

How do I Winsorize data in Excel?

Authored by
stats writer

April 24, 2024

RECOMMENDED CITATION

stats writer (2024). *How do I Winsorize data in Excel?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=138944>

Winsorizing data in Excel is a statistical method used to limit the impact of extreme values on a data set. This process involves replacing the extreme values with the nearest non-extreme value, thereby reducing the effect of outliers on the overall data. To Winsorize data in Excel, one can use the built-in functions such as "TRIMMEAN" or "TRIM" to calculate the trimmed mean or remove the top and bottom percentages of data, respectively. Alternatively, the "IF" function can be used to manually identify and replace extreme values. This technique is beneficial in ensuring a more accurate and robust analysis of data.

Winsorize Data in Excel

To winsorize data means to set extreme outliers equal to a specified percentile of the data.

For example, a 90% winsorization sets all greater than the 95th percentile equal to the value at the 95th percentile and all observations less than the 5th percentile equal to the value at the 5th percentile.

This tutorial provides a step-by-step example of how to winsorize a dataset in Excel.

Step 1: Create the Data

First, we'll create the following dataset:

	A	B	C	D	E	F	
1	Data						
2	3						
3	14						
4	16						
5	16						
6	17						
7	29						
8	34						
9	36						
10	39						
11	47						
12	59						
13	64						
14	65						
15	66						
16	68						
17	79						
18	91						
19	98						
20							
21							
22							
23							
24							
25							
26							
27							

Step 2: Calculate the Upper and Lower Percentiles

For this example, we'll perform a 90% winsorization. This means we'll set all values greater than the 95th percentile equal to the 95th percentile and all values less than the 5th percentile equal to the 5th percentile.

The following formulas show how to find the 5th and 95th percentiles:

	A	B	C	D	E	F	G	H
1	Data				Formula used			
2	3		5th percentile	12.35	=PERCENTILE(A2:A19, 0.05)			
3	14		95th percentile	92.05	=PERCENTILE(A2:A19, 0.95)			
4	16							
5	16							
6	17							
7	29							
8	34							
9	36							
10	39							
11	47							
12	59							
13	64							
14	65							
15	66							
16	68							
17	79							
18	91							
19	98							
20								
21								
22								
23								
24								
25								
26								
27								

The 5th percentile turns out to be 12.35 and the 95th percentile turns out to be 92.05.

Step 3: Winsorize the Data

Lastly, we'll use the following formula to winsorize the data:

	A	B	C	D	E	F	G	H	I	J
1	Data					Winsorized Data				
2	3		5th percentile	12.35		12.35	=IF(A2<\$D\$2,\$D\$2,IF(A2>\$D\$3,\$D\$3,A2))			
3	14		95th percentile	92.05		14				
4	16					16				
5	16					16				
6	17					17				
7	29					29				
8	34					34				
9	36					36				
10	39					39				
11	47					47				
12	59					59				
13	64					64				
14	65					65				
15	66					66				
16	68					68				
17	79					79				
18	91					91				
19	98					92.05				
20										
21										
22										
23										
24										
25										

Note that we just copy and pasted the formula in cell F2 down to the remaining cells in column F.

In this case, the value 3 became changed to 12.35 and the value 98 became changed to 92.05.

Note that in this example we performed a 90% winsorization, but it's possible to also perform an 80% winsorization, 95% winsorization, 99% winsorization, etc. by simply calculating different upper and lower percentiles.