Using Regression Analysis for Forecasting

Linear Trend Projection

Seasonality without Trend

Seasonality with Trend

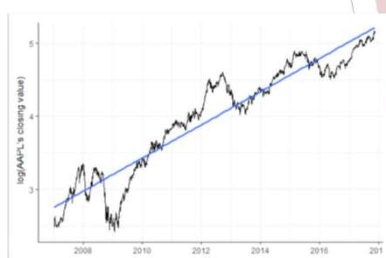
Using Regression Analysis as a Causal Forecasting Method

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: y_t
  - Independent Variable: t (time period)



Using Regression Analysis for Forecasting

Linear Trend Projection (cont.):

Equation for the trendline:

$$\hat{y}_t = b_0 + b_1 t$$

Trend equation for the bicycle sales time series:

$$\hat{y}_t = 20.4 + 1.1t$$

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

$$\hat{y}_{11} = 20.4 + 1.1(11) = 32.5$$

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

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	*Regression Statistics*								
4	Multiple R	0.874526167							
5	R Square	0.764796016							
6	Adjusted R Square	0.735395518							
7	Standard Error	1.958953802							
8	Observations	10							
9									
10	ANOVA								
11		*df*	*SS*	*MS*	*F*	*Significance F*			
12	Regression	1	99.825	99.825	26.01302932	0.000929509			
13	Residual	8	30.7	3.8375					
14	Total	9	130.525						
15									
16		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 99.0%*	*Upper 99.0%*
17	Intercept	20.4	1.338220211	15.24412786	3.39989E-07	17.31405866	23.48594134	15.90975286	24.89024714
18	Year	1.1	0.215673715	5.100296983	0.000929509	0.60265552	1.59734448	0.376331148	1.823668852

Using Regression Analysis for Forecasting

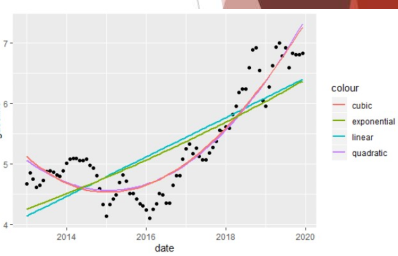
Linear Trend Projection (cont.):

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

$$\hat{y}_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$$

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

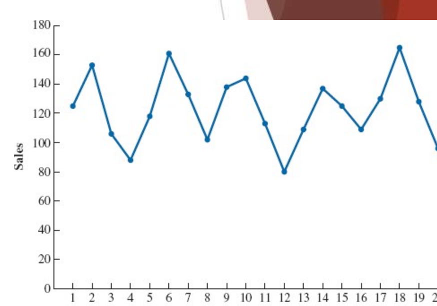
$$\hat{y}_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + b_3 y_{t-3}$$



Using Regression Analysis for Forecasting

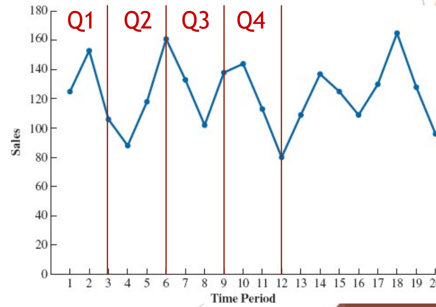
Seasonality without Trend:

- 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
  - 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:

$$\hat{y}_t = b_0 + b_1\text{Qtr1}_t + b_2\text{Qtr2}_t + b_3\text{Qtr3}_t$$

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

$$\text{Qtr1}_t = 1 \text{ if period } t \text{ is quarter 1; 0 otherwise.}$$
$$\text{Qtr2}_t = 1 \text{ if period } t \text{ is quarter 2; 0 otherwise.}$$
$$\text{Qtr3}_t = 1 \text{ if period } t \text{ is quarter 3; 0 otherwise.}$$

Using Regression Analysis for Forecasting

Table 8.11: Umbrella Sales Time Series with Dummy Variables

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

Using Regression Analysis for Forecasting

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

$$\hat{y}_t = 95.0 + 29.0Qtr1_t + 57.9Qtr_2 + 26.0Qtr_3$$

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 = $95.0 + 29.0(0) + 57.9(0) + 26.0(0) = 95$

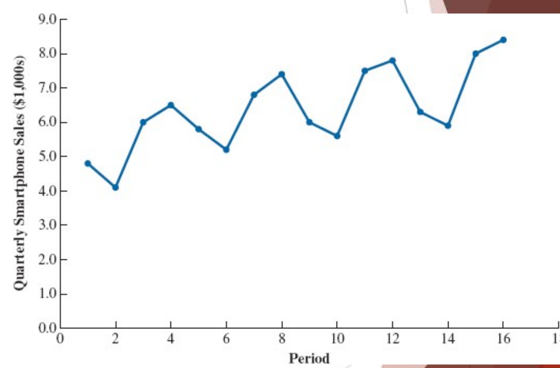
Using Regression Analysis for Forecasting

Seasonality with Trend:

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

$$\hat{y}_t = b_0 + b_1\text{Qtr1}_t + b_2\text{Qtr2}_t + b_3\text{Qtr3}_t + b_4 t$$

Quarterly Smart Phones



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.

$$\hat{y}_t = 6.07 - 1.36Qtr1_t - 2.03Qtr_2 - 0.304Qtr_3 + 0.146t$$

- Quarter 1: Sales = 4.71 + 0.146t.
  - *t = 17 - Quarter 1 in year 5 = 7.19 (7,190 sales)*
- Quarter 2: Sales = 4.04 + 0.146t.
  - *t = 18 - Quarter 2 in year 5 = 6.67 (6,670 sales)*
- Quarter 3: Sales = 5.77 + 0.146t.
  - *t = 19 - Quarter 3 in year 5 = 8.54 (8,540 sales)*
- Quarter 4: Sales = 6.07 + 0.146t.
  - *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	X Independent
Sales	Advertising Expenses
New Housing Construction	Mortgage Rate
Starting Salaries	Grade Point Average
Demand for Product	Price of Product
Value of Stock	Value of DOW JONES
Daily Temperature	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:

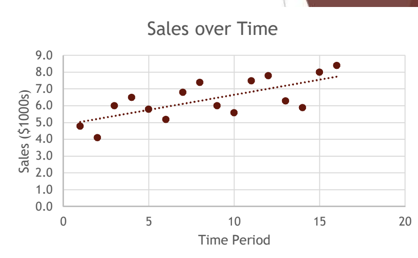
- Regression models are very flexible and can incorporate both causal variables and time series effects.

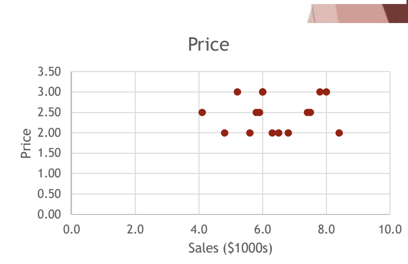
Considerations in Using Regression in Forecasting:

- 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

- LOOK AT THE DATA
  - Is there seasonality?
  - Is there a trend?
- 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.





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!

Sales (1000s)	Year	Quarter	Time Period	Quarter1	Quarter2	Quarter3	Price*Ad	Price	Advertising Expenditures	Model 1 Expected (Seasonal)	Model 2 Expected (Seasonal and Trend)	Model 3 Expected (Seasonal + Trend + Interaction)
4.8	1	1	1					2.00	50	5.725	4.816	4.8257
4.1	1	2	2					2.50	50	5.2	4.332	4.3314
6.0	1	3	3					3.00	50	7.075	6.208	6.2071
6.5	1	4	4					2.00	50	7.525	6.654	6.6528
5.8	2	1	5					2.50	50	5.725	4.77	4.7685
5.2	2	2	6					3.00	50	5.2	5.546	5.5442
6.8	2	3	7					2.00	50	7.075	6.792	6.7899
7.4	2	4	8					2.50	50	7.525	7.238	7.2356
6.0	3	1	9					3.00	50	5.725	5.984	5.9813
5.6	3	2	10					2.00	100	5.2	5.5	5.497
7.5	3	3	11					2.50	100	7.075	7.376	7.3727
7.8	3	4	12					3.00	100	7.525	7.822	7.8184
6.3	4	1	13					2.00	100	5.725	6.568	6.5641
5.9	4	2	14					2.50	100	5.2	6.084	6.0798
8.0	4	3	15					3.00	100	7.075	7.96	7.9555
8.4	4	4	16					2.00	100	7.525	8.406	8.4012

MSE	0.45640625	0.0913385	0.091350065
MAPE	9.418841568	3.154929807	3.165616222

Model 1 Abs % Error	Model 1 Square Error	Model 2 Abs % Error	Model 2 Square Error	Model 2 Abs % Error	Model 2 Square Error
19.27083333	0.855625	0.333333333	0.000256	0.535416667	0.00066049
26.82926829	1.21	5.658536585	0.053824	5.643902439	0.05354596
17.91666667	1.155625	3.466666667	0.043264	3.451666667	0.04289041
15.76923077	1.050625	2.369230769	0.023716	2.350769231	0.02334784
1.293103448	0.005625	17.75862069	1.0609	17.78448276	1.06399225
0	0	6.653846154	0.119716	6.619230769	0.11847364
4.044117647	0.075625	0.117647059	6.4E-05	0.148529412	0.00010201
1.689189189	0.015625	2.189189189	0.026244	2.221621622	0.02702736
4.583333333	0.075625	0.266666667	0.000256	0.311666667	0.00034969
7.142857143	0.16	1.785714286	0.01	1.839285714	0.010609
5.666666667	0.180625	1.653333333	0.015376	1.697333333	0.01620529
3.525641026	0.075625	0.282051282	0.000484	0.235897436	0.00033856
9.126984127	0.330625	4.253968254	0.071824	4.192063492	0.06974881
11.86440678	0.49	3.118644068	0.033856	3.047457627	0.03232804
11.5625	0.855625	0.5	0.0016	0.55625	0.00198025
10.41666667	0.765625	0.071428571	3.6E-05	0.014285714	1.44E-06