

How to Easily Interpolate Missing Data in Google Sheets

Authored by
stats writer

November 28, 2025

RECOMMENDED CITATION

stats writer (2025). *How to Easily Interpolate Missing Data in Google Sheets*.

PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=101155>

Introduction to Data Interpolation in Google Sheets

Interpolation is a critical process in Data Analysis, involving the estimation of unknown values that fall within a range of known data points. When working with time series data or large datasets in applications like Google Sheets, it is common to encounter missing entries. These gaps, if left untreated, can skew results, hinder trend analysis, and compromise the integrity of statistical modeling. Effective data cleaning techniques often require filling these voids using mathematically sound methods, and linear interpolation provides a robust, straightforward approach for many practical scenarios.

In the context of spreadsheets, performing interpolation manually can be tedious and prone to human error, especially across extensive datasets. Fortunately, Google Sheets offers powerful built-in functions that, when combined creatively, automate this process. While complex methods might require scripting languages like Apps Script, simple linear interpolation can often be achieved using standard spreadsheet formulas like IFERROR function, along with mathematical functions such as AVERAGE function or `COUNTBLANK`. This detailed guide will walk you through a practical, step-by-step methodology to calculate and fill these missing observations precisely and efficiently, ensuring your data remains continuous and ready for advanced analysis.

Why Linear Interpolation is Necessary for Data Integrity

Missing data points, often referred to as 'Not Available' (NA) or blank cells, are endemic to real-world data collection, whether due to sensor malfunctions, data entry errors, or skipped observations in a time series. Ignoring these missing values by simply deleting the rows containing them can lead to a significant reduction in dataset size and potentially introduce selection bias, thereby undermining the validity of subsequent analyses. Conversely, replacing missing data with arbitrary constants, such as zero or the overall mean, tends to compress the variance and distort the underlying distribution of the data, failing to capture the natural progression or trend of the phenomenon being measured.

Linear Interpolation offers a superior solution, particularly when the data is expected to change at a relatively constant rate between known points. This method assumes that the relationship between two adjacent known data points is linear, allowing us to estimate the intermediate values by drawing a straight line between them. This approach minimizes the disturbance to the data's inherent trend, making the resulting dataset more representative of the true process. This technique is especially useful in scenarios involving financial metrics, environmental monitoring readings, or sequential experimental results where continuity is paramount for accurate modeling and forecasting.

While more sophisticated methods exist, such as polynomial or spline interpolation, the linear approach strikes an optimal balance between mathematical simplicity and practical effectiveness

within the constrained environment of a standard spreadsheet application like [Google Sheets](#). The subsequent steps will demonstrate how to leverage its functionality to precisely calculate the 'step value' necessary for consistent, linear increments across the gaps.

Setting Up Your Dataset with Missing Values

The foundation of successful interpolation lies in properly structuring your data and identifying the exact locations and boundaries of the missing entries. For this tutorial, we will focus on a single column or series where data is sequential. Before proceeding, ensure that the known start and end points of the gap are clearly identified, as these anchor points define the range over which the interpolation will occur. The process requires at least one known value preceding the gap and one known value succeeding the gap.

To demonstrate this methodology, we begin by creating a simple dataset within a [Google Sheets](#) workbook that includes several sequential values but also deliberately incorporates a block of blank cells. This represents a typical scenario encountered in real-world data handling. The initial dataset setup is crucial for establishing the variables we will use in our calculation formulas.

As shown in the image below, our data resides in column A. Notice that cells A9 through A12 are currently empty, representing the missing observations we intend to estimate. We have established clear anchor points: the starting known value (A8) and the ending known value (A13).

	A	B	C	D	E
1	Data				
2		3			
3		6			
4		8			
5		10			
6		14			
7		17			
8		20			
9					
10					
11					
12					
13		35			
14		39			
15		44			
16		49			
17		52			
18		57			
19		61			
20		65			
21		70			
22					

Understanding Linear Interpolation Methodology

Linear Interpolation works by determining the constant rate of change, or 'step value,' required to move from the starting known value (Start) to the ending known value (End) across the series of missing observations. This step value acts as the increment added to each preceding value to sequentially fill the blanks, ensuring a smooth, linear transition across the gap.

The mathematical principle behind calculating this step value is simple division: we determine the total difference between the End and Start values, and then distribute that difference equally across the total number of intervals in the gap. Importantly, if there are 'N' missing observations, there will be 'N + 1' intervals or steps required to connect the start and end points. This relationship is formalized by the following equation, which forms the basis of our spreadsheet calculation:

$$\text{Step} = (\text{End Value} - \text{Start Value}) / (\text{Number of Missing Observations} + 1)$$

Applying this formula to our specific example dataset provides a tangible illustration. The known start value is 20 (Cell A8), and the known end value is 35 (Cell A13). The number of missing

observations is four (Cells A9, A10, A11, A12). Therefore, the calculation proceeds as follows, resulting in a constant step size of 3:

$$\text{Step} = (35 - 20) / (4 + 1) = 15 / 5 = 3$$

Implementing the Step Value Formula in Google Sheets

While calculating the step value manually is straightforward for small gaps, using a dynamic formula ensures scalability and accuracy, especially if the data changes or the number of missing points varies. We utilize the [Google Sheets](#) `COUNTBLANK` function, which is designed specifically for counting empty cells within a specified range, making it ideal for determining the 'Number of Missing Observations' dynamically.

To automate the step value calculation, we input the following comprehensive formula into a helper cell, such as **D1**. This formula directly translates the methodology derived in the previous section into a usable spreadsheet command, referencing the specific cell locations for the start, end, and missing range:

$$=(A13-A8)/(COUNTBLANK(A9:A12)+1)$$

This formula first calculates the difference between A13 (End) and A8 (Start). It then divides this difference by the count of blank cells in the range A9:A12 (which is 4), plus one (to account for the total number of steps). The resulting value in cell D1 is the precise step size needed for linear interpolation. Utilizing this method ensures that the calculation is robust; should the boundaries (A8 or A13) or the number of missing observations change, the step value in D1 automatically updates, reducing the need for manual recalculations.

The visual representation below confirms the successful implementation of the formula, showing the calculated step value displayed in cell D1 based on our sample data. This value is the critical coefficient that will be used in the next phase to systematically fill every missing cell in the series.

	A	B	C	D
D1				$= (A13 - A8) / (\text{COUNTBLANK}(A9:A12) + 1)$
1	Data			3
2		3		
3		6		
4		8		
5		10		
6		14		
7		17		
8		20		
9				
10				
11				
12				
13		35		
14		39		
15		44		
16		49		
17		52		
18		57		
19		61		
20		65		
21		70		
22				

Executing the Interpolation Process

With the constant step value now calculated and stored (in D1), the final stage involves integrating this increment into the missing data cells. The goal is to generate a new value in each blank cell by simply adding the step value to the value immediately preceding it. This recursive process ensures linearity and consistency throughout the interpolated segment.

We begin with the first missing cell, which is A9. Since the value in A9 must be the starting value (A8) plus one step, we enter the following formula into cell A9:

`=A8 + D1`

The formula must reference cell D1 using absolute cell referencing (`'D1'`) if you plan to drag the formula down, ensuring the step value remains constant. However, for simplicity and manual entry in this tutorial, we will use relative referencing initially and modify the approach for subsequent

cells. Since A8 contains 20 and D1 contains 3, cell A9 should now display 23. The visual evidence of this initial calculation is provided below:

	A	B	C
1	Data		
2	3		
3	6		
4	8		
5	10		
6	14		
7	17		
8	20		
9	23		
10			
11			
12			
13	35		
14	39		
15	44		
16	49		
17	52		
18	57		
19	61		
20	65		
21	70		
22			
23			

We repeat this process for the subsequent missing cells, always referencing the cell directly above and adding the fixed step value. For cell A10, the formula becomes $=A9 + D1$, adding 3 to 23 to yield 26. For A11, it is $=A10 + D1$, resulting in 29, and finally, for A12, the formula is $=A11 + D1$, resulting in 32. This cascade of additions systematically fills the gap. When cell A12 (32) is added to the step value (3), it equals the known end value (A13, which is 35), confirming the accuracy of the interpolation calculation and the smooth continuity of the series.

A9:A12		fx	=A8+3
	A	B	C
1	Data		
2		3	
3		6	
4		8	
5		10	
6		14	
7		17	
8		20	
9		23	
10		26	
11		29	
12		32	
13		35	
14		39	
15		44	
16		49	
17		52	
18		57	
19		61	
20		65	
21		70	
22			
23			

Visual Confirmation of the Interpolated Data

Once the missing values have been filled using the calculated step size, the next crucial phase in Data Analysis is validation. While the mathematical check ($A_{12} + \text{Step} = A_{13}$) confirms computational accuracy, a visual inspection using a chart provides an intuitive understanding of how well the interpolated points align with the overall dataset trend. A successful linear interpolation should result in the newly added points falling perfectly along the straight line defined by the surrounding known data points, maintaining the expected trajectory of the series.

To perform this visualization in Google Sheets, follow these steps to generate a Line Chart:

Highlight the entire range of your data, including the headers and all values (in this case, cells **A2:A21**, assuming A1 is a header).

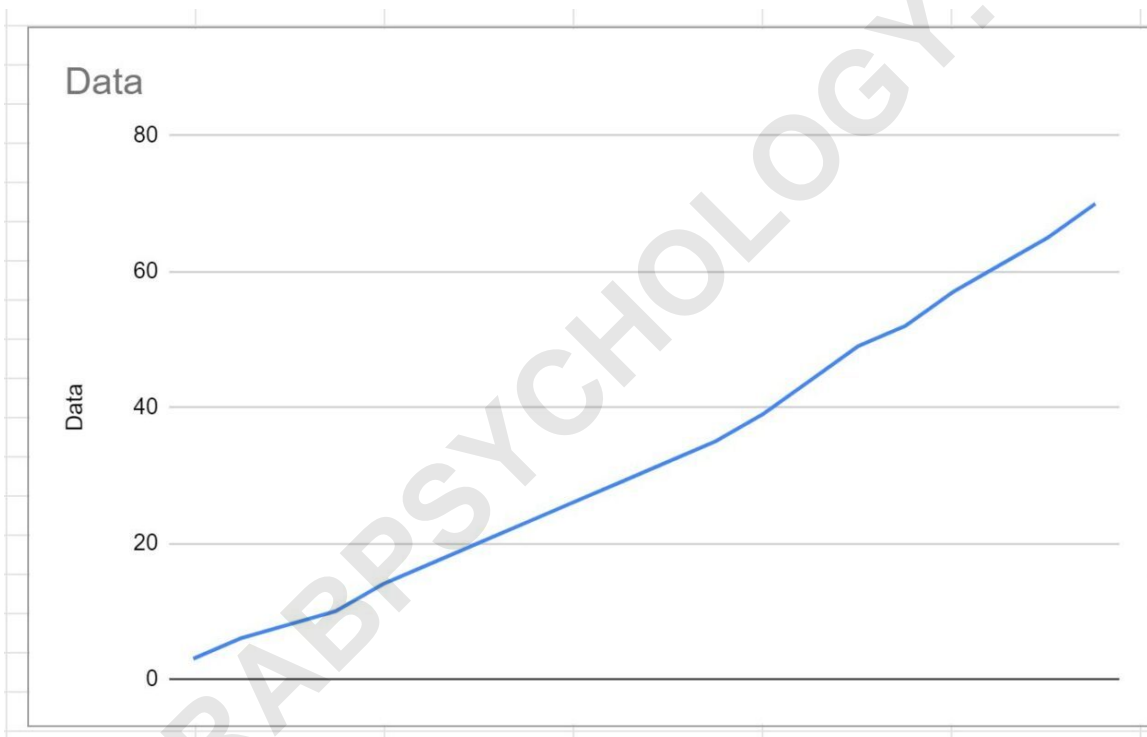
Navigate to the **Insert** tab in the Sheets menu.

Click on **Chart** to open the Chart editor panel on the right side of the screen.

In the **Chart editor**, ensure the selected **Chart type** is set to **Line chart**.

The resulting line chart immediately provides a graphical assessment of the interpolation quality. If the original data points already demonstrated a reasonably linear trend, the interpolated section should appear as a perfectly straight segment connecting the two known anchor points. This confirms that the estimated values seamlessly bridge the data gap without introducing artificial volatility or sharp breaks.

The visual output, as depicted below, clearly demonstrates that the interpolated values (23, 26, 29, 32) fit cohesively within the established trend of the dataset. This graphical verification is the final step in confirming that the chosen linear interpolation method was appropriate and successfully executed, yielding a complete and continuous data series ready for further processing or reporting.



Advanced Considerations and Alternative Techniques

While linear interpolation is effective and computationally cheap, it is important to recognize its limitations and when alternative approaches might be necessary. Linear interpolation assumes a constant rate of change; if the underlying data trend is non-linear (e.g., exponential growth, decay, or cyclical patterns), using this method will introduce systematic error, potentially smoothing out critical inflection points. In such cases, more complex estimation techniques are required.

For datasets exhibiting curved or cyclical behavior, specialized techniques such as polynomial

interpolation or cubic spline interpolation may offer better accuracy, though these are typically implemented using statistical software (like R or Python) rather than standard Google Sheets formulas. Furthermore, for highly volatile time series data, methods based on sophisticated statistical models, such as ARIMA models or Kalman filtering, are often employed to account for seasonality and autocorrelation.

Another commonly encountered approach in spreadsheet applications is using the AVERAGE function or median imputation. While simpler to execute, these statistical imputation methods do not respect the temporal sequence or trend of the data points; they merely replace the missing value with a static measure of central tendency. If the goal is to maintain the integrity of a sequential series, as is the case in time series Data Analysis, linear interpolation is vastly superior to simple average imputation. The use of conditional functions like IFERROR function can also be integrated to create complex formulas that only apply the interpolation logic to cells that are confirmed blank, adding another layer of robustness to large-scale data cleaning operations.

Conclusion and Further Resources

Mastering the technique of linear interpolation in Google Sheets provides an essential skill for any serious data handler. By systematically calculating the required step value using functions like `COUNTBLANK` and then recursively applying that increment across the missing range, you can transform discontinuous data into a smooth, analytical series. This methodology is simple, transparent, and mathematically grounded, ensuring that your data preparation phase is thorough and accurate.

For those seeking to explore more advanced techniques or functional programming within the Google ecosystem, investigating Google Apps Script allows for the creation of custom functions capable of handling more complex interpolation algorithms (e.g., non-linear estimation) directly within your spreadsheet environment, bypassing the inherent limitations of standard cell formulas. Regardless of complexity, the principle remains the same: ensuring data continuity is paramount for reliable statistical inference and accurate trend forecasting.