

# How to Create and Interpret a Relative Frequency Histogram

Authored by  
**stats writer**

March 7, 2026

## RECOMMENDED CITATION

stats writer (2026). *How to Create and Interpret a Relative Frequency Histogram*.  
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=134346>

## Foundations of Statistical Data Representation and Frequency Analysis

In the vast field of **statistics**, the ability to organize and interpret raw data is a fundamental skill that allows researchers to draw meaningful conclusions from seemingly chaotic numbers. One of the primary methods for achieving this is through the use of a **frequency** distribution, which serves as the bedrock for more complex visual tools. At its core, frequency refers to the number of times a specific value or a range of values occurs within a given **dataset**. By categorizing data into these specific occurrences, we can begin to see the underlying structure of the information, whether we are analyzing consumer behavior, scientific measurements, or financial trends. This process of categorization is the first step toward building a comprehensive **histogram**, a graphical tool that has been utilized by statisticians for over a century to provide a visual summary of data density and distribution.

A **relative frequency** histogram takes this analysis a step further by shifting the focus from absolute counts to proportions. Instead of merely stating how many times an event occurred, a relative frequency approach tells us what percentage of the total observations fall within a specific category. This transformation is crucial for comparing datasets of different sizes; for instance, comparing the test scores of a class of 30 students to a university-wide exam involving thousands of participants would be nearly impossible using raw frequencies alone. By utilizing **percentages** or proportions, the **data visualization** becomes standardized, allowing for a clear and concise comparison of the shape, center, and spread of the data regardless of the sample size. This article will explore the definition, construction, and practical application of these histograms to enhance your analytical capabilities.

The importance of choosing the correct visualization method cannot be overstated in the context of modern data science. While a standard histogram provides a clear view of the volume of data, the relative frequency histogram offers a more nuanced perspective on the **probability distribution** of the variables involved. Understanding how to calculate these values and represent them graphically is essential for anyone looking to present data in a way that is both professional and academically rigorous. In the following sections, we will delve into the specific mechanics of **frequency tables**, the transition to graphical formats, and the distinct advantages of using relative frequencies in various analytical scenarios. Through a detailed example involving retail sales data, we will illustrate exactly how these concepts are applied in a real-world setting.

### The Role of Frequency Tables in Data Organization

To begin our exploration, we must first examine the **frequency table**, which acts as the organizational precursor to any histogram. A frequency table is a structured method of displaying the number of observations that fall within various non-overlapping intervals, often referred to as "classes" or "bins." This format is particularly useful when dealing with **continuous data** or large

sets of discrete data where individual values might repeat infrequently. By grouping these values into intervals, such as price ranges or age groups, we can simplify the dataset without losing the essential characteristics of the distribution. This simplification is vital for identifying the mode, or the most frequent class, and for understanding the overall range and **variance** within the data.

Consider a practical scenario where a retail shop tracks the prices of items sold during a standard business week. Without organization, a list of hundreds of individual transaction prices would be difficult to interpret at a glance. However, by constructing a frequency table, the shop owner can quickly identify which price points are driving the most traffic. In this table, the first column defines the "class" (the range of prices), while the second column provides the **frequency** (the number of items sold within that range). This systematic approach allows for the identification of trends, such as whether customers prefer budget-friendly items or if there is a significant market for premium products. The clarity provided by such a table is the first step in effective **data analysis**.

The following table illustrates how a particular shop might categorize its weekly sales based on the price of the items sold, providing a clear example of how raw data is transformed into a structured **frequency distribution**:

Item Price	Frequency
\$1 - \$10	20
\$11 - \$20	21
\$21 - \$30	13
\$31 - \$40	8
\$41 - \$50	4

In this specific example, the classes are defined by ten-dollar increments, which provides a balanced level of detail for the shop's inventory. We can see that the most frequent sales occur in the \$11 - \$20 range, with 21 items sold, while the \$41 - \$50 range sees the fewest sales. This tabular data is functional, but it lacks the immediate impact of a visual representation. To truly grasp the "shape" of these sales--whether they are skewed toward lower prices or follow a **normal distribution**--we must move beyond the table and into the realm of **data visualization** through the creation of a frequency histogram.

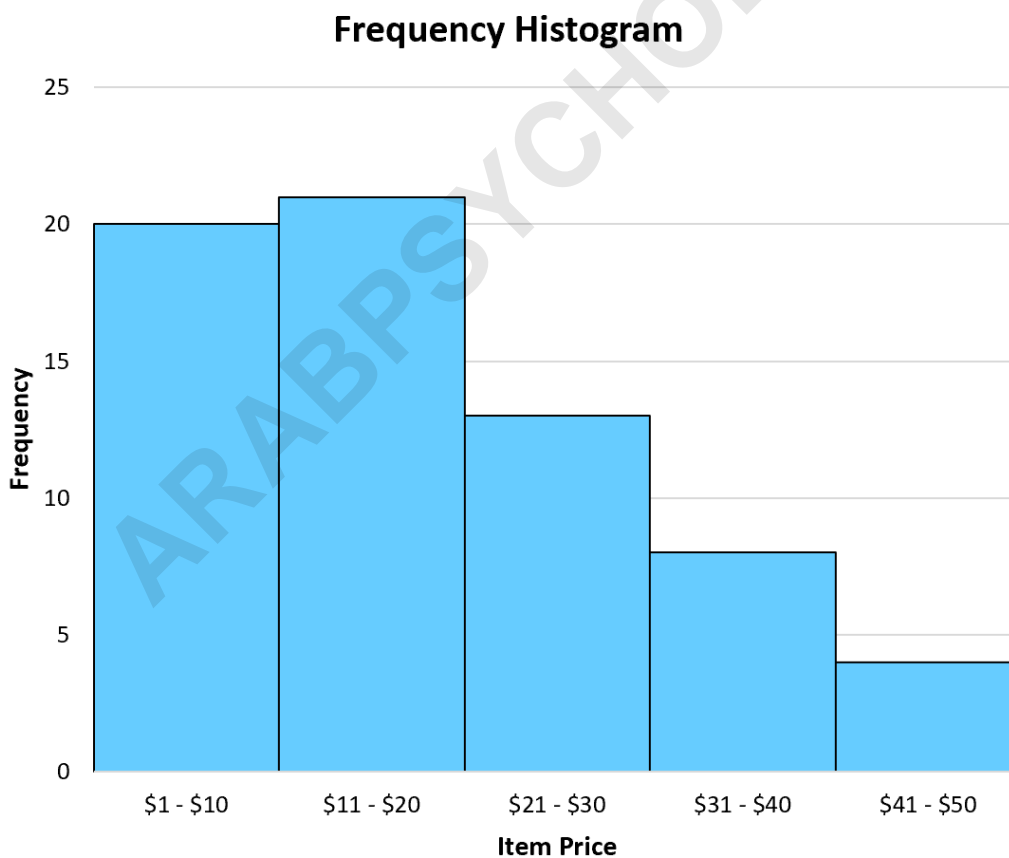
## Transitioning from Tabular Data to Frequency Histograms

While tables are excellent for precision, the human brain is naturally wired to process visual information more efficiently than raw numbers. This is where the **histogram** becomes an indispensable tool. A frequency histogram is a graphical version of the frequency table, where the horizontal **x-axis** represents the classes (the price ranges in our example) and the vertical **y-axis**

represents the absolute frequency of each class. The height of each bar corresponds directly to the number of observations within that interval. Unlike a bar chart, which is used for categorical data, a histogram is used for quantitative data, and the bars are typically drawn touching each other to indicate the continuous nature of the underlying scale.

By visualizing the retail sales data, the shop owner can instantly see the distribution's skewness. If the bars are higher on the left side of the graph and taper off to the right, the data is **positively skewed**, indicating that lower-priced items are sold more frequently. This visual insight is much more powerful than a table for presentations or quick decision-making. The histogram provides a "big picture" view of the **dataset**, highlighting outliers or gaps in the data that might not be immediately obvious in a list of numbers. It allows for the identification of patterns, such as a bimodal distribution where two different price points are equally popular, which could suggest two distinct customer bases.

The following frequency histogram provides a visual representation of the sales data categorized in the table above, demonstrating how the heights of the bars make the **frequency** of each class easily comparable at a glance:



In this graph, the **frequencies** are clearly mapped, showing a peak in the second bin. However, a

significant limitation of this standard frequency histogram is that it only tells us about this specific week's sales in absolute terms. If we wanted to compare this week's performance to a holiday week where 1,000 items were sold, the bars on this graph would look insignificant by comparison. To solve this problem and gain a deeper understanding of the distribution's internal proportions, we must transition to the **relative frequency** histogram, which standardizes the data and provides a more universal perspective on the information.

## The Mathematical Concept of Relative Frequency

The transition from raw frequencies to **relative frequency** involves a simple but powerful mathematical calculation. To find the relative frequency of a class, you must divide the frequency of that specific class by the total number of observations in the entire **dataset**. The formula is expressed as: **Relative Frequency = Class Frequency / Total Sample Size**. The resulting value is a decimal or a fraction between 0 and 1, which can also be expressed as a **percentage** by multiplying the result by 100. This process effectively normalizes the data, making it possible to understand the "weight" of each category relative to the whole, rather than focusing on the raw count.

Applying this to our retail example, the first step is to calculate the total number of items sold across all price ranges. By summing the frequencies (20 + 21 + 13 + 8 + 4), we determine that the total sample size is 66. Once this total is established, we can calculate the relative frequency for each interval. This calculation is essential because it provides a **probability**-based view of the data. For instance, if we were to pick a random transaction from that week, the relative frequency tells us the likelihood that the transaction falls within a certain price range. This shift in perspective is what distinguishes descriptive **statistics** from more advanced inferential techniques.

Understanding **relative frequency** is also vital for the creation of **cumulative frequency** distributions, which track the sum of frequencies as you move across the classes. However, for the purpose of constructing a relative frequency histogram, the primary goal is to ensure that the sum of all relative frequencies equals exactly 1 (or 100% if using percentages). This ensures that the entire dataset is accounted for and that the proportions are accurate. By using these proportions, statisticians can compare different groups--such as sales at different store locations--without being misled by the total volume of sales at each location, thereby providing a much more accurate comparison of consumer trends.

## Constructing and Interpreting Relative Frequency Tables

Once the mathematical calculations are complete, the data is organized into a **relative frequency table**. This table is identical in structure to a standard frequency table, but it includes an additional column for the calculated proportions. This column is the key to creating a relative frequency

histogram. By presenting the data in this way, we can see exactly how much each "class" contributes to the total. This is particularly useful in business environments for **market segmentation** and resource allocation, as it highlights which segments are the most significant in terms of their overall share of the market.

Let us look at the updated table for our retail shop, which now includes the relative frequencies for each of the item price ranges. These values provide a clearer picture of the percentage of total sales each range represents:

Item Price	Frequency	Relative Frequency
\$1 - \$10	20	0.303
\$11 - \$20	21	0.318
\$21 - \$30	13	0.197
\$31 - \$40	8	0.121
\$41 - \$50	4	0.061

To interpret this table, we look at the relative frequency of the \$1 - \$10 price range. With 20 items sold out of a total of 66, the calculation is  $20 / 66$ , which equals approximately **0.303**, or 30.3%. Similarly, the \$11 - \$20 range, with 21 items sold, has a relative frequency of  $21 / 66$ , which is approximately **0.318**, or 31.8%. This tells the shop owner that over 60% of their sales are concentrated in items priced under \$20. Without these calculations, the owner might know that those items are popular, but they wouldn't know exactly what portion of the business they represent. This level of detail is crucial for effective **financial analysis** and inventory management.

By performing these calculations for every class, the **relative frequency distribution** is completed. It is important to note that due to rounding, the sum of these decimals might occasionally be slightly off from 1.000 (e.g., 0.999 or 1.001), but for all practical **statistical** purposes, they represent the total whole. This table now serves as the direct data source for the vertical bars in our next visualization step: the relative frequency histogram. This transition ensures that the visual output is a direct reflection of the proportional data, providing a standardized look at the **dataset**.

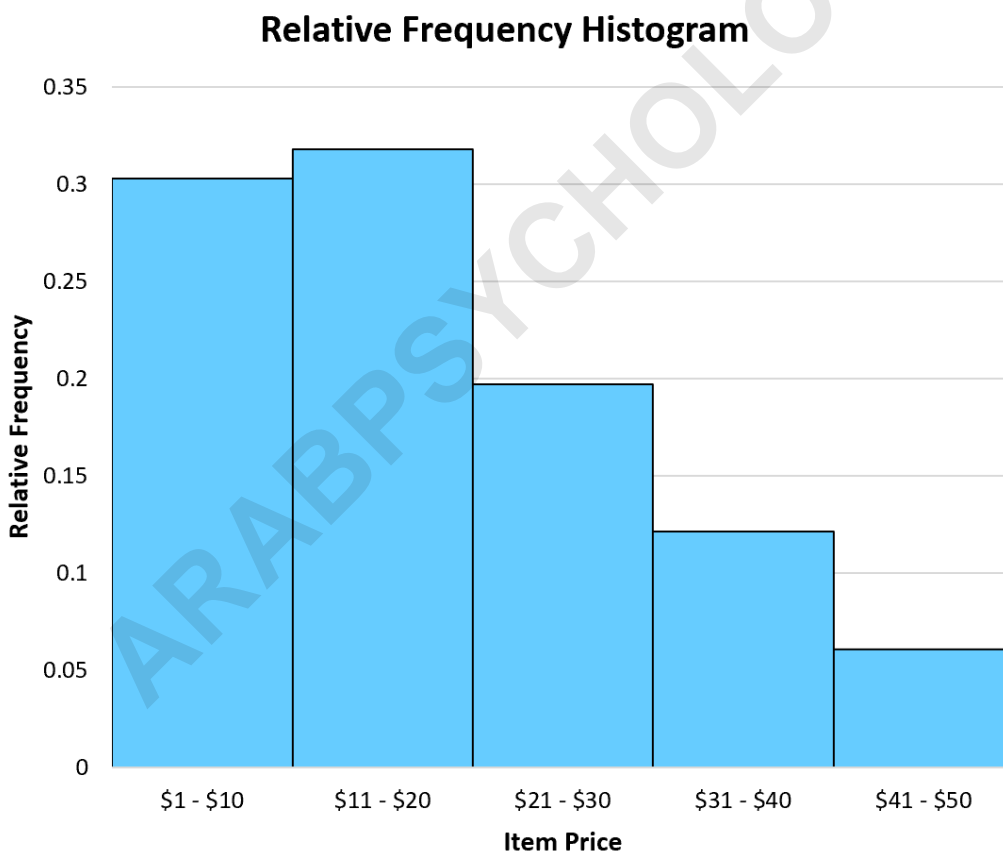
## Visualizing Proportions through Relative Frequency Histograms

The **relative frequency** histogram is the visual culmination of the data organization process. Structurally, it is identical to the standard frequency histogram, with the same horizontal **x-axis** showing the classes. However, the vertical **y-axis** is now labeled with percentages or decimals rather than raw counts. This shift is subtle but transformative. The primary benefit of this **data visualization** is that it allows the viewer to interpret each bar as a portion of the whole. If a bar

reaches the 0.30 mark on the y-axis, the viewer knows immediately that 30% of the data falls within that interval, regardless of whether the total sample size was 66 or 66,000.

In a relative frequency histogram, the total area of all the bars combined is equal to 1 (or 100%). This property makes the histogram a visual representation of a **probability density function** for discrete intervals. It allows researchers to visualize the "concentration" of data. For example, a very tall, narrow bar indicates a high concentration of data in a small range, while short, wide bars indicate that the data is spread out more evenly. This ability to visualize the "density" of the **dataset** is why the relative frequency histogram is a staple in fields like economics, sociology, and quality control, where understanding the distribution of a population is more important than knowing individual counts.

The image below displays the relative frequency histogram for the retail shop sales. Notice how the shape of the distribution remains the same as the previous histogram, but the scale on the left side has been updated to reflect proportions:



This graph is particularly powerful for setting and monitoring business goals. Instead of aiming for a specific number of sales, which can fluctuate based on foot traffic or seasonality, a manager can set a goal based on the **relative frequency**. For instance, if the goal is to have high-end items

(\$41 - \$50) make up at least 5% of total sales, the relative frequency histogram provides an immediate visual confirmation of whether that goal is being met. In our example, the bar for the \$41 - \$50 range sits at 0.061, or 6.1%, meaning the shop is successfully exceeding its target proportion for premium items.

## Comparative Analysis: Frequency vs. Relative Frequency Histograms

When deciding whether to use a standard **frequency** histogram or a **relative frequency** histogram, the choice depends entirely on the goals of the analysis. A standard frequency histogram is most appropriate when the absolute volume of the data is the primary concern. For instance, a warehouse manager needs to know exactly how many units are in stock to manage physical space. In this case, knowing that "10% of items are large" is less helpful than knowing "there are 500 large items," because the physical space is limited by count, not by percentage. Frequency histograms are essential for operational logistics and capacity planning.

Conversely, relative frequency histograms are the preferred choice when the focus is on the distribution's **shape** or when comparing multiple **datasets**. If you are comparing the income distribution of two different countries, using raw frequencies would be misleading because their population sizes are different. By using relative frequencies, you can overlay the two histograms to see which country has a higher proportion of middle-income earners, regardless of population. This standardizing effect is what makes the relative frequency histogram a more versatile tool for **statistical** comparison and theoretical modeling.

It is fascinating to note that the physical "silhouette" or shape of both histograms will be identical if the same bin widths are used. The bars will rise and fall in the exact same pattern. The only difference lies in the **y-axis** scaling. This means that a relative frequency histogram preserves all the visual information about the distribution's skewness, **kurtosis**, and modality, while adding a layer of proportional context. For most academic and research purposes, the relative frequency version is considered more sophisticated because it facilitates a deeper understanding of the **probability** of observing specific values within the population being studied.

## Practical Use Cases and Decision Making in Data Science

In the real world, the application of relative frequency histograms extends far beyond simple retail examples. In **quality control**, for instance, manufacturers use these histograms to monitor the proportions of products that fall outside of acceptable tolerances. If a relative frequency histogram of bolt diameters shows that 2% of the products are too small, the company can calculate the potential cost of waste relative to their total production volume. This allows for data-driven decisions regarding machine maintenance or process adjustments. The ability to see these errors as a **percentage** helps management understand the severity of the issue in the context of the

entire operation.

Furthermore, in the field of **epidemiology**, researchers use relative frequency histograms to visualize the age distribution of patients affected by a specific disease. By comparing the relative frequencies of age groups across different regions, health officials can identify which demographic is most at risk. This information is vital for the targeted allocation of vaccines or public health resources. In this context, raw numbers might be confusing if one region has a much larger population than another, but the relative frequencies provide a clear, standardized signal that guides critical healthcare **decision-making**.

Ultimately, mastering the use of these **visual tools** is about more than just drawing graphs; it is about choosing the most effective way to communicate complex information. Whether you are a student, a researcher, or a business professional, understanding the nuances of **relative frequency** histograms allows you to present data that is not only accurate but also highly insightful. By focusing on proportions and percentages, you provide your audience with a standardized framework for understanding the world, making your findings more accessible and your arguments more persuasive. As data continues to grow in volume and importance, the ability to visualize it proportionally will remain a cornerstone of effective **statistical** literacy.

## Software Implementation and Best Practices

In the modern era, creating a relative frequency histogram is typically handled by specialized software, which automates the calculation of bins and proportions. Programs like **Microsoft Excel**, **R**, and **Python** (using libraries like Matplotlib or Seaborn) have built-in functions to generate these charts from raw **datasets**. When using these tools, it is important to follow best practices, such as choosing an appropriate number of bins. Too few bins can oversimplify the data and hide important patterns, while too many bins can make the histogram look jagged and difficult to interpret. A common rule of thumb is the **Sturges' formula**, which helps determine the optimal number of intervals based on the size of the dataset.

Another best practice is to always clearly label your axes and provide a descriptive title. In a **relative frequency** histogram, it is especially important to specify whether the y-axis represents decimals or **percentages**. This ensures that the viewer can accurately read the values. Additionally, if you are comparing two different distributions, using semi-transparent bars or "step" histograms can help the viewer see where the distributions overlap and where they differ. These technical adjustments can significantly enhance the clarity and professional quality of your **data visualizations**, ensuring that your message is communicated effectively to your audience.

The following tutorials and guides offer in-depth explanations on how to leverage various statistical software packages to create high-quality relative frequency histograms. By mastering these tools, you can move from manual calculations to sophisticated **data analysis**, allowing you to handle

larger datasets and more complex distributions with ease. Whether you are conducting academic research or providing business insights, these software skills are a valuable asset in the data-driven landscape of the 21st century.

ARABPSYCHOLOGY.COM