Using Regression Analysis for Forecasting



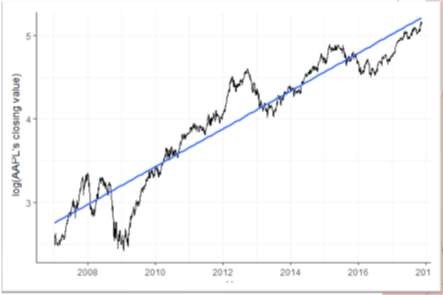
Linear Trend Projection Seasonality without Trend Seasonality with Trend

Using Regression Analysis as a Causal

Forecasting Method

Combining Causal Variables with Trend and Seasonality Effects

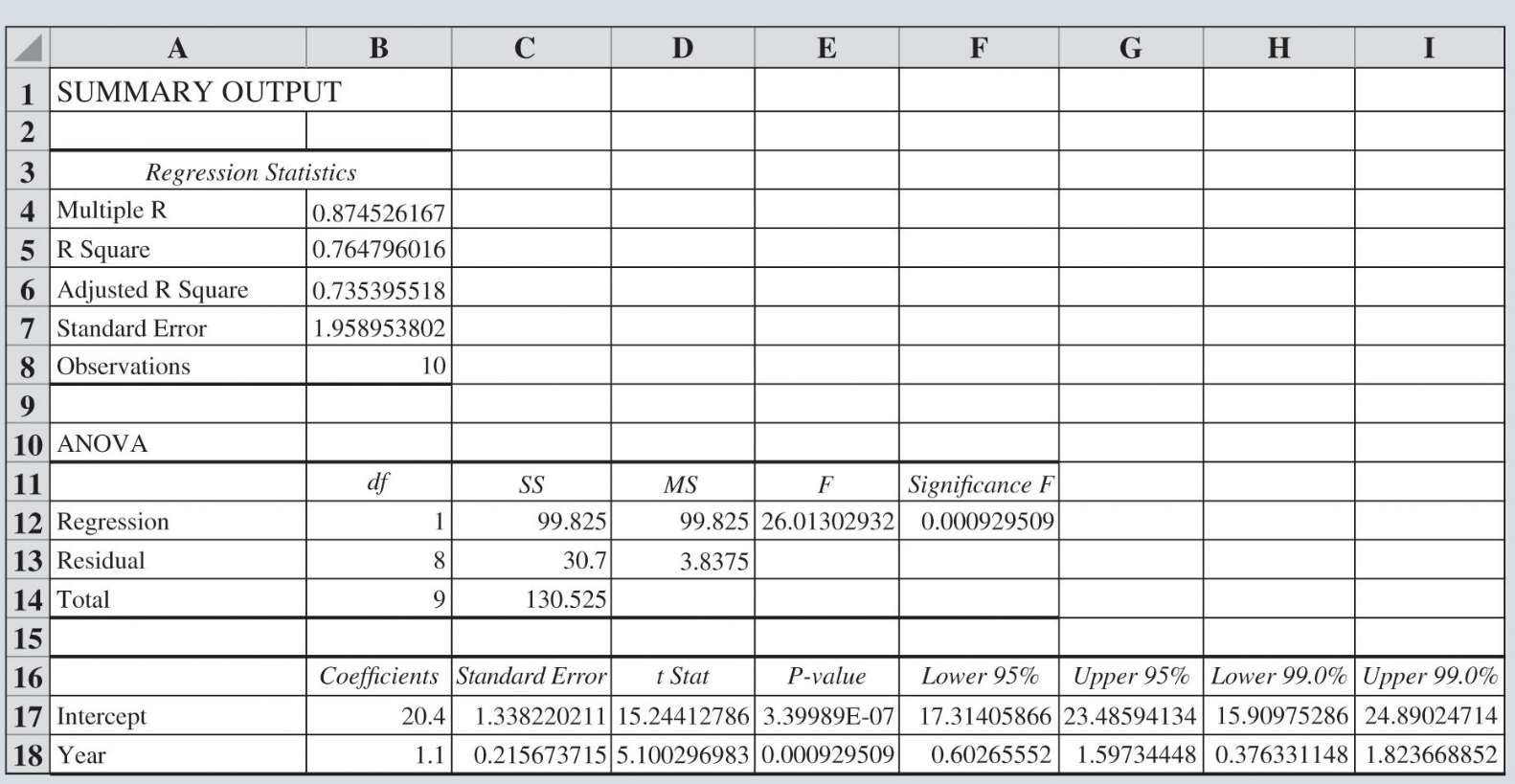
Considerations in Using Regression in Forecasting



Using Regression Analysis for Forecasting

* **Linear Trend Projection:**
  + Regression analysis can be used to forecast a time series with a linear trend.
  + Minimizes the MSE.
  + Dependent Variable: 𝑦t
  + Independent Variable: t (time period)





Using Regression Analysis for Forecasting

**Linear Trend Projection (cont.):**

Equation for the trendline:

Trend equation for the bicycle sales time series:



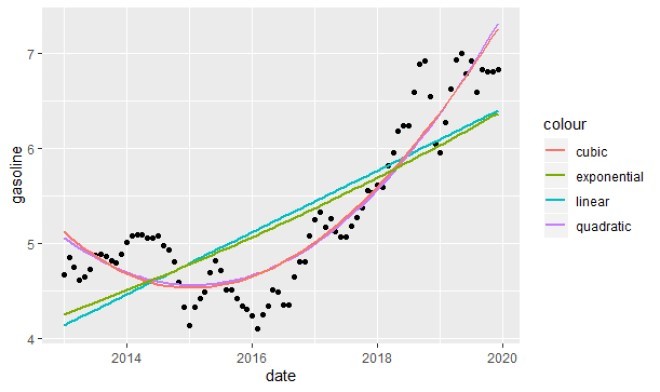
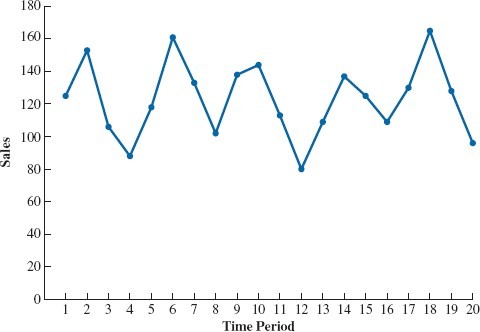
Substituting 𝑡 = 11 into the above equation yields next year’s trend projection,

Thus, the linear trend model yields a sales forecast of 32,500 bicycles for the next year.



Using Regression Analysis for Forecasting

Figure 8.15: Excel Simple Linear Regression Output for Trendline Model for Bicycle Sales Data



Using Regression Analysis for Forecasting

**Linear Trend Projection (cont.):**

We can also use more complex regression models to fit nonlinear trends:

*y*ˆ = *b* + *b t* + *b t*2 + *b t*3

*t*

0 1

2

3

**Autoregressive models**: Regression models in which the independent variables are previous values of the time series.

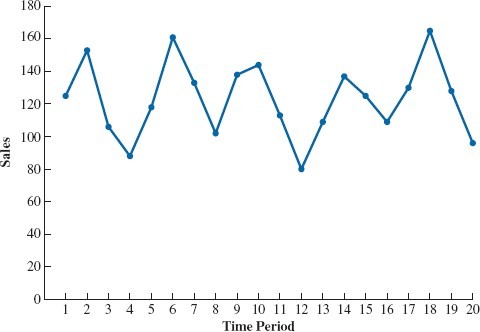
*y*ˆ*t* = *b*0 + *b*1*yt* 1 + *b*2 *yt* 2 + *b*3*yt* 3



Using Regression Analysis for Forecasting

**Seasonality without Trend:**

* We can model a time series with a seasonal pattern by treating the season as a dummy variable.
* **Illustration:**
  + Consider the data on the number of umbrellas sold.
  + The time series plot corresponding to this data does not suggest any long-term trend in sales.



Using Regression Analysis for Forecasting

* **Illustration (cont.):**
  + Closer inspection of the time series plot suggests that a quarterly seasonal pattern is present.
  + k – 1 dummy variables are required to model a categorical variable that has k levels Q1 Q2 Q3 Q4
  + We need 4 – 1 = 3 dummy variables



Using Regression Analysis for Forecasting

**Seasonality without Trend Illustration (cont.):**

* The three dummy variables can be coded as follows:



* General form of the equation relating the number of umbrellas sold to the quarter the sales take place:

Qtr1*t* = 1 if period *t* is quarter 1; 0 otherwise. Qtr2*t* = 1 if period *t* is quarter 2; 0 otherwise. Qtr3*t* = 1 if period *t* is quarter 3; 0 otherwise.



Using Regression Analysis for Forecasting

Table 8.11: Umbrella Sales Time Series with Dummy Variables



Using Regression Analysis for Forecasting

* **Seasonality without Trend Illustration (cont.):**

𝑦^t = 95.0 + 29.0𝑄𝑡𝑟1t + 57.9𝑄𝑡𝑟2 + 26.0𝑄𝑡𝑟3

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Period** | **Year** | **Quarter** | **Qtr1** | **Qtr2** | **Qtr3** | **Sales** |
| 1 | 1 | 1 | 1 | 0 | 0 | 125 |
| 2 |  | 2 | 0 | 1 | 0 | 153 |
| 3 |  | 3 | 0 | 0 | 1 | 106 |
| 4 |  | 4 | 0 | 0 | 0 | 88 |
| 5 | 2 | 1 | 1 | 0 | 0 | 118 |
| 6 |  | 2 | 0 | 1 | 0 | 161 |
| 7 |  | 3 | 0 | 0 | 1 | 133 |
| 8 |  | 4 | 0 | 0 | 0 | 102 |

|  |  |  |  |
| --- | --- | --- | --- |
| Quarter 1: Sales = 95.0 + 29.0 1 | + 57.9 0 | + 26.0 0 | = 124 |
| Quarter 2: Sales = 95.0 + 29.0 0 | + 57.9 1 | + 26.0 0 | = 152 |
| Quarter 3: Sales = 95.0 + 29.0 0 | + 57.9 0 | + 26.0 1 | = 121 |
| Quarter 4: Sales = 95.0 + 29.0 0 | + 57.9 0 | + 26.0 0 | = 95 |



Using Regression Analysis for Forecasting

Quarterly Smart Phones

**Seasonality with Trend:**

* The time series contains both seasonal effects and a linear trend.
* The general form of the regression equation takes the form.

Using Regression Analysis for Forecasting

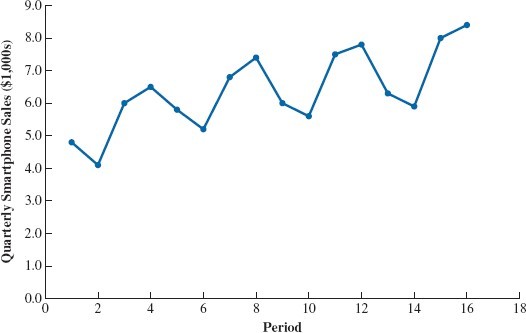


**Seasonality with Trend (cont.):**

* The dummy variables in the equation for Smartphone Sales time series provide four equations given time period *t* corresponds to quarters 1, 2, 3, and 4.

𝑦^t = 6.07 − 1.36𝑄𝑡𝑟1t − 2.03𝑄𝑡𝑟2 − 0.304𝑄𝑡𝑟3 + 0.146𝑡

* + Quarter 1: Sales = 4.71 + 0.146*t.*



* + - *t = 17 – Quarter 1 in year 5 = 7.19 (7,190 sales)*
  + Quarter 2: Sales = 4.04 + 0.146*t.*
    - *t = 18 – Quarter 2 in year 5 = 6.67 (6,670 sales)*
  + Quarter 3: Sales = 5.77 + 0.146*t.*
    - *t = 19 – Quarter 3 in year 5 = 8.54 (8,540 sales)*
  + Quarter 4: Sales = 6.07 + 0.146*t.*
    - *t = 20 – Quarter 4 in year 5 = 8.99 (8,990 sales)*

When t is in Quarter 1, all others are 0.

*When t* is in Quarter 2, all other are 0…



Using Regression Analysis for Forecasting

**Using Regression Analysis as a Causal Forecasting Method:**

* Can we include any other variables besides dummy variables and time?
  + YES!

**Y Dependent Variable**

Sales

New Housing Construction Starting Salaries

Demand for Product Value of Stock

Daily Temperature

**X Independent** Advertising Expenses Mortgage Rate Grade Point Average Price of Product Value of DOW JONES

Electricity Usage

**Causal models**: Models that include only variables that are believed to cause changes in the variable to be forecast.



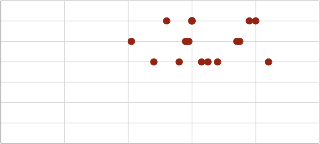
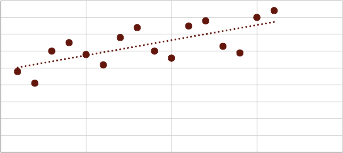
Using Regression Analysis for Forecasting

**Combining Causal Variables with Trend and Seasonality Effects:**

* Regression models are very flexible and can incorporate both causal variables and time series effects.

**Considerations in Using Regression in Forecasting:**

* Can we identify and obtain data for independent variables that are closely related to the time series?
* Focus on the selection of the set of independent variables that provides the best forecasting model.



Determining the Best Forecasting Model to Use

9.0

8.0

Sales over Time

* LOOK AT THE DATA
  + Is there seasonality?
  + Is there a trend?

7.0

6.0

5.0

4.0

3.0

2.0

1.0

0.0

0

5

10

Time Period

15

20

* USE SCATTER CHARTS
  + Determine strong linear or nonlinear relationships between the independent and dependent variables.
  + If certain relationships appear totally random, this may lead you to exclude these variables from the model.

Price

3.50

3.00

2.50

2.00

1.50

1.00

0.50

0.00

0.0 2.0 4.0 6.0 8.0 10.0

Sales ($1000s)



Sales ($1000s)

Price



Determining the Best Forecasting Model to Use

* For Large Data Sets
  + Don’t Forget using Training and Validation Sets
  + Be careful where you divide the two data sets
* Pick the Model that minimizes: MAE, MSE or MAPE.
* There are software packages that will automatically select the best model to use.
* Ultimately, you choose!



Putting It All Together!