

May 2007

Forecasting with Excel: Suggestions for Managers

Scott Nadler

East Carolina University, nadlers@mail.ecu.edu

John F. Kros

East Carolina University, krosj@ecu.edu

Follow this and additional works at: <http://epublications.bond.edu.au/ejsie>



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Recommended Citation

Nadler, Scott and Kros, John F. (2007) Forecasting with Excel: Suggestions for Managers, *Spreadsheets in Education (eJSiE)*: Vol. 2: Iss. 2, Article 5.

Available at: <http://epublications.bond.edu.au/ejsie/vol2/iss2/5>

This In the Classroom Article is brought to you by the Bond Business School at [ePublications@bond](mailto:epublications@bond.edu.au). It has been accepted for inclusion in *Spreadsheets in Education (eJSiE)* by an authorized administrator of [ePublications@bond](mailto:epublications@bond.edu.au). For more information, please contact [Bond University's Repository Coordinator](#).

Forecasting with Excel: Suggestions for Managers

Abstract

This article provides readers and more especially business managers with an overview of moving average, exponential smoothing, trend analysis, and linear regression approaches to forecasting. The authors then provide specific examples for each approach and the Excel formulas necessary to develop effective forecasts. This is an important contribution to the literature because it demonstrates that businesses with limited resources can develop reliable and accurate forecasts in a timely and cost effective manner using readily available software.

Keywords

forecasting, quantitative, Excel

Distribution License



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Forecasting with Excel: Suggestions for Managers

Scott Nadler

Assistant Professor of Marketing and Supply Chain Management
Department of Marketing and Supply Chain Management
College of Business
3102 Bate Building
East Carolina University
Greenville, NC 27858
nadlers@mail.ecu.edu
252-328-5328

John F. Kros*

Assistant Professor of Marketing and Supply Chain Management
Department of Marketing and Supply Chain Management
College of Business
East Carolina University
3121 Bate Building
Greenville, NC 27858-4353
krosj@mail.ecu.edu
(252) 328-6364

Abstract

This article provides readers and more especially business managers with an overview of moving average, exponential smoothing, trend analysis, and linear regression approaches to forecasting. The authors then provide specific examples for each approach and the Excel formulas necessary to develop effective forecasts. This is an important contribution to the literature because it demonstrates that businesses with limited resources can develop reliable and accurate forecasts in a timely and cost effective manner using readily available software.

Keywords: forecasting, quantitative, Excel

1. Introduction

In the ever-changing world of business being able to accurately forecast customer demand is of the utmost importance. Without accurate forecasts companies essentially waste resources by carrying either too much (i.e. overstocks) or too little (i.e. stock outs) inventory [2]. To remedy this problem companies use many different methods of forecasting in an attempt to predict customer demand. Forecasting can be either quantitatively or qualitatively based. Quantitative forecasting uses historical data to project demand. This article focuses solely on the four quantitative approaches that businesses can develop using Excel spreadsheets.

The quantitative methods that will be addressed in this article include: moving averages, exponential smoothing, trend analysis, and linear regression. These approaches can be further aggregated into two groups, time-series models and associative models, with all methods presented falling under time-series except for linear regression. Time-series analysis utilizes historic data in series to make accurate forecasts. Associative models incorporate one or more extraneous variables to accomplish this purpose.

The purpose of this article is to provide readers and more specifically managers of small to medium sized businesses with a better understanding of how businesses with limited access to sophisticated forecasting tools can more accurately forecast demand. To accomplish the stated purpose the authors will provide an overview of each method and an example of how it can be implemented. Each example will be illustrated using manually inputted cell references in Excel as well as using Excel's Data Analysis Add-In package. This will be followed by a summary with recommendations that will enable readers to more effectively implement a forecasting program.

2. Quantitative Forecasting

2.1 Moving Averages

Moving averages are considered to be the easiest method of quantitative forecasting to develop. This particular method of forecasting is essentially a smoothing model, which makes forecasts more accurate. The basic premise behind the use of moving averages is to take data averages from x periods in order to forecast demand for the upcoming period. Most businesses use either a 3 month (quarterly) or 6 month (biannual) moving average because sales data is recorded on a monthly basis [1]. The longer a moving average is, the smoother the data set becomes. In fact, one can control the effect of smoothing on the time series data by using a longer or shorter averaging time period. The formula for calculating a moving average is as follows:

$$MA_n = \frac{\sum_{t=1}^n D_t}{n} \quad (1)$$

where n = number of periods in moving average and D_t = data in period t .

2.1.1 Moving Average Example Using Excel Functions

For example, a restaurant that features smoked ribs wants to forecast sales for the next year. In this example the restaurant manager has actual sales data for each of the last twenty-two years. Using this scenario the manager employs a three and five-year moving average to forecast the next year's sales figure. The mathematical equation for the three-year moving average is expressed as:

$$MA(3) = M_{t+1} = \frac{(D_t + D_{t-1} + D_{t-2})}{3}$$

whereas the five-year moving average is

$$MA(5) = M_{t+1} = \frac{(D_t + D_{t-1} + D_{t-2} + D_{t-3} + D_{t-4})}{5}$$

where $MA(3)$ is the moving average using 3 periods of data, $MA(5)$ is the moving average using 5 periods of data and D is demand and t is time.

Table 1: Sales Data Set for Rib Sales

Year	Sales (\$000)	Year	Sales (\$000)
1980	52.04	1991	83.18
1981	59.42	1992	87.05
1982	55.66	1993	84.79
1983	53.86	1994	73.49
1984	64.59	1995	76.23
1985	75.28	1996	96.54
1986	61.89	1997	95.08
1987	73.74	1998	87.05
1988	81.19	1999	96.02
1989	97.52	2000	98.90
1990	86.50	2001	83.23

The $MA(3)$ and $MA(5)$ forecasts are created manually via the following formulae in Excel (see Figure 1).

	A	B	C	D
1	Year	Sales (\$000)	3-Year MA	5-Year MA
2	1980	52.04		
3	1981	59.42		
4	1982	55.66		
5	1983	53.86	= (B4+B3+B2)/3	
6	1984	64.59	= (B5+B4+B3)/3	
7	1985	75.28	= (B6+B5+B4)/3	= (B6+B5+B4+B3+B2)/5
8	1986	61.89	= (B7+B6+B5)/3	= (B7+B6+B5+B4+B3)/5
9	1987	73.74	= (B8+B7+B6)/3	= (B8+B7+B6+B5+B4)/5
10	1988	81.19	= (B9+B8+B7)/3	= (B9+B8+B7+B6+B5)/5
11	1989	97.52	= (B10+B9+B8)/3	= (B10+B9+B8+B7+B6)/5
12	1990	86.5	= (B11+B10+B9)/3	= (B11+B10+B9+B8+B7)/5
13	1991	83.18	= (B12+B11+B10)/3	= (B12+B11+B10+B9+B8)/5
14	1992	87.05	= (B13+B12+B11)/3	= (B13+B12+B11+B10+B9)/5
15	1993	84.79	= (B14+B13+B12)/3	= (B14+B13+B12+B11+B10)/5
16	1994	73.49	= (B15+B14+B13)/3	= (B15+B14+B13+B12+B11)/5
17	1995	76.23	= (B16+B15+B14)/3	= (B16+B15+B14+B13+B12)/5
18	1996	96.54	= (B17+B16+B15)/3	= (B17+B16+B15+B14+B13)/5
19	1997	95.08	= (B18+B17+B16)/3	= (B18+B17+B16+B15+B14)/5
20	1998	87.05	= (B19+B18+B17)/3	= (B19+B18+B17+B16+B15)/5
21	1999	96.02	= (B20+B19+B18)/3	= (B20+B19+B18+B17+B16)/5
22	2000	98.9	= (B21+B20+B19)/3	= (B21+B20+B19+B18+B17)/5
23	2001	83.23	= (B22+B21+B20)/3	= (B22+B21+B20+B19+B18)/5
24	2002		= (B23+B22+B21)/3	= (B23+B22+B21+B20+B19)/5

Figure 1: Moving Average Formulae in Excel

The actual forecasts can be found in Figure 2 and Figure 3. Based on this analysis the manager would forecast \$92,720 in sales using the three-year moving average and \$92,060 in sales using the five-year moving average.

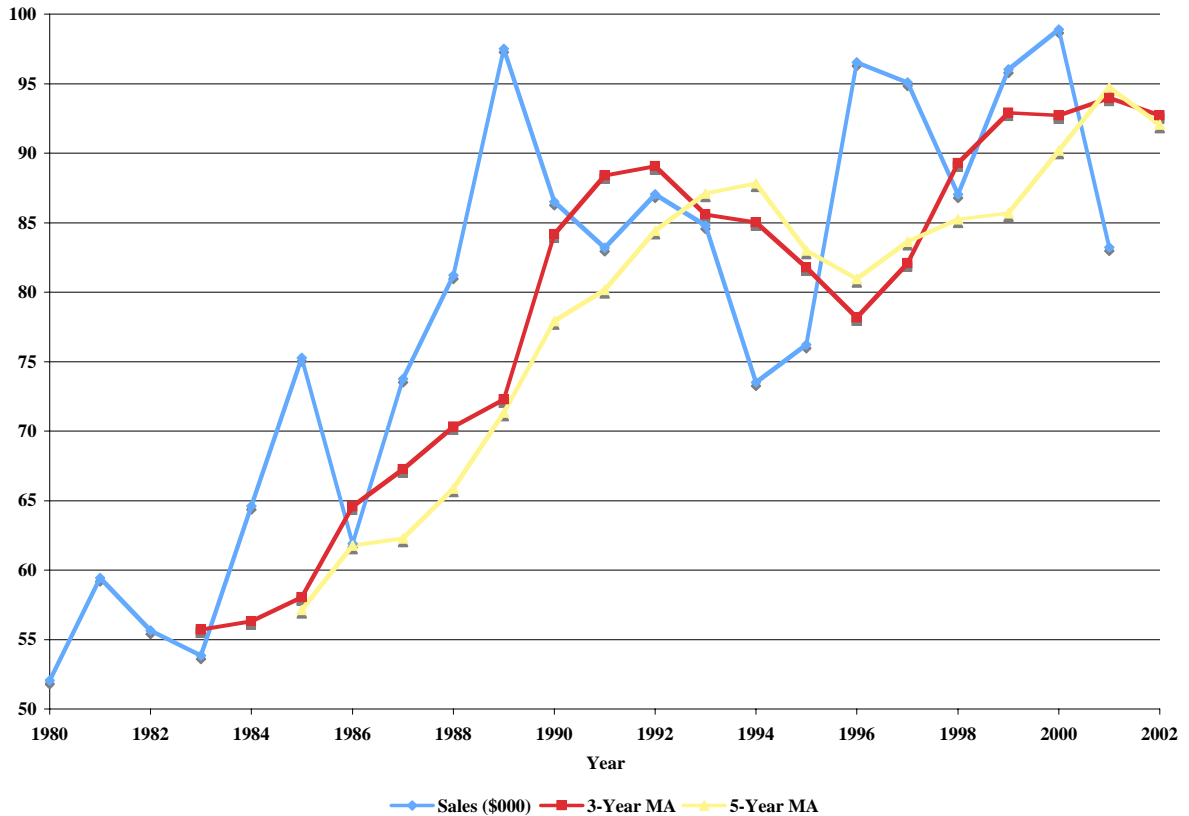


Figure 2: Graph of Actual Sales, MA(3)Sales Forecast, and MA(5) Sales Forecast

2.1.2 Moving Average Example Using Excel’s Data Analysis Add-In

To illustrate the concept of forecasting using moving averages, let’s analyze the data set listed in Table 1. We will demonstrate a three-year and five-year moving using Excel’s Moving Average tool. The following steps should be completed to create a moving average in Excel:

1. Be sure to have Excel’s Data Analysis package loaded (see Microsoft Help menu)
2. Choose the Tools menu in Excel and then the Data Analysis option (see Figure 4), a new menu resembling Figure 5 will appear on screen.

	A	B	C	D
1	Year	Sales (\$000)	3-Year MA	5-Year MA
2	1980	52.04		
3	1981	59.42		
4	1982	55.66		
5	1983	53.86	55.71	
6	1984	64.59	56.31	
7	1985	75.28	58.04	57.11
8	1986	61.89	64.58	61.76
9	1987	73.74	67.25	62.26
10	1988	81.19	70.30	65.87
11	1989	97.52	72.27	71.34
12	1990	86.50	84.15	77.92
13	1991	83.18	88.40	80.17
14	1992	87.05	89.07	84.43
15	1993	84.79	85.58	87.09
16	1994	73.49	85.01	87.81
17	1995	76.23	81.78	83.00
18	1996	96.54	78.17	80.95
19	1997	95.08	82.09	83.62
20	1998	87.05	89.28	85.23
21	1999	96.02	92.89	85.68
22	2000	98.90	92.72	90.18
23	2001	83.23	93.99	94.72
24	2002		92.72	92.06

Figure 3: MA(3) and MA(5) Sales Forecasts

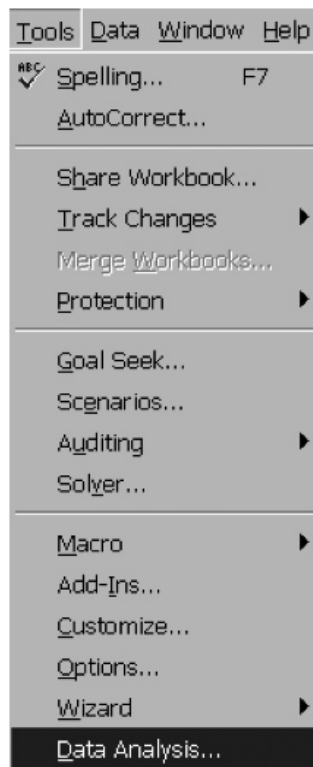


Figure 4: Tools Menu in Excel

3. From the dialog box displayed in Figure 5, choose Moving Average and click OK.
4. After clicking OK, Excel displays a dialog box that prompts you for an Input range, Interval, and Output range. This dialog box accompanied by the time series data is shown in Figure 6.
5. Enter the sales data range in the Input range edit box by highlighting the data or typing in its reference.
6. Click on the Labels in first row box if the data set contains a label in the first row.
7. Enter the number of periods to be included in the moving average in the Interval edit box.
8. Click on the Output range edit box and enter the address of the cell or click on the cell where the output should start. The authors suggest juxtaposing your forecasted values with your actual values. For example, for a MA(3) cell C3 should be entered (refer to Figure 6 and Figure 7).
9. Click OK.

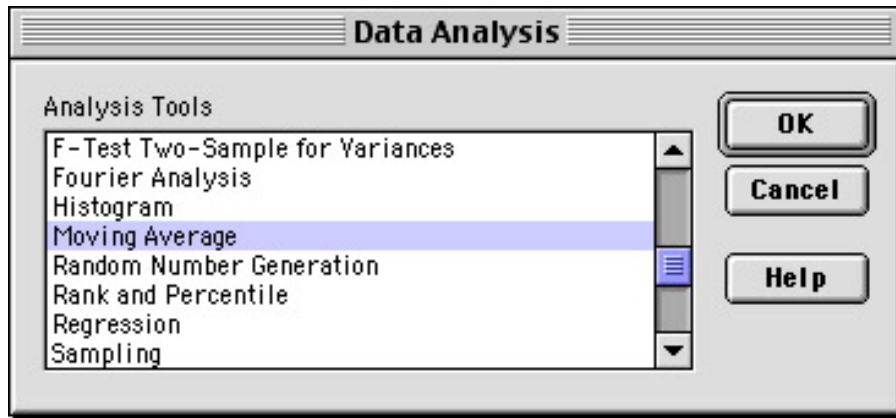


Figure 5: Data Analysis Dialog Box

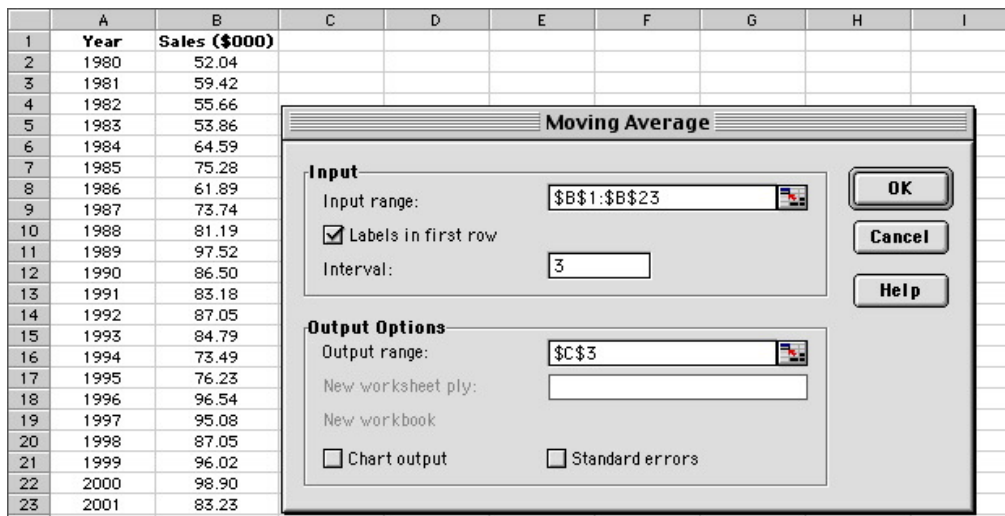


Figure 6: Excel's Moving Average Dialog Box and Time Series Data

Excel will fill in the formulas for the moving average. The moving average will begin with the term $\#N/A$. There will be as many $\#N/A$ terms as the interval you specified, minus one. Excel does this because there is not enough data to calculate an average for those observations that number less than the interval. Figure 7 displays the results of the three and five-month moving averages.

From cells C24 and D24 in Figure 6, it can be noted that the three-year and five-year moving average forecasts for 2002 are 92.72 and 92.06. These values are identical to the forecasts in Figure 4. Forecasts are generally compared using a measure of error. Many forecasters use mean absolute deviation (MAD) as a measure of model error. Equation (2) displays the formula for MAD. MAD for the three-year and five-year moving averages are 8.09 and 9.90 respectively. It could be said that relative to the each other the three-year moving average model does a better job of forecasting than the five-year moving average model.

$$MAD = \frac{\sum_{t=1}^n (D_t - F_t)}{n} \tag{2}$$

◇	A	B	C	D
1	Year	Sales (\$000)	3-Year MA	5-Year MA
2	1980	52.04		
3	1981	59.42	#N/A	#N/A
4	1982	55.66	#N/A	#N/A
5	1983	53.86	55.71	#N/A
6	1984	64.59	56.31	#N/A
7	1985	75.28	58.04	57.11
8	1986	61.89	64.58	61.76
9	1987	73.74	67.25	62.26
10	1988	81.19	70.30	65.87
11	1989	97.52	72.27	71.34
12	1990	86.50	84.15	77.92
13	1991	83.18	88.40	80.17
14	1992	87.05	89.07	84.43
15	1993	84.79	85.58	87.09
16	1994	73.49	85.01	87.81
17	1995	76.23	81.78	83.00
18	1996	96.54	78.17	80.95
19	1997	95.08	82.09	83.62
20	1998	87.05	89.28	85.23
21	1999	96.02	92.89	85.68
22	2000	98.90	92.72	90.18
23	2001	83.23	93.99	94.72
24	2002		92.72	92.06

Figure 7: Three and Five-Year Moving Averages for Sales Time Series Data

2.2 Exponential Smoothing

Exponential smoothing is another method of quantitative forecasting that some businesses have found useful due to its short-term precision and cost effectiveness. The use of exponential smoothing as a means of forecasting is slightly more complex than traditional moving average forecasts. Specifically, exponential smoothing combines a prior forecast with a weighted error measurement. This error measurement comes from the difference between the prior forecast and the actual observation at the time of the prior forecast. It also requires managers to arbitrarily select a value for alpha [4]. It must be stated that although any value of alpha between 0 and 1.0 inclusive is acceptable, certain values may yield better forecasts. By rewriting the basic exponential smoothing model shown in (3) to resemble (4), insight into good values of alpha can be obtained.

The formula that the exponential smoothing approach is based on can be expressed as

$$F_{t+1} = F_t + \alpha * e_t \quad (3)$$

Given that, t is time, F_t is the forecast at time t , and α is a smoothing constant (a value between 0.0 and 1.0 inclusive), and $e_t = D_t - F_t$, where D_t denotes the actual observation at time t , equation (2) can be re-written as follows:

$$F_{t+1} = F_t + \alpha * (D_t - F_t) \quad (4)$$

From (4) it can be seen that the new forecast, F_{t+1} , is equal to the previous forecast, F_t , plus an adjustment (i.e., $\alpha * (D_t - F_t)$). Specifically, the forecast for period $t+1$ is obtained by adjusting the forecast in period t by a fraction of the forecast error (i.e.,

$D_t - F_t$). If the time series contains substantial variability, a small value of the smoothing constant is preferred so not to overreact and adjust forecasts too quickly. For a time series with small variability a larger value of alpha is preferred so the forecast can quickly adjust when forecasting errors due occur and therefore allow the model to react faster to changing conditions. Overall, a user should choose an alpha value that minimizes the error of the model.

With some mathematical manipulation, equation (4) can be rearranged to yield the following equations.

$$F_{t+1} = F_t + \alpha * D_t - \alpha * F_t \quad (5a)$$

$$F_{t+1} = (1 - \alpha) * F_t + \alpha * D_t \quad (5b)$$

Equation (5b) describes the forecast at period $t+1$ as a proportion of the forecast at period t plus a proportion of the actual observation at period t . If α is 1.0 the forecast at period $t+1$ is the actual observation, D_t , from period t . Conversely, if α is 0.0 the forecast at period $t+1$ is the forecast, F_t , from period t .

2.2.1 Exponential Smoothing Example Using Excel Functions

The data from Table 1 will be used to demonstrate exponential smoothing. For this example, alpha is arbitrarily set at two levels, 0.1 and 0.3. In addition, a common method of bootstrapping is used to initialize the first period's forecast. This bootstrap method is to assume the first year's forecast is the same as the first year's sales demand (see Figure 8, cell C3 and D3). Using the same data set as presented previously (see Table 1) the exponential smoothing forecast is created in Excel using the following spreadsheet formulae (see Figure 8).

The actual forecasts can be found in Figure 9. Based on this analysis the manager would forecast \$81,720 in sales using the exponential smoothing method with an alpha of 0.1 and \$89,860 in sales using the exponential smoothing method with an alpha of 0.3. The model using an alpha of 0.1 has an MAD of 14.13 whereas the model using an alpha of 0.3 has an MAD of 8.52. Comparing the MAD of each model it is seen that the model that utilizes an alpha of 0.3 does a better job of forecasting.

FORECASTING WITH EXCEL: SUGGESTIONS FOR MANAGERS

	A	B	C	D
1	Year	Sales (\$000)	Exp. Smoothed Forecast - alpha=0.1	Exp. Smoothed Forecast - alpha=0.3
2			0.1	0.3
3	1980	52.04	=B3	=B3
4	1981	59.42	=SC\$2*B3+(1-SC\$2)*C3	=DS\$2*B3+(1-DS\$2)*D3
5	1982	55.66	=SC\$2*B4+(1-SC\$2)*C4	=DS\$2*B4+(1-DS\$2)*D4
6	1983	53.86	=SC\$2*B5+(1-SC\$2)*C5	=DS\$2*B5+(1-DS\$2)*D5
7	1984	64.59	=SC\$2*B6+(1-SC\$2)*C6	=DS\$2*B6+(1-DS\$2)*D6
8	1985	75.28	=SC\$2*B7+(1-SC\$2)*C7	=DS\$2*B7+(1-DS\$2)*D7
9	1986	61.89	=SC\$2*B8+(1-SC\$2)*C8	=DS\$2*B8+(1-DS\$2)*D8
10	1987	73.74	=SC\$2*B9+(1-SC\$2)*C9	=DS\$2*B9+(1-DS\$2)*D9
11	1988	81.19	=SC\$2*B10+(1-SC\$2)*C10	=DS\$2*B10+(1-DS\$2)*D10
12	1989	97.52	=SC\$2*B11+(1-SC\$2)*C11	=DS\$2*B11+(1-DS\$2)*D11
13	1990	86.5	=SC\$2*B12+(1-SC\$2)*C12	=DS\$2*B12+(1-DS\$2)*D12
14	1991	83.18	=SC\$2*B13+(1-SC\$2)*C13	=DS\$2*B13+(1-DS\$2)*D13
15	1992	87.05	=SC\$2*B14+(1-SC\$2)*C14	=DS\$2*B14+(1-DS\$2)*D14
16	1993	84.79	=SC\$2*B15+(1-SC\$2)*C15	=DS\$2*B15+(1-DS\$2)*D15
17	1994	73.49	=SC\$2*B16+(1-SC\$2)*C16	=DS\$2*B16+(1-DS\$2)*D16
18	1995	76.23	=SC\$2*B17+(1-SC\$2)*C17	=DS\$2*B17+(1-DS\$2)*D17
19	1996	96.54	=SC\$2*B18+(1-SC\$2)*C18	=DS\$2*B18+(1-DS\$2)*D18
20	1997	95.08	=SC\$2*B19+(1-SC\$2)*C19	=DS\$2*B19+(1-DS\$2)*D19
21	1998	87.05	=SC\$2*B20+(1-SC\$2)*C20	=DS\$2*B20+(1-DS\$2)*D20
22	1999	96.02	=SC\$2*B21+(1-SC\$2)*C21	=DS\$2*B21+(1-DS\$2)*D21
23	2000	98.9	=SC\$2*B22+(1-SC\$2)*C22	=DS\$2*B22+(1-DS\$2)*D22
24	2001	83.23	=SC\$2*B23+(1-SC\$2)*C23	=DS\$2*B23+(1-DS\$2)*D23
25	2002		=SC\$2*B24+(1-SC\$2)*C24	=DS\$2*B24+(1-DS\$2)*D24

Figure 8: Exponential Smoothing Formulae in Excel

	A	B	C	D
1	Year	Sales (\$000)	Exp. Smoothed Forecast - alpha=0.1	Exp. Smoothed Forecast - alpha=0.3
2			0.1	0.3
3	1980	52.04	52.04	52.04
4	1981	59.42	52.04	52.04
5	1982	55.66	52.78	54.25
6	1983	53.86	53.07	54.68
7	1984	64.59	53.15	54.43
8	1985	75.28	54.29	57.48
9	1986	61.89	56.39	62.82
10	1987	73.74	56.94	62.54
11	1988	81.19	58.62	65.90
12	1989	97.52	60.88	70.49
13	1990	86.50	64.54	78.60
14	1991	83.18	66.74	80.97
15	1992	87.05	68.38	81.63
16	1993	84.79	70.25	83.26
17	1994	73.49	71.70	83.72
18	1995	76.23	71.88	80.65
19	1996	96.54	72.32	79.32
20	1997	95.08	74.74	84.49
21	1998	87.05	76.77	87.67
22	1999	96.02	77.80	87.48
23	2000	98.90	79.62	90.04
24	2001	83.23	81.55	92.70
25	2002		81.72	89.86

Figure 9: Exponentially Smoothed Forecasts with a = 0.1 and a = 0.3

2.2.2 Exponential Smoothing Example Using Excel's Data Analysis Add-In

The same data set will be used here to illustrate Excel's Data Analysis Add-In for Exponential Smoothing. Again for consistency, we will use $\alpha = 0.1$ and $\alpha = 0.3$. Take note that Excel's Data Analysis Add-In uses what is called a "damping factor." The damping factor is equal to 1 minus the smoothing constant. See step 6 for further discussion. The following steps should be completed to create an exponentially smoothed model in Excel:

1. Be sure to have Excel's Data Analysis package loaded (see Excel Help Menu)
2. Choose the Tools menu in Excel and then the Data Analysis
3. From the dialog box displayed in Figure 11, choose Exponential Smoothing and click OK.

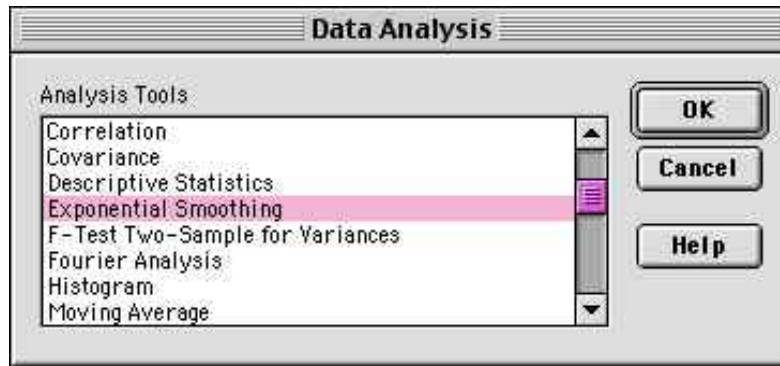


Figure 10: Data Analysis Dialog Box

4. Excel displays a dialog box that prompts you for an Input range, Damping factor, and Output range. This dialog box accompanied by the time series data is shown in Figure 11.

FORECASTING WITH EXCEL: SUGGESTIONS FOR MANAGERS

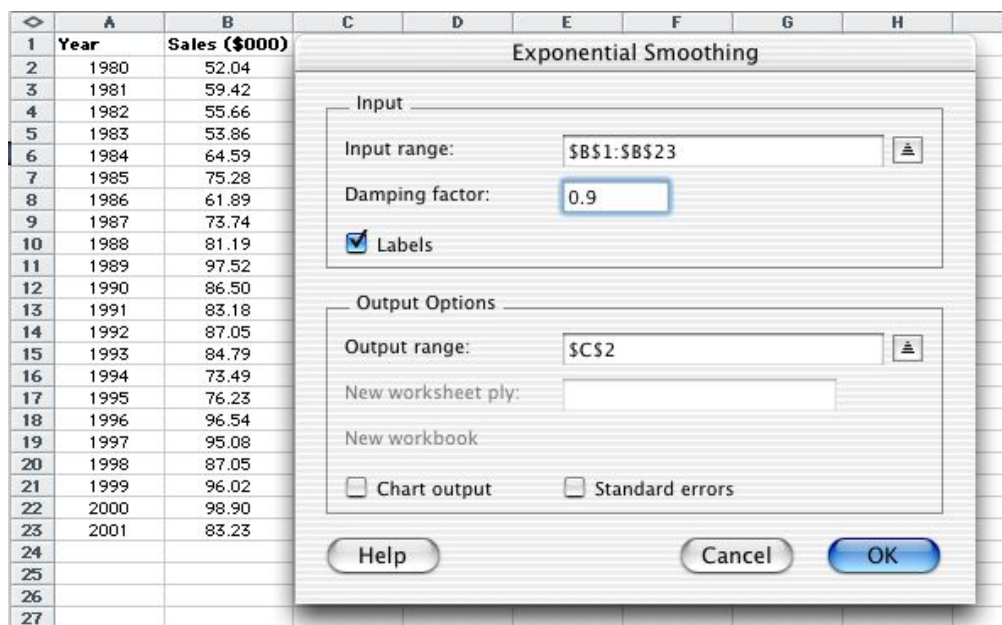


Figure 11: Excel's Exponential Smoothing Dialog Box

5. Enter the sales data range in the Input range edit box by highlighting the data or typing in its reference.
6. Enter the value of $1-\alpha$ in the Damping factor edit box. Excel requests a damping factor instead of a smoothing constant. The damping factor is equal to 1 minus the smoothing constant (i.e., $1-\text{smoothing constant} = \text{damping factor}$).
7. Click on the Labels in first row box if the data set contains a label in the first row.
8. Click on the Output range edit box and enter the address of the cell or click on the cell where the output should start. The authors suggest juxtaposing your forecasted values with your actual values (i.e., cell C2 should be entered for this example so the forecasted values match up to the actual values, see Figure 11 and Figure 12).
9. Click OK.

Excel will fill in the formulas for the exponentially smoothed model. The exponentially smoothed forecasts will begin with the term #N/A. There will be one #N/A terms. Excel does this because there is not enough data to calculate an exponentially smoothed forecast for the first observation. Figure 12 contains the Data Analysis package exponential smoothing results.

	A	B	C	D
1	Year	Sales (\$000)	Exp. Smoothed Forecast alpha=0.1	Exp. Smoothed Forecast - alpha=0.3
2	1980	52.04	#N/A	#N/A
3	1981	59.42	52.04	52.04
4	1982	55.66	52.78	54.25
5	1983	53.86	53.07	54.68
6	1984	64.59	53.15	54.43
7	1985	75.28	54.29	57.48
8	1986	61.89	56.39	62.82
9	1987	73.74	56.94	62.54
10	1988	81.19	58.62	65.90
11	1989	97.52	60.88	70.49
12	1990	86.50	64.54	78.60
13	1991	83.18	66.74	80.97
14	1992	87.05	68.38	81.63
15	1993	84.79	70.25	83.26
16	1994	73.49	71.70	83.72
17	1995	76.23	71.88	80.65
18	1996	96.54	72.32	79.32
19	1997	95.08	74.74	84.49
20	1998	87.05	76.77	87.67
21	1999	96.02	77.80	87.48
22	2000	98.90	79.62	90.04
23	2001	83.23	81.55	92.70
24	2002		81.72	89.86

Figure 12: Exponential Smoothing Results

It should be noted that cells C24 and D24 in Figure 12 were copied and pasted from cells C23 and D23 respectively. Although the exponential smoothing procedure in Excel's Data Analysis procedure does fill in the applicable formulae it does not automatically create a one-step ahead post hoc forecast. In addition, from cells C24 and D24 in Figure 12, it can be noted that the exponentially smoothed forecasts for 2002 are 81.72 and 89.86. These values are identical to the forecasts in Figure 9.

A question that is often asked is "Which type of time series model should be used?". Overall, which type of model used depends on the forecasting goal (e.g., tracking long term trends or looking for quick reactions to change). Forecasters must examine the trade-off between model sensitivity and model reliability. Moving averages lag the actual data but exponential smoothing can be prone to hyperactive movement. Some prefer to use exponential smoothing models for shorter time periods attempting to capture changes in the data more quickly whereas some prefer using moving averages over longer time periods to identify long-term trend changes.

It is a common misconception among novice forecasters that greater sensitivity and quicker response to actual data changes are utmost beneficial. However, this misconception leads directly to the question about the trade off between sensitivity and reliability. The more sensitive a model is to the actual data the more volatile the forecasts will be. Although this volatility may prove timely it can lead to an increase in model overreaction. Conversely, the less sensitive a model is to the actual data the volatility of the forecast may decrease but at the expense of precision.

Therefore, when using moving averages the shorter the moving average is the more sensitive it will be to actual data and when using exponential smoothing models the larger alpha is the more sensitive it will be to the actual data. In sum, a forecaster must examine the trade-off between sensitivity and reliability depending on the data set at hand and personal preference.

2.3 Linear Trend Analysis

Linear trend is a method of forecasting variations in a dependent variable, Y , employing one independent variable, X . When graphed the linear factor is used to determine if there is an increase or decrease in the dependent variable, Y , as the independent variable, X , increases [3].

An example of trend analysis is graphically illustrated in the following example. Using the data from Table 1 the following graph can be created.

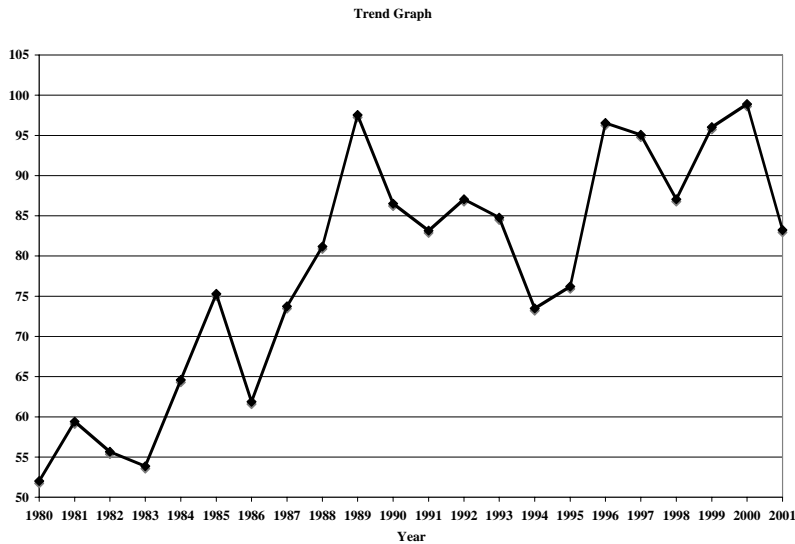


Figure 13: Graph of Year Versus Sales

In this example it is evident that there is an upward trend. However, to analyze a graph only by “eyeballing” the results is insufficient for most managers. In order to obtain a better notion of the trend associated with the data one can use Excel to construct a trend equation that explains the relation of the data. The next section presents two methods within Excel for analyzing linear trend.

2.3.1 Linear Trend Analysis Example Using Excel’s Trend Function

Once again we will enter the data into the Excel spreadsheet using the same sales numbers as in the previous example. Figure 14 contains the data and the trend formulae entered into the spreadsheet.

	A	B	C
	Year	Sales (\$000)	Predicted Sales Trend (\$000)
1			
2	1980	52.04	=TREND(B\$2:B\$23,A\$2:A\$23,A2)
3	1981	59.42	=TREND(B\$2:B\$23,A\$2:A\$23,A3)
4	1982	55.66	=TREND(B\$2:B\$23,A\$2:A\$23,A4)
5	1983	53.86	=TREND(B\$2:B\$23,A\$2:A\$23,A5)
6	1984	64.59	=TREND(B\$2:B\$23,A\$2:A\$23,A6)
7	1985	75.28	=TREND(B\$2:B\$23,A\$2:A\$23,A7)
8	1986	61.89	=TREND(B\$2:B\$23,A\$2:A\$23,A8)
9	1987	73.74	=TREND(B\$2:B\$23,A\$2:A\$23,A9)
10	1988	81.19	=TREND(B\$2:B\$23,A\$2:A\$23,A10)
11	1989	97.52	=TREND(B\$2:B\$23,A\$2:A\$23,A11)
12	1990	86.5	=TREND(B\$2:B\$23,A\$2:A\$23,A12)
13	1991	83.18	=TREND(B\$2:B\$23,A\$2:A\$23,A13)
14	1992	87.05	=TREND(B\$2:B\$23,A\$2:A\$23,A14)
15	1993	84.79	=TREND(B\$2:B\$23,A\$2:A\$23,A15)
16	1994	73.49	=TREND(B\$2:B\$23,A\$2:A\$23,A16)
17	1995	76.23	=TREND(B\$2:B\$23,A\$2:A\$23,A17)
18	1996	96.54	=TREND(B\$2:B\$23,A\$2:A\$23,A18)
19	1997	95.08	=TREND(B\$2:B\$23,A\$2:A\$23,A19)
20	1998	87.05	=TREND(B\$2:B\$23,A\$2:A\$23,A20)
21	1999	96.02	=TREND(B\$2:B\$23,A\$2:A\$23,A21)
22	2000	98.9	=TREND(B\$2:B\$23,A\$2:A\$23,A22)
23	2001	83.23	=TREND(B\$2:B\$23,A\$2:A\$23,A23)

Figure 14: Linear Trend Formulae in Excel

The TREND function takes the form =TREND(*known_y's*, *known_x's*, *new_x's*, *const*). The first two arguments represent the known values of your dependent and independent variables. As in LINEST, the *known_y's* argument is a single column, a single row, or a rectangular range. The *known_x's* argument also follows the pattern described for LINEST. The third and fourth arguments are optional. If you omit *new_x's*, the TREND function considers *new_x's* to be identical to *known_x's*. If you include *const*, the value of that argument must be TRUE or FALSE (or 1 or 0). If *const* is TRUE, TREND forces *b* to be 0. In the example provided we have omitted the fourth argument and Excel defaults the value to FALSE, which allows for a constant, *b*, to be calculated.

To extrapolate from existing data, you must supply a range for *new_x's*. You can supply as many or as few cells for *new_x's* as you want. The resultant array will be the same size as the *new_x's* range. To arrive at these values, we added the third argument within the TREND function. The predicted sales values as well as the forecast for 2002 are displayed in Figure 15. For the new *x* value that was used, refer to the value in cell A24, Figure 15. The predicted sales trend number for 2002, 99.95, is found in cell C24, Figure 15.

FORECASTING WITH EXCEL: SUGGESTIONS FOR MANAGERS

◇	A	B	C
1	Year	Sales (\$000)	Predicted Sales Trend (\$000)
2	1980	52.04	58.59
3	1981	59.42	60.47
4	1982	55.66	62.35
5	1983	53.86	64.23
6	1984	64.59	66.11
7	1985	75.28	67.99
8	1986	61.89	69.87
9	1987	73.74	71.75
10	1988	81.19	73.63
11	1989	97.52	75.51
12	1990	86.50	77.39
13	1991	83.18	79.27
14	1992	87.05	81.15
15	1993	84.79	83.03
16	1994	73.49	84.91
17	1995	76.23	86.79
18	1996	96.54	88.67
19	1997	95.08	90.55
20	1998	87.05	92.43
21	1999	96.02	94.31
22	2000	98.90	96.19
23	2001	83.23	98.07
24	2002		99.95

Figure 15: Linear Trend Forecasts in Excel

The trend function in Excel does a good job a providing the predicted sales forecasts. However, it does little to explain the regression itself or inform the user how the regression is performing. The next section details the use of the Linear Regression feature of Excel's Data Analysis Package, to perform linear trend analysis. It provides not only predicted sales forecasts but also provides explanation and inferential statistics regarding the linear trend analysis.

2.3.2 Linear Trend Via Linear Regression Feature of Excel's Data Analysis Package

Linear regression is a method of forecasting that is used by firms operating in a stable environment and which have access to historical data. Regression analysis is used to test the relationship between two or more variables. More specifically, linear regression is designed to estimate the conditional expected value of the variable or interest Y (the dependent variable), given the values of some other variable or variables X (the independent variable). The goal of linear regression is to find the best straight line that will predict X and Y when the values of X and Y are known. One important parameter in simple linear regression is the slope. If the slope is zero, the line is flat. If the slope is negative, the line slopes downward from left to right. Conversely, if the line is positive the line will slope upward from left to right [5]. Simple linear regression analysis is relatively easy to calculate using Microsoft Excel's Data Analysis Package. A detailed tutorial on Excel's regression feature is next. To begin proceed with steps 1 and 2

1. Enter the data into MS Excel, see Figure 6 column A and B.
2. On the main menu bar in MS Excel, click on Tools and Data Analysis (see Figure 3).

Select Regression from the list and click OK (see Figure 16). In turn a screen resembling Figure 17 will appear.

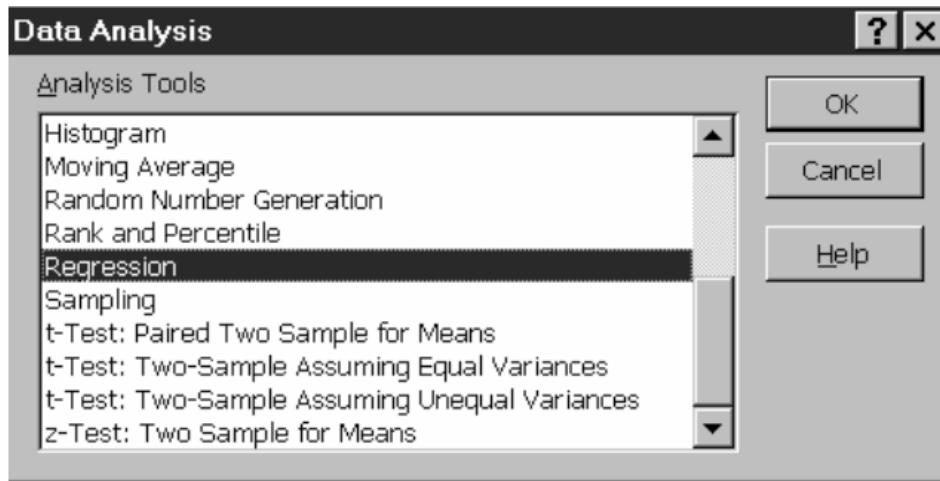


Figure 16: Data Analysis Menu in Excel

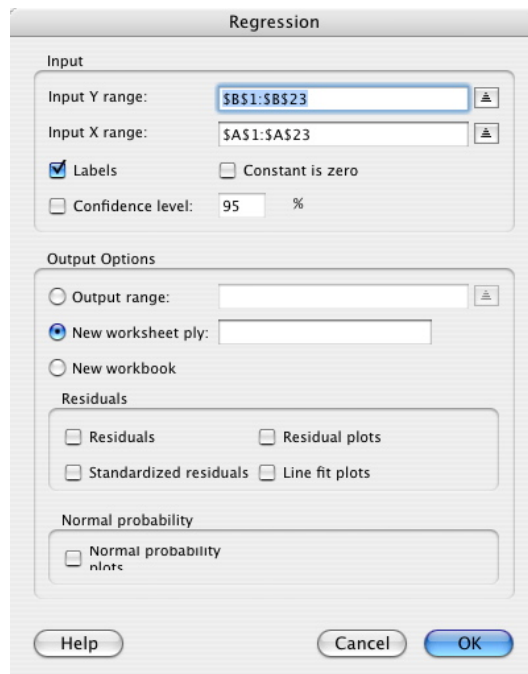


Figure 17: Regression Menu in Excels Data Analysis Package

In Figure 17, it should be noted that the cells B1:B23 were entered in the Input Y range menu and that the cells A1:A23 were entered in the Input X range by highlighting the appropriate cells (the \$ are added by Excel). The Labels check box and New worksheet ply menu were also chosen. Explanations of the other options available in the Regression menu can be found in the Excel Data Analysis Help Menu.

2.3.3 Interpretation of Regression Output from Excel's Linear Regression Feature

When finished entering data and choosing the appropriate options, click OK and the regression result will be generated. The results of this analysis are provided in Figure 18.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	<i>Regression Statistics</i>								
4	Multiple R	0.815129889							
5	R Square	0.664436735							
6	Adjusted R Square	0.647658572							
7	Standard Error	8.891007952							
8	Observations	22							
9									
10	<i>ANOVA</i>								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
12	Regression	1	3130.482047	3130.482047	39.6012797	3.82841E-06			
13	Residual	20	1581.000448	79.0500224					
14	Total	21	4711.482495						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
17	Intercept	-3664.27127	594.7317221	-6.161217124	5.09309E-06	-4904.8599	-2423.682641	-4904.8599	-2423.682641
18	Year	1.880231508	0.298783573	6.292954767	3.82841E-06	1.256979898	2.503483117	1.256979898	2.503483117

Figure 18: Regression Results from Excel's Linear Regression Feature

The regression equation is created using the values found under the *Coefficients* column (see cells B17 and B18 in Figure 18). Thus the regression equation is as follows:

$$\hat{y} \approx -3664.27 + 1.8802x_1 \quad (6)$$

This equation is the regression equation using all the variables. It should be noted that the regression equation given contains rounded values for the intercept and year coefficient terms. Therefore, allowing $x_1 = 2002$, the following prediction for sales in 2002 can be calculated as follows via equations (7) and (8):

$$\hat{y} \approx -3664.27 + 1.8802(2002) \quad (7)$$

$$\hat{y} \approx 99.90 \quad (8)$$

Take note that any difference between the predicted sales value given here, 99.90, and the predicted sales value given in Figure 14, cell C24, is due to rounding. The MAD for the linear regression model is 6.94. The following section interprets the regression statistics.

2.3.4 Interpretation of Statistical Output from Excel's Linear Regression Feature

There are three areas from Figure 18 that must be interpreted in regards to statistical significance. These areas are labeled in Figure 18 as *Regression Statistics* (Rows 3 through 8), ANOVA (Rows 10 through 14), and the section we refer to as the coefficients (Rows 16 through 8). Table 2 presents a detailed analysis each of these areas next.

Table 2: Explanation of Excel's Regression Menu Options

<p><u>Multiple R</u></p> <p>The multiple R represents the strength of the linear relationship between the actual and the estimated values for the dependent variables. The scale ranges from -1.0 to 1.0 where 1.0 indicates a good fit. The multiple R is found to be 81.51% (refer to cell B4, Figure 18), which means there is a strong linear relationship.</p>
<p><u>R square</u></p> <p>R^2 is a symbol for a coefficient of multiple determination between a dependent variable and the independent variables. It tells how much of the variability the dependent variable is explained by the independent variables. R^2 is a goodness-of-fit-measure and the scale ranges from 0.0 to 1.0.</p> <p>This study shows R^2 is equal to 0.6644. This means that 66.44 % (refer to cell B5, Figure 18) of the total variance in the dependent variable around its mean has been accounted for by the independent variables in the estimated regression function.</p>
<p><u>Adjusted R square</u></p> <p>The adjusted R-square, is adjusted to give a truer estimate of how much the independent variables in a regression analysis explains the dependent variable. Taking into account the number of independent variables makes the adjustment and in turn defines adjusted R-square as a measure of the proportion of the variation in the dependent variable accounted for by the explanatory variables. The adjusted R-square is found to be 64.77% (refer to cell B6, Figure 18). In general, the adjusted R-square should not be significantly different in magnitude than R^2 or it may be concluded that major explanatory variables are missing from the model.</p>
<p><u>Standard Error of Estimates</u></p> <p>The standard error of estimates is a regression line. The error is how much the research is off when using the regression line to predict particular scores. The standard error is the standard deviation of those errors from the regression line. The standard error of estimate is thus a measure of the variability of the errors. It measures the average error over the entire scatter plot.</p> <p>The lower the standard error of estimate, the higher the degree of linear relationship between the two variables in the regression. The larger the standard error, the less confidence can be put in the estimate. For this data set, the standard error of estimates equals to 8.8910 (refer to cell B7, Figure 18). This indicates that the Y value falls 8.8910 units away from the regression line.</p>

t-test

The rule of thumb says the t-value should be greater than a t critical value 2. To be more accurate the t-table has been used. Using a two-tailed t test at a 5% level of significance with 21 degrees of freedom, the t-critical value is 2.08. Critical values for t can be found in the back of most statistical textbooks. The critical value is the value that determines the critical regions in a sampling distribution. The critical values separate the obtained values that will and will not result in rejecting the null hypothesis.

Referring back to the regression results (Figure 18, cell D18), it needs to be determined if the coefficients are significantly different from zero at the 5% significance level. From the data, it can be seen that the t stat for the year variable equals 6.2930. Since this value is bigger than our t critical value, 2.080, we then can generally conclude that year is significant in predicting sales. From this analysis, we are able to show that chance variation is not a reasonable explanation for the result for the year variable. This is a classical approach of assessing the statistical significance of findings, which involves comparing empirically observed sample findings with theoretically expected findings. In sum, we can conclude that year does a good job of explaining the sales data.

F-Statistic

A good rule of thumb is for the F-statistic to be greater than 4. From Figure 18, cell E12, it can be seen that the F statistic is 39.6013. This value is clearly above 4. However, it is important to analyze the F statistic with regards to the actual F-distribution. For our example the F distribution shows 1 and 20 degrees of freedom. Looking at the F distribution table at the 5% level of significance, the value 4.3512 can be found. In the results, it can be seen that the critical value of 39.6013 and is greater than the F-statistic of 4.3512. This implies that in this example the intercept and year are doing a good job explaining the variation in sales.

3. Summary

Overall, three forecasting models have been presented: moving averages, exponential smoothing, and linear regression. Each model has advantages and disadvantages. Simply comparing MAD across all three models as seen in Table 3, it can be concluded that the linear regression model tend to predict sales with more accuracy. In addition, the rules of thumb regarding regression performance all indicate the regression model is performing well.

Table 3: Sales Data Set for Rib Sales

Model	MAD
3 Year Moving Average	8.09
5 Year Moving Average	9.90
Exponential Smoothing Alpha = 0.1	8.52
Exponential Smoothing Alpha = 0.3	14.13
Linear Regression	6.94

However, it should be noted that forecasting is an imperfect science and that is as much “art” that accompanies that science. It must also be noted that moving averages

and exponential smoothing models do work well for stationary data (i.e., data that does not contain trends or seasonality) but other methods such as regression must be used when data is not stationary.

This article has attempted to provide readers with an overview of some of the many methods of forecasting available to businesses and demonstrates how each can be implemented using a simple Excel spreadsheet. The authors make important contributions to both the academic literature and to business managers. From an academic perspective this article makes an important contribution to the literature because it demonstrates how easily business forecasting can be taught to either students or business managers. Meanwhile, from a managerial perspective, this article demonstrates that developing obtaining accurate forecasts need not be expensive or overly time consuming.

References

- [1] Bigalow, S. W. and Elliot, D. (2004) *Day Trading With Candlesticks And Moving Averages*. *Futures: News, Analysis & Strategies for Futures*. Options & Derivatives Traders 33 (14): 40-42.
- [2] Bowersox, D. J., Closs, D. J., and Bixby Cooper, M. (2002) *Supply Chain Logistics Management*. New York: McGraw-Hill Irwin.
- [3] D'Attilio, D. F. (1989) *Practical Applications of Trend Analysis in Business Forecasting*. *Journal of the Academy of Marketing Science* 25 (3): 9-11.
- [4] Taylor, J. W. (2004) *Volatility Forecasting With Smooth Transition Exponential Smoothing*. *International Journal of Forecasting*, 20 (2): 273-284.
- [5] Vasilopoulos, A. (2005) *Regression Analysis Revisited*. *Review of Business*, 26 (3): 36-46.