Learning to Calculate Moving Averages Using SAS

Authored by

Mohammed looti

October 27, 2025

Understanding Moving Averages in Data Analysis

In the fields of statistics and time series data analysis, a **moving average** (MA) serves as an essential technical indicator designed to smooth out fluctuating data. This technique works by creating a constantly recalculated average value based on a specified number of preceding data points. By averaging the previous *n* values from a dataset, the moving average effectively minimizes the impact of short-term noise and random volatility, thereby making underlying trends and cyclical patterns much easier to identify and interpret.

The concept of the moving average is versatile, finding widespread application across numerous professional disciplines. It is critical in financial markets for trend analysis, allowing traders and analysts to gauge market direction. Similarly, it is fundamental in signal processing to filter high-frequency interference, and in industrial quality control to ensure manufacturing processes remain stable over time. While various forms exist--most notably the Simple Moving Average (SMA) and the Exponential Moving Average (EMA)--our focus here is on efficiently calculating the Simple Moving Average using the powerful capabilities of SAS software.

Calculating a moving average in the SAS environment is typically streamlined through specialized, built-in procedures. The most straightforward approach involves utilizing the PROC EXPAND statement. This procedure is specifically engineered for transforming time series data, allowing users to apply complex transformations, including the calculation of moving averages, with remarkable efficiency and minimal programming effort. Understanding the structure and syntax of PROC EXPAND is key to mastering this fundamental analytical operation in SAS.

Leveraging PROC EXPAND for Time Series Transformation

The PROC EXPAND procedure is a robust component of the SAS/ETS (Econometric and Time Series) product suite. Its core functionality revolves around manipulating time series data by processes such as interpolation, extrapolation, aggregation, and, crucially for our application, transformation. When the objective is to compute moving averages, PROC EXPAND provides a dedicated, high-performance transformation option that simplifies the task compared to manual DATA step logic.

To effectively utilize PROC EXPAND for calculating moving averages, two primary statements are required. First, the ID statement is mandatory, as it specifies the time or sequence variable that dictates the order of observations. Second, the powerful CONVERT statement defines the transformation itself. Within the CONVERT statement, you specify the input variable, the desired name for the output (transformed) variable, and the specific transformation method using the TRANSOUT option.

For the calculation of moving averages, the specific transformation applied via the TRANSOUT

option is MOVAVE. The MOVAVE transformation requires a single numerical argument: *n*, which represents the number of periods (the window size) over which the average should be computed. This function automatically handles the complex aggregation logic and the sliding window calculation, making it the preferred method for rapidly generating smoothed time series in SAS.

Step-by-Step: Preparing Data and Calculating a 3-Period MA

To clearly demonstrate the practical application of PROC EXPAND, we will first establish a working sample dataset. This initial step is vital for ensuring that the data structure is correctly formatted for subsequent time series analysis within SAS. Our sample data will include a variable representing the sequential time periods and a numerical variable containing the observations we wish to smooth.

The following DATA step is used to create our input data table, named original_data. This table contains two critical variables: time, which serves as our sequential time identifier, and values, which holds the raw numerical observations. We employ the DATALINES statement to embed the data directly into the program for convenience. After the data creation is complete, a standard PROC PRINT statement is included to display and verify the structure and contents of the newly generated data set.

```
/*create dataset*/
data original_data;
input time values;
datalines;
1 7
2 12
3 14
4 12
5 16
6 18
7 11
8 10
9 14
10 17
;
run;

/*view dataset*/
proc print data=original_data;
```

Obs	time	values
1	1	7
2	2	12
3	3	14
4	4	12
5	5	16
6	6	18
7	7	11
8	8	10
9	9	14
10	10	17

Once the input data is successfully prepared and verified, the next step is to calculate the 3-period moving average for the values column. A 3-period moving average means that, for any given observation, the average is computed using the current value and the two data points immediately preceding it. This calculation will result in the creation of a new variable that represents the smoothed time series.

The following PROC EXPAND code executes this calculation. We specify the input data (original_data) and define the output data (out_data). Crucially, we use METHOD=NONE because we are performing a direct transformation (moving average) and not interpolation. The ID statement designates time, and the CONVERT statement instructs the procedure to generate a new column, values_ma3, by applying the MOVAVE transformation over a window of 3 periods to the source values variable.

```
/*calculate 3-period moving average for values*/
proc expand data=original_data out=out_data method=none;
id time;
convert values = values_ma3 / transout=(movave 3);
run;

/*view results*/
proc print data=out_data;
```

Obs	time	values_ma3	values
1	1	7.0000	7
2	2	9.5000	12
3	3	11.0000	14
4	4	12.6667	12
5	5	14.0000	16
6	6	15.3333	18
7	7	15.0000	11
8	8	13.0000	10
9	9	11.6667	14
10	10	13.6667	17

Interpreting the Sliding Window Calculation

After executing the PROC EXPAND code, the output dataset, out_data, now features a new series named values_ma3. This column holds the 3-period moving average for the original data. A key characteristic of the Simple Moving Average calculation is the initialization phase: the first two entries in values_ma3 are assigned missing values. This occurs because the procedure requires three preceding data points to compute the average, and these points are not available until the third observation.

The fundamental principle of the moving average is the sliding window. As the window moves forward one period at a time, the average is recalculated. We can illustrate this by manually verifying the values generated by SAS. The first calculated average appears at time period 3, derived from the original values at periods 1, 2, and 3. This smoothing process continues sequentially throughout the dataset.

Moving Average calculation at time 3: (Value at T1 + Value at T2 + Value at T3) / 3 = (7 + 12 + 14) / 3 = **11.0000**

The subsequent calculation at time period 4 demonstrates the sliding window effect. The average shifts forward, dropping the value from period 1 and including the value from period 4. This process is repeated for every observation, providing a continuous, smoothed representation of the data trend.

Moving Average calculation at time 4: (Value at T2 + Value at T3 + Value at T4) / 3 = (12 + 14 + 12) / 3 = **12.6667**

Extending the Analysis: Adjusting the Moving Average Period

The efficiency of using PROC EXPAND lies in its flexibility. If the analytical requirement changes--for instance, if we needed to calculate a 4-period moving average to achieve a greater degree of smoothing--the modification required is minimal. We only need to adjust the numerical argument within the MOVAVE transformation from 3 to 4.

The choice of the period length (*n*) is a critical modeling decision. A larger period (e.g., 10 periods) results in a smoother line that is highly resistant to short-term volatility but introduces significant lag, meaning the average reacts slowly to genuine changes in trend. Conversely, a smaller period (e.g., 3 periods) produces an average that is more responsive to recent data but retains more noise, making it less effective for long-term trend identification.

The following code snippet demonstrates how to modify the CONVERT statement to calculate a 4-period moving average, creating a new variable named values_ma4.

```
/*calculate 4-period moving average for values*/
proc expand data=original_data out=out_data method=none;
id time;
convert values = values_ma4 / transout=(movave 4);
run;

/*view results*/
proc print data=out_data;
```

Obs	time	values_ma4	values
1	1	7.00	7
2	2	9.50	12
3	3	11.00	14
4	4	11.25	12
5	5	13.50	16
6	6	15.00	18
7	7	14.25	11
8	8	13.75	10
9	9	13.25	14
10	10	13.00	17

As expected, the resulting dataset now includes the values_ma4 column. Notice that, for a 4-period average, the first three values in the new column are missing, as four data points are required for the calculation to begin. This example clearly demonstrates the ease with which analysts can iterate on different smoothing parameters using PROC EXPAND.

Advanced Considerations and Real-World Applications in SAS

While PROC EXPAND is the most efficient method for calculating standard Simple Moving Averages, analysts working within the SAS ecosystem should be aware of alternative procedures for more complex scenarios. When dealing with irregular time intervals, requiring forecasting capabilities, or implementing weighted averages (like the EMA), procedures such as PROC TIMESERIES offer greater control and functionality. Furthermore, for those requiring maximum control over the window calculation or handling specific boundary conditions, custom logic can be implemented using the DATA step in combination with iterative loops or LAG functions.

Moving averages are not merely academic tools; they are invaluable for deriving actionable insights from noisy, real-world data. In financial markets, they are used to validate momentum: a moving average sloping upward confirms an uptrend, while a downward slope signals a downtrend. More sophisticated technical analysis often relies on the crossover of short-term and long-term moving averages--a short MA crossing above a long MA is frequently interpreted as a strong bullish signal.

Beyond finance, moving averages have powerful applications across various sectors. Manufacturers use them to monitor quality control metrics and detect subtle shifts in production standards before defects become systemic. Public health officials rely on them to smooth daily reported case counts, revealing the true trajectory of disease spread. Similarly, meteorologists use MAs to analyze climate data, filtering out daily volatility to highlight long-term changes in temperature or precipitation patterns. Their universal simplicity and effectiveness cement the moving average as a foundational and essential technique in any comprehensive data analysis toolkit.

Additional Resources

The following articles explain how to perform other common tasks in SAS: