIXED(B9, -1)
11		5123.63		5,100	=FIXED(B10, -2)
12		5123.63		5,000	=FIXED(B11, -3)

DOLLAR

Finally, the DOLLAR function converts numbers to text and applies currency formatting.

	A	B	C	D
1	Data	Formula	Result	Description
2	1234.567	=DOLLAR(A2,2)	\$1,234.57	Displays the first number in a currency format, 2 digits to the right of the decimal point.
3	-1234.567	=DOLLAR(A3,-2)	(\$1,200)	Displays the second number in a currency format, 2 digits to the left of the decimal point.
4	-0.123	=DOLLAR(A4,4)	(\$0.1230)	Displays the third number in a currency format, 4 digits to the right of the decimal point.
5	99.888	=DOLLAR(A5)	\$99.89	Displays the fourth number in a currency format, 2 digits to the right of the decimal point.

	A	B	C	D	E
13					
14	DOLLAR	1234.567		\$1,234.57	=DOLLAR(B14,2)
15		-1234.567		(\$1,200)	=DOLLAR(B15,-2)
16		-0.123		(\$0.1230)	=DOLLAR(B16,4)
17		99.888		\$99.89	=DOLLAR(B17)
18					

5. Text manipulation

Excel provides many functions that help with “manipulating” text data.

PROPER

UPPER and LOWER help you set the case of your text strings. But have you heard of the PROPER function? It capitalizes the first letter in a text string and any other letters in the text that follow any character other than a letter (also known as title case). All other letters are converted to lowercase letters.

Original	Academy to Innovate HR
=UPPER(E3)	ACADEMY TO INNOVATE HR
=LOWER(E3)	academy to innovate hr
=PROPER(E3)	Academy To Innovate Hr

	A	B	C	D	E
18					
19	PROPER	Academy to Innovative HR		ACADEMY TO INNOVATIVE HR	=UPPER(B19)
20		Academy to Innovative HR		academy to innovative hr	=LOWER(B20)
21		Academy to Innovative HR		Academy To Innovative Hr	=PROPER(B21)
22					

LEFT & RIGHT

Sometimes you only need part of a text string, such as the first couple of characters or the last ones.

In those cases, you can use LEFT and RIGHT, which respectively allows you to extract any number of characters from the start (LEFT) or end (RIGHT) of a text string.

For instance, maybe you don't want to use the full gender labels but only the first letter.

fx =LEFT([@Gender]; 1)							
	F	G	H	I	J	K	L
d	Gender	JobLevel	JobRole	MonthlyIncome	PerformanceRating	YearsAtCompany	Column
	Female	2	Sales Executive	5993	3	6	F
	Male	2	Research Scientist	5130	4	10	M
	Male	1	Laboratory Technician	2090	3	0	M
	Female	1	Research Scientist	2909	3	8	F
	Male	1	Laboratory Technician	3468	3	2	M
	Male	1	Laboratory Technician	3068	3	7	M

	A	B	C	D	E
22					
23	LEFT & RIGHT	Female		F	=LEFT(B23,1)
24		Female		ale	=RIGHT(B24,3)
25					

SUBSTITUTE

Sometimes you want to replace specific parts of a text. For instance, you might want to rename the “Sales” department to the “Commercial” department.

In Excel, you can use the SUBSTITUTE function to replace specific parts of a text string with an alternative text.

SUM		=SUBSTITUTE([@JobRole],"Sales","Commercial")													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Age	BusinessTravel	Department	Education	EducationField	Gender	JobLevel	JobRole	MonthlyIncome	PerformanceRating	YearsAtCompany	Column	Column1		
2	26	Travel_Rarely	Human Resources	1	Life Sciences	Female	1	Human Resources	2942	4	8	F	=SUBSTITUTE([@JobRole],"Sales","Commercial")		
3	36	Travel_Frequently	Human Resources	1	Human Resources	Male	2	Human Resources	3886	4	10	M	=SUBSTITUTE(text,old_text,new_text,[instance_num])		
4	46	Travel_Rarely	Human Resources	2	Medical	Male	2	Human Resources	5021	4	4	M	Human Resources		
5	40	Travel_Rarely	Human Resources	2	Medical	Male	4	Manager	16437	4	21	M	Manager		
6	36	Travel_Rarely	Human Resources	3	Life Sciences	Male	1	Human Resources	2342	4	4	S	M	Human Resources	

As you can see, this also changed the Job Role for the “Sales representative” on row 1026.

SUM		=SUBSTITUTE([@JobRole],"Sales","Commercial")													
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	Age	BusinessTravel	Department	Education	EducationField	Gender	JobLevel	JobRole	MonthlyIncome	PerformanceRating	YearsAtCompany	Column	Column1		
1024	35	Travel_Rarely	Research & Devel	5	Life Sciences	Male	2	Research Scientist	5208	3	16	M	Research Scientist		
1025	50	Travel_Rarely	Research & Devel	5	Life Sciences	Female	3	Research Director	13269	3	14	F	Research Director		
1026	19	Travel_Frequently	Sales	1	Technical Degree	Female	1	Sales Representative	2325	4	0	F	=SUBSTITUTE([@JobRole],"Sales","Commercial")		
1027	42	Travel_Rarely	Sales	1	Life Sciences	Male	2	Sales Executive	4907	4	20	M	=SUBSTITUTE(text,old_text,new_text,[instance_num])		
1028	48	Travel_Rarely	Sales	1	Medical	Female	3	Manager	12504	4	0	F	Manager		
1029	53	Travel_Rarely	Sales	1	Medical	Female	2	Sales Executive	8381	4	14	F	Commercial Executive		

6. Find and Replace

You might encounter situations where you need to find and/or replace data values in your workbook. For instance, you might be looking for a specific employee, department, or job role.

You can use the shortcut **CTRL + F** to open up the FIND menu. This allows you to input the data value you are looking for, and Excel will show you all occurrences of it in your worksheet.

If you press **CTRL + H**, this opens the FIND & REPLACE menu.

This allows you to not only locate the occurrences of data values but also enables you to replace them with another value. You could try to rename the ‘Sales’ department with this approach as well.

SUM		=SUBSTITUTE([@JobRole],"Sales","Commercial")													
	A	B	C	D	E	F	G	H	I	J	K	L	M		
1	Age	BusinessTravel	Department	Education	EducationField	Gender	JobLevel	JobRole	MonthlyIncome	PerformanceRating	YearsAtCompany	Column	Column1		
2	26	Travel_Rarely	Human Resources	1	Life Sciences	Female	1	Human Resources	2942	4	8	F	Human Resources		
3	36	Travel_Frequently	Human Resources	1	Human Resources	Male	2	Human Resources	3886	4	10	M	Human Resources		
4	46	Travel_Rarely	Human Resources	2	Medical	Male	2	Human Resources	5021	4	4	M	Human Resources		
5	40	Travel_Rarely	Human Resources	2	Medical	Male	4	Manager	16437	4	21	M	Manager		
6	36	Travel_Rarely	Human Resources	3	Life Sciences	Male	1	Human Resources	2342	4	4	S	M	Human Resources	
7	30	Travel_Rarely	Human Resources	3	Life Sciences	Male	2	Human Resources	5208	3	16	M	Research Scientist		
8	45	Travel_Rarely	Human Resources	3	Life Sciences	Female	3	Research Director	13269	3	14	F	Research Director		
9	37	Travel_Rarely	Human Resources	4	Human Resources	Male	4	Manager	12504	4	0	F	Manager		
10	59	Non-Travel	Human Resources	4	Human Resources	Female	2	Sales Executive	4907	4	20	M	Commercial Executive		
11	42	Travel_Rarely	Human Resources	1	Technical Degree	Male	1	Sales Representative	2325	4	0	F	Commercial Executive		
12	24	Travel_Rarely	Human Resources	1	Human Resources	Male	2	Human Resources	3886	4	10	M	Human Resources		
13	29	Travel_Rarely	Human Resources	1	Medical	Male	1	Human Resources	2804	3	1	M	Human Resources		

If you open the advanced options, you can even replace all occurrences of a value across your entire workbook.

SUM		=SUBSTITUTE([@JobRole],"Sales","Commercial")													
	A	B	C	D	E	F	G	H	I	J	K	L	M		
1	Age	BusinessTravel	Department	Education	EducationField	Gender	JobLevel	JobRole	MonthlyIncome	PerformanceRating	YearsAtCompany	Column	Column1		
2	26	Travel_Rarely	Human Resources	1	Life Sciences	Female	1	Human Resources	2942	4	8	F	Human Resources		
3	36	Travel_Frequently	Human Resources	1	Human Resources	Male	2	Human Resources	3886	4	10	M	Human Resources		
4	46	Travel_Rarely	Human Resources	2	Medical	Male	2	Human Resources	5021	4	4	M	Human Resources		
5	40	Travel_Rarely	Human Resources	2	Medical	Male	4	Manager	16437	4	21	M	Manager		
6	36	Travel_Rarely	Human Resources	3	Life Sciences	Male	1	Human Resources	2342	4	4	S	M	Human Resources	
7	30	Travel_Rarely	Human Resources	3	Life Sciences	Male	2	Human Resources	5208	3	16	M	Research Scientist		
8	45	Travel_Rarely	Human Resources	3	Life Sciences	Female	3	Research Director	13269	3	14	F	Research Director		
9	37	Travel_Rarely	Human Resources	4	Human Resources	Male	4	Manager	12504	4	0	F	Manager		
10	59	Non-Travel	Human Resources	4	Human Resources	Female	2	Sales Executive	4907	4	20	M	Commercial Executive		
11	42	Travel_Rarely	Human Resources	1	Technical Degree	Male	1	Sales Representative	2325	4	0	F	Commercial Executive		
12	24	Travel_Rarely	Human Resources	1	Human Resources	Male	2	Human Resources	3886	4	10	M	Human Resources		
13	29	Travel_Rarely	Human Resources	1	Medical	Male	1	Human Resources	2804	3	1	M	Human Resources		
14	59	Travel_Rarely	Human Resources	2	Medical	Male	2	Human Resources	5021	4	4	M	Human Resources		
15	19	Travel_Rarely	Human Resources	2	Technical Degree	Male	1	Human Resources	2564	3	1	M	Human Resources		
16	34	Travel_Rarely	Human Resources	2	Human Resources	Male	1	Human Resources	3737	3	3	M	Human Resources		

7. Paste special

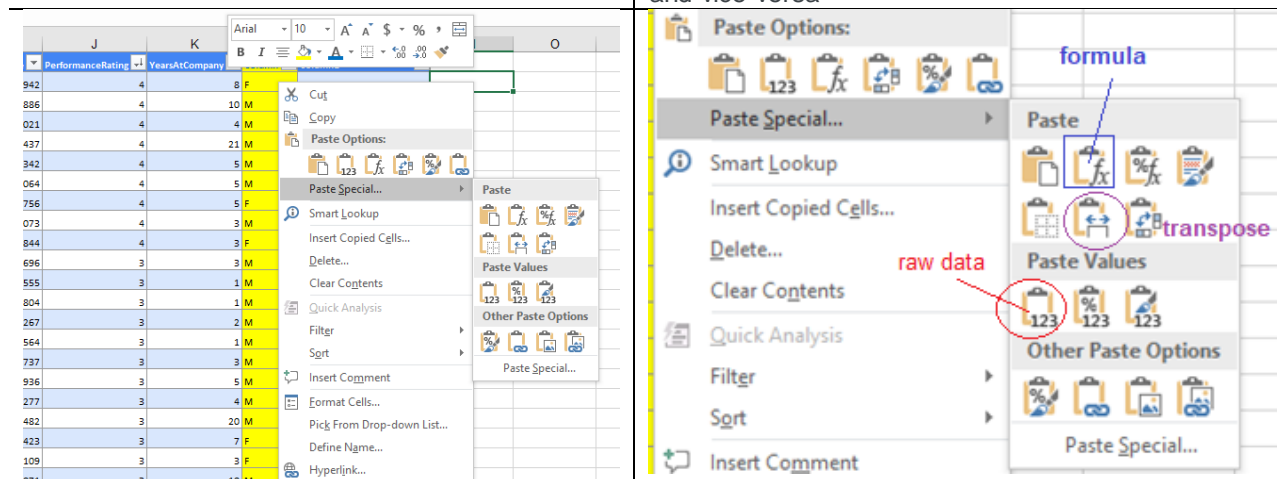
Another great feature of Excel is the Paste Special functionality.

Data values in Excel are often formatted in an explicit way — think of typeface, color, or font size. When copying and pasting values, sometimes you want to **retain this formatting**. At other times, you want to keep **only the raw data**. And again, in another situation, you might only want the formatting.

However, you can use Paste Special to instruct Excel to do exactly what you want with the copied data and its formatting.

You will find the functionality when you click the right mouse button after you've copied any data.

For instance, you only want to keep the **raw data**: (Value). Another handy trick is to paste the data **transposed**, which converts the rows into columns and vice versa



Did you know you can even paste the raw results with any formulas you have copied?

Or you can copy your selection as an image, including the formatting. This great functionality is handy when you want to add an Excel Table to an email or a presentation.

Step 2: Data analysis

Once you have cleaned your data and are confident in its quality, you are ready to start analyzing your data. First, you might want to calculate [descriptive analytics](#) that summarizes your data to assist in uncovering insightful patterns. These insights can help to make important decisions on how to improve your HR processes and policies as well as evaluate their business impact.

Excel offers a lot of functions for data analysis. PivotTables are one such function that can really take your data analysis capabilities to the next level. If you want to find out more about how to use PivotTables, AIHR has provided step-by-step instructions on [how to work with PivotTables](#).

Below, let's unpack some additional tools and functions you can use in Excel and pivot tables to help you with your data analysis:

1. Slicers and Filters

Slicers are clickable buttons that filter your Excel PivotTables. You can add a slicer by selecting any cell in a table, then going to Table Design > Insert Slicer.

Step 3: Storytelling

Step 4: Dashboarding