How can I use the GROUP BY function in Google Sheets Query to organize my data?

Authored by

stats writer

April 29, 2024

ARABPSYCHOLOGY.COM

RECOMMENDED CITATION

stats writer (2024). *How can I use the GROUP BY function in Google Sheets Query to organize my data?*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=141058

The GROUP BY function in Google Sheets Query allows users to organize their data by grouping similar values together. This feature is particularly useful when dealing with large data sets, as it enables users to quickly summarize and analyze their data. By using the GROUP BY function, users can easily identify patterns and trends within their data, making it easier to draw meaningful insights and make informed decisions. This function is a valuable tool for efficiently managing and analyzing data in Google Sheets.

Google Sheets Query: Use Group By

You can use the following syntax to group and aggregate data in a Google Sheets Query:

=query(A1:D12, "select B, avg(D) group by B",1)

In this example, we select columns B and D in cell range A1:D12. We then find the average of column D, grouped by column B. We also specify a 1 to indicate that there is 1 header row at the top of the dataset.

In this example, we used the avg() aggregate function, but we can use any of the following aggregate functions:

avg()sum()count()min()max()

The following examples show how to group and aggregate data in practice.

ARABPSYCHOLOGY.COM

Example 1: Group & Aggregate by One Column

We can use the following formula to select the Team and Points columns, then find the average of the Points column, grouped by Team:

F1 fx =query(A1:D12, "select B, avg(D) group by B", 1)

	A	B	C	D	E	F	G
1	**Player**	**Team**	**Conference**	**Points**		Team	avg Points
2	Andy	Cavs	East	13.4		Cavs	10.8
3	Bob	Mavericks	West	7.8		Celtics	16.26666667
4	Carl	Celtics	East	13.7		Mavericks	18.8
5	Dave	Warriors	West	22.3		Warriors	21.66666667
6	Eric	Mavericks	West	27.8			
7	Fred	Mavericks	West	20.8			
8	George	Celtics	East	12.7			
9	Harold	Cavs	East	8.2			
10	Isaiah	Warriors	West	12.5			
11	Joe	Warriors	West	30.2			
12	Ken	Celtics	East	22.4			
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							

ARABPSYCHOLOGY.COM

This tells us that:

The average points scored by players on the Cavs is 10.8.The average points scored by players on the Celtics is 12.7.

And so on.

Example 2: Group & Aggregate by Multiple Columns

We can use the following formula to select the Team, Conference and Points columns, then find the maximum of the Points column, grouped by Team and Conference:

F1 =query(A1:D12, "select B, C, max(D) group by B, C", 1)

	A	B	C	D	E	F	G	H
1	**Player**	**Team**	**Conference**	**Points**		Team	Conference	max Points
2	Andy	Cavs	East	13.4		Cavs	East	13.4
3	Bob	Mavericks	West	7.8		Celtics	East	22.4
4	Carl	Celtics	East	13.7		Mavericks	West	27.8
5	Dave	Warriors	West	22.3		Warriors	West	30.2
6	Eric	Mavericks	West	27.8				
7	Fred	Mavericks	West	20.8				
8	George	Celtics	East	12.7				
9	Harold	Cavs	East	8.2				
10	Isaiah	Warriors	West	12.5				
11	Joe	Warriors	West	30.2				
12	Ken	Celtics	East	22.4				
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								

ARABPSYCHOLOGY.COM

This tells us that:

The max points scored by any player on the Cavs team in the East Conference is 13.4.The max points scored by any player on the Celtics team in the East

Conference is 22.4.

How to Avoid #VALUE! Errors

To avoid #VALUE! errors when grouping and aggregating data, make sure that every column included in the select statement meets one of the following criteria:

Has an aggregate function applied to it.Is included in the group by statement.

As long as each column in the select statement meets one of these criteria, then you can avoid a #VALUE! error.

ARABPSYCHOLOGY.COM