Understanding Quartiles: A Step-by-Step Guide for Even and Odd Datasets

Authored by

Mohammed looti

March 2, 2026

RECOMMENDED CITATION

Mohammed looti (2026). *Understanding Quartiles: A Step-by-Step Guide for Even and Odd Datasets*. PSYCHOLOGICAL STATISTICS. Retrieved from https://statistics.arabpsychology.com/?p=3163

The Essential Role of Quartiles and the Foundation of Median Calculation

In the broad field of descriptive statistics, **quartiles** stand out as fundamental measures that provide critical insights into the spread, variability, and central tendency of a **dataset**. They are highly specialized positional values that systematically segment an ordered distribution into four perfectly equal parts, guaranteeing that each segment contains exactly 25% of the total observations. This structured division offers a considerably more granular perspective on data distribution compared to relying solely on basic measures like the arithmetic mean, thereby enabling precise identification of potential outliers and a clearer understanding of the data's overall shape and skewness.

The structure of quartile analysis relies on three principal markers. The **first quartile** (Q1) defines the 25th percentile, establishing a boundary below which 25% of the data points reside. The critical second quartile (Q2) is, by definition, the median of the entire dataset, representing the definitive 50th percentile mark. Completing the set, the **third quartile** (Q3) indicates the 75th percentile, meaning 75% of all observations fall below this value. The distance spanning from Q1 to Q3 is formally termed the **interquartile range** (IQR), which functions as a robust, non-parametric measure of statistical dispersion, highly resistant to the influence of extreme values.

A prerequisite for accurate quartile calculation is the precise determination of the median. The median, or Q2, is the absolute midpoint of an ordered sequence of numbers. The method used to locate this pivot point depends entirely on the size, or length, of the data collection. If a dataset possesses an **odd number** of elements, the median is simply the value that occupies the central position. Conversely, if the dataset contains an **even number** of values, the median must be calculated as the average of the two innermost values. Establishing this central median is paramount, as it serves as the definitive point for partitioning the data into the lower and upper halves from which Q1 and Q3 are subsequently derived.

The Precise Methodology for Even-Length Datasets

The computation of **quartiles** for a dataset containing an **even number** of observations necessitates a meticulous, step-by-step approach to ensure statistical accuracy. The foundational step, universally applicable across all quartile calculations, involves arranging every single data point in strictly ascending order--from the minimum observed value to the maximum. This initial sorting process is non-negotiable because quartiles are inherently positional measures, meaning their values are entirely dependent upon the rank and sequence of the data points within the ordered list.

Following the ordering of the data, the subsequent crucial task is locating the overall median (Q2). Because we are dealing with an even number of data points, there is no single middle value that acts as the median. Instead, Q2 is determined by calculating the average of the two observations situated symmetrically around the center of the dataset. For instance, if a dataset contains 20 observations, the 10th and 11th values would be summed and divided by two to yield the median. This calculated median value then conceptually divides the entire dataset into two perfectly balanced groups: the lower half (LHS) and the upper half (RHS), which will be used for the subsequent quartile calculations.

To determine the **first quartile** (Q1) and the **third quartile** (Q3), we analyze these two newly defined halves independently. Q1 is defined as the median of the lower half of the dataset. Critically, because the median (Q2) for even-length data is a calculated value and not an actual data point, it is naturally **not included** in either the lower or upper halves. Similarly, Q3 is calculated as the median of the upper half of the dataset. This strict exclusion of the conceptual median from the sub-datasets is fundamental to the method detailed here, ensuring consistency, especially when utilizing standardized statistical tools such as **TI-84 calculators**.

Executing Quartile Calculation for Odd-Length Datasets

When the analysis shifts to a dataset characterized by an **odd number** of values, the methodology for identifying **quartiles** differs slightly, predominantly concerning the treatment of the overall median. As always, the initial and most crucial action is the meticulous arrangement of all numerical data points in a single, ascending sequence. This ordered arrangement is the immutable basis for all subsequent positional calculations that define the quartiles.

For a dataset with an odd number of observations, the determination of the overall median (Q2) is straightforward and unambiguous: it is the single, unique data point located precisely in the middle of the ordered list. This specific central value serves as the definitive pivot point for the entire distribution. For example, if a list contains 15 values, the 8th value is the median. Once identified, this median value conceptually separates the dataset into two distinct sub-datasets: a lower half consisting of all values preceding the median, and an upper half comprising all values succeeding the median.

The next step involves calculating the **first quartile** (Q1) and the **third quartile** (Q3) by finding the median of these newly isolated halves. Q1 is the median of the lower half. Crucially, in the methodology favored by many educational institutions and software (such as **TI-84 calculators**), the overall median value of the full dataset is strictly **excluded** from both the lower and upper halves when calculating Q1 and Q3. Similarly, Q3 is the median of the upper half, with the overall median point rigorously excluded from consideration. This consistent exclusion ensures adherence to the specific convention utilized in these popular tools.

Understanding Methodological Variations and Conventions in Quartile Calculation

It is essential for any professional dealing with statistics or data analysis to acknowledge a critical fact: there is no single, globally standardized method for calculating **quartiles**, particularly when applied to **discrete distributions**. Discrepancies often arise because different statistical software packages, academic curricula, and authoritative textbooks employ slightly different interpolation formulas, which can result in minor variations in the calculated Q1 and Q3 values for the identical dataset. This lack of strict universal agreement fundamentally revolves around how the overall median (Q2) is treated when the data is partitioned.

Specifically, the variation centers on whether the overall median should be included or excluded from the lower and upper sub-datasets when calculating the quartiles. Some methodologies, often referred to as the "inclusive method," may dictate that the overall median point--especially in odd-length datasets where Q2 is an observed data point--should be included in both the lower half (for Q1 calculation) and the upper half (for Q3 calculation). This approach contrasts sharply with the "exclusive method," which ensures the median is always treated as the separation point and never included in the sub-datasets used for finding the first quartile (Q1) and third quartile (Q3).

The calculation steps and formulas presented throughout this guide strictly adhere to the exclusive methodology, which is the convention adopted by **TI-84 calculators** and is widely utilized in secondary and collegiate educational settings. This specific convention provides a robust and repeatable method for students and analysts. While acknowledging the existence of alternative methods (such as the Mendenhall and Sincich method or the Tukey method), mastering this TI-84 convention provides a strong, practical foundation. When presenting or comparing statistical results, it is imperative to clearly state the specific quartile calculation method used to ensure accurate communication and prevent misinterpretation.

Worked Example 1: Calculating Quartiles for an Even-Length Dataset

To solidify the theoretical understanding, we now apply the exclusive methodology to a concrete example involving an **even-length dataset**. Consider a scenario where we have collected ten distinct numerical values. Our first, non-negotiable step is to arrange these raw values in ascending sequence to prepare them for positional analysis, ensuring all subsequent calculations are based on the correct ranks.

The ordered data points are presented below:
Data: 3, 3, 6, 8, 10, 14, 16, 16, 19, 24

Since the count of observations is ten (an even number), the overall median (Q2) must be calculated as the average of the two central values. In this list, these are the 5th value (10) and the

6th value (14). The calculation is (10 + 14) / 2 = **12**. This calculated median of 12 acts as the conceptual dividing line between the lower and upper halves of the data. Consistent with the exclusive method, this value of 12 is **not included** in either of the sub-datasets used for the quartile calculations.

The lower half of the dataset, consisting of the five values preceding the median, is: 3, 3, 6, 8, 10. To find the **first quartile** (Q1), we determine the median of this sub-list. Since this lower half has five values (an odd number), the median is the single middle value, which is the 3rd value. Thus, Q1 is calculated as **6**.
Q1 sub-dataset: 3, 3, **6**, 8, 10

The upper half of the dataset, comprising the five values following the median, is: 14, 16, 16, 19, 24. To find the **third quartile** (Q3), we calculate the median of this upper sub-list. With five values, the median is the single middle value, which is the 3rd value in this sub-list. Hence, Q3 equals **16**.
Q3 sub-dataset: 14, 16, **16**, 19, 24

Worked Example 2: Calculating Quartiles for an Odd-Length Dataset

We now examine a scenario involving a **quartile** calculation for a dataset with an odd number of observations. For this illustration, we use a dataset containing nine values, ensuring they are already correctly sorted in ascending order, ready for analysis. This example highlights the differences in treating the median when it is an actual data point.

The ordered data points are:
Data: 3, 3, 6, 8, 10, 14, 16, 16, 19

Given that there are nine observations (an odd number), the overall median (Q2) is the value positioned exactly in the middle. This is the 5th value in the ordered list, which is the number **10**. Following the exclusive methodology adopted in this guide (aligned with **TI-84 calculators**), this specific median data point (10) is **excluded** when the dataset is partitioned into the lower and upper halves for subsequent quartile calculation.

The lower half of the dataset, excluding the median, consists of the four values: 3, 3, 6, 8. To calculate the **first quartile** (Q1), we must find the median of this lower half. Since this sub-dataset has an even number of values (four), its median is the average of the two middle values (the 2nd and 3rd values). Consequently, Q1 = (3 + 6) / 2 = **4.5**.
Q1 sub-dataset: 3, **3**, **6**, 8

The upper half of the dataset, also excluding the overall median, includes the four values: 14, 16, 16, 19. Similarly, the **third quartile** (Q3) is the median of this upper half. With four values, its median is the average of its two central values (the 2nd and 3rd values). Therefore, Q3 = (16 + 16)

/ 2 = **16**.
Q3 sub-dataset: 14, **16**, **16**, 19

These two detailed examples demonstrate that while the core principle of finding the median of the halves remains constant, the initial determination and subsequent exclusion of the overall median (Q2) must be carefully handled depending on whether the original dataset length is even or odd.

Expanding Your Statistical Toolkit: Recommended Resources and Further Study

Achieving proficiency in statistical analysis requires more than just understanding a single method; it demands a broad capability to utilize various tools and conventions. To further enhance your analytical skills and deepen your comprehension of **quartiles** and other fundamental descriptive statistics, we strongly advise exploring supplementary tutorials that illustrate how to calculate these positional measures using different software environments (such as R, Python's NumPy/SciPy, or Excel). Exposure to multiple platforms will inevitably strengthen your ability to interpret results derived from varying methodological conventions.

The specific methodology detailed in this guide--the exclusive method--is highly relevant to educational contexts due to its alignment with standard handheld graphing calculators. However, analysts should be aware that enterprise-level statistical packages may default to different calculation types (e.g., methods 6, 7, or 8 in R's `quantile()` function). Understanding these nuances is crucial for accurate professional reporting and collaborative data analysis.

The following resources are recommended to continue your statistical journey and explore alternative quartile calculation techniques:

Review official documentation for statistical software like R or Python libraries to understand their default percentile calculation methods.
Consult advanced statistics textbooks focusing on non-parametric methods and robust measures of dispersion.
Practice calculating the Interquartile Range (IQR) and creating **box plots**, as these visualizations rely entirely on the accurately determined quartiles (Q1, Q2, Q3).
Explore tutorials demonstrating the use of **TI-84 calculators** for calculating one-variable statistics, ensuring the results match the manual calculations derived using the exclusive method.