Chapter 7

Forecasting

Principles of Forecasting

- Many types of forecasting models
- Each differ in complexity and amount of data
- Forecasts are perfect only by accident
- Forecasts are more accurate for grouped data than for individual items
- Forecast are more accurate for shorter than longer time periods

Forecasting Steps

- Decide what needs to be forecast
  - Level of detail, units of analysis & time horizon required
- Evaluate and analyze appropriate data
  - Identify needed data & whether it's available
- Select and test the forecasting model
  - Cost, ease of use & accuracy
- Generate the forecast
- Monitor forecast accuracy over time

Types of Forecasting Models

- Qualitative methods – judgmental methods
  - Forecasts generated subjectively by the forecaster
  - Educated guesses

- Quantitative methods:
  - Forecasts generated through mathematical modeling

Qualitative Methods

Type	Characteristics	Strengths	Weaknesses
Executive opinion	A group of managers meet & come up with a forecast	Good for strategic or new-product forecasting	One person's opinion can dominate the forecast
Market research	Uses surveys & interviews to identify customer preferences	Good determinant of customer preferences	It can be difficult to develop a good questionnaire
Delphi method	Seeks to develop a consensus among a group of experts	Excellent for forecasting long-term product demand, technological	Time consuming to develop

Quantitative Methods

- **Time Series Models:**
  - Assumes information needed to generate a forecast is contained in a time series of data
  - Assumes the future will follow same patterns as the past

- **Causal Models or Associative Models**
  - Explores cause-and-effect relationships
  - Uses leading indicators to predict the future
  - E.g. housing starts and appliance sales

Causal Models

- Causal models establish a cause-and-effect relationship between dependent variable to be forecast (Y) and independent variables (x_i)
- A common tool of causal modeling is multiple linear regression:

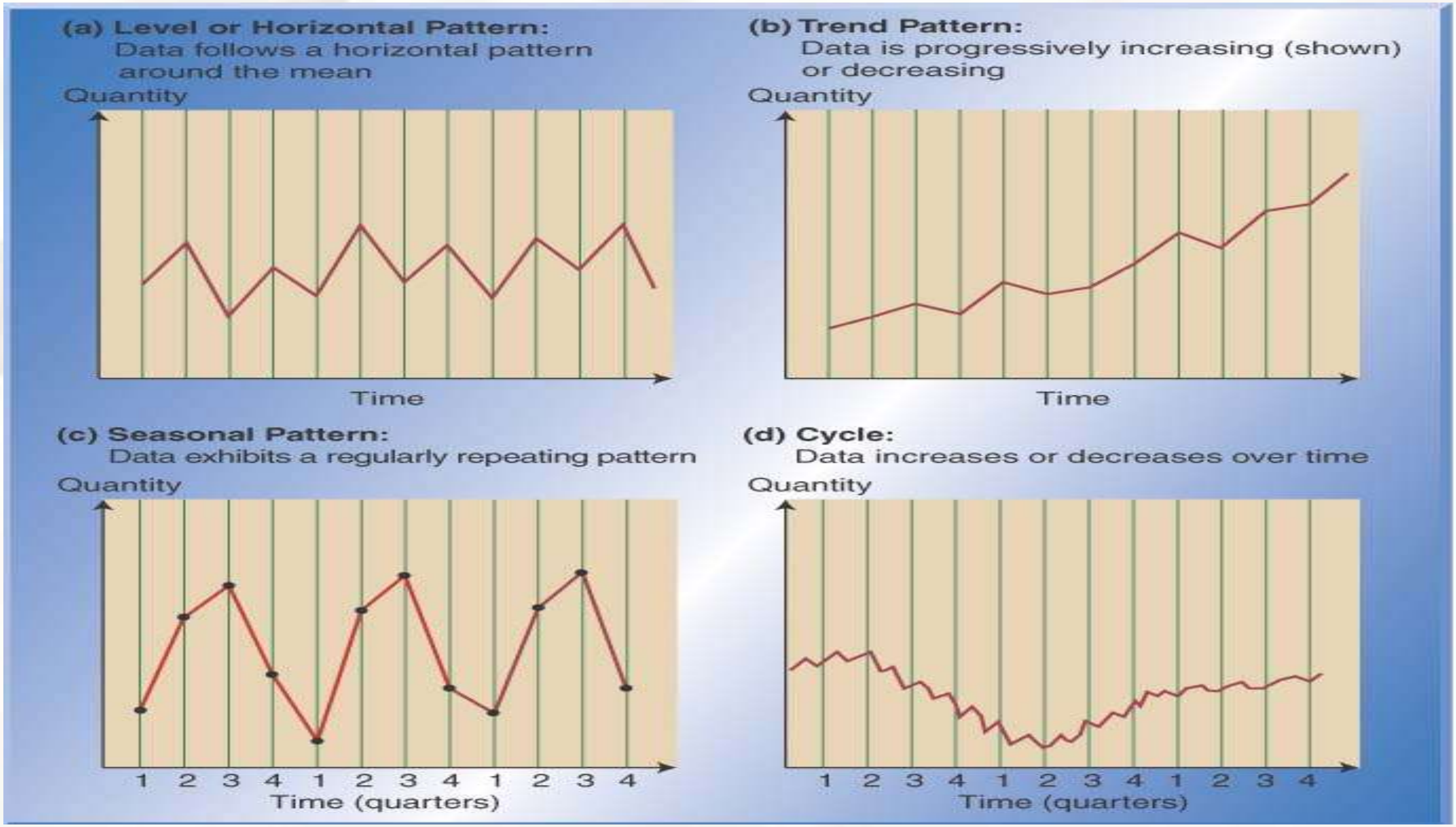
$$Y = a + b_1x_1 + b_2x_2 + \ldots + b_kx_k$$

- Often, leading indicators can be included to help predict changes in future demand e.g. housing starts

Time Series Models

- Forecaster looks for data patterns as
  - Data = historic pattern + random variation
- Historic pattern to be forecasted:
  - Level (long-term average) – data fluctuates around a constant mean
  - Trend – data exhibits an increasing or decreasing pattern
  - Seasonality – any pattern that regularly repeats itself and is of a constant length
  - Cycle – patterns created by economic fluctuations
- Random Variation cannot be predicted

Time Series Patterns



Time Series Models

- Naive: $F_{t+1} = A_t$
  - The forecast is equal to the actual value observed during the last period – good for level patterns
- Simple Mean: $F_{t+1} = \sum A_t / n$
  - The average of all available data - good for level patterns
- Moving Average: $F_{t+1} = \sum A_t / n$
  - The average value over a set time period (e.g.: the last four weeks)
  - Each new forecast drops the oldest data point & adds a new observation
  - More responsive to a trend but still lags behind actual data

Time Series Problem Solution

Period	Actual	2-Period	4-Period
1	300		
2	315		
3	290		
4	345		
5	320		
6	360		
7	375	**340.0**	**328.8**
8		**367.5**	**350.0**

Time Series Models (continued)

- **Weighted Moving Average:** $$F_{t+1} = \sum C_t A_t$$

- All weights must add to 100% or 1.00
  e.g. C_t **.5**, C_{t-1} **.3**, C_{t-2} **.2** (weights add to 1.0)

- Allows emphasizing one period over others; above indicates more weight on recent data (C_t=.5)

- Differs from the simple moving average that weighs all periods equally - more responsive to trends

Time Series Models (continued)

- **Exponential Smoothing:**

$$F_{t+1} = \alpha A_t + (1-\alpha)F_t$$

Most frequently used time series method because of ease of use and minimal amount of data needed

- Need just three pieces α of data to start:
  - Last period's forecast (F_t)
  - Last periods actual value (A_t)
  - Select value of smoothing coefficient, α, between 0 and 1.0
- If no last period forecast is available, average the last few periods or use naive method
- Higher α values (e.g. .7 or .8) may place too much weight on last period's random variation

Forecasting Trends

- Basic forecasting models for trends compensate for the lagging that would otherwise occur
- One model, **trend-adjusted exponential smoothing** uses a three step process
  - Step 1 - Smoothing the level of the series

$$S_t = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1})$$

  - Step 2 – Smoothing the trend

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1}$$

  - Forecast including the trend

$$FIT_{t+1} = S_t + T_t$$

Forecasting trend problem: a company uses exponential smoothing with trend to forecast usage of its lawn care products. At the end of July the company wishes to forecast sales for August. July demand was 62. The trend through June has been 15 additional gallons of product sold per month. Average sales have been 57 gallons per month. The company uses alpha+0.2 and beta +0.10. Forecast for August.

- **Smooth the level of the series:**

$$S_{July} = \alpha A_t + (1-\alpha)(S_{t-1} + T_{t-1}) = (0.2)(62) + (0.8)(57 + 15) = 70$$

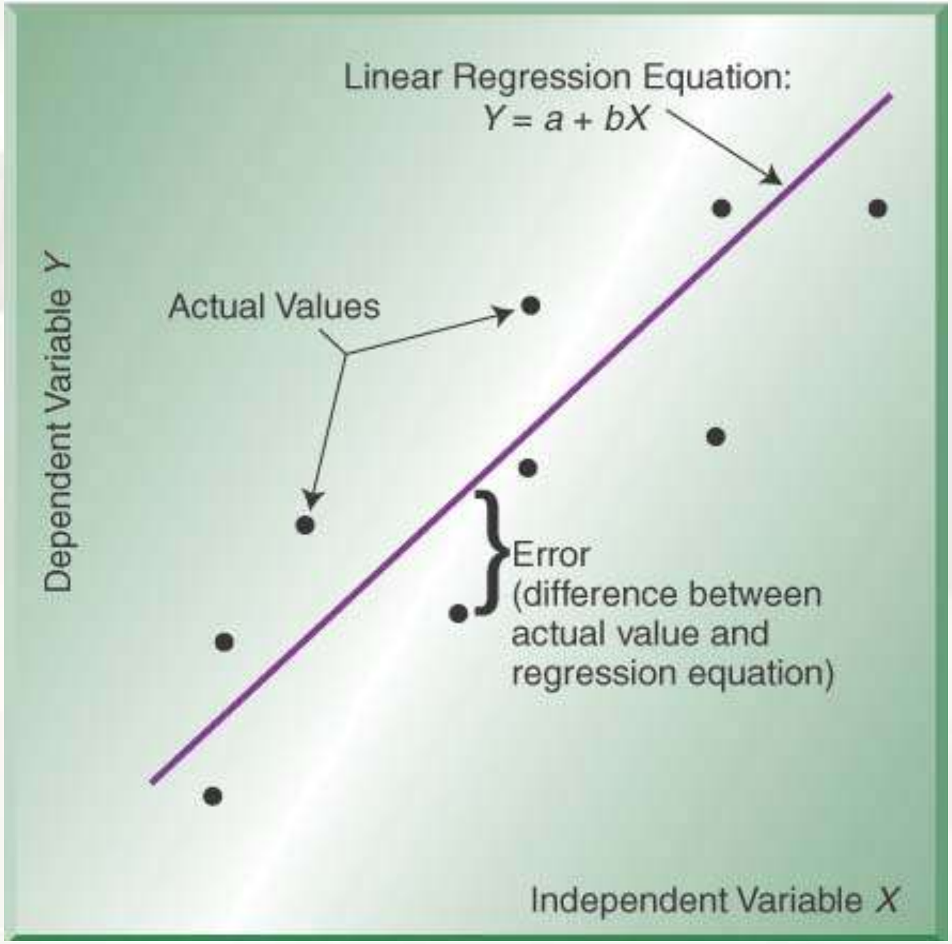
- **Smooth the trend:**

$$T_{July} = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} = (0.1)(70 - 57) + (0.9)(15) = 14.8$$

- **Forecast including trend:**

$$FIT_{August} = S_t + T_t = 70 + 14.8 = 84.8 \text{ gallons}$$

Linear Regression



- Identify dependent **(y)** and independent **(x)** variables
- Solve for the slope of the line

$$\mathbf{b} = \frac{\sum \mathbf{XY} - \mathbf{n}\overline{\mathbf{X}}\,\overline{\mathbf{Y}}}{\sum \mathbf{X}^2 - \mathbf{n}\overline{\mathbf{X}}^2}$$

- Solve for the y intercept

$$\mathbf{a} = \overline{\mathbf{Y}} - \mathbf{b}\overline{\mathbf{X}}$$

- Develop your equation for the trend line

Y=a + bX

Linear Regression Problem: A maker of golf shirts has been tracking the relationship between sales and advertising dollars. Use linear regression to find out what sales might be if the company invested $53,000 in advertising next year.

	Sales $ (Y)	Adv.$ (X)	XY	X^2	Y^2
1	130	32	4160	2304	16,900
2	151	52	7852	2704	22,801
3	150	50	7500	2500	22,500
4	158	55	8690	3025	24964
5	**153.85**	53			
Tot	589	189	28202	9253	87165
Avg	**147.25**	**47.25**			

$$b = \frac{\sum XY - n\overline{X}\overline{Y}}{\sum X^2 - n\overline{X}^2}$$

$$b = \frac{28202 - 4(47.25)(147.25)}{9253 - 4(47.25)^2} = 1.15$$

$$a = \overline{Y} - b\overline{X} = 147.25 - 1.15(47.25)$$

$$a = 92.9$$

$$Y = a + bX = 92.9 + 1.15X$$

$$Y = 92.9 + 1.15(53) = 153.85$$

How Good is the Fit? – Correlation Coefficient

- Correlation coefficient **(r)** measures the direction and strength of the linear relationship between two variables. The closer the r value is to 1.0 the better the regression line fits the data points.

$$\mathbf{r} = \frac{n\left(\sum XY\right) - \left(\sum X\right)\left(\sum Y\right)}{\sqrt{n\left(\sum X^2\right) - \left(\sum X\right)^2} * \sqrt{n\left(\sum Y^2\right) - \left(Y\right)^2}}$$

$$\mathbf{r} = \frac{4(28{,}202) - 189(589)}{\sqrt{4(9253) - (189)^2} * \sqrt{4(87{,}165) - (589)^2}} = .982$$

$$\mathbf{r}^2 = (.982)^2 = .964$$

- Coefficient of determination (r^2) measures the amount of variation in the dependent variable about its mean that is explained by the regression line. Values of (r^2) close to 1.0 are desirable.

Measuring Forecast Error

- Forecasts are never perfect
- Need to know how much we should rely on our chosen forecasting method
- Measuring forecast error:

$$\mathbf{E_t = A_t - F_t}$$

- Note that over-forecasts = negative errors and under-forecasts = positive errors

Measuring Forecasting Accuracy

- Mean Absolute Deviation (MAD)
  - measures the total error in a forecast without regard to sign
- Cumulative Forecast Error (CFE)
  - Measures any bias in the forecast
- Mean Square Error (MSE)
  - Penalizes larger errors
- Tracking Signal
  - Measures if your model is working

$$\mathbf{MAD} = \frac{\sum \left|\mathbf{actual} - \mathbf{forecast}\right|}{\mathbf{n}}$$

$$\mathbf{CFE} = \sum \left(\mathbf{actual} - \mathbf{forecast}\right)$$

$$\mathbf{MSE} = \frac{\sum \left(\mathbf{actual\ \text{-}\ forecast}\right)^2}{\mathbf{n}}$$

$$\mathbf{TS} = \frac{\mathbf{CFE}}{\mathbf{MAD}}$$

Selecting the Right Forecasting Model

- The amount & type of available data
  - Some methods require more data than others
- Degree of accuracy required
  - Increasing accuracy means more data
- Length of forecast horizon
  - Different models for 3 month vs. 10 years
- Presence of data patterns
  - Lagging will occur when a forecasting model meant for a level pattern is applied with a trend

Chapter 8 Highlights

- Three basic principles of forecasting are: forecasts are rarely perfect, are more accurate for groups than individual items, and are more accurate in the shorter term than longer time horizons.
- The forecasting process involves five steps: decide what to forecast, evaluate and analyze appropriate data, select and test model, generate forecast, and monitor accuracy.
- Forecasting methods can be classified into two groups: qualitative and quantitative. Qualitative methods are based on the subjective opinion of the forecaster and quantitative methods are based on mathematical modeling.
- Time series models are based on the assumption that all information needed is contained in the time series of data. Causal models assume that the variable being forecast is related to other variables in the environment.

Highlights (continued)

- There are four basic patterns of data: level or horizontal, trend, seasonality, and cycles. In addition, data usually contain random variation. Some forecast models used to forecast the level of a time series are: naïve, simple mean, simple moving average, weighted moving average, and exponential smoothing. Separate models are used to forecast trends and seasonality.
- A simple causal model is linear regression in which a straight-line relationship is modeled between the variable we are forecasting and another variable in the environment. The correlation is used to measure the strength of the linear relationship between these two variables.
- Three useful measures of forecast error are mean absolute deviation (MAD), mean square error (MSE) and tracking signal.
- There are four factors to consider when selecting a model: amount and type of data available, degree of accuracy required, length of forecast horizon, and patterns present in the data.