Building a Simple Forecasting Model

In this exercise, we will build a series of forecasting model for sales data and compare them. The data are contained in the Excel file. Weekly sales and the week number of the observation period are provided. Monthly dummy variables have also been created, which we will use to account for seasonality if it necessary. Note that there are dummy variables for February through December. We do not include a dummy variable for January, as doing so would result in multicollinearity. The interpretation of the dummy variables’ coefficients is the difference compared to January sales, holding all other variables constant. Lastly, in column O, we have calculated ln(week number) so that we may consider a non-linear trend.

We will use the regression tool within Excel. If you do not have the Analysis Toolpak enabled, you can enable it by going to File 🡪 Options 🡪 Add-Ins, and clicking “Go” next to Manage: Excel Add-Ins. You should enable both the Analysis Toolpak and Solver Add-In. These will appear on the Data ribbon.

**Model 1: Linear Trend**

Clicking on the Analysis Toolpak, select Regression. For the dependent variable (Y), select the range A1:A106. We will use the remaining rows of data to evaluate the model. For the independent variables, select the range C1:C106. Be sure that the “Labels” box is checked, as the first row of our data contains variable names. Click OK. Rename this worksheet “linear trend”.

To make predictions using the regression results, in cell P2 of the Data sheet, enter the following formula:

='linear trend'!$B$17+('linear trend'!$B$18\*data!C2)

Copy this formula down the worksheet by double-clicking in the lower right corner of the cell.

**Model 2: Log-linear Trend**

As an alternative to the linear trend, we will instead include a log-linear trend. Clicking on the Analysis Toolpak, select Regression. For the dependent variable (Y), select the range A1:A106. We will use the remaining rows of data to evaluate the model. For the independent variables, select the range O1:O106. Be sure that the “Labels” box is checked, as the first row of our data contains variable names. Click OK. Rename this worksheet “log-linear trend”.

To make predictions using the regression results, in cell Q2 of the Data sheet, enter the following formula:

='linear trend'!$B$17+('linear trend'!$B$18\*data!C2)

Copy this formula down the worksheet by double-clicking in the lower right corner of the cell.

**Model 3: Linear Trend with Monthly Dummy Variables**

Clicking on the Analysis Toolpak, select Regression. For the dependent variable (Y), select the range A1:A106. We will use the remaining rows of data to evaluate the model. For the independent variables, select the range C1:N106. Be sure that the “Labels” box is checked, as the first row of our data contains variable names. Click OK. Rename this worksheet “linear with dummies”.

To make predictions using the regression results, in cell R2 of the Data sheet, enter the following formula:

='linear with dummies'!$B$17+('linear with dummies'!$B$18\*data!C2)+('linear with dummies'!$B$19\*data!D2)+('linear with dummies'!$B$20\*data!E2)+('linear with dummies'!$B$21\*data!F2)+('linear with dummies'!$B$22\*data!G2)+('linear with dummies'!$B$23\*data!H2)+('linear with dummies'!$B$24\*data!I2)+('linear with dummies'!$B$25\*data!J2)+('linear with dummies'!$B$26\*data!K2)+('linear with dummies'!$B$27\*data!L2)+('linear with dummies'!$B$28\*data!M2)+('linear with dummies'!$B$29\*data!N2)

Copy this formula down the worksheet by double-clicking in the lower right corner of the cell.

**Model 4: Log-linear Trend with Monthly Dummy Variables**

Clicking on the Analysis Toolpak, select Regression. For the dependent variable (Y), select the range A1:A106. We will use the remaining rows of data to evaluate the model. For the independent variables, select the range D1:O106. Be sure that the “Labels” box is checked, as the first row of our data contains variable names. Click OK. Rename this worksheet “log-linear with dummies”.

To make predictions using the regression results, in cell S2 of the Data sheet, enter the following formula:

='log-linear with dummies'!$B$17+('log-linear with dummies'!$B$29\*data!O2)+('log-linear with dummies'!$B$18\*data!D2)+('log-linear with dummies'!$B$19\*data!E2)+('log-linear with dummies'!$B$20\*data!F2)+('log-linear with dummies'!$B$21\*data!G2)+('log-linear with dummies'!$B$22\*data!H2)+('log-linear with dummies'!$B$23\*data!I2)+('log-linear with dummies'!$B$24\*data!J2)+('log-linear with dummies'!$B$25\*data!K2)+('log-linear with dummies'!$B$26\*data!L2)+('log-linear with dummies'!$B$27\*data!M2)+('log-linear with dummies'!$B$28\*data!N2)

Copy this formula down the worksheet by double-clicking in the lower right corner of the cell.

**Evaluating Performance**

In T2, calculate the absolute difference between the observed outcome and that predicted by Model 1:

=abs(P2-$A2)

To calculate the absolute error for Models 2-4, you can drag this formula across cells U2-W2. Then, copy the formula down the worksheet. To summarize the performance, we will calculate the mean absolute error (MAE) during the forecasting period. In cell T152, calculate the average absolute error during the forecasting period as:

=AVERAGE(T107:T151)

Copy this formula into cells U152-W152.

Based on your analysis, which forecasting model would you recommend going forward?