

Interpolate Missing Values in Google Sheets

Authored by
Mohammed loot

October 29, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Interpolate Missing Values in Google Sheets*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=5221>

In the challenging and dynamic landscape of [data analysis](#), practitioners frequently encounter sequential or time-series data plagued by gaps. The presence of [missing values](#) within a critical [dataset](#) can severely compromise the accuracy of subsequent calculations, visualizations, and predictive models. To overcome this common obstacle, advanced data cleaning techniques are necessary, chief among them being [interpolation](#), a powerful statistical method used to estimate unknown data points based on surrounding known observations.

This authoritative guide provides a practical, detailed walkthrough demonstrating how to efficiently execute [interpolation](#) directly within the accessible environment of [Google Sheets](#). We will concentrate specifically on [linear interpolation](#), a robust and simple technique that assumes a constant or steady rate of change between two established data points. By the conclusion of this tutorial, you will possess the requisite knowledge to reliably manage and fill similar data deficiencies in your own projects, ensuring your analytical inputs are complete and trustworthy.

The Rationale: Understanding and Addressing Data Gaps

The integrity of any analysis relies heavily on the completeness of the source data. Imperfections in data collection, whether resulting from sensor failure, human transcription errors, or intermittent recording periods, inevitably lead to the creation of [missing values](#). Ignoring these omissions, or attempting simplistic fixes like zero imputation, often introduces significant bias, leading to distorted statistical summaries, inaccurate trend forecasting, and fundamentally flawed visual reporting.

While various methodologies exist for handling data deficiencies--such as listwise deletion, which removes entire records, or mean/median imputation--[interpolation](#) stands out as a far more sophisticated and context-preserving solution. Rather than discarding information or relying on central tendency estimates, interpolation infers the likely value of the missing observation by using the existing trend and structure of the surrounding data. This method is particularly valuable when analyzing time series where the underlying process is known to be continuous.

For the practical implementation demonstrated here using [Google Sheets](#), we utilize the fundamental principles of [linear interpolation](#). This approach posits that the missing points should lie precisely on the straight line connecting the immediate preceding and succeeding known data points. Its simplicity and effectiveness make it the ideal starting point for data scientists and analysts requiring a rapid, accurate method to smooth out minor data inconsistencies within numerical sequences.

Step 1: Structuring the Dataset and Identifying Gaps

The foundation of successful data remediation begins with meticulous preparation of the [dataset](#) itself. To clearly illustrate the process, we will deliberately construct a sample series within a fresh [Google Sheets](#) workbook that contains several intentional [missing values](#). This setup is crucial

for demonstrating the precise mechanics of the interpolation method.

Begin by opening a new sheet and inputting your sequential numerical data into a single column. It is paramount to clearly define the boundary conditions--that is, the last known value before the gap begins (the start point) and the first known value after the gap ends (the end point). For linear interpolation to perform optimally, the data segment should ideally exhibit a numerical and sequential pattern, suggesting a somewhat consistent rate of change. Below, we present the visual representation of our example data structure, highlighting the segment requiring remediation:

	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					

In this specific scenario, we observe a sequential numerical series. The known starting value is 20, located in cell A8, and the known ending value is 35, located in cell A13. The critical task is to accurately and logically populate the four consecutive blank cells--A9, A10, A11, and A12--thereby ensuring a seamless numerical transition between the known endpoints of the series.

Step 2: Calculating the Consistent Interpolation Step Value

The core mechanism of [linear interpolation](#) requires establishing a single, consistent value--the

"step value"--which dictates the incremental change between every consecutive data point across the missing segment. This value ensures the linear assumption is met, distributing the total change observed between the start and end points uniformly across all intervening intervals.

The calculation relies on quantifying the total difference between the known boundaries and dividing this difference by the total number of intervals present. Note that the number of intervals is always one greater than the number of missing observations. The formula governing this essential calculation is formally defined as:

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

Applying this formula to our sample [dataset](#), where the **Start Value** (A8) is 20 and the **End Value** (A13) is 35, and we have 4 blank cells between them (A9 through A12), the manual calculation is straightforward: $\text{Step} = (35 - 20) / (4 + 1) = 15 / 5 = 3$. This result dictates that each missing cell must increase by exactly 3 from the preceding value.

To implement this robustly within [Google Sheets](#), it is advisable to use a dynamic formula that automatically references the cell locations and counts the empty cells. We will place this formula in cell **D1** (or any designated utility cell) to compute the step value automatically:

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

The integral use of the [COUNTBLANK](#) function ensures the formula remains flexible; it dynamically counts the empty cells (A9:A12), making the calculation adaptable even if the size of the gap changes. The addition of `+1` in the denominator correctly establishes the total number of segments across which the change must be distributed. The result of this formula, 3, is stored in D1, ready for the next step.

D1	fx	$= (A13 - A8) / (\text{COUNTBLANK}(A9:A12) + 1)$			
	A	B	C	D	
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					

Step 3: Systematically Populating the Missing Values

Once the precise **step value** has been calculated and stored, the subsequent stage involves the systematic deployment of formulas to fill the **missing values** sequentially. This process mandates starting with the very first blank cell immediately succeeding the known start point and iteratively adding the calculated step value to the previous cell's content.

For our specific dataset, the interpolation process begins in cell **A9**. Since the preceding known value is A8 (which holds 20) and our step value is located in D1 (which holds 3), the formula for A9 must reference these locations: $=A8 + D1$. Executing this formula yields the first interpolated value: $20 + 3 = 23$. It is often useful to ensure the reference to the step value cell (D1) is absolute if you plan to drag the formula, though in this simple example, sequential manual entry is straightforward.

We proceed by applying this additive logic to the remaining blank cells. For cell **A10**, the formula becomes $=A9 + D1$ ($23 + 3 = 26$). Continuing this linear progression, cell **A11** is calculated as $=A10 + D1$ ($26 + 3 = 29$), and finally, cell **A12** receives the formula $=A11 + D1$ ($29 + 3 = 32$). This

disciplined application ensures that the sequence maintains the assumed constant rate of change across the entire interpolated segment.

A9	∇	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			
11			
12			
13	35		
14	39		
15	44		
16	49		
17	52		
18	57		
19	61		
20	65		
21	70		
22			
23			

Upon completing these entries, you will observe the previously vacant cells being populated with the inferred values. This successfully bridges the gap, transforming the incomplete series into a complete and continuous **dataset**. The final, continuous sequence should now flawlessly connect the initial value of 20 to the concluding value of 35, validating the accuracy of the sequential step calculation.

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				

Step 4: Validating Interpolation Through Data Visualization

After successfully generating the interpolated values, the final and most vital stage is the visual validation of the updated [dataset](#). Visualization provides an immediate and intuitive check, confirming that the newly inserted values align logically and seamlessly with the overall trend established by the original data points. For sequential numerical data, a [line chart](#) is the most appropriate tool, as it excels at illustrating continuity and trend progression over ordered categories or time.

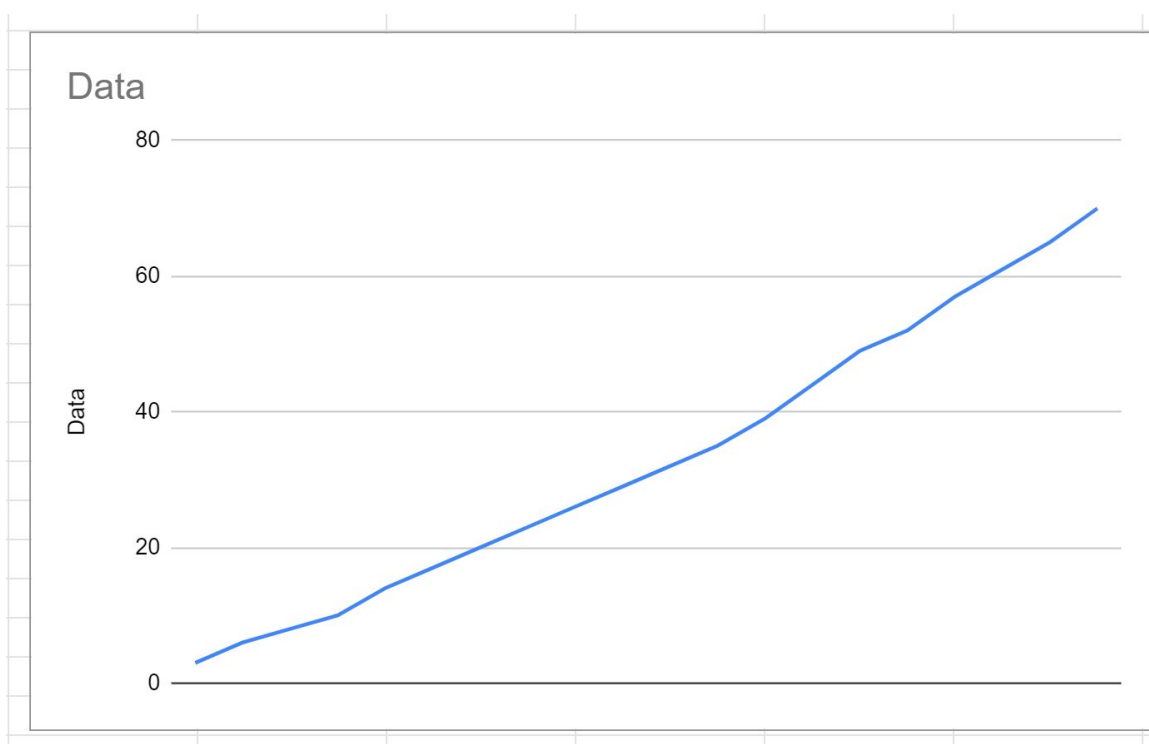
To generate this crucial visual proof within [Google Sheets](#), adhere to the following precise steps:

Data Selection: Carefully highlight the entire range of data encompassing both the original known values and the newly interpolated values. For our specific example, this range covers cells **A2:A21**.

Chart Insertion: Access the main menu toolbar, click on the **Insert** option, and then select **Chart** from the subsequent dropdown menu.

Configuration: The dedicated **Chart editor** panel will load automatically on the right-hand side of the interface. Within the "Chart type" configuration area, ensure that the **Line chart** option is selected. Google Sheets will instantly render the visual representation based on your now complete dataset.

The resulting [line chart](#) offers immediate confirmation of the interpolation's success. Critically examine the segment that corresponds to the previously missing data. The line should exhibit a smooth, constant slope, gracefully connecting the two endpoints (20 and 35) without any sharp deviations or breaks. This visual evidence is invaluable for verifying the integrity and logical coherence of your data imputation strategy.



As clearly depicted in the resulting visualization, the interpolated points integrate flawlessly, maintaining the overall upward trajectory of the series. This final validation step confirms that our application of [linear interpolation](#) successfully provided meaningful estimates for the [missing values](#), preparing the series for advanced analytical procedures.

Conclusion: Ensuring Data Integrity and Analytical Readiness

Mastering the effective management of [missing values](#) remains a cornerstone skill for any data professional. As demonstrated in this thorough guide, employing [linear interpolation](#) within [Google Sheets](#) furnishes a reliable and accessible methodology for seamlessly closing data gaps, especially when dealing with sequential data where proportional change is a reasonable

assumption. By precisely calculating the necessary step value and applying it systematically, we transform an incomplete series into a coherent, fully analyzable [dataset](#).

The disciplined process detailed here not only achieves data completeness but fundamentally enhances the overall integrity of the data, making it suitable for high-stakes analytical tasks, including trend forecasting, performance monitoring, and comprehensive visual reporting. It is imperative to remember the core assumption inherent in this method: linearity. For datasets exhibiting non-linear or highly volatile patterns, more complex statistical interpolation techniques--such as polynomial or cubic spline interpolation--would be required, though these typically necessitate specialized software beyond the native capabilities of basic spreadsheet functions.

We strongly advocate for practicing this foundational data cleaning technique across diverse personal and professional datasets. The ability to confidently and accurately address data deficiencies is an invaluable asset that significantly improves the reliability and depth of insights generated from your data analysis endeavors in today's increasingly data-centric world.