

Learning Linear Interpolation with Excel: A Step-by-Step Guide

Authored by
Mohammed loot

November 6, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Learning Linear Interpolation with Excel: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=11568>

Introduction to Linear Interpolation

The core concept of [interpolation](#) serves as a cornerstone in numerous disciplines, including advanced data analysis, engineering, and predictive statistics. At its heart, interpolation is a powerful mathematical technique used to estimate an unknown data point that lies within the range of a set of known, discrete data points. In practical applications, especially when dealing with real-world measurements or experimental results, analysts frequently encounter gaps in their data. Interpolation provides a reliable method to infer values where direct observation or measurement was not conducted.

Specifically, **linear interpolation** simplifies this estimation by assuming a fundamental, straight-line relationship between the two nearest known data points. This assumption allows for a quick and computationally efficient way to bridge data gaps, enabling analysts to derive reasonable predictions based on established trends within the dataset. While sophisticated alternatives exist, such as [spline interpolation](#) or polynomial methods, the linear approach remains highly favored due to its simplicity and its exceptional ease of integration into common spreadsheet environments, most notably [Microsoft Excel](#).

This comprehensive tutorial is designed to guide you through the precise mechanics of executing **linear interpolation** directly within **Excel**. We will construct a robust, dynamic formula that automates the typically tedious process of selecting the necessary bounding points from your larger dataset, ensuring accurate and efficient results for any intermediate value you wish to estimate.

The Mathematical Foundation of Linear Interpolation

Before translating this technique into a spreadsheet formula, it is essential to revisit the underlying mathematical model. Linear interpolation relies on the fundamental principles of coordinate geometry. Given two known coordinate pairs, conventionally labeled as (x_1, y_1) and (x_2, y_2) , our objective is to determine the unknown y -value (y) that corresponds to a specific intermediate x -value (x).

The core derivation is built upon the property of [similar triangles](#). If the relationship is linear, the rate of change, or [slope](#), must be constant across the line segment. Therefore, the slope calculated using the unknown point (x, y) and the lower known point (x_1, y_1) must be identical to the slope calculated using the two known endpoints (x_2, y_2) and (x_1, y_1) . This equality yields the standardized formula for **linear interpolation**:

$$y = y_1 + (x - x_1) \cdot \frac{y_2 - y_1}{x_2 - x_1}$$

Within this powerful equation, the fraction $(y_2 - y_1) / (x_2 - x_1)$ explicitly defines the constant **slope**

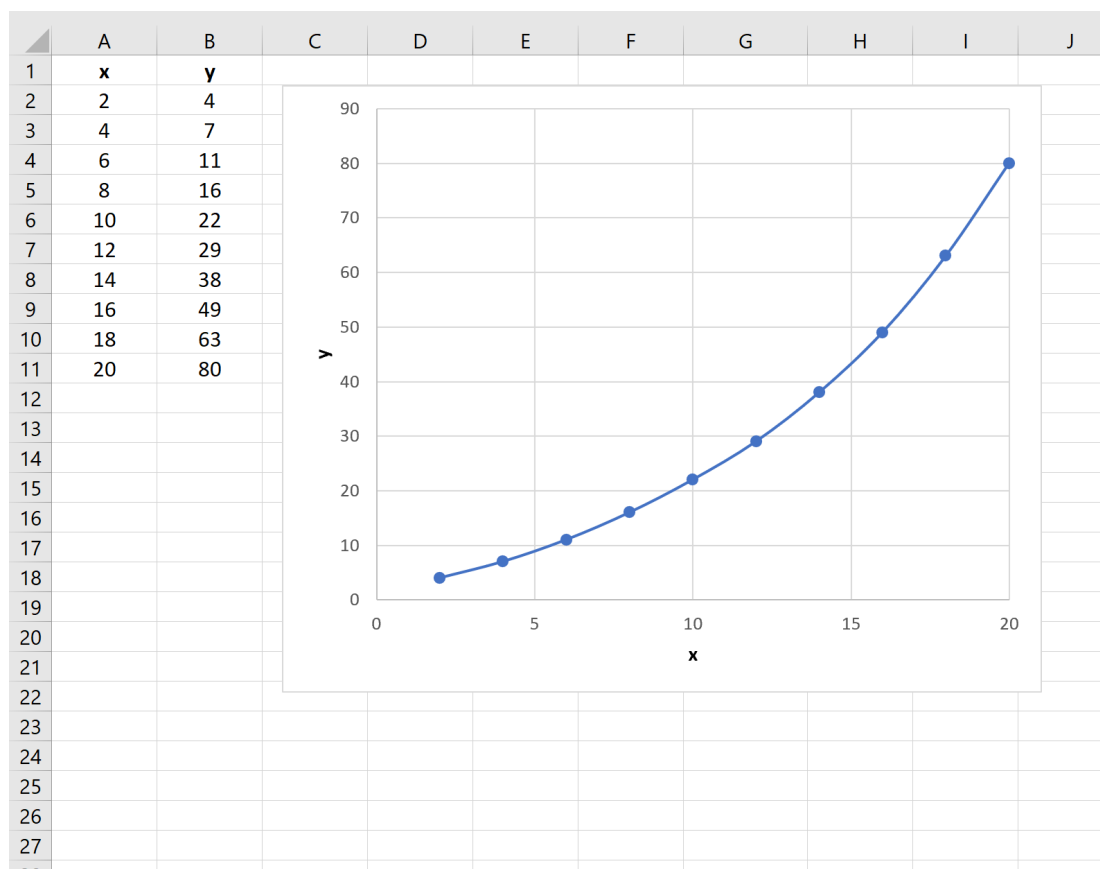
of the line segment connecting the two boundaries. This slope is then multiplied by the horizontal distance from the lower bound ($x - x_1$), and the starting y-value (y_1) is added to determine the accurately estimated height, y . The subsequent sections will detail how this precise mathematical logic can be seamlessly implemented into a singular, highly functional formula within **Excel**.

Preparing Your Dataset and Defining the Goal

To illustrate the application of this technique, we will use a common scenario involving a sequence of measurements that establish a clear relationship between an independent variable (X) and a dependent variable (Y). For demonstration purposes, assume the following raw data points have been entered into two organized columns within an **Excel** workbook:

	A	B	C	D	E	F
1	x	y				
2	2	4				
3	4	7				
4	6	11				
5	8	16				
6	10	22				
7	12	29				
8	14	38				
9	16	49				
10	18	63				
11	20	80				
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						

It is always beneficial to visualize the data before attempting interpolation. Plotting these points helps confirm the assumed linearity of the relationship between X and Y. As demonstrated in the chart below, the measured points exhibit a distinct upward progression, validating that **linear interpolation** is an appropriate and robust method for estimating values within this specific range:



Our immediate objective is to calculate the corresponding Y-value for a new, specific X-value: **13**. While our existing measurements provide Y-values for $X = 12$ ($Y = 32$) and $X = 14$ ($Y = 35$), we must accurately estimate the intermediate value at $X = 13$. To achieve maximum efficiency, we require an **Excel** formula capable of automatically identifying the correct bounding points--(x_1, y_1) = (12, 32) and (x_2, y_2) = (14, 35)--and then applying the interpolation logic without manual intervention.

Implementing the Dynamic Excel Formula

Although the underlying mathematical formula for **linear interpolation** is simple, translating this into a dynamic function within **Excel**--one that automatically finds the correct boundaries for any target X-value--requires the use of advanced range definition and lookup functions. The most reliable and robust solution employs a powerful combination of three essential functions: **FORECAST** (or the modern **FORECAST.LINEAR**), **OFFSET**, and **MATCH**.

The structure of this integrated formula is specifically engineered to perform three distinct yet interconnected operations necessary for dynamic interpolation:

The **MATCH** function first identifies the precise row position of the lower bounding X-value (x_1)

relative to the known X-range.

The **OFFSET** function then utilizes the position identified in Step 1 to dynamically define a two-cell range containing the two required known X-values (x1 and x2) and the corresponding Y-values (y1 and y2).

Finally, the **FORECAST** function receives the target X-value (x) and the dynamically defined known ranges (Known Ys and Known Xs) to calculate the resulting interpolated Y-value.

Assuming 'NewX' references the cell containing our target value (13), 'KnownY' is the range of measured Y-values, and 'KnownX' is the range of corresponding X-values, the complete dynamic formula for **linear interpolation** in **Excel** is constructed as follows:

```
=FORECAST(NewX,OFFSET(KnownY,MATCH(NewX,KnownX,1)-1,0,2),  
OFFSET(KnownX,MATCH(NewX,KnownX,1)-1,0,2))
```

This single-cell formula encapsulates the entire logic required for accurate linear estimation, irrespective of where the target X-value falls within the existing dataset.

Deconstructing the Advanced Formula Logic

Understanding the dynamic formula requires a detailed breakdown of how the nested functions interact, particularly the crucial role played by the **MATCH** and **OFFSET** combination:

MATCH(NewX, KnownX, 1): This is the anchor of the dynamic formula. The critical third argument, '1', instructs the **MATCH** function to find the position of the largest value in the 'KnownX' range that is less than or equal to 'NewX' (13). This precisely identifies the row number corresponding to the lower bounding point, x1.

MATCH(...)-1: The **OFFSET** function calculates its shift relative to the starting cell of the defined range (e.g., the first cell of KnownY). If our data starts in row 2, and **MATCH** returns 5 (for the 5th row in the range), we need to adjust this offset to ensure it correctly points to the starting position of the two required bounding points.

OFFSET(KnownY, , 0, 2): This section dynamically defines the two required Y-values (y1 and y2). It begins at the 'KnownY' range, shifts down by the calculated row offset, moves zero columns horizontally, and critically, specifies a height of 2 rows. This action returns the exact pair of Y-values needed for the interpolation calculation. The identical logic is applied to **OFFSET(KnownX, ...,)**, which extracts the corresponding pair of X-values (x1 and x2).

FORECAST(NewX, Known_Ys, Known_Xs): The final layer is the **FORECAST** function. This function is designed specifically to perform linear regression and [linear interpolation](#). By receiving the target X-value (13) and the dynamically supplied pairs of known data points, it executes the

mathematical formula derived from the similar triangles principle, computing the accurate resulting Y-value.

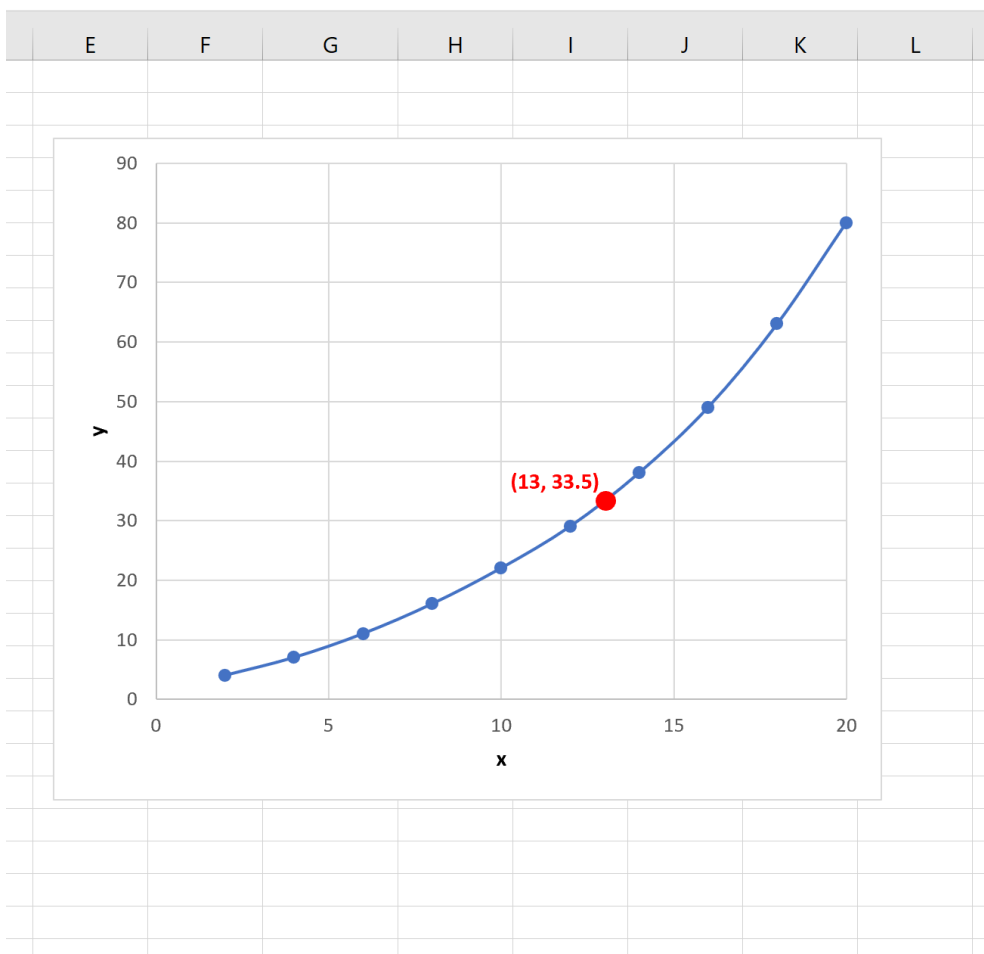
Executing the Calculation and Reviewing Results

When the dynamic formula is applied to our structured dataset, **Excel** efficiently locates the precise boundaries and calculates the intermediate estimate. In a typical sheet setup, if the X-values are located in column A (A2:A10), the Y-values in column B (B2:B10), and the new X-value (13) is placed in cell D2, the formula adjusts automatically to reference these specific cell ranges:

	A	B	C	D	E	F	G	H	I	J	K	L
1	x	y										
2	2	4										
3	4	7										
4	6	11										
5	8	16										
6	10	22										
7	12	29										
8	14	38										
9	16	49										
10	18	63										
11	20	80										
12												
13	=FORECAST (13,OFFSET (B2:B11,MATCH (13,A2:A11,1)-1,0,2), OFFSET (A2:A11,MATCH (13,A2:A11,1)-1,0,2))											
14												
15												
16												
17												
18												
19												
20												
21												
22												
23												
24												
25												
26												

Upon successful execution, **Excel** determines the estimated Y-value to be exactly **33.5**. This result is mathematically expected, as the target X-value (13) is precisely centered between the known bounding X-values (12 and 14). Consequently, the estimated Y-value is the arithmetic mean of the two known Y-values (32 and 35).

To definitively validate the accuracy of this estimation, the newly interpolated point (13, 33.5) can be integrated back into the original data plot. As clearly illustrated below, the interpolated point resides perfectly along the established linear path connecting the surrounding measured data points, providing visual confirmation of the successful and accurate application of the dynamic **linear interpolation** method:



Important Considerations and Limitations

While the combined **FORECAST**, **OFFSET**, and **MATCH** formula offers a powerful, automated solution for interpolation in **Excel**, maintaining accuracy and functionality requires strict adherence to several key data preparation rules and conceptual understandings:

Data Ordering is Crucial: For the **MATCH** function to operate correctly with the lookup type argument '1' (which finds the next smallest value), the known X-values (KnownX) must be rigorously sorted in **ascending order**. If the data remains unsorted, the formula will fail to identify the correct bounding points, resulting in a fundamentally inaccurate interpolation result.

Interpolation vs. Extrapolation: The dynamic technique demonstrated here is strictly intended for **interpolation**. This means the target X-value must fall strictly **within** the range defined by the minimum and maximum existing X-values in the source dataset. If the target X-value lies outside this range, the process is known as **extrapolation**. While the formula will still produce a numerical result during **extrapolation**, that result is based on extending an observed trend beyond the scope of the measured data and is often misleading or statistically unsound.

Function Compatibility and Updates: Users working with modern versions of **Excel**, such as Excel 2016 or Microsoft 365, are strongly encouraged to utilize the updated function **FORECAST.LINEAR** instead of the legacy **FORECAST** function. Although both functions are identical in their mathematical logic and required arguments, **FORECAST.LINEAR** offers improved clarity and better long-term compatibility within the Microsoft ecosystem.

Mastering the creation and implementation of this dynamic formula provides users with the ability to efficiently and accurately calculate intermediate values across extensive datasets, fully automating a process that would otherwise necessitate laborious manual calculations and lookups.

*You can find more **Excel** tutorials and advanced resources dedicated to [data analysis](#) and sophisticated spreadsheet operations on our site.*