How can I remove duplicate rows in Excel while also retaining the row with the newest date?

Authored by

stats writer

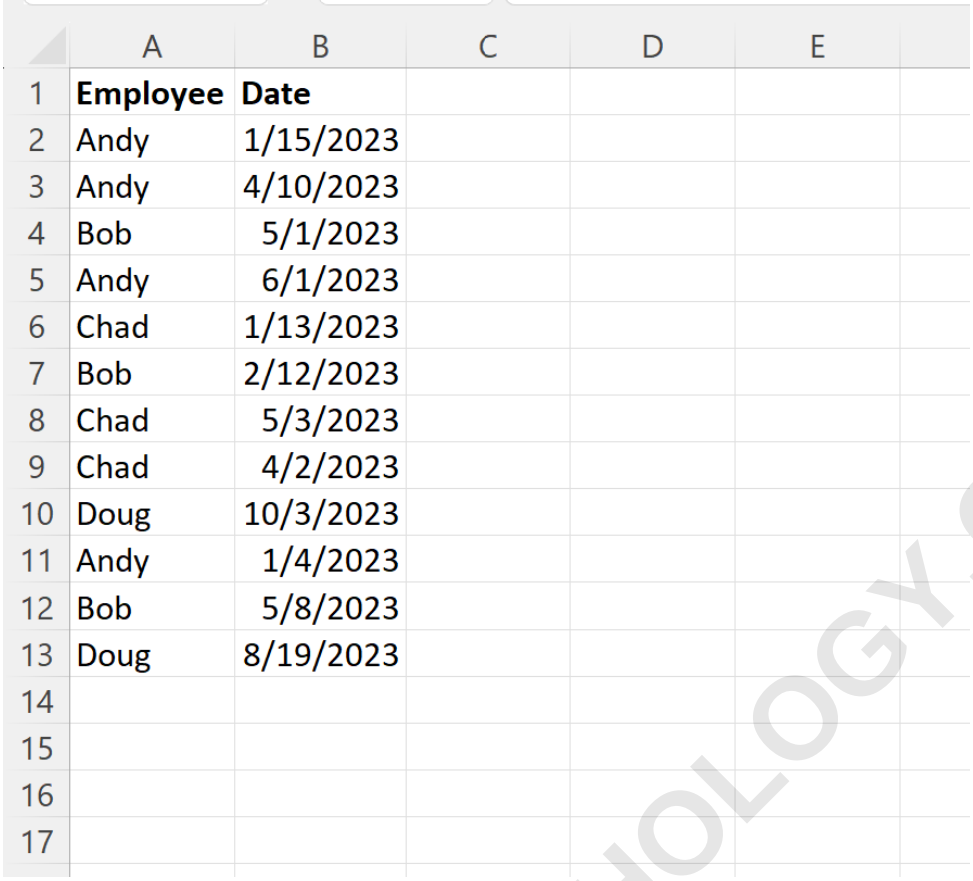
June 22, 2024

Removing duplicate rows in Excel can be easily achieved by using the "Remove Duplicates" function. However, in cases where there are multiple rows with the same data but different dates, it is important to retain the row with the newest date. To do this, the "Remove Duplicates" function can be combined with the "Sort" function to sort the data by date in descending order before removing duplicates. This will ensure that the row with the newest date is kept while eliminating any duplicate rows.

Excel: Remove Duplicates but Keep Row with Newest Date

Often you may want to remove duplicate rows in Excel but keep the row that contains the newest date in a particular column.

For example, suppose you have the following dataset that contains information about sales made on various dates by employees at some company:



	A	B	C	D	E
1	**Employee**	**Date**			
2	Andy	1/15/2023			
3	Andy	4/10/2023			
4	Bob	5/1/2023			
5	Andy	6/1/2023			
6	Chad	1/13/2023			
7	Bob	2/12/2023			
8	Chad	5/3/2023			
9	Chad	4/2/2023			
10	Doug	10/3/2023			
11	Andy	1/4/2023			
12	Bob	5/8/2023			
13	Doug	8/19/2023			
14					
15					
16					
17					

Suppose you'd like to remove rows with duplicate values in the Employee column but keep the row with the newest date in the Date column:

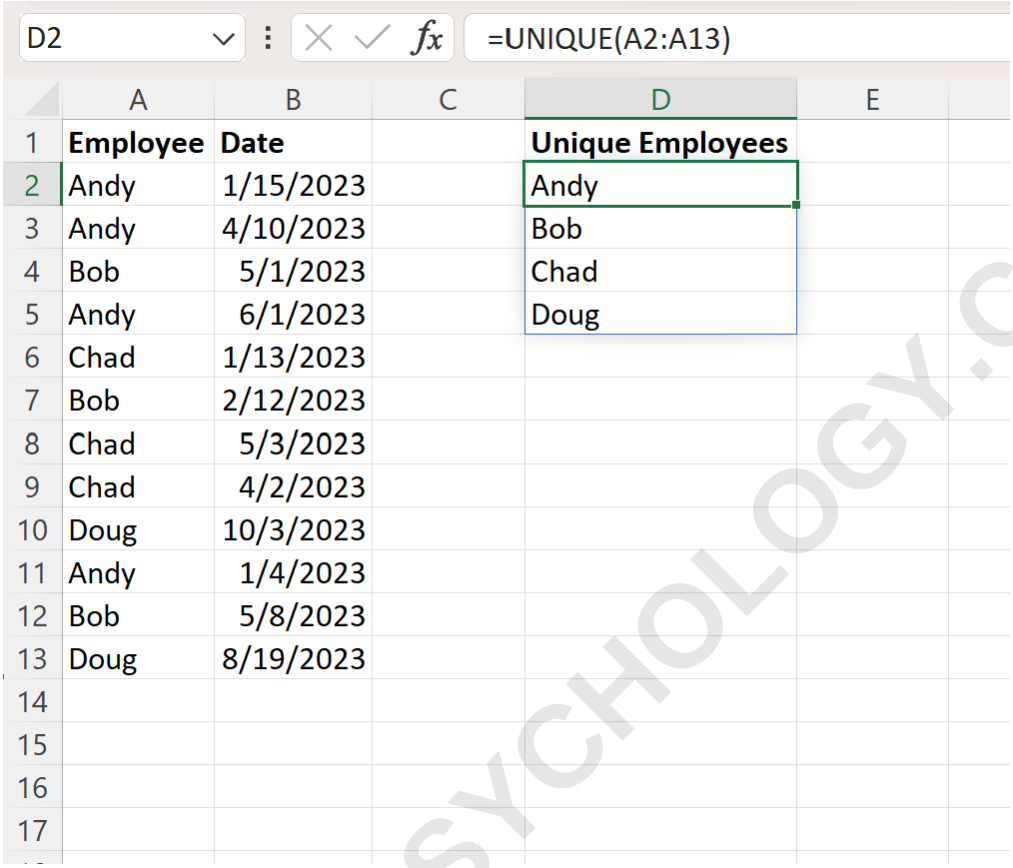
The following example shows exactly how to do so.

Example: Remove Duplicates but Keep Row with Newest Date in Excel

First, type the following formula into cell D2:

=UNIQUE(A2:A13)

This produces a list of unique employee names from the dataset:

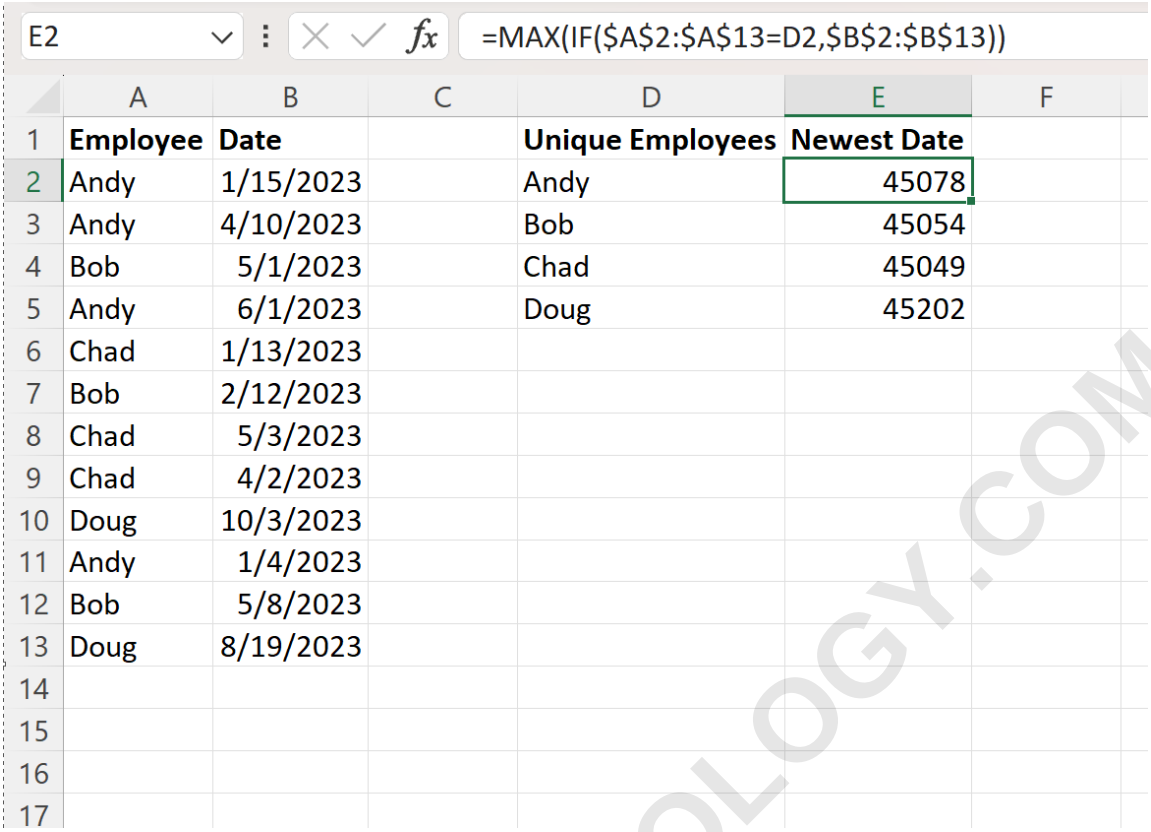


	A	B	C	D	E
1	**Employee**	**Date**		**Unique Employees**	
2	Andy	1/15/2023		Andy	
3	Andy	4/10/2023		Bob	
4	Bob	5/1/2023		Chad	
5	Andy	6/1/2023		Doug	
6	Chad	1/13/2023			
7	Bob	2/12/2023			
8	Chad	5/3/2023			
9	Chad	4/2/2023			
10	Doug	10/3/2023			
11	Andy	1/4/2023			
12	Bob	5/8/2023			
13	Doug	8/19/2023			
14					
15					
16					
17					

Next, type the following formula into cell E2 to find the newest date in the Date column associated with each unique employee name:

=MAX(IF($A$2:$A$13=D2,$B$2:$B$13))

Then click and drag this formula down to the remaining cells in column E:

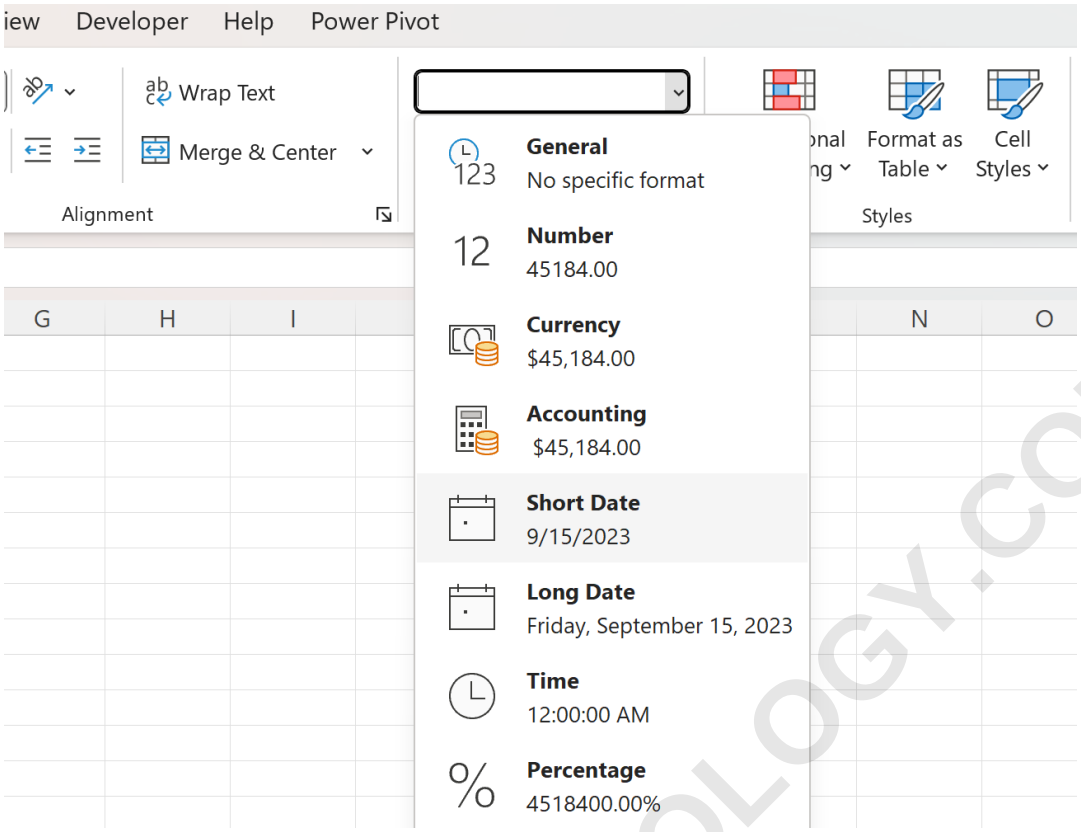


E2 =MAX(IF($A$2:$A$13=D2,$B$2:$B$13))

	A	B	C	D	E	F
1	**Employee**	**Date**		**Unique Employees**	**Newest Date**	
2	Andy	1/15/2023		Andy	45078	
3	Andy	4/10/2023		Bob	45054	
4	Bob	5/1/2023		Chad	45049	
5	Andy	6/1/2023		Doug	45202	
6	Chad	1/13/2023				
7	Bob	2/12/2023				
8	Chad	5/3/2023				
9	Chad	4/2/2023				
10	Doug	10/3/2023				
11	Andy	1/4/2023				
12	Bob	5/8/2023				
13	Doug	8/19/2023				
14						
15						
16						
17						

By default, Excel displays these dates as serial numbers.

To convert these numbers to dates, highlight the range E2:E5, then click the Number Format dropdown menu in the Number group of the Home tab, then click Short Date:



	A	B	C	D	E
1	**Employee**	**Date**		**Unique Employees**	**Newest Date**
2	Andy	1/15/2023		Andy	6/1/2023
3	Andy	4/10/2023		Bob	5/8/2023
4	Bob	5/1/2023		Chad	5/3/2023
5	Andy	6/1/2023		Doug	10/3/2023
6	Chad	1/13/2023			
7	Bob	2/12/2023			
8	Chad	5/3/2023			
9	Chad	4/2/2023			
10	Doug	10/3/2023			
11	Andy	1/4/2023			
12	Bob	5/8/2023			
13	Doug	8/19/2023			
14					
15					
16					
17					
18					

Columns D and E now show the original dataset with the rows removed that had duplicate values in the Employee column and only kept the rows with the newest date in the Date column.

For example, there are four rows with a duplicate "Andy" value in the Employee column but the row with the newest date among these four rows is the one with 6/1/2023:

	A	B	C	D	E	F
1	**Employee**	**Date**		**Unique Employees**	**Newest Date**	
2	Andy	1/15/2023		Andy	6/1/2023	
3	Andy	4/10/2023		Bob	5/8/2023	
4	Bob	5/1/2023		Chad	5/3/2023	
5	Andy	6/1/2023		Doug	10/3/2023	
6	Chad	1/13/2023				
7	Bob	2/12/2023				
8	Chad	5/3/2023				
9	Chad	4/2/2023				
10	Doug	10/3/2023				
11	Andy	1/4/2023				
12	Bob	5/8/2023				
13	Doug	8/19/2023				
14						
15						
16						
17						
18						

This is the one row that is kept among all possible rows with "Andy" in the filtered dataset.

The same process is repeated for every other unique employee name.

The following tutorials explain how to perform other common tasks in Excel: