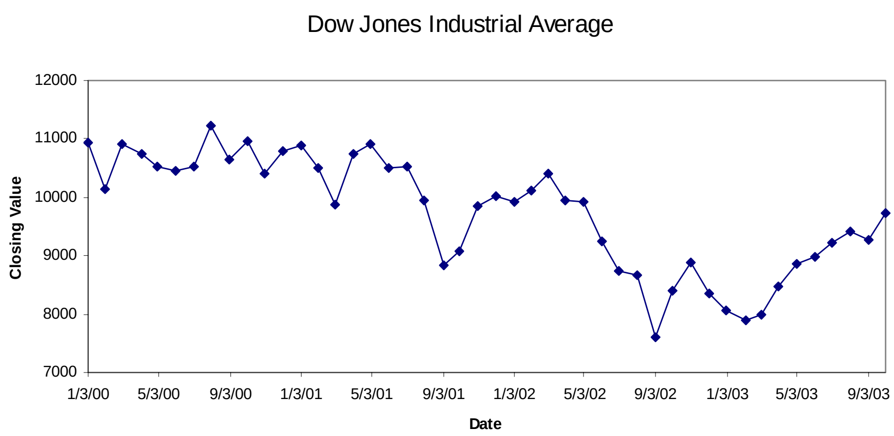
Time Series and Forecasting

Time Series

- A time series is a sequence of measurements over time, usually obtained at equally spaced intervals
  - Daily
  - Monthly
  - Quarterly
  - Yearly

Time Series Example



Components of a Time Series

- Secular Trend
  - Linear
  - Nonlinear
- Cyclical Variation
  - Rises and Falls over periods longer than one year
- Seasonal Variation
  - Patterns of change within a year, typically repeating themselves
- Residual Variation

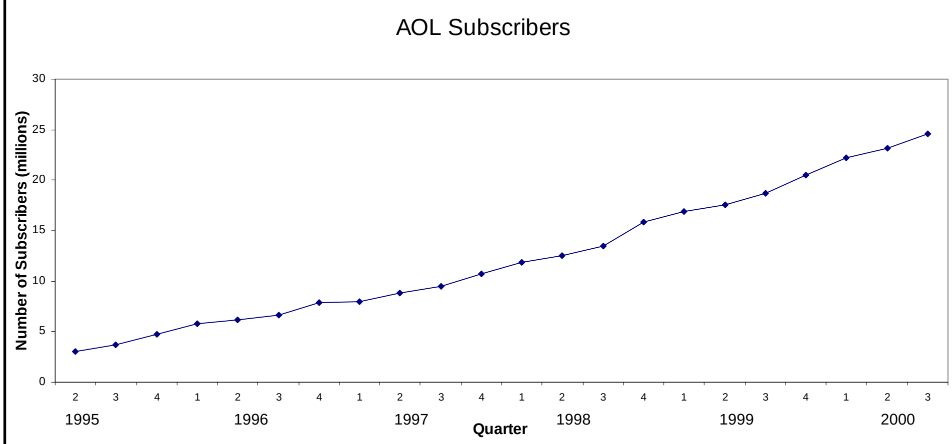
Components of a Time Series

$$Y_t = T_t + C_t + S_t + R_t$$

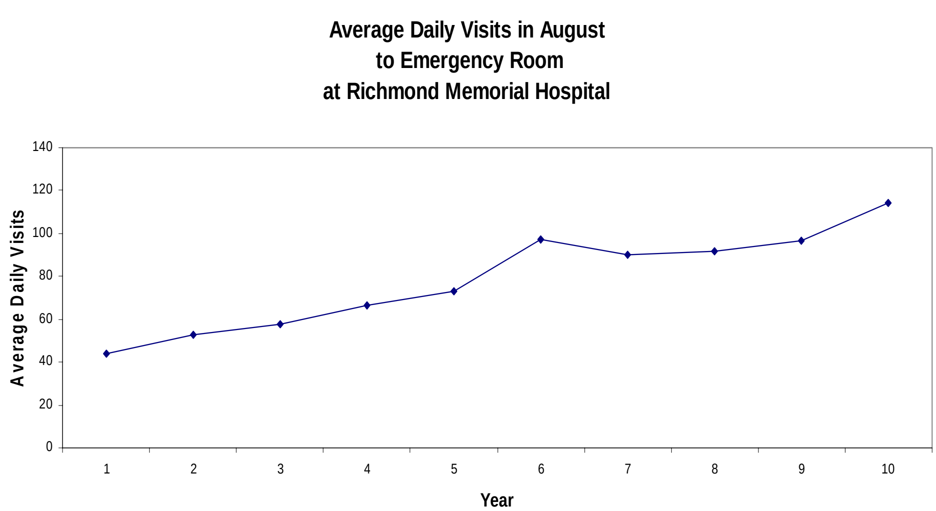
Time Series with Linear Trend

$$Y_t = a + b\,t + e_t$$

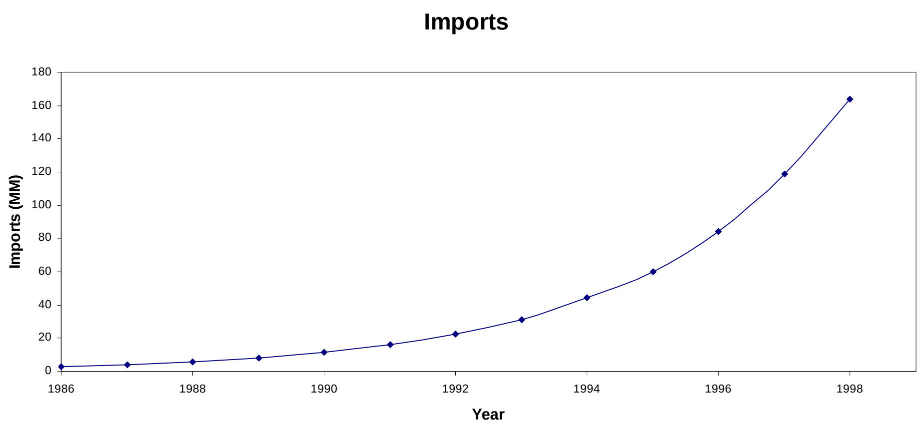
Time Series with Linear Trend



Time Series with Linear Trend



Time Series with Nonlinear Trend



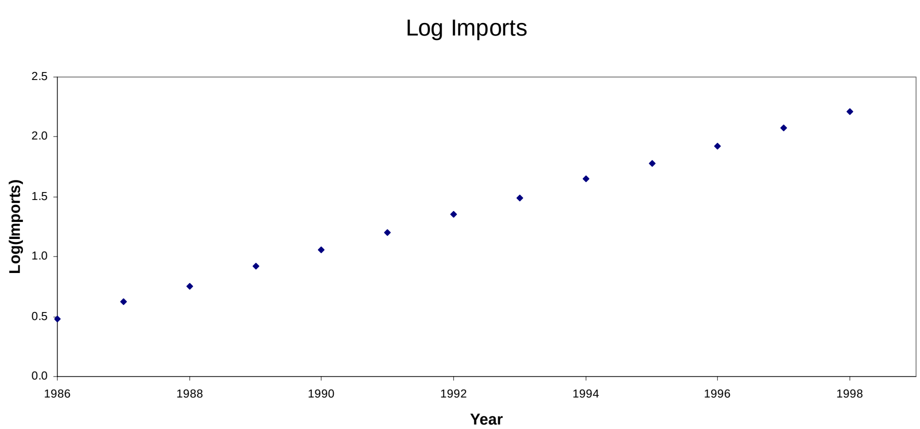
Time Series with Nonlinear Trend

- Data that increase by a constant amount at each successive time period show a linear trend.
- Data that increase by increasing amounts at each successive time period show a curvilinear trend.
- Data that increase by an equal percentage at each successive time period can be made linear by applying a logarithmic transformation.

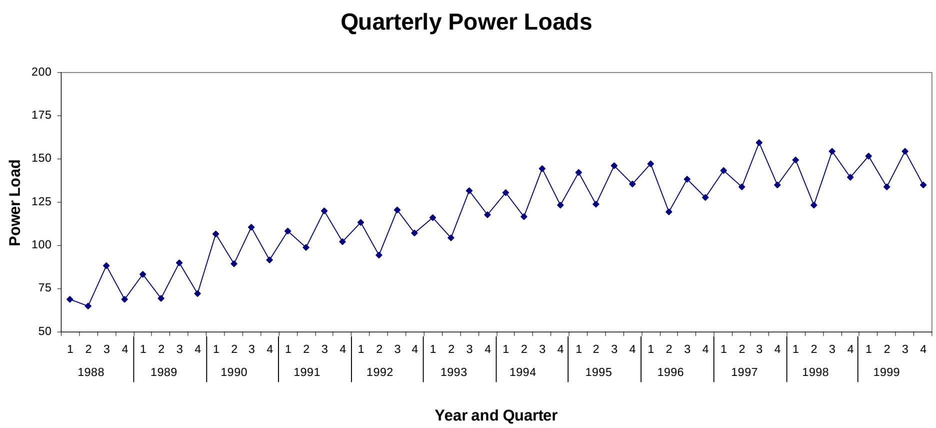
Nonlinear Time Series transformed to a Linear Time Series with a Logarithmic Transformation

$$\log(Y_t) = a + b\,t + e_t$$

Transformed Time Series



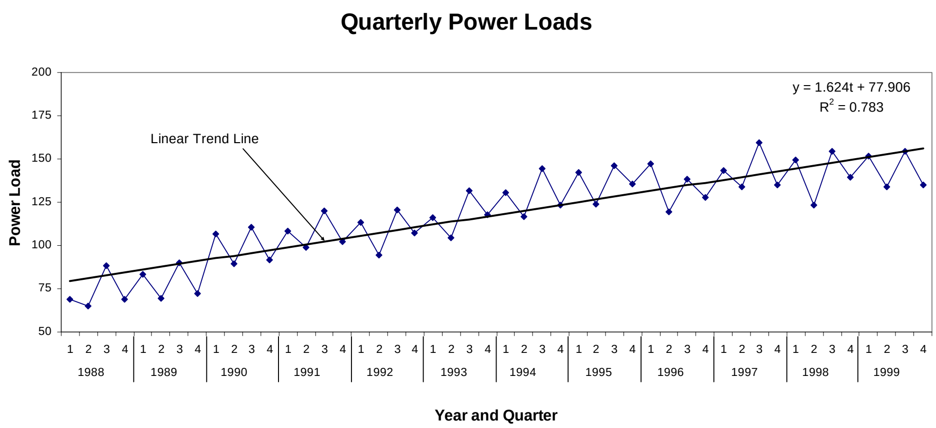
Time Series with both Trend and Seasonal Pattern



Model Building

- For the Power Load data
  – What kind of trend are we seeing?
    - Linear
    - Logarithmic
    - Polynomial
  – How can we smooth the data?
  – How do we model the distinct seasonal pattern?

Power Load Data with Linear Trend



Modeling a Nonlinear Trend

- If the time series appears to be changing at an increasing rate over time, a logarithmic model in Y may work:

$$\ln(Y_t) = a + b\,t + e_t$$

or

$$Y_t = \exp\{a + b\,t + e_t\}$$

- In Excel, this is called an exponential model

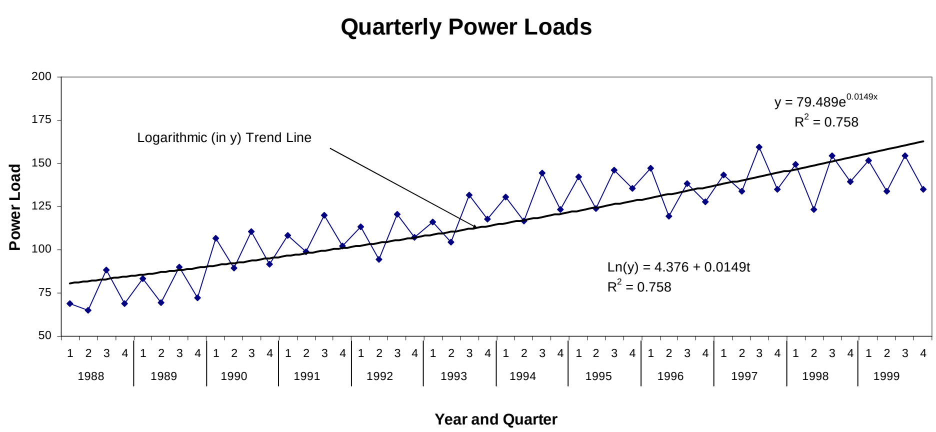
Modeling a Nonlinear Trend

- If the time series appears to be changing at a decreasing rate over time, a logarithmic model in t may work:

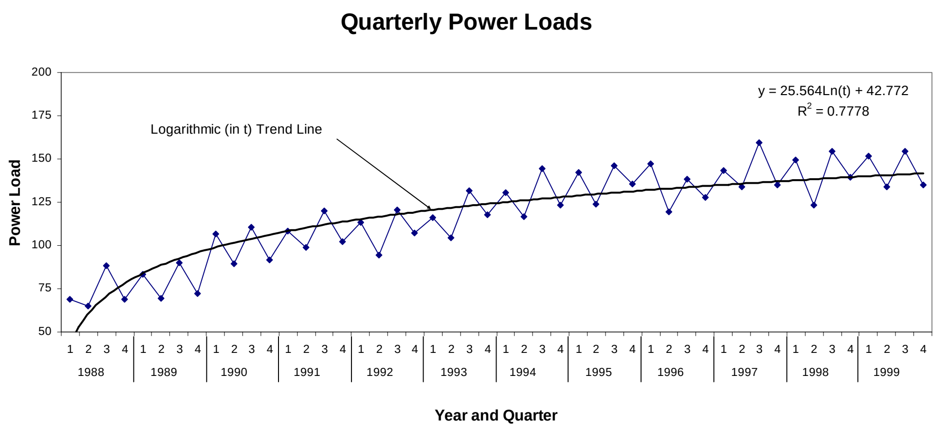
$$Y_t = a + b\ \ln(t) + e_t$$

- In Excel, this is called a logarithmic model

Power Load Data with Exponential Trend



Power Load Data with Logarithmic Trend



Modeling a Nonlinear Trend

- General curvilinear trends can often be model with a polynomial:
  - Linear (first order)

    $$Y_t = a + b\,t + e_t$$

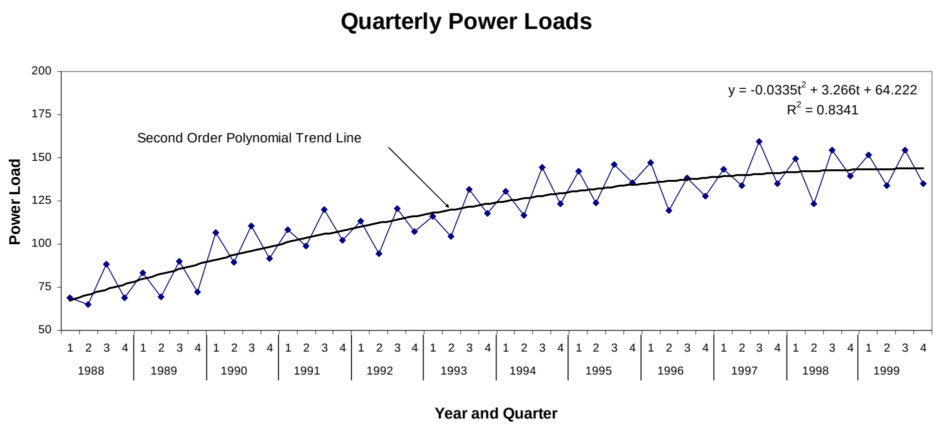
  - Quadratic (second order)

    $$Y_t = a + b_1\,t + b_2\,t^2 + e_t$$

  - Cubic (third order)

    $$Y_t = a + b_1\,t + b_2\,t^2 + b_3\,t^3 + e_t$$

Power Load Data modeled with Second Degree Polynomial Trend



Moving Average

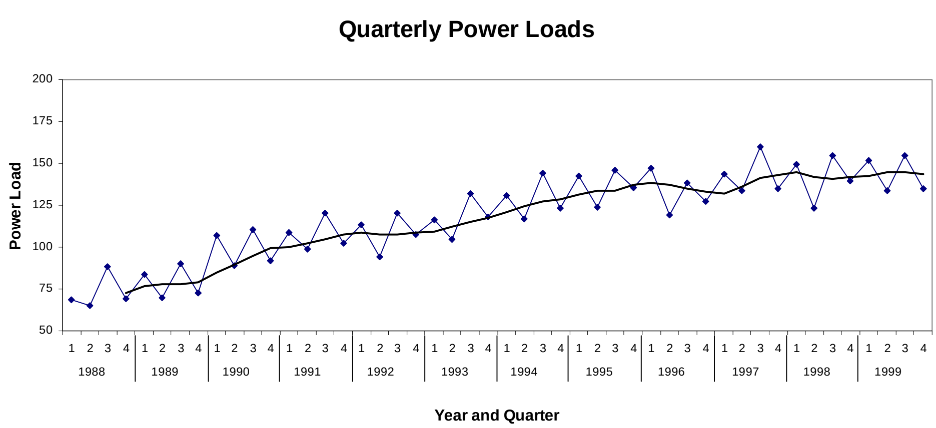
- Another way to examine trends in time series is to compute an average of the last m consecutive observations
- A 4-point moving average would be:

$$\bar{y}_{MA(4)} = \frac{(y_t + y_{t-1} + y_{t-2} + y_{t-3})}{4}$$

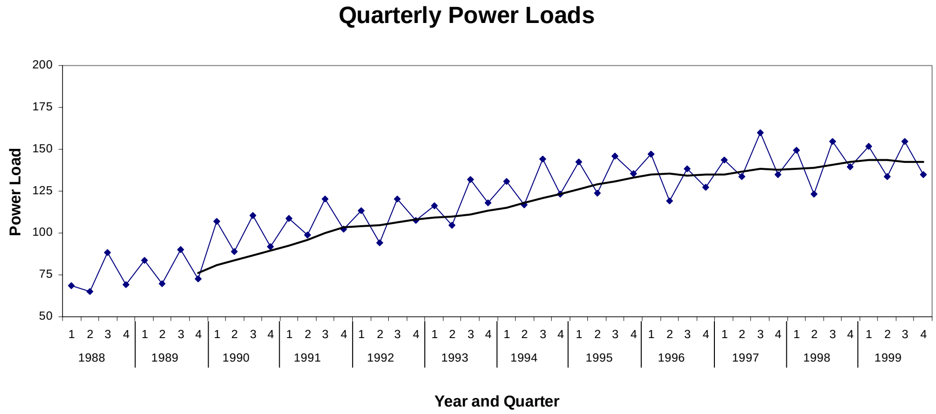
Moving Average

- In contrast to modeling in terms of a mathematical equation, the moving average merely smooths the fluctuations in the data.
- A moving average works well when the data have
  - a fairly linear trend
  - a definite rhythmic pattern of fluctuations

Power Load Data with 4-point Moving Average



Power Load Data with 8-point Moving Average



Exponential Smoothing

- An exponential moving average is a weighted average that assigns positive weights to the current value and to past values of the time series.
- It gives greater weight to more recent values, and the weights decrease exponentially as the series goes farther back in time.

Exponentially Weighted Moving Average

$$S_1 = Y_1$$

$$S_t = wY_t + (1-w)S_{t-1}$$
$$= wY_t + w(1-w)Y_{t-1} + w(1-w)^2 Y_{t-2} + \cdots$$

Exponentially Weighted Moving Average

Let w=0.5

$$S_1 = Y_1$$
$$S_2 = 0.5Y_2 + (1-0.5)S_1 = 0.5Y_2 + 0.5Y_1$$
$$S_3 = 0.5Y_3 + (1-0.5)S_2 = 0.5Y_3 + 0.25Y_2 + 0.25Y_1$$
$$S_4 = 0.5Y_4 + (1-0.5)S_3 = 0.5Y_4 + 0.25Y_3 + 0.125Y_2 + 0.125Y_1$$

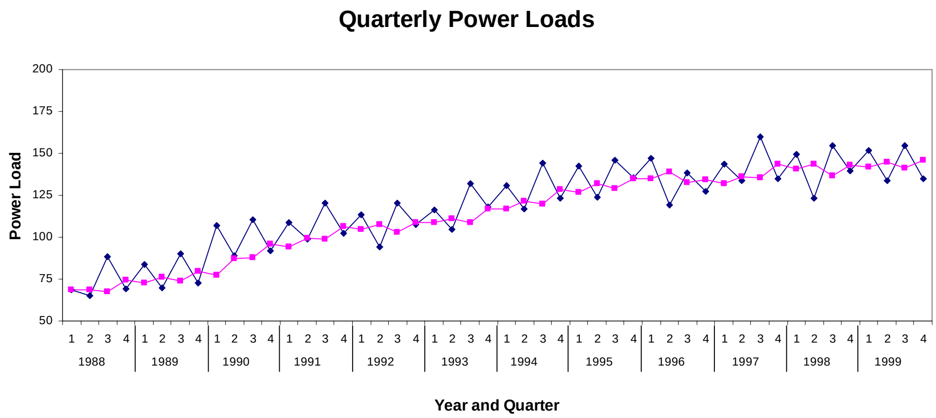
Exponential Weights

w	w*(1-w)	w*(1-w)^2	w*(1-w)^3	w*(1-w)^4
0.01	0.0099	0.0098	0.0097	0.0096
0.05	0.0475	0.0451	0.0429	0.0407
0.1	0.0900	0.0810	0.0729	0.0656
0.2	0.1600	0.1280	0.1024	0.0819
0.3	0.2100	0.1470	0.1029	0.0720
0.4	0.2400	0.1440	0.0864	0.0518
0.5	0.2500	0.1250	0.0625	0.0313
0.6	0.2400	0.0960	0.0384	0.0154
0.7	0.2100	0.0630	0.0189	0.0057
0.8	0.1600	0.0320	0.0064	0.0013
0.9	0.0900	0.0090	0.0009	0.0001
0.95	0.0475	0.0024	0.0001	0.0000
0.99	0.0099	0.0001	0.0000	0.0000

Exponential Smoothing

- The choice of w affects the smoothness of E_t.
  - The smaller the value of w, the smoother the plot of E_t.
  - Choosing w close to 1 yields a series much like the original series.

Power Load Data with Exponentially Weighted Moving Average (w=.34)



Forecasting with Exponential Smoothing

- The predicted value of the next observation is the exponentially weighted average corresponding to the current observation.

$$\hat{y}_{t+1} = S_t$$

Assessing the Accuracy of the Forecast

- Accuracy is typically assessed using either the Mean Squared Error or the Mean Absolute Deviation

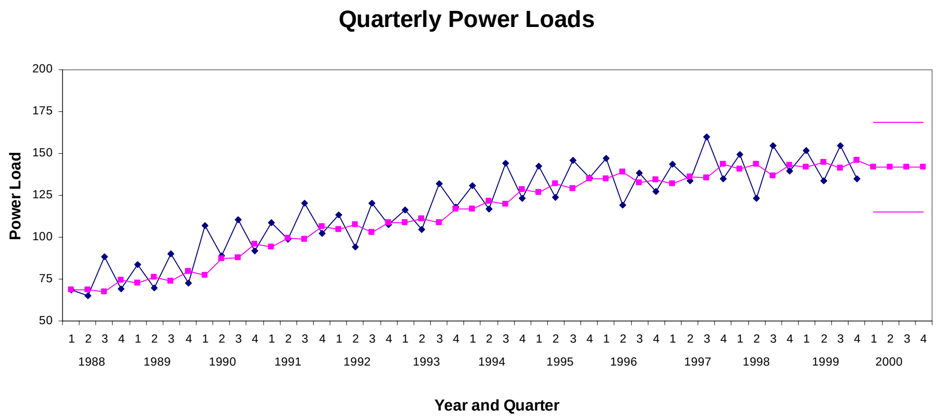
$$\text{MSE} = \frac{\sum_{t=1}^{n} \left(y_t - \hat{y}_t\right)^2}{n}$$

$$\text{MAD} = \frac{\sum_{t=1}^{n} \left|y_t - \hat{y}_t\right|}{n}$$

Assessing the Accuracy of the Forecast

- It is usually desirable to choose the weight w to minimize MSE or MAD.
- For the Power Load Data, the choice of w = .34 was based upon the minimization of MSE.

Power Load Data with Forecast for 2000 using Exponentially Weighted Moving Average (w=.34)



Exponential Smoothing

- One Parameter (w)
  - Forecasted Values beyond the range of the data, into the future, remain the same.
- Two Parameter
  - Adds a parameter (v) that accounts for trend in the data.

2 Parameter Exponential Smoothing

$$S_t = wY_t + (1-w)\left(S_{t-1} + T_{t-1}\right)$$
$$T_t = v\left(S_t - S_{t-1}\right) + (1-v)T_{t-1}$$

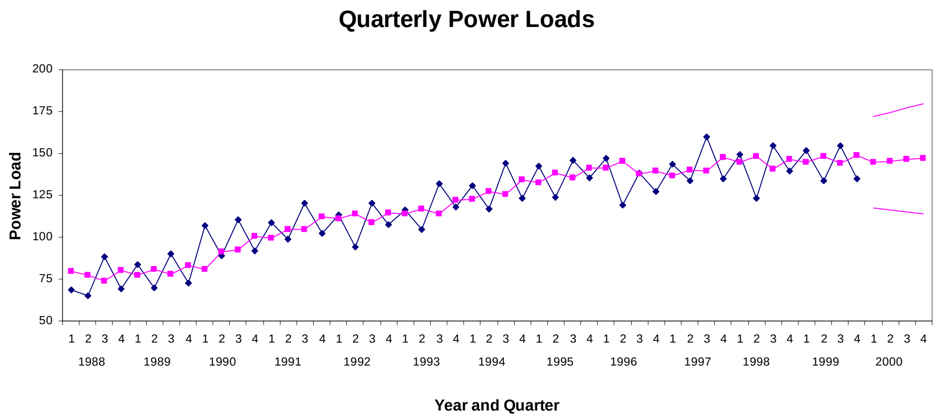
The forecasted value of y is

$$\hat{y}_{t+1} = S_t + T_t$$

2 Parameter Exponential Smoothing

- The value of S_t is a weighted average of the current observation and the previous forecast value.
- The value of T_t is a weighted average of the change in S_t and the previous estimate of the trend parameter.

Power Load Data with Forecast for 2000 using 2 Parameter Exponential Smoothing with w=.34 and v=.08



Exponential Smoothing

- One Parameter (w)
  - Determines location.
- Two Parameter
  - Adds a parameter (v) that accounts for trend in the data.
- Three Parameter
  - Adds a parameter (c) that accounts for seasonality.

3 Parameter Exponential Smoothing

$$S_t = w\frac{Y_t}{I_{t-p}} + (1-w)(S_{t-1} + T_{t-1})$$

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

$$I_p = c\frac{Y_t}{S_n} + (1-c)I_{t-p}$$

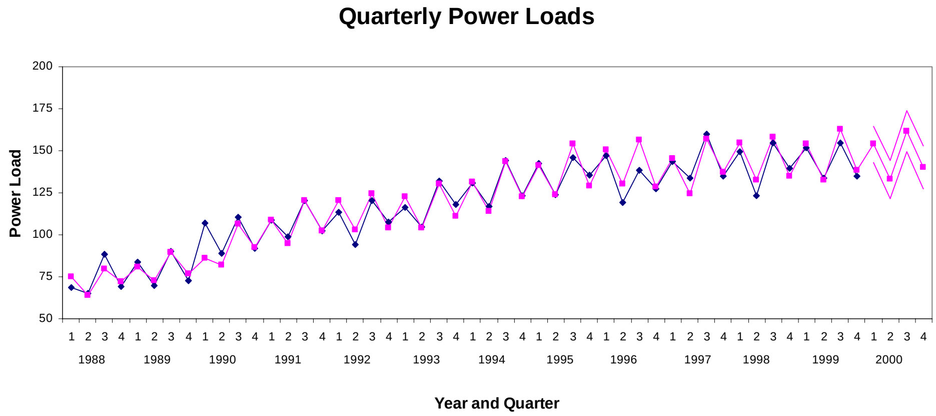
The forecasted value of y is

$$\hat{y}_{t+1} = S_t + T_t + I_t$$

3 Parameter Exponential Smoothing

- The value of I_t represents a seasonal index at point p in the season.

Power Load Data with Forecast for 2000 using 3 Parameter Exponential Smoothing with w=.34, v=.08, c=.15



Modeling Seasonality

- Seasonality can also be modeled using dummy (or indicator) variables in a regression model.

Modeling Seasonality

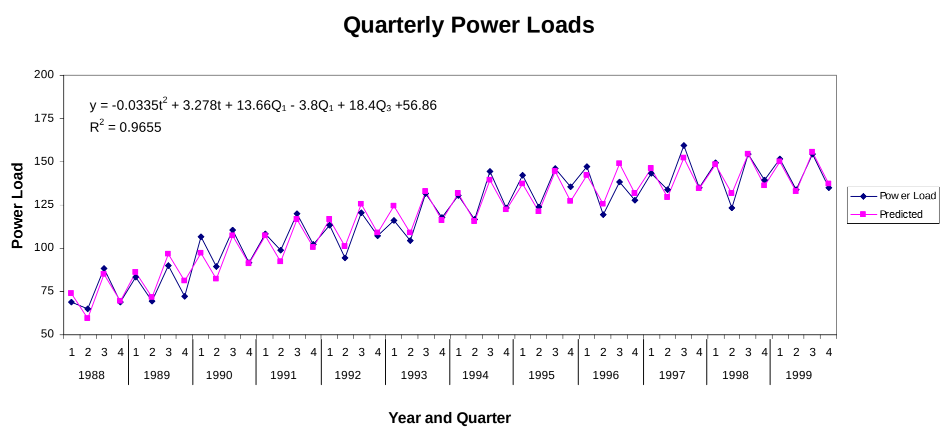
- For the Power Load data, both trend and seasonality can be modeled as follows:
  $Y_t = a + b_1 t + b_2 t^2 + b_3 Q_1 + b_4 Q_2 + b_5 Q_3 + e_t$
  where

$$Q_1 = \begin{cases} 1 & \text{if quarter 1} \\ 0 & \text{if quarters 2, 3, 4} \end{cases}$$

$$Q_2 = \begin{cases} 1 & \text{if quarter 2} \\ 0 & \text{if quarters 1, 3, 4} \end{cases}$$

$$Q_3 = \begin{cases} 1 & \text{if quarter 3} \\ 0 & \text{if quarters 1, 2, 4} \end{cases}$$

Power Load Data modeled for both Trend and Seasonality



Predicted Power Loads