

# How can quantiles be calculated by group in R, and can you provide some examples?

Authored by  
**stats writer**

April 28, 2024

## RECOMMENDED CITATION

stats writer (2024). *How can quantiles be calculated by group in R, and can you provide some examples?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=140515>

Quantiles are statistical measures that divide a dataset into equal-sized groups, with each group containing an equal number of values. In R, quantiles can be calculated by group using the "quantile" function. This function takes in a dataset and a grouping variable and returns the quantile values for each group. For example, if we have a dataset of students' exam scores and we want to calculate the 25th, 50th, and 75th quantiles for each grade level, we can use the "quantile" function with the grade level variable as the grouping variable. This will give us the quantile values for each grade level separately. Other examples of grouping variables could be gender, race, or age. By calculating quantiles by group, we can gain insights into the distribution of a variable within different subgroups of a dataset.

## Calculate Quantiles by Group in R (With Examples)

**In statistics, quantiles are values that divide a ranked dataset into equal groups.**

**To calculate the quantiles grouped by a certain variable in R, we can use the following functions from the package in R:**

```
library(dplyr)
```

```
#define quantiles of interest
```

```
q = c(.25, .5, .75)
```

```
#calculate quantiles by grouping variable
```

```
df %>%
```

```
group_by(grouping_variable) %>%
```

```
summarize(quant25 = quantile(numeric_variable, probs  
= q),
```

```
quant50 = quantile(numeric_variable, probs = q),  
quant75 = quantile(numeric_variable, probs = q))
```

The following examples show how to use this syntax in practice.

Examples: Quantiles by Group in R

The following code shows how to calculate the quantiles for the number of wins grouped by team for a dataset in R:

```
library(dplyr)  
  
#create data  
df <- data.frame(team=c('A', 'A', 'A', 'A', 'A', 'A', 'A', 'A',  
'B', 'B', 'B', 'B', 'B', 'B', 'B', 'B',  
'C', 'C', 'C', 'C', 'C', 'C', 'C', 'C'),  
wins=c(2, 4, 4, 5, 7, 9, 13, 13, 15, 15, 14, 13,  
11, 9, 9, 8, 8, 16, 19, 21, 24, 20, 19, 18))
```

#view first six rows of data

```
head(df)
```

```
team wins
```

```
1 A 2
```

2 A 4

3 A 4

4 A 5

5 A 7

6 A 9

```
#define quantiles of interest
```

```
q = c(.25, .5, .75)
```

```
#calculate quantiles by grouping variable
```

```
df %>%
```

```
group_by(team) %>%
```

```
summarize(quant25 = quantile(wins, probs = q),
```

```
quant50 = quantile(wins, probs = q),
```

```
quant75 = quantile(wins, probs = q))
```

```
team quant25 quant50 quant75
```

```
1 A 4 6 10
```

```
2 B 9 12 14.2
```

```
3 C 17.5 19 20.2
```

Note that we can also specify any number of quantiles that we'd like:

```
#define quantiles of interest
```

```
q = c(.2, .4, .6, .8)
```

```
#calculate quantiles by grouping variable
```

```
df %>%
```

```
group_by(team) %>%
```

```
summarize(quant20 = quantile(wins, probs = q),
```

```
quant40 = quantile(wins, probs = q),
```

```
quant60 = quantile(wins, probs = q),
```

```
quant80 = quantile(wins, probs = q))
```

```
team quant20 quant40 quant60 quant80
```

```
1 A 4 4.8 7.4 11.4
```

```
2 B 9 10.6 13.2 14.6
```

```
3 C 16.8 18.8 19.2 20.6
```

We can also choose to calculate just one quantile by group. For example, here's how to calculate the 90th percentile of the number of wins for each team:

```
#calculate 90th percentile of wins by team
```

```
df %>%
```

```
group_by(team) %>%
```

```
summarize(quant90 = quantile(wins, probs = 0.9))
```

## team quant90

1 A 13

2 B 15

3 C 21.9

ARABPSYCHOLOGY.COM