

How to naive forecasting in Excel?

Authored by
stats writer

December 14, 2025

RECOMMENDED CITATION

stats writer (2025). *How to naive forecasting in Excel?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=107441>

Introduction to Naive Forecasting

Naive forecasting represents one of the most fundamental and straightforward approaches utilized in time series analysis. This technique operates on the principle that the future value of a variable will simply be equal to its most recently observed past value. While incredibly simplistic, this method serves as a crucial benchmark against which more complex, sophisticated forecasting models are often measured. It requires minimal computational effort and is exceptionally easy to implement, making it a powerful tool for initial assessments or when dealing with datasets that exhibit random walk characteristics.

The core concept of naive forecasting dictates that if sales in March were 100 units, the forecast for April will also be 100 units. This approach inherently assumes that the system generating the data is stable and lacks predictable structures such as underlying growth patterns or regular cyclical movements. Because of this inherent simplicity, it is often dismissed in favor of methods like exponential smoothing or ARIMA models. However, its effectiveness in many real-world scenarios, particularly in finance and inventory management where volatility is high, often surprises analysts due to its robustness against complex noise.

Implementing this method within a spreadsheet environment like Microsoft Excel is efficient and accessible, even for users without deep statistical knowledge. The process involves identifying the latest data point and using a simple cell reference formula to project this value forward across the future time periods. While powerful for establishing a baseline, it is essential to remember its major limitation: it completely ignores structural changes such as underlying trend or seasonality. If your data exhibits strong cyclical patterns or clear growth, the naive forecast will likely prove inaccurate beyond the immediate next period.

Understanding the Naive Forecasting Mechanism

The conceptual clarity behind the naive forecast is its greatest strength. By definition, a naive projection (often denoted as F_{t+1}) for the upcoming period ($t+1$) is set precisely equal to the actual observed value (Y_t) from the previous period (t). This model can be mathematically expressed as: **$F_{t+1} = Y_t$** . This instantaneous carry-over of the last known value ensures that the forecast is always reactive to the most recent data, providing a baseline prediction that reflects current performance without averaging or weighting past observations.

To illustrate this mechanism practically, consider a scenario involving the quarterly sales figures for a specific consumer product. If January sales were 200 units, February sales were 215 units, and March sales reached 250 units, the naive projection relies solely on the March figure. The accompanying image depicts this initial setup and how the actual data is arranged prior to the forecasting calculation, showcasing the three historical data points available to us.

Month	Actual Sales
January	34
February	37
March	44

Following the naive rule, the forecast for sales in April, the next designated time period, must mirror the sales volume achieved in March. Therefore, the April forecast is 44 units. This direct inheritance of the prior month's value is what distinguishes the naive method. This simplicity is often a benefit, as studies in forecasting accuracy have demonstrated that for certain types of data--especially financial data exhibiting random walk behavior--the naive forecast performs remarkably well, sometimes outperforming overly complex models that attempt to fit noise.

Month	Actual Sales	Forecasted Sales
January	34	
February	37	
March	44	
April	?	44

This step-by-step guide will now transition into the detailed process of performing and evaluating naive forecasting specifically within the structured environment of Microsoft Excel, ensuring the generated forecasts are robust and measurable through subsequent error analysis.

Step 1: Entering Historical Data in Excel

The initial requirement for any forecasting exercise is the accurate compilation of historical, time-sequenced data. For this tutorial, we will utilize a dataset representing monthly sales figures over a 12-month period for a fictional corporation. Organizing the data correctly in Excel is paramount, typically involving two dedicated columns: one for the time index (e.g., Month or Date) and another for the corresponding observed value (e.g., Sales Units or Revenue).

Begin by opening a new worksheet in Excel. Label the first column 'Month' or 'Period' and the second column 'Actual Sales' or 'Observed Value'. Input the 12 data points sequentially, ensuring that each observation is correctly associated with its respective time marker. This precise

arrangement facilitates the subsequent application of formulas and ensures that the temporal relationship between observations is maintained throughout the analysis.

This step requires meticulous data entry. Ensure that the data is sorted chronologically, as the time sequence is intrinsic to the validity of any time series forecasting model. A single misplaced data point can significantly skew the resulting error calculations and subsequent interpretation of the model's performance.

As depicted in the image below, our dataset spans from January to December, providing a full year of historical performance. It is crucial to verify the integrity and accuracy of this historical data, as any errors introduced at this stage will cascade directly into the subsequent forecasting and accuracy calculation steps.

	A	B	C	D	E
1	Month	Actual Sales			
2	January	34			
3	February	37			
4	March	44			
5	April	47			
6	May	48			
7	June	48			
8	July	46			
9	August	43			
10	September	32			
11	October	27			
12	November	26			
13	December	24			
14					
15					
16					
17					
18					

Step 2: Generating Naive Forecasts

Once the historical data is meticulously entered, the next step involves generating the naive forecasts. This process leverages the core principle of the model: the forecast for the current period is simply the actual observation from the preceding period. In Excel, this is achieved through the use of direct cell referencing, a simple yet highly effective technique that avoids complex mathematical functions.

Create a new column adjacent to the 'Actual Sales' column and label it 'Naive Forecast'. Since

forecasting requires prior data, the first available forecast will be for the second period (February), based on the actual sales from the first period (January). Thus, the cell corresponding to the February forecast will contain a formula referencing the January actual sales cell. For instance, if the actual January sales value is located in cell B2, the forecast for February, located in cell C3, will contain the simple formula:

=B2

This formula can then be efficiently copied down the entire 'Naive Forecast' column. By dragging the fill handle, Excel automatically adjusts the cell reference for each subsequent period (e.g., the March forecast references February sales, the April forecast references March sales, and so forth). This action generates a completed forecast column where the forecast value is lagged by one period relative to the actual observation it is attempting to predict. This visual setup, including the lagged data, is critical for the subsequent accuracy calculations.

The resultant table, displaying both the actual historical sales and the corresponding naive forecasts, should appear as follows. Note that the final forecast, which projects into the first future period (Month 13), uses the last known actual value (December sales). This single calculation completes the projection phase of the time series analysis.

	A	B	C	D	E
1	Month	Actual Sales	Forecasted Sales	Formula used	
2	January	34			
3	February	37	34	=B2	
4	March	44	37	=B3	
5	April	47	44	=B4	
6	May	48	47	=B5	
7	June	48	48	=B6	
8	July	46	48	=B7	
9	August	43	46	=B8	
10	September	32	43	=B9	
11	October	27	32	=B10	
12	November	26	27	=B11	
13	December	24	26	=B12	
14					
15					
16					
17					
18					

Step 3: Evaluating Forecast Accuracy Metrics

The true utility of any forecasting model, including the naive approach, is determined by its accuracy when compared against actual outcomes. Consequently, the final and most critical phase of this process involves calculating robust error metrics. These metrics quantify the discrepancy between the projected values and the subsequent actual values, providing objective data on the model's performance. Without these measurements, it is impossible to gauge whether the naive method is a suitable forecasting strategy for the specific data under analysis, especially relative to alternative, more complex methods.

Before calculating the summary statistics, two intermediate columns must be added to the Excel sheet to facilitate error measurement: one for the 'Error' (or Residual) and one for the 'Absolute Error'. The raw error is calculated as the Actual Value minus the Forecasted Value ($E_t = Y_t - F_t$). The absolute error, crucial for ensuring positive aggregation of errors regardless of direction, is the absolute value of the raw error ($|E_t|$). These foundational columns are necessary inputs for deriving more complex measures.

Two primary metrics are commonly employed in business and quantitative analysis to assess forecasting effectiveness, each offering a slightly different perspective on the magnitude of the errors. These metrics are particularly valuable because they are easy to interpret and widely recognized across various industries. They are the **Mean Absolute Percentage Error (MAPE)** and the **Mean Absolute Deviation (MAD)**. We will now detail the calculation and interpretation of each metric in the context of our Excel model.

It is important to emphasize that accuracy assessment is inherently comparative. The goal is not simply to calculate a MAPE or MAD value, but to use these statistics to benchmark the naive model against more advanced forecasting techniques (e.g., Holt-Winters or ARIMA). A model is deemed "useful" only if its error metrics are lower than, or competitive with, those derived from alternative models, especially when considering the trade-off between complexity and accuracy.

Calculating Mean Absolute Percentage Error (MAPE)

The **Mean Absolute Percentage Error (MAPE)** is a highly popular measure because it expresses the forecast error as a percentage, making it easy to interpret and understand in a managerial context. A MAPE of 10% signifies that, on average, the forecast deviated from the actual outcome by 10%. It is calculated by taking the average of the absolute percentage errors for each forecast period. The percentage error for a single period ($\%PE_t$) is calculated as the absolute difference between the actual and forecasted values, divided by the actual value, multiplied by 100.

To implement this in Excel, you must first create a column titled 'Absolute Percentage Error'. For each relevant row, the formula calculates the absolute percentage difference. For example, if

Actual Sales are in column B and Naive Forecasts are in column C, the formula in the percentage error column (say, column E) would be:

$$=ABS((B3-C3)/B3)$$

It is essential to note the constraints of MAPE: it should ideally only be used when the actual values (the denominator) are strictly non-zero, as division by zero renders the metric undefined. Furthermore, MAPE tends to place a heavier penalty on negative errors (overestimates) than on positive errors (underestimates).

After calculating the individual absolute percentage errors for all periods where a forecast is available, the final step involves averaging this column using Excel's `AVERAGE()` function. This average yields the MAPE, which is conventionally presented as a percentage. The image below illustrates the necessary intermediate calculations within the spreadsheet, leading to the final MAPE value. This comprehensive breakdown confirms the high level of detail required for robust analysis.

	A	B	C	D	E
1	Month	Actual Sales	Forecasted Sales	Absolute Percent Error	Formula used
2	January	34			
3	February	37	34	8.11	=ABS(B3-C3)/B3*100
4	March	44	37	15.91	=ABS(B4-C4)/B4*100
5	April	47	44	6.38	=ABS(B5-C5)/B5*100
6	May	48	47	2.08	=ABS(B6-C6)/B6*100
7	June	48	48	0.00	=ABS(B7-C7)/B7*100
8	July	46	48	4.35	=ABS(B8-C8)/B8*100
9	August	43	46	6.98	=ABS(B9-C9)/B9*100
10	September	32	43	34.38	=ABS(B10-C10)/B10*100
11	October	27	32	18.52	=ABS(B11-C11)/B11*100
12	November	26	27	3.85	=ABS(B12-C12)/B12*100
13	December	24	26	8.33	=ABS(B13-C13)/B13*100
14			MAPE	9.90	=AVERAGE(D3:D13)
15					
16					
17					
18					

Based on our calculated figures, the overall Mean Absolute Percentage Error for the naive model applied to this 12-month dataset is found to be **9.9%**. This result provides a quantifiable, scale-independent measure of the model's accuracy, suggesting that, on average, the naive forecast missed the actual sales target by just under ten percent.

Calculating Mean Absolute Deviation (MAD)

The **Mean Absolute Deviation (MAD)**, also known as the Average Absolute Error, is another vital metric for assessing forecast accuracy. Unlike MAPE, MAD provides the error in the original units of measurement (e.g., units sold, dollars, or tons), making it easier for operational managers to understand the physical magnitude of the forecasting mistakes. It is calculated by taking the average of the absolute errors (the absolute difference between the actual value and the forecasted value) across all time periods.

To compute MAD, we rely on the 'Absolute Error' column created earlier (Actual Sales minus Naive Forecast, all in absolute terms). If Actual Sales are in column B and Forecasts are in column C, the Absolute Error column (say, column D) uses the formula:

`=ABS(B3-C3)`

for each period where both an actual value and a forecast exist. This calculation isolates the size of the error, regardless of whether the forecast was an overestimate or an underestimate, effectively treating all errors equally in magnitude.

The final MAD value is simply the arithmetic average of all values contained within the 'Absolute Error' column. In Excel, this can be quickly calculated using the

`AVERAGE()`

function applied to the entire range of absolute errors. Because MAD is measured in the same unit as the data, it is excellent for comparing different models applied to the same dataset, but it is less useful for comparing accuracy across different product lines or datasets with vastly different scales or magnitudes.

	A	B	C	D	E
1	Month	Actual Sales	Forecasted Sales	Absolute Deviation	Formula Used
2	January	34			
3	February	37	34	3.00	=ABS(B3-C3)
4	March	44	37	7.00	=ABS(B4-C4)
5	April	47	44	3.00	=ABS(B5-C5)
6	May	48	47	1.00	=ABS(B6-C6)
7	June	48	48	0.00	=ABS(B7-C7)
8	July	46	48	2.00	=ABS(B8-C8)
9	August	43	46	3.00	=ABS(B9-C9)
10	September	32	43	11.00	=ABS(B10-C10)
11	October	27	32	5.00	=ABS(B11-C11)
12	November	26	27	1.00	=ABS(B12-C12)
13	December	24	26	2.00	=ABS(B13-C13)
14			Mean Abs. Deviation	3.45	=AVERAGE(D3:D13)
15					
16					
17					
18					
19					

As shown in the calculated results, the **Mean Absolute Deviation** for this naive forecast model is determined to be **3.45**. This means that, on average, our naive prediction was off by 3.45 units of sales. This absolute measure is highly useful when managing inventory or production planning, where the physical quantity of the error directly impacts holding or stockout costs.

Conclusion: Interpreting and Contextualizing the Results

Successfully implementing naive forecasting in Excel is just the first step; the real value lies in the interpretation and contextualization of the resulting accuracy metrics. The calculated MAPE of 9.9% and the MAD of 3.45 units provide a clear quantitative baseline for the data's predictability. If these error values are significantly high, it is a strong indicator that the time series contains pronounced structural elements, such as underlying growth trend or seasonality, which the naive method inherently fails to capture.

For time series data exhibiting clear patterns, the naive forecast typically serves as a lower bound for acceptable performance. If a more complex model cannot outperform the naive benchmark, it suggests that the added complexity is not justified by the data, or that the model itself is incorrectly specified. For example, if a sophisticated ARIMA model yields a MAPE of 15%, the naive model (at 9.9%) is demonstrably superior for this dataset.

If, however, these accuracy metrics are acceptably low, or if they are comparable to results obtained from more sophisticated models (like Simple Exponential Smoothing), then the naive

approach proves to be the superior choice due to its minimal complexity and computational demands. The principle of parsimony--preferring the simplest model that adequately explains the data--is highly relevant here. The naive forecast is particularly robust in environments where data volatility is high and the signal-to-noise ratio is low, as attempting to fit a complex structure to random data often leads to dangerous overfitting.

Ultimately, the decision of whether this naive forecast is "useful" rests on comparing its performance against other forecasting models. Analysts should proceed by calculating the MAPE and MAD for alternative models and selecting the one that achieves the lowest error metrics while balancing model complexity. This rigorous comparative evaluation ensures that the chosen forecasting method provides the most reliable projections for future planning and decision-making.

ARABPSYCHOLOGY.COM