Introduction to Time Series Analysis

- A *time-series* is a set of observations on a quantitative variable collected over time.
- Examples
  - Dow Jones Industrial Averages
  - Historical data on sales, inventory, customer counts, interest rates, costs, etc
- Businesses are often very interested in forecasting time series variables.
- Often, independent variables are not available to build a regression model of a time series variable.
- In time series analysis, we analyze the past behavior of a variable in order to predict its future behavior.

Methods used in Forecasting

- Regression Analysis
- Time Series Analysis (TSA)
  - A statistical technique that uses time-series data for explaining the past or forecasting future events.
  - The prediction is a function of time (days, months, years, etc.)
  - No *causal* variable; examine past behavior of a variable and and attempt to predict future behavior

Components of TSA

- Time Frame (How far can we predict?)
  - short-term (1 - 2 periods)
  - medium-term (5 - 10 periods)
  - long-term (12+ periods)
  - No line of demarcation
- Trend
  - Gradual, long-term movement (up or down) of demand.
  - Easiest to detect

Components of TSA (Cont.)

- Cycle
  - An up-and-down repetitive movement in demand.
  - repeats itself over a long period of time
- Seasonal Variation
  - An up-and-down repetitive movement within a trend occurring periodically.
  - Often weather related but could be daily or weekly occurrence
- Random Variations
  - Erratic movements that are not predictable because they do not follow a pattern

Combination of Four Components

- Considering the effects of these four components, two different types of models are generally used for a time series.
  - Additive Model

$$Y(t) = T(t) + S(t) + C(t) + I(t)$$

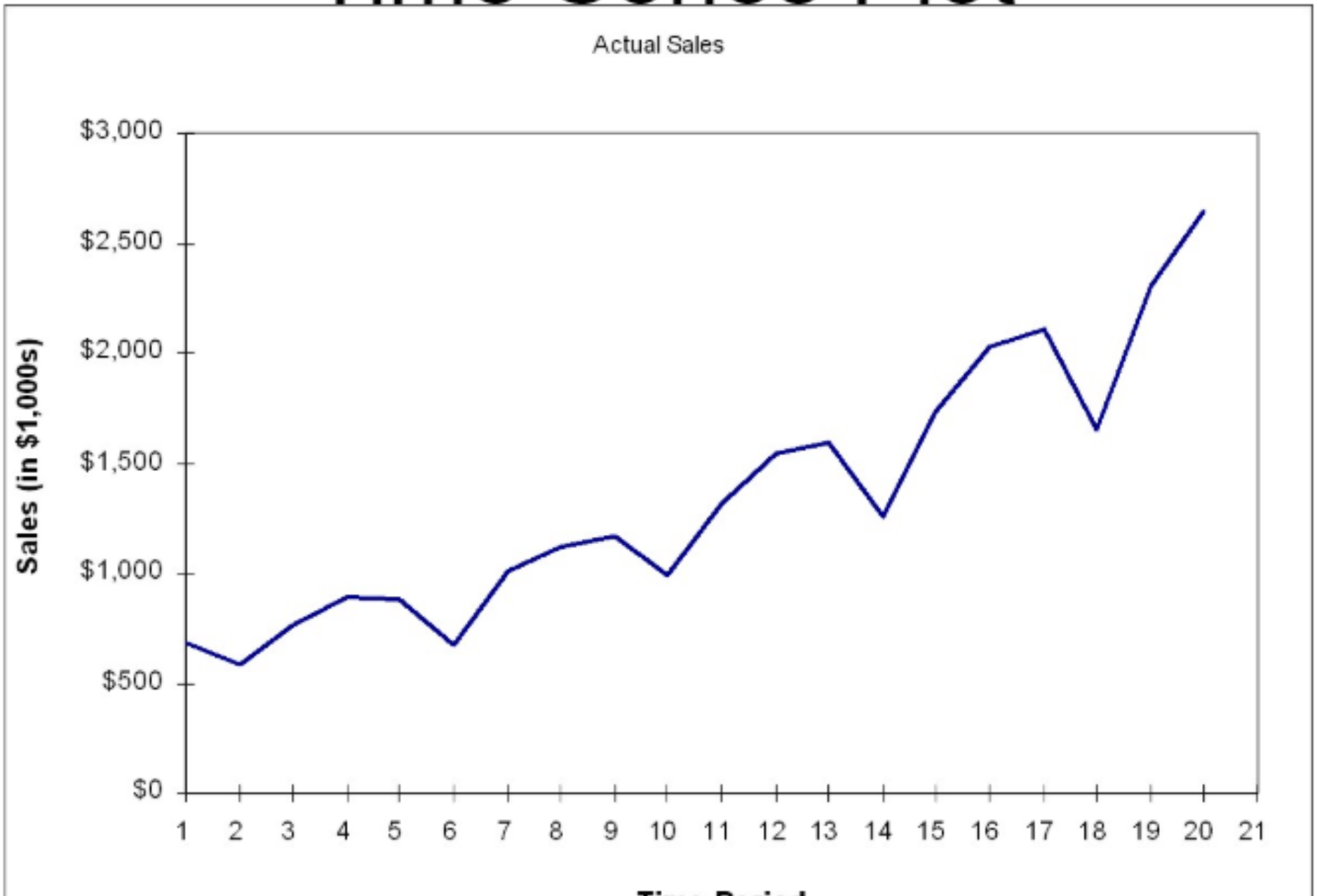
Assumption: These four components are independent of each other.

  - Multiplicative Model

$$Y(t) = T(t) \times S(t) \times C(t) \times I(t)$$

Assumption: These four components of a time series are not necessarily independent and they can affect one another.

Time Series Plot



Moving Averages

$$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + Y_{t-k+1}}{k}$$

- No general method exists for determining k.
- We must try out several k values to see what works best.

Forecasts for time periods 25 and 26 at time period 24:

$$\hat{Y}_{25} = \frac{Y_{24} + Y_{23}}{2} = \frac{36+35}{2} = 35.5$$

$$\hat{Y}_{26} = \frac{\hat{Y}_{25} + Y_{24}}{2} = \frac{35.5+36}{2} = 35.75$$

Weighted Moving Average

- The moving average technique assigns equal weight to all previous observations

$$\hat{Y}_{t+1} = \frac{1}{k} Y_t + \frac{1}{k} Y_{t-1} + \cdots + \frac{1}{k} Y_{t-k-1}$$

- The weighted moving average technique allows for different weights to be assigned to previous observations.

$$\hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + \cdots + w_k Y_{t-k-1}$$

$$\text{where } 0 \le w_i \le 1 \text{ and } \sum w_i = 1$$

- We must determine values for k and the w_i

Forecasts for time periods 25 and 26 at time period 24:

$$\hat{Y}_{25} = w_1 Y_{24} + w_2 Y_{23} = 0.291 \times 36 + 0.709 \times 35 = 35.3$$

$$\hat{Y}_{26} = w_1 \hat{Y}_{25} + w_2 Y_{24} = 0.291 \times 35.3 + 0.709 \times 36 = 35.80$$

Exponential Smoothing

$$\hat{Y}_{t+1} = \hat{Y}_t + \alpha(Y_t - \hat{Y}_t)$$

$$\text{where } 0 \le \alpha \le 1$$

- It can be shown that the above equation is equivalent to:

$$\hat{Y}_{t+1} = \alpha Y_t + \alpha(1-\alpha)Y_{t-1} + \alpha(1-\alpha)^2 Y_{t-2} + \cdots + \alpha(1-\alpha)^n Y_{t-n} + \cdots$$

Forecasting With The Exponential Smoothing Model

Forecasts for time periods 25 and 26 at time period 24

$$\hat{Y}_{25} = \hat{Y}_{24} + \alpha(Y_{24} - \hat{Y}_{24}) = 35.74 + 0.268(36 - 35.74) = 35.81$$

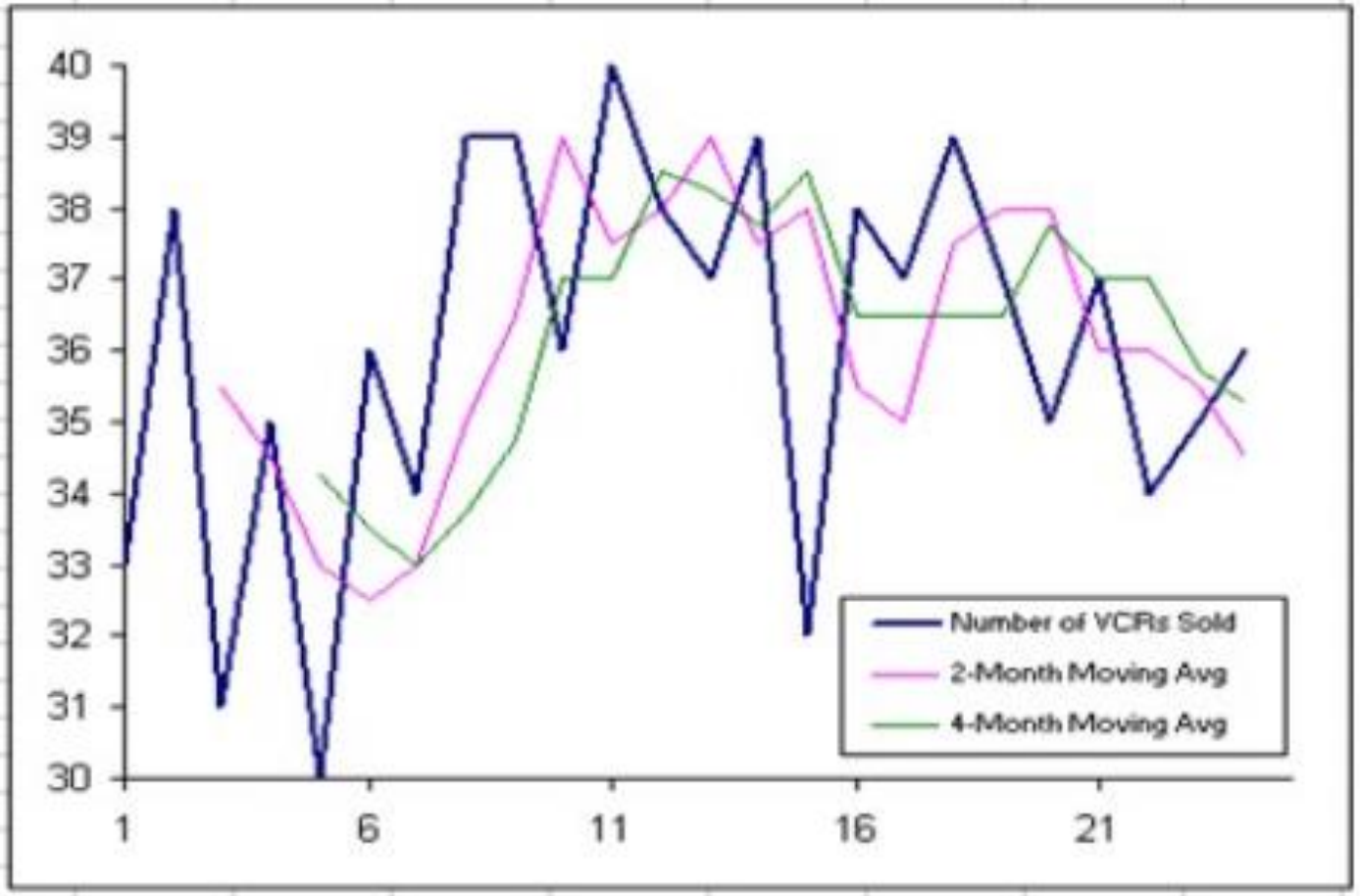
$$\hat{Y}_{26} = \hat{Y}_{25} + \alpha(Y_{25} - \hat{Y}_{25}) \approx \hat{Y}_{25} + \alpha(\hat{Y}_{25} - \hat{Y}_{25}) = \hat{Y}_{25} = 35.81$$

Note that,

$$\hat{Y}_t = 35.81, \text{ for } t = 25,\ 26,\ 27,\ \ldots$$

Implementing the Model

Time Period	Number of VCRs Sold	2-Month Moving Avg	4-Month Moving Avg
1	33	--	--
2	38	--	--
3	31	35.50	--
4	35	34.50	--
5	30	33.00	34.25
6	36	32.50	33.50
7	34	33.00	33.00
8	39	35.00	33.75
9	39	36.50	34.75
10	36	39.00	37.00
11	40	37.50	37.00
12	38	38.00	38.50
13	37	39.00	38.25
14	39	37.50	37.75
15	32	38.00	38.50
16	38	35.50	36.50
17	37	35.00	36.50
18	39	37.50	36.50
19	37	38.00	36.50
20	35	38.00	37.75
21	37	36.00	37.00
22	34	36.00	37.00
23	35	35.50	35.75
24	36	34.50	35.25



ARIMA Models

- Auto Regressive (AR) Model:
  - Value of a variable in one period is related to the values in previous period.
  - AR(p) - Current values depend on its own p-previous values
  - P is the order of AR process
  - Ex : AR(1,0,0) or AR(1)
- Moving Average (MA) Model:
  - Accounts for possibility of a relationship b/w a variable & residuals from previous period.

- ARIMA with environmental variable is very important in the case when external variable start impacting the series
- Ex. Flight delay prediction depends not only historical time series data but external variables like weather condition (temperature , pressure, humidity, visibility, arrival of other flights, weighting time etc.)