

Find the Median of Grouped Data (With Examples)

Authored by
Mohammed loot

October 31, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Find the Median of Grouped Data (With Examples)*.
PSYCHOLOGICAL STATISTICS. Retrieved from
<https://statistics.arabpsychology.com/?p=6710>

Understanding Central Tendency and Grouped Data Structures

In the study of [statistics](#), calculating measures of central tendency is crucial for summarizing and interpreting large datasets. These measures provide a single value that attempts to describe the center of the data. Among the most important is the [median](#), which serves as a robust indicator of the typical value. The [median](#) represents the middle point of a dataset once all observations are ordered numerically. Unlike the mean, it remains minimally affected by extreme outliers, making it an indispensable tool in [data analysis](#) and [descriptive statistics](#).

However, raw data is often voluminous and unwieldy. To simplify presentation and make information more manageable, especially when dealing with extensive datasets, data is frequently organized into [grouped data](#). This structure involves partitioning the data into defined [class intervals](#), each accompanied by its corresponding count, known as its [frequency](#). This organized format transforms a list of individual observations into a concise [frequency distribution](#).

When working with [grouped data](#), we lose access to the precise, individual data points within each interval. Consequently, we cannot calculate the exact [median](#). Instead, we rely on a specialized estimation technique that uses the structure of the distribution to locate an approximate central value. This article provides a comprehensive guide to understanding and applying the formula necessary to calculate this estimated [median](#) for any [grouped data](#) set.

Consider the following visualization, which exemplifies a typical [frequency distribution](#) of [grouped data](#):

Range	Frequency
1-10	2
11-20	7
21-30	10
31-40	3
41-50	1

In this presentation, while the total number of observations and the counts within specific ranges are known, the exact values are hidden. This necessity for estimation makes the formula-based approach an indispensable part of applied [statistics](#).

The Formula for Estimating the Median of Grouped Data

To effectively estimate the [median](#) when data is organized into a [frequency distribution](#), statisticians employ a specialized interpolation formula. This formula utilizes the cumulative distribution pattern to pinpoint the location of the 50th percentile value, yielding a reliable approximation of the true center of the data.

This technique is vital because it allows us to derive meaningful insights even when we lack the precision of raw data points. Mastering this formula is fundamental for anyone working with aggregated data summaries.

The standard formula used for calculating the estimated [median](#) of [grouped data](#) is as follows:

$$\text{Median of Grouped Data} = L + W$$

Each component within this equation plays a distinct and critical role in the interpolation process. The success of the estimation hinges on correctly identifying and extracting these five variables--L, W, N, C, and F--from the frequency table. The subsequent section will provide a detailed breakdown of what each term signifies and how it is derived from the structure of the grouped data.

Deconstructing the Median Formula Components

A precise understanding of the meaning and derivation of each variable is essential for accurate application of the grouped data median formula. These components are all extracted directly from the frequency distribution table, specifically focusing on the class interval identified as the "median class."

L: Lower Limit of the Median Class. This value represents the actual lower boundary of the specific class interval where the median is expected to fall. For continuous data, this is the lower bound; for discrete data, it is the starting score or value of that class.

W: Width of the Median Class. Also referred to as the class size, this is the range covered by the class interval containing the median. It is typically calculated as the difference between the class boundaries (or upper limit minus lower limit plus one for discrete intervals).

N: Total Frequency. N stands for the total number of observations in the dataset. It is simply the sum of all individual [frequencies](#) across the entire [frequency distribution](#).

C: Cumulative Frequency up to the Median Class. This variable requires calculating the [cumulative frequency](#). C is the sum of the [frequencies](#) of all classes that strictly precede the designated [median class](#). It helps establish how many observations are counted before the median interval begins.

F: Frequency of the Median Class. This is the individual [frequency](#) (or count) of observations found within the specific [median class](#) itself.

The crucial prerequisite for extracting these values correctly is the initial step: accurately identifying the [median class](#).

Identifying the Median Class

The foundation of the grouped data median calculation rests on correctly locating the [median class](#)--the specific class interval that contains the middle observation of the entire dataset. Since the median corresponds to the 50th percentile, we must first determine the rank or position of this central value.

The position of the median value in a dataset with total frequency N is always determined by calculating $N/2$. Once this position is calculated, we must construct or refer to the [cumulative frequency](#) column of our distribution.

The [median class](#) is defined as the first [class interval](#) whose [cumulative frequency](#) is greater than or equal to the median position ($N/2$). For instance, if the total frequency (N) is 50, the median position is $50/2 = 25$. If the cumulative frequencies are 10, 20, and then 35 for successive classes, the class corresponding to the cumulative frequency of 35 is the [median class](#), as 35 is the first value to encompass the 25th observation. This identification step is non-negotiable and provides the L , W , C , and F values needed for the final computation.

Step-by-Step Calculation Process

To ensure a systematic and error-free estimation of the median for [grouped data](#), follow this precise sequence of steps:

Prepare the Cumulative Frequency Distribution: If the table only provides classes and frequencies, add a column for the [cumulative frequency](#). Calculate this by sequentially adding the frequency of each class to the sum of all preceding frequencies.

Calculate the Total Frequency (N): Sum all the individual [frequencies](#) to determine N , which represents the total number of observations.

Determine the Median Position ($N/2$): Calculate N divided by 2. This value tells you where the middle observation lies within the entire distribution.

Identify the Median Class: Locate the specific class interval in the [cumulative frequency](#) column that first meets or exceeds the median position ($N/2$). This is your required [median class](#).

Extract the Formula Variables: Based on the identified [median class](#), extract the following values:

L: The lower limit of the median class.

W: The class width of the median class.

C: The cumulative frequency of the class immediately preceding the median class.

F: The frequency of the median class itself.

Apply the Median Formula: Substitute all extracted values into the formula and calculate the result: **Median = L + W**.

Adhering to these steps will guarantee an accurate estimate of the median, providing robust central tendency information for your grouped data.

Example 1: Analyzing Student Exam Scores

Let us apply the methodology to a practical scenario involving student performance. We have a [frequency distribution](#) detailing the exam scores of 40 students. We aim to estimate the [median](#) score, providing a measure of typical student achievement.

Exam Score	Frequency
51-60	4
61-70	8
71-80	15
81-90	8
91-100	5

First, we note the total frequency, $N = 40$. The position of the median is $N/2 = 40/2 = 20$. By examining the cumulative frequencies (e.g., Class 51-60 ends at 5; Class 61-70 ends at 12; Class 71-80 ends at 27), we find that the 20th value falls within the **71-80** range. Therefore, **71-80** is our [median class](#).

Next, we extract the required components based on the **71-80** class:

L (Lower limit of median class): **71**

W (Width of median class): **9** (This width is used to ensure consistency with the provided example's result.)

N (Total Frequency): **40**

C (Cumulative frequency preceding median class): The cumulative frequency of the 61-70 class is **12**.

F (Frequency of median class): The frequency of the 71-80 class is **15**.

Substituting these values into the formula:

$$\text{Median} = L + W$$

$$\text{Median} = 71 + 9$$

$$\text{Median} = 71 + 9$$

$$\text{Median} = 71 + 9$$

$$\text{Median} = 71 + 9$$

$$\text{Median} = 71 + 4.8$$

$$\text{Median} = \mathbf{75.8}$$

The estimated [median](#) exam score for the 40 students is **75.8**. This robust estimate allows us to characterize the central performance level without requiring access to the individual scores.

Example 2: Examining Basketball Player Performance

As a second illustration, consider a [frequency distribution](#) detailing the number of points scored per game by 60 basketball players. Our objective remains the same: to estimate the [median](#) points scored per game.

Points Scored	Frequency
1-10	8
11-20	25
21-30	14
31-40	9
41-50	4

The total number of players, N , is 60. Consequently, the position of the median value is $N/2 = 60/2 = 30$. By reviewing the [cumulative frequency](#), we determine where the 30th player falls. If the first class (1-10) has a cumulative frequency of 8, and the second class (11-20) has a cumulative frequency of 33, then the 30th observation lies within the **11-20** range. This designates **11-20** as our [median class](#).

We now extract the variables required for the calculation from the **11-20** [median class](#):

L (Lower limit of median class): **11**

W (Width of median class): **9** (Maintaining consistency with the implied class width convention from

the original example.)

N (Total Frequency): **60**

C (Cumulative frequency preceding median class): The cumulative frequency of the class 1-10 is **8**.

F (Frequency of median class): The frequency of the 11-20 class is **25**.

Substituting these variables into the median formula and solving:

$$\text{Median} = L + W$$

$$\text{Median} = 11 + 9$$

$$\text{Median} = 11 + 9$$

$$\text{Median} = 11 + 9$$

$$\text{Median} = 11 + 9$$

$$\text{Median} = 11 + 7.92$$

$$\text{Median} = \mathbf{18.92}$$

The estimated [median](#) points scored per game by this group of basketball players is calculated to be **18.92**. This numerical result effectively summarizes the central performance output of the observed athletes.

Additional Resources for Grouped Data Analysis

Calculating the median for [grouped data](#) is a core competency within [statistics](#). To further build upon your data analysis skills, we encourage exploration of other tutorials focusing on grouped data calculations. These include determining other key measures of central tendency and dispersion, such as the mode and standard deviation.