Time Series Analysis and Forecasting

**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 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

### Some Time Series Terms

- **Stationary Data** - A time series variable exhibiting no significant upward or downward trend over time.
- **Nonstationary Data** - A time series variable exhibiting a significant upward or downward trend over time.
- **Seasonal Data** - A time series variable exhibiting a repeating pattern at regular intervals over time.
**Approaching Time Series Analysis**

- There are many, many different time series techniques.
- It is usually impossible to know which technique will be best for a particular data set.
- It is customary to try out several different techniques and select the one that seems to work best.
- To be an effective time series modeler, you need to keep several time series techniques in your "tool box."

**Measuring Accuracy**

- We need a way to compare different time series techniques for a given data set.
- Four common techniques are the:
  - mean absolute deviation, \( \text{MAD} = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i| \)
  - mean absolute percent error, \( \text{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \)
  - the mean square error, \( \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \)
  - root mean square error, \( \text{RMSE} = \sqrt{\text{MSE}} \)
- We will focus on MSE.

**Extrapolation Models**

- Extrapolation models try to account for the past behavior of a time series variable in an effort to predict the future behavior of the variable.
- \( \hat{Y}_{t+1} = f(Y_t, Y_{t-1}, Y_{t-2}, \ldots) \)

**Moving Averages**

- \( \hat{Y}_{t+1} = \frac{Y_t + Y_{t+1} + Y_{t+k}}{k} \)
  - No general method exists for determining \( k \).
  - We must try out several \( k \) values to see what works best.
**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.
  \[ \hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + \cdots + w_k Y_{t-k+1} \]
  where \( 0 \leq w_j \leq 1 \) and \( \sum w_j = 1 \)

- We must determine values for \( k \) and the \( w_j \)

---

**Exponential Smoothing**

\[ \hat{Y}_{t+1} = \hat{Y}_t + \alpha (Y_t - \hat{Y}_t) \]
where \( 0 \leq \alpha \leq 1 \)

- It can be shown that the above equation is equivalent to:
  \[ \hat{Y}_{t+1} = \alpha Y_t + \alpha (1-\alpha) Y_{t-1} + \alpha (1-\alpha)^2 Y_{t-2} + \cdots + \alpha (1-\alpha)^t Y_{t-t} + \cdots \]

---

**Seasonality**

- Seasonality is a regular, repeating pattern in time series data.
- May be additive or multiplicative in nature...

---

**Stationary Seasonal Effects**

- [Additive Seasonal Effects]
- [Multiplicative Seasonal Effects]
**Trend Models**

- Trend is the long-term sweep or general direction of movement in a time series.
- We’ll now consider some nonstationary time series techniques that are appropriate for data exhibiting upward or downward trends.

**The Linear Trend Model**

\[
\hat{Y}_t = b_0 + b_1 X_{t,1}
\]

where \( X_{t,1} = t \)

For example:

\( X_1 = 1, X_2 = 2, X_3 = 3, \ldots \)

**The TREND() Function**

\[
\text{TREND(Y-range, X-range, X-value for prediction)}
\]

where:

- **Y-range** is the spreadsheet range containing the dependent Y variable.
- **X-range** is the spreadsheet range containing the independent X variable(s).
- **X-value for prediction** is a cell (or cells) containing the values for the independent X variable(s) for which we want an estimated value of Y.

Note: The TREND() function is dynamically updated whenever any inputs to the function change. However, it does not provide the statistical information provided by the regression tool. It is best two use these two different approaches to doing regression in conjunction with one another.

**The Quadratic Trend Model**

\[
\hat{Y}_t = b_0 + b_1 X_{t,1} + b_2 X_{t,2}
\]

where \( X_{t,1} = t \) and \( X_{t,2} = t^2 \)
Combining Forecasts

- It is also possible to combine forecasts to create a composite forecast.
- Suppose we used three different forecasting methods on a given data set.
  - Denote the predicted value of time period $t$ using each method as follows: $F_1$, $F_2$, and $F_3$.
  - We could create a composite forecast as follows:
    $\hat{Y}_t = b_0 + b_1F_{1,t} + b_2F_{2,t} + b_3F_{3,t}$.