Find the Range of Grouped Data (With Examples)

Authored by

Mohammed looti

October 29, 2025

Estimating Dispersion: The Range of Grouped Data

In statistical analysis, large collections of observations are often organized into grouped data, where individual measurements are summarized into distinct class intervals instead of being listed separately. This practice streamlines the handling of voluminous datasets, making complex statistical operations more feasible. A fundamental metric for assessing the variability or spread of data is the range, traditionally calculated by finding the difference between the maximum and minimum values observed within a given collection of data points.

For raw, ungrouped data, calculating the range is straightforward: maximum value minus minimum value. However, when working with grouped data, the precise individual data points are obscured; we only know the boundaries within which the data falls. Consequently, determining the exact range becomes impossible. To overcome this limitation and still provide a valuable measure of data dispersion, statisticians employ specialized estimation techniques.

Despite the lack of precise observations, it is entirely possible to derive a meaningful estimate of the range for grouped datasets. Statisticians primarily utilize two distinct estimation methodologies, each providing a unique perspective on the data's overall variability. These methods allow analysts to quantify the potential spread of the dataset, offering crucial insights even without access to every single observation.

To illustrate the context of our calculations, consider the following example of grouped data, categorized into various classes alongside their associated frequencies:

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

As the table demonstrates, while the class intervals are clearly defined, the specific values within each interval remain unknown. To accurately estimate the range, we must apply one of the specialized formulas designed for this purpose. The selection of the appropriate formula often depends on the desired analytical goal and the assumptions made about the data's distribution within the intervals. We will now explore these two principal formulas in detail.

Formula 1: Utilizing Extreme Class Limits

The first standard method for calculating the estimated range of grouped data relies on identifying and using the absolute extreme class limits of the entire frequency distribution. This approach establishes the widest possible boundaries for the dataset. It operates under the principle that the smallest possible value in the dataset is represented by the lower boundary of the lowest class interval, and conversely, the largest possible value is represented by the upper boundary of the highest class interval.

The fundamental formula for this conservative estimation method is defined as follows:

Range of Grouped Data = Umax - Lmin

The components of this formula are precisely defined by the boundaries of the dataset:

Umax: Denotes the upper limit of the maximum (highest) class interval within the frequency distribution. This value signifies the highest potential observation, defining the upper bound of the range.
Lmin: Represents the lower limit of the minimum (lowest) class interval in the frequency distribution. This value signifies the lowest potential observation, establishing the lower bound of the range.

This formula typically results in the broadest possible estimate of the range, as it spans the entire observable extent defined by the class boundaries. It is often favored when analysts wish to account for the full potential spread of the data, assuming that values could exist right up to the outermost boundaries of the initial and final classes.

Formula 2: Leveraging Class Midpoints

An alternative and frequently applied methodology for estimating the range of grouped data involves using the midpoints of the extreme class intervals. This approach is grounded in the premise that data points within any class interval often cluster around the central value, or midpoint, of that interval. By focusing on these central tendencies rather than the absolute boundaries, we aim to provide a more representative or typical estimate of the data's overall spread.

The specific formula for this midpoint-based estimation is presented below:

Range of Grouped Data = Midpointmax - Midpointmin

The required terms for this calculation are derived from the center of the extreme classes:

Midpointmax: Refers to the midpoint of the highest class interval. This is calculated by averaging the upper and lower limits of that specific highest interval, providing the center of the highest data group.
Midpointmin: Refers to the midpoint of the lowest class interval. Similarly, this is calculated by averaging the upper and lower limits of the lowest interval, representing the center of the lowest data group.

This formula offers a distinct analytical perspective, often resulting in a narrower estimated range compared to the method utilizing extreme class limits. It suggests that the spread is determined by where the majority of the data clusters within the extreme intervals, making it particularly useful when data is not expected to be uniformly spread across the class boundaries.

Example 1: Calculating the Range of Exam Scores

To solidify our understanding, let us apply both formulas to a practical scenario involving the exam scores obtained by 40 students. This performance data is efficiently summarized in the following frequency distribution table:

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

Our objective is to calculate and compare the estimated range for these exam scores using the two primary methodologies: Formula 1 (Extreme Class Limits) and Formula 2 (Class Midpoints).

Using Formula 1: Extreme Class Limits (Exam Scores)

To apply Formula 1, we must first locate the absolute outermost boundaries of the dataset within the frequency table:

The highest class interval is 91-100. Therefore, the maximum possible score (**Umax**) is **100**.
The lowest class interval is 51-60. Therefore, the minimum possible score (**Lmin**) is **51**.

Substituting these values into the formula yields the following result:

Range of Grouped Data = Umax - Lmin
Range of Grouped Data = 100 - 51
Range of Grouped Data = 49

Using the extreme class limits, the estimated range for the student exam scores is **49**. This suggests a maximum potential spread of 49 points across the entire class performance.

Using Formula 2: Class Midpoints (Exam Scores)

For Formula 2, we calculate the midpoints for the highest and lowest class intervals:

For the highest class interval (91-100), the **Midpointmax = (91 + 100) / 2 = 95.5**.
For the lowest class interval (51-60), the **Midpointmin = (51 + 60) / 2 = 55.5**.

Substituting these midpoints into the second formula:

Range of Grouped Data = Midpointmax - Midpointmin
Range of Grouped Data = 95.5 - 55.5
Range of Grouped Data = 40

Based on the class midpoints, the estimated range for the exam scores is **40**. This result is narrower, reflecting the assumption that the actual scores are concentrated closer to the center of their respective extreme intervals.

Example 2: Analyzing Basketball Player Points

Our second example focuses on quantifying the scoring variability of basketball players. We examine the number of points scored per game by 60 players, summarized in the following frequency distribution:

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

We will once again apply both estimation formulas to this dataset to reinforce the calculation process and demonstrate the consistent application of these statistical methods across different data types.

Using Formula 1: Extreme Class Limits (Basketball Points)

Applying Formula 1 requires identifying the extreme class limits for the basketball points distribution:

The highest class interval is 41-50. Thus, **Umax = 50**.
The lowest class interval is 1-10. Thus, **Lmin = 1**.

The resulting calculation is as follows:

Range of Grouped Data = Umax - Lmin
Range of Grouped Data = 50 - 1
Range of Grouped Data = 49

Based on the extreme class limits, the estimated range of points scored is **49**. This represents the widest potential spread between the lowest and highest scoring performances.

Using Formula 2: Class Midpoints (Basketball Points)

Next, we apply Formula 2 by calculating the midpoints of the extreme intervals:

For the highest class interval (41-50), the **Midpointmax = (41 + 50) / 2 = 45.5**.
For the lowest class interval (1-10), the **Midpointmin = (1 + 10) / 2 = 5.5**.

Substituting these midpoints into the formula:

Range of Grouped Data = Midpointmax - Midpointmin

Range of Grouped Data = 45.5 - 5.5
Range of Grouped Data = 40

Using the midpoints, the estimated range for the basketball player points is **40**. This narrower estimate suggests that the central values of the extreme scoring groups are 40 points apart.

Interpreting the Estimated Range Values

As clearly illustrated by both preceding examples, the two methods for estimating the range of grouped data consistently yield different values. Formula 1, which relies on the extreme class limits (Umax - Lmin), invariably produces a broader estimate. This is because it defines the range based on the absolute outermost potential boundaries of the data, offering a maximally conservative estimate of the data's entire possible spread.

In contrast, Formula 2, which utilizes class midpoints (Midpointmax - Midpointmin), results in a narrower estimated range. This methodology is predicated on the assumption that data points within each interval are either uniformly distributed or, more commonly, cluster significantly around the midpoint. Thus, it offers an estimate that reflects the spread based on the central tendencies of the extreme classes, which may provide a more realistic view if data is not expected to be spread evenly across the class boundaries.

The choice between these two estimation techniques must be consciously guided by the specific analytical goals and the underlying assumptions an analyst is willing to make regarding the distribution of data within each interval. Both are statistically valid approaches for estimating the range, but they communicate subtly different interpretations of the data's dispersion. A comprehensive understanding of these differences is essential for accurate statistical reporting and informed decision-making when working with grouped data.

Additional Resources for Grouped Data Analysis

To further enhance your expertise and proficiency in performing statistical calculations with grouped data, we strongly recommend exploring tutorials on related measures of dispersion and central tendency. These resources can provide deeper insights into other common operations and analyses that are frequently applied to grouped datasets, such as finding the variance or standard deviation.

How to Find the Variance of Grouped Data