**Seasonal Decomposition using Moving Averages (Ratio to Moving Average Method)**

An alternative to Winters’ method consists of first obtaining de-seasonalized or seasonally adjusted data. Time series consisting of weekly, monthly, or quarterly observations often exhibit seasonality. In multiplicative decomposition model, ratio to moving average is a popular method for measuring seasonal variation. In this method, the trend is estimated using a centered moving average.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Year-q** | **t** | **Yt** | **MA(4)** | **CMA** | **S,I=Y/CMA** | **S** | **DSY=Y/S** | **T** | |  | | --- | | **= S\*T** | | **et** | **PE** |
| **2010-q1** | **1** | **48.0** |  |  |  | **0.932** | **51.491** | **52.467** | **48.910** | **-0.910** | **-1.896** |
| **q2** | **2** | **41.0** |  |  |  | **0.838** | **48.940** | **53.939** | **45.188** | **-4.188** | **-10.214** |
| **q3** | **3** | **60.0** | **53.500** | **54.750** | **1.096** | **1.093** | **54.877** | **55.410** | **60.583** | **-0.583** | **-0.971** |
| **q4** | **4** | **65.0** | **56.000** | **57.375** | **1.133** | **1.143** | **56.853** | **56.882** | **65.033** | **-0.033** | **-0.051** |
| **2011-q1** | **5** | **58.0** | **58.750** | **59.750** | **0.971** | **0.932** | **62.218** | **58.353** | **54.397** | **3.603** | **6.213** |
| **q2** | **6** | **52.0** | **60.750** | **61.875** | **0.840** | **0.838** | **62.070** | **59.824** | **50.118** | **1.882** | **3.618** |
| **q3** | **7** | **68.0** | **63.000** | **63.250** | **1.075** | **1.093** | **62.194** | **61.296** | **67.018** | **0.982** | **1.445** |
| **q4** | **8** | **74.0** | **63.500** | **64.000** | **1.156** | **1.143** | **64.725** | **62.767** | **71.762** | **2.238** | **3.024** |
| **2012-q1** | **9** | **60.0** | **64.500** | **65.375** | **0.918** | **0.932** | **64.364** | **64.239** | **59.883** | **0.117** | **0.195** |
| **q2** | **10** | **56.0** | **66.250** | **66.750** | **0.839** | **0.838** | **66.845** | **65.710** | **55.049** | **0.951** | **1.698** |
| **q3** | **11** | **75.0** | **67.250** | **67.625** | **1.109** | **1.093** | **68.597** | **67.181** | **73.453** | **1.547** | **2.063** |
| **q4** | **12** | **78.0** | **68.000** | **68.375** | **1.141** | **1.143** | **68.223** | **68.653** | **78.491** | **-0.491** | **-0.630** |
| **2013-q1** | **13** | **63.0** | **68.750** | **69.375** | **0.908** | **0.932** | **67.582** | **70.124** | **65.370** | **-2.370** | **-3.762** |
| **q2** | **14** | **59.0** | **70.000** | **70.750** | **0.834** | **0.838** | **70.426** | **71.596** | **59.980** | **-0.980** | **-1.661** |
| **q3** | **15** | **80.0** | **71.500** |  |  | **1.093** | **73.170** | **73.067** | **79.888** | **0.112** | **0.140** |
| **q4** | **16** | **84.0** |  |  |  | **1.143** | **73.471** | **74.538** | **85.220** | **-1.220** | **-1.452** |
| **2014-q1** | **17** |  |  |  |  | **0.932** |  | **76.010** | **70.856** | **52.834** | **-0.140** |
| **q2** | **18** |  |  |  |  | **0.838** |  | **77.481** | **64.910** | **Sumsq** | **MPE** |
| **q3** | **19** |  |  |  |  | **1.093** |  | **78.952** | **86.323** | **3.302** |  |
| **q4** | **20** |  |  |  | **Y/CMA** | **1.143** |  | **80.424** | **91.949** | **MSE** |  |
|  |  |  |  |  |  |  |  |  |  | **RMSE** |  |
|  |  |  |  |  |  |  |  |  |  | **1.8172** |  |

The above table shows the four years quarterly data of sales (Yt) ($millions) from 2010 to 2013. Our goal is to predict sales for four quarters of 2014 accounting for the trend and seasonality factors. The plot of the data below clearly shown a quarterly seasonal pattern with a peak in the fourth quarter every year. Moreover, there is also a clear upward trend from year to year.

We converted the first column with year and quarters to a coded time t = 1,2…16…20 in the second column. The data is available only up to t = 16. The rest four quarters are future periods for which we want to predict/forecast the Y-values or **t** for the four quarters of 2014 (or t = 17 to 20).

The first step is to calculate the moving average (with 4 periods of seasonal interval) denoted as MA(4) in the fourth column. We put the first value in the third period although the position should be 2.5 or between second and third period. This is because we have the seasonal lag of 4, an even number. Since there is no position in the spreadsheet for 2.5 row, we put it in the third row. We will correct this issue in the next step when we calculate Centered moving Average CMA in the fifth column. For monthly data with lag of 12 we have similar issue because the MA(12) for the first 12 months belongs to the 6.5 potion but will be put in the 7th month and corrected by calculating CMA. This issue does not arise if we have odd numbered seasonal lag such as 3, 55, 7, 9, etc.

The formula we applied to get the first entry for MA(4) is =AVERAGE(C2:C5) which averaged the first four values of Yt. We copied this formula down (by dragging the corner) all the way up to t = 15 (we do not have four values to average after that). The formula for the last entry in column MA(4) becomes =AVERAGE(C14:C17) giving the value 7.15.

The next step is to calculate CMA. This is done simply by taking the average of two MA values. The first CMA entry is obtained by averaging 5.35 and 5.60 (MA(4) entry for t = 3 and 4) and reporting it in t = 3. The formula applied is =AVERAGE(D4:D5). We copy it down to t = 14, because there are not two values of MA available after that. The CMA is like smoothed data which has ironed out the fluctuations and mainly includes the Trend, as shown below.

Since the CMA reflects mainly the trend, we can use it to extract the Seasonal and Irregular component from the actual data by diving the Yt by CMA because Yt = T\*S\*I in the multiplicative model (ignoring the cyclical component for data with only few years as the present data). So, we call Yt/CMA as S,I and report it in the sixth column. The formula for the first entry is =C4/E4 which is copied down from t = 3 to t = 14. The first entry of S,I is for q3 of the first year. It shows that q3 value is about 10% higher than the first-year average. Similarly, the second entry shows that the fourth quarter value is about 13% higher than the first-year average. But these percentages change from year to year due to “Irregularities”. To remove the annual irregularities, we take the average by quarter for the various years (the textbook suggests using Median instead of Mean, but we follow the more popular method). There are three values for q1 in column S,I: 0.97, 0.92 and 0.97. The average rounded to decimal is 0.93 reported in the seventh column S, which has purged the irregularities from S,I. You can use the excel function for average such as =AVERAGE(F6,F10,F14) for quarter 1 and =AVERAGE(F7,F11,F15) for quarter 2, =AVERAGE(F4,F8,F12) for quarter 3, and =AVERAGE(F5,F9,F13) for quarter 4. But the problem is that it is very likely to mess up in selecting the appropriate cells by quarters. In case of monthly data, it will be even more difficult to keep track. You can use the “If” function of Excel if you know how. Otherwise, you can first create another table to facilitate this process. It does not matter whether you show the quarters in columns or rows. The values in the following table come from column S,I.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Seasonal Index |  |  |  |  |
|  |  |  |  |  |
| Year\Quarter | Q1 | Q2 | Q3 | Q4 |
| 2010.0000 |  |  | 1.0960 | 1.1330 |
| 2011.0000 | 0.9710 | 0.8400 | 1.0750 | 1.1560 |
| 2012.0000 | 0.9180 | 0.8390 | 1.1090 | 1.1410 |
| 2013.0000 | 0.9080 | 0.8340 |  |  |
| Average | 0.9320 | 0.8380 | 1.0930 | 1.1430 |

Now we can type these four average numbers in column S (seasonal indexes) for the four quarters (same value for a quarter each year) from t= 1 to t = 20 (we extend for four quarters beyond the sample period for prediction/forecasting). For example, the value 0.932 is reported for the first quarter repeatedly for the five years.

Next, we create the “de-seasonalized” or seasonally adjusted data denoted here as DSYt = Yt/S for t =1 to 16 in column 8. The following plot shows that it follows the linear trendline very closely. It is largely devoid of seasonal fluctuations. It has one advantage over the CMA that it covers all the sample period while CMA is missing for two observations in the beginning and two at the end. Besides, the Seasonal Indexes calculated in the process of deriving DSY will be useful for future forecasting periods.

The next step is to use the de-seasonalized data to estimate the trend line. We will use the Regression function of Excel. Click on data, then Data Analysis, and select regression from the menu. In the dialogue box select the column for DSY for the dependent (Y) variable and the column for t as the explanatory (X) variable. Check, labels and as many other options as you want. The resulting computer output is given below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |
|  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |
| Multiple R | 0.959579 |  |  |  |  |  |
| R Square | 0.920791 |  |  |  |  |  |
| Adjusted R Square | 0.915133 |  |  |  |  |  |
| Standard Error | 2.126712 |  |  |  |  |  |
| Observations | 16 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 736.0933 | 736.093258 | 162.7478588 | 4.248E-09 |  |
| Residual | 14 | 63.32068 | 4.522905945 |  |  |  |
| Total | 15 | 799.4139 |  |  |  |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 50.9961 | 1.115257 | 45.72585628 | 1.20987E-16 | 48.604112 | 53.38809 |
| t | 1.471387 | 0.115337 | 12.7572669 | 4.24772E-09 | 1.2240135 | 1.718761 |

The summary output shows a very successful regression with R-square 0.9208 implying that the linear model captures over 92% of variations in DSY across time (also confirmed by the high F-value and high t-values of the coefficients. The numbers of particular interest to us are the value of Intercept = 5.0996 and slope with respect to t = 0.1471. We will use these numbers to derive the Trend values for Y from t = 1 to t = 20 reported in column T. The first entry was obtained by using the Excel formula =$Q$31+$Q$32\*B2 which means Intercept plus slope times the t-value (=1). The intercept and slope coefficients were reported in the Q column at rows 31 and 32.We put $ signs in the cell reference for intercept and slope to lock the cells when we copy the formula for changing value of t. Copying this formula down the T column gives all the 20 values for T.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | |  |  |  |  |  |
|  |  |  |  |  |  |  |

The residuals and their plot show very low errors in the regression estimation and errors following a nonsystematic or random feature around zero.

|  |  |  |
| --- | --- | --- |
| RESIDUAL OUTPUT | |  |
|  |  |  |
| *Observation* | *Predicted DSY=Y/S* | *Residuals* |
| 1 | 52.46749 | -0.97642 |
| 2 | 53.93888 | -4.9988 |
| 3 | 55.41026 | -0.53299 |
| 4 | 56.88165 | -0.02894 |
| 5 | 58.35304 | 3.865339 |
| 6 | 59.82442 | 2.245918 |
| 7 | 61.29581 | 0.898429 |
| 8 | 62.7672 | 1.95743 |
| 9 | 64.23859 | 0.125252 |
| 10 | 65.70997 | 1.135011 |
| 11 | 67.18136 | 1.415228 |
| 12 | 68.65275 | -0.42949 |
| 13 | 70.12413 | -2.5421 |
| 14 | 71.59552 | -1.16956 |
| 15 | 73.06691 | 0.102785 |
| 16 | 74.5383 | -1.0671 |

Finally, we derive the Predicted/forecasted values of Y for all 20 periods returning the seasonality into the forecasted by multiplying the trend values with seasonal corresponding quarter seasonal indexes =S\*T. The forecasted values for the four quarters of 2014 are the results for which we performed all the above exercises. Let us see how our forecast incorporation trend and seasonality looks compared to the actual data. The plot shows a very good fit and the future quarter forecasts exhibit a trend and seasonality similar to the actual data.

**Example 2: Seasonal decomposition (using monthly data): solved Example: Forecast for 6 Months of 2006 using Ratio to Moving Average**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year-mo | t | Sales-Y | MA(12) | CMA | S,I=Y/CMA | S | DSY=Y/S | T | |  | | --- | | =S\*T | |
| 2004Jan | 1 | 180 |  |  |  | 1.30 | 138.33 | 123.56 | 160.78 |
| 2004Feb | 2 | 170 |  |  |  | 1.24 | 137.38 | 125.99 | 155.91 |
| 2004Mar | 3 | 150 |  |  |  | 1.13 | 132.55 | 128.42 | 145.32 |
| 2004Apr | 4 | 140 |  |  |  | 1.05 | 132.97 | 130.84 | 137.76 |
| 2004May | 5 | 130 |  |  |  | 0.95 | 136.73 | 133.27 | 126.71 |
| 2004Jun | 6 | 115 |  |  |  | 0.88 | 130.79 | 135.69 | 119.31 |
| 2004Jul | 7 | 105 | 138.33 | 139.17 | 0.75 | 0.76 | 137.37 | 138.12 | 105.57 |
| 2004Aug | 8 | 100 | 140.00 | 140.83 | 0.71 | 0.72 | 138.61 | 140.55 | 101.40 |
| 2004Sep | 9 | 120 | 141.67 | 142.71 | 0.84 | 0.84 | 142.60 | 142.97 | 120.31 |
| 2004Oct | 10 | 130 | 143.75 | 144.58 | 0.90 | 0.92 | 140.93 | 145.40 | 134.13 |
| 2004Nov | 11 | 150 | 145.42 | 146.25 | 1.03 | 1.02 | 146.87 | 147.82 | 150.97 |
| 2004Dec | 12 | 170 | 147.08 | 148.13 | 1.15 | 1.12 | 152.06 | 150.25 | 167.98 |
| 2005Jan | 13 | 200 | 149.17 | 150.00 | 1.33 | 1.30 | 153.70 | 152.68 | 198.66 |
| 2005Feb | 14 | 190 | 150.83 | 151.67 | 1.25 | 1.24 | 153.54 | 155.10 | 191.94 |
| 2005Mar | 15 | 175 | 152.50 | 153.33 | 1.14 | 1.13 | 154.64 | 157.53 | 178.27 |
| 2005Apr | 16 | 160 | 154.17 | 155.42 | 1.03 | 1.05 | 151.97 | 159.95 | 168.41 |
| 2005May | 17 | 150 | 156.67 | 157.71 | 0.95 | 0.95 | 157.76 | 162.38 | 154.39 |
| 2005Jun | 18 | 140 | 158.75 | 159.58 | 0.88 | 0.88 | 159.23 | 164.80 | 144.90 |
| 2005Jul | 19 | 125 | 160.42 | 161.46 | 0.77 | 0.76 | 163.54 | 167.23 | 127.82 |
| 2005Aug | 20 | 120 | 162.50 | 163.75 | 0.73 | 0.72 | 166.33 | 169.66 | 122.40 |
| 2005Sep | 21 | 140 | 165.00 | 166.25 | 0.84 | 0.84 | 166.37 | 172.08 | 144.81 |
| 2005Oct | 22 | 160 | 167.50 | 169.17 | 0.95 | 0.92 | 173.45 | 174.51 | 160.98 |
| 2005Nov | 23 | 175 | 170.83 | 172.08 | 1.02 | 1.02 | 171.35 | 176.93 | 180.70 |
| 2005Dec | 24 | 190 | 173.33 | 174.58 | 1.09 | 1.12 | 169.95 | 179.36 | 200.52 |
| 2006Jan | 25 | 225 | 175.83 | 177.29 | 1.27 | 1.30 | 172.92 | 181.79 | 236.54 |
| 2006Feb | 26 | 220 | 178.75 | 180.00 | 1.22 | 1.24 | 177.78 | 184.21 | 227.96 |
| 2006Mar | 27 | 205 | 181.25 | 182.71 | 1.12 | 1.13 | 181.15 | 186.64 | 211.21 |
| 2006Apr | 28 | 200 | 184.17 | 185.83 | 1.08 | 1.05 | 189.96 | 189.06 | 199.06 |
| 2006May | 29 | 180 | 187.50 | 189.38 | 0.95 | 0.95 | 189.31 | 191.49 | 182.07 |
| 2006Jun | 30 | 170 | 191.25 | 192.92 | 0.88 | 0.88 | 193.35 | 193.92 | 170.50 |
| 2006Jul | 31 | 160 | 194.58 |  |  | 0.76 | 209.33 | 196.34 | 150.07 |
| 2006Aug | 32 | 150 |  |  |  | 0.72 | 207.92 | 198.77 | 143.40 |
| 2006Sep | 33 | 175 |  |  |  | 0.84 | 207.96 | 201.19 | 169.30 |
| 2006Oct | 34 | 200 |  |  |  | 0.92 | 216.81 | 203.62 | 187.83 |
| 2006Nov | 35 | 220 |  |  |  | 1.02 | 215.41 | 206.05 | 210.43 |
| 2006Dec | 36 | 230 |  |  |  | 1.12 | 205.73 | 208.47 | 233.07 |
| 2007Jan | 37 |  |  |  |  | 1.30 |  | 210.90 | 274.42 |
| 2007Feb | 38 |  |  |  |  | 1.24 |  | 213.32 | 263.98 |
| 2007Mar | 39 |  |  |  |  | 1.13 |  | 215.75 | 244.15 |
| 2007Apr | 40 |  |  |  |  | 1.05 |  | 218.18 | 229.71 |
| 2007May | 41 |  |  |  |  | 0.95 |  | 220.60 | 209.75 |
| 2007Jun | 42 |  |  |  |  | 0.88 |  | 223.03 | 196.10 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | |  |  |  |  |  |
|  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |
| Multiple R | 0.969 |  |  |  |  |  |
| R Square | 0.940 |  |  |  |  |  |
| Adjusted R Square | 0.938 |  |  |  |  |  |
| Standard Error | 6.561 |  |  |  |  |  |
| Observations | 36 |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 22864.03307 | 22864.03 | 531.1104 | 2.48078E-22 |  |
| Residual | 34 | 1463.682875 | 43.0495 |  |  |  |
| Total | 35 | 24327.71594 |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 121.1378 | 2.2334 | 54.2381 | 0.0000 | 116.5989 | 125.6767 |
| t | 2.4259 | 0.1053 | 23.0458 | 0.0000 | 2.2120 | 2.6399 |



|  |  |  |
| --- | --- | --- |
| RESIDUAL OUTPUT | |  |
|  |  |  |
| *Observation* | *Predicted DSY=Y/S* | *Residuals* |
| 1 | 123.564 | 14.769 |
| 2 | 125.990 | 11.386 |
| 3 | 128.416 | 4.134 |
| 4 | 130.842 | 2.129 |
| 5 | 133.267 | 3.458 |
| 6 | 135.693 | -4.900 |
| 7 | 138.119 | -0.746 |
| 8 | 140.545 | -1.934 |
| 9 | 142.971 | -0.367 |
| 10 | 145.397 | -4.472 |
| 11 | 147.823 | -0.951 |
| 12 | 150.249 | 1.809 |
| 13 | 152.675 | 1.028 |
| 14 | 155.101 | -1.564 |
| 15 | 157.527 | -2.886 |
| 16 | 159.953 | -7.986 |
| 17 | 162.379 | -4.618 |
| 18 | 164.805 | -5.578 |
| 19 | 167.231 | -3.691 |
| 20 | 169.657 | -3.323 |
| 21 | 172.083 | -5.711 |
| 22 | 174.509 | -1.062 |
| 23 | 176.934 | -5.583 |
| 24 | 179.360 | -9.413 |
| 25 | 181.786 | -8.871 |
| 26 | 184.212 | -6.432 |
| 27 | 186.638 | -5.488 |
| 28 | 189.064 | 0.894 |
| 29 | 191.490 | -2.178 |
| 30 | 193.916 | -0.569 |
| 31 | 196.342 | 12.988 |
| 32 | 198.768 | 9.149 |
| 33 | 201.194 | 6.770 |
| 34 | 203.620 | 13.188 |
| 35 | 206.046 | 9.367 |
| 36 | 208.472 | -2.746 |