Given the limited background from the surveys and that Chapter 7 in the book is complex, we will cover less material.

- The role of forecasting in the chain
- Characteristics of forecasts
- Basic approach to demand forecasting
- Measures of forecast error

- We will skip seasonality and Holt’s and Winter’s methods. We do not spend any time on static forecasts
Role of Forecasting in a Supply Chain

◆ The basis for all strategic and planning decisions in a supply chain
◆ Used for both push and pull processes
◆ Examples:
  – Production: scheduling, inventory, aggregate planning
  – Marketing: sales force allocation, promotions, new production introduction
  – Finance: plant/equipment investment, budgetary planning
  – Personnel: workforce planning, hiring, layoffs
◆ All of these decisions are interrelated
Characteristics of Forecasts

- A FORECAST is a statement about the future
  - Absatzprognose, Vorhersage
- Forecasts are always wrong. Report both the expected value of the forecast and the measure of error
- Long-term forecasts are less accurate than short-term forecasts (forecast horizon is important)
- Aggregate forecasts are more accurate than disaggregate forecasts
- In order to forecast, the past has to have some relevance to the future
Forecasting Methods

◆ Qualitative: primarily subjective; use judgment/opinion

◆ Time Series: use historical demand only
  – Static
  – Adaptive  We will only consider adaptive

◆ Causal (or associative): use the relationship between demand and a factor other than pure time to develop forecast

◆ Simulation
  – Imitate consumer choices that give rise to demand
  – Can combine time series and causal methods
Basic Approach to Demand Forecasting

- Understand the objectives of forecasting
- Integrate demand planning and forecasting
- Identify major factors that influence the demand forecast
- Understand and identify customer segments
- Determine the appropriate forecasting technique
- Establish performance and error measures for the forecast
Components of an Observation

Observed demand (O) =
Systematic component (S) + Random component (R)

- **Level** (current deseasonalized demand)
- **Trend** (growth or decline in demand)
- **Seasonality** (predictable seasonal fluctuation)

- Systematic component: Expected value of demand
- Random component: The part of the forecast that deviates from the systematic component
- Forecast error: difference between forecast and actual demand
Steps in the Forecasting Process

**Step 1** Determine the purpose of forecast

**Step 2** Pick an appropriate time horizon

**Step 3** Select a forecasting technique
- Plotting data may reveal patterns

**Step 4** Gather and analyze data *in detail*
- State assumptions
- Validate Data: May need to cleanse or filter for past events

**Step 5** Calculate forecast

**Step 6** Analyze/Monitor the forecast- Measure Accuracy

*Are results acceptable?*

No: Return to **Step 3**, revising forecast technique
Yes: Publish forecast

◆ For ongoing forecasting: repeat Steps 4 through 6
Time Series Forecasts

**Trend** - long-term movement in data
- Linear: steady increase (or decrease) over time
- Not all trends are linear. Demand may be exponential, may both increase and decrease over product life cycle: VHS players

**Seasonality** - short-term *regular* variations in data
- Example: Walgreen’s sales of cold medications over the year
- Not just limited to Fall/Winter/Spring/Summer variations
  - Weekly demand for reservations at expensive restaurants
  - Daily cycle of coffee sales at Starbucks

**Irregular variations** - caused by unusual circumstances
- Infrequent spikes
- i.e. a stock market crash, 9/11 catastrophe

**Random variations** - caused by chance
Time Series Forecasting
Techniques Covered in this Class

◆ Averaging (or Smoothing)
  – Moving Average
  – Weighted Moving Average
  – Exponential Smoothing
    - Trend-Adjusted Exponential Smoothing (Holt’s) is not covered, nor is Winters (trend + seasonality)

◆ Linear Trend Analysis
Moving Average

◆ The average of the N most recent observations:

\[
MA_n(t) = \frac{\sum_{i=t-n}^{t-1} A_i}{n}
\]

◆ Example: a 4-period MA for time period 7 would be

\[
F_7 = \frac{(A_6 + A_5 + A_4 + A_3)}{4}
\]

◆ The larger N, the smoother the forecast, but the greater the Lag (ability to respond to “real” changes)
General Note: To Write Formula with Respect to \(t\) or \(t+1\)

◆ Which is correct?

1. \(F_{t+1} = (\text{blah})A_t + \text{blah} \ldots\)
2. \(F_t = (\text{blah})A_{t-1} + \text{blah} \ldots\)

◆ Both are! As long as it’s clear just what \(t\) (or \(t+1\)) represents, these can be usable interchangeably

- If you need to calculate the forecast for period 5, be clear whether \(t=5\) or \(t+1=5\)

◆ But most important, don’t write:

\[ F_t = (\text{blah})A_t + \text{blah} \]
Weighted Moving Average

- *Premise*: The most recent observations might have the best predictive value. Yet for simple moving averages older data points have same importance as most recent.

- We modify to give greater weight to more recent observations. (Remember: $\sum W_i = 1$ or you get bias!)

$$WMA_n(t) = \sum_{i=t-n}^{t-1} W_i A_i$$

- Advantages: Avoids “oversmoothing” and lag time is decreased.

- Weights are arbitrary, often found through trial and error!
Exponential Smoothing

- Exponential Smoothing is a type of Weighted Moving Average:

\[ F_t = \alpha A_{t-1} + \alpha (1-\alpha) A_{t-2} + \alpha (1-\alpha)^2 A_{t-3} + \alpha (1-\alpha)^3 A_{t-4} + \ldots \]

- However, it is much more easily written, computed and understood as:

\[ F_t = \alpha A_{t-1} + (1-\alpha) F_{t-1} \]

Where \( \alpha \) is between 0 and 1
More Exponential Smoothing

- Rearranging the terms shows this method can be viewed as the previous period’s forecast adjusted by a percentage of the previous period’s error \(\varepsilon_{t-1} = A_{t-1} - F_{t-1}\)

\[
F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1})
\]

- The quickness of forecast adjustment is determined by the smoothing constant, \(\alpha\)
  - The closer \(\alpha\) to zero, the greater the smoothing
  - How do we pick \(\alpha\)? Trial and error or can even optimize for it!
  - How to initialize the forecast? Many ways!
    » Book: average all the data we have so far
    » More practical and repeatable? Start it with \(F_2 = A_1\)
How to Select a Forecast

- To forecast data without trends, we could use a simple naïve forecast, a MA (still need to pick the window), a WMA (need to pick both the window and the weights) or an exponential smoother (need to pick $\alpha$).

- We will get many different answers—how do we pick the one we feel will have the best chance of being close to what will happen?

- We calculate past forecast accuracy, and we then pick the one that is most accurate.
Measuring Forecast Accuracy

• Error: difference between actual value and predicted value. Many different measures exist (we use MAD only)

\[ \text{MAD}_t = \frac{1}{t} \sum_{i=1}^{t} |A_i - F_i| \]

• Mean Absolute Deviation (MAD)

• Mean Squared Error (MSE)

\[ \text{MSE}_t = \frac{1}{t-1} \sum_{i=1}^{t} (A_i - F_i)^2 \]

• Standard Error (Standard Deviation)

\[ \sigma_t = \sqrt{\text{MSE}_t} \]

rule of thumb \( \sigma_t \approx 1.25 \ast \text{MAD}_t \)
If there is a trend, the smoothing filters we have covered will LAG, resulting in bad forecasts.

Plot the data first- in this example, do you see a trend?

If the trend is linear, we will use **Linear Trend Analysis**

- Caveat: Not all trends are linear! (We do not cover curvilinear regression in this class)
Linear Trend Equation

\[ Y_t = a + b \times t \]

- **Where**
  - \( a \) = intercept
  - \( b \) = slope

- Looks like a simple line equation, but \( a \) and \( b \) are determined to minimize Mean Squared Error (MSE)
Graphing the Results

Regressing Sales Against Week

Sales

145 150 155 160 165 170 175 180 185

Week

1 2 3 4 5 6

Actual
Forecast

Graph showing the regression of sales against week.
An intern at Netflix has been tracking weekly rentals of the direct-to-DVD movie “Barney Meets Vin_Diesel: Fossilized!” and, after noticing a pattern, has modeled them with a linear trend analysis. The Excel model gives $a = 400$, $b = -10$, where $t = 0$ was the DVD release week. It is currently week 12.

- What is the linear trend equation?
- What are sales predicted to be this week?
- What does the model predict sales will be a half year (26 weeks) after the release?
- When does the model predict sales to stop? Do you believe this model is likely to be accurate for making long term predictions?
Associative Forecasting

What if we think there is a better indicator of future behavior than time?

- Use explanatory or predictor variables to predict the future
  - E.g. Using interest rates to predict home purchases

- The associative technique used in this class is Simple **Linear Regression**
  - Linear Trend Analysis was an example where time \( t \) was used at the dependent variable. But now we can use factors other than time

\[
Y_t = a + bt \quad \text{Linear Trend Analysis}
\]

\[
Y_t = a + bX_t \quad \text{Associative Forecast}
\]

  ( \( X_t \) is used instead of \( t \))

- Again, we use Excel to determine \( a \) and \( b \)
Associative Forecasting: Linear Regression

Regression equation:
\[ Y = a + bX \]
Associative Forecasting: Linear Regression

- Dependent variable
- Independent variable
- Actual value of $Y$
- Estimate of $Y$ from regression equation
- Deviation, or error

Regression equation:
\[ Y = a + bX \]
Example: Associative Forecast

<table>
<thead>
<tr>
<th>Month</th>
<th>Sales (000 units)</th>
<th>Advertising (000 $)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>264</td>
<td>2.5</td>
</tr>
<tr>
<td>2</td>
<td>116</td>
<td>1.3</td>
</tr>
<tr>
<td>3</td>
<td>165</td>
<td>1.4</td>
</tr>
<tr>
<td>4</td>
<td>101</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>209</td>
<td>2.0</td>
</tr>
</tbody>
</table>

A rotor manufacturer has recorded these sales figures over the past 5 months. They also know the budget spent promoting these parts.

1) Is there wide variation in sales?
2) Does this appear to be a good candidate for linear trend analysis?
Now Consider Sales versus Advertising Spend

\[ Y = -8.136 + 109.229(X) \]
How Strong is the Relationship?

**Correlation** (r), measures strength and direction of the forecast of Y with respect to X (or t for linear trend analysis)
- 0 < r < 1 Positive correlation (sales of ice cream cones vs temperature)
- -1 < r < 0 Negative correlation (sales of sweatshirts vs temperature)

You are not expected to compute r by hand! Instead use Regression Analysis in Excel (the “multiple R entry”)

**R-Squared** (r^2) measures the *percentage of variation* in y that is “explained” by x
- If .8 < r^2 < 1, X is a very good predictor
- For r^2 < .25, X is a poor predictor- look for something else!

Even if you have a good predictor, remember correlation is not causation
### Example: Excel Solution to Linear Regression

#### SUMMARY OUTPUT

<table>
<thead>
<tr>
<th>Regression Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
</tr>
<tr>
<td>R Square</td>
</tr>
<tr>
<td>Adjusted R Square</td>
</tr>
<tr>
<td>Standard Error</td>
</tr>
<tr>
<td>Observations</td>
</tr>
</tbody>
</table>

#### ANOVA

<table>
<thead>
<tr>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Significance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>1</td>
<td>17323.66</td>
<td>17323.66</td>
<td>71.1604</td>
</tr>
<tr>
<td>Residual</td>
<td>3</td>
<td>730.3361</td>
<td>243.4454</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>18054</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Coefficients

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t Stat</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
<th>Lower 95.0%</th>
<th>Upper 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-8.13</td>
<td>-0.36</td>
<td>0.74</td>
<td>-79.27</td>
<td>63.00</td>
<td>-79.27</td>
<td>63.00</td>
</tr>
<tr>
<td>X Variable 1</td>
<td>109.23</td>
<td>8.44</td>
<td>0.00</td>
<td>68.02</td>
<td>150.44</td>
<td>68.02</td>
<td>150.44</td>
</tr>
</tbody>
</table>
Correlation is not Causation: An Example

The following scatter-plot shows that the average *life expectancy* for a country is related to the number of *doctors* per person in that country. We could come up with all sorts of reasonable explanations justifying this, but…
Lurking Variables and Causation
Another Example

◆ This new scatter-plot shows that the average life expectancy for is also related to the number of televisions per person in that country.
  – And the relationship is even stronger: $R^2$ of 72% instead of 62%

◆ Since TVs are cheaper than doctors, Why don’t we send TVs to countries with low life expectancies in order to extend lifetimes. Right?
How about considering a lurking variable? That makes more sense…

- Countries with higher standards of living have both longer life expectancies \textit{and} more doctors (and TVs!).
- If higher living standards \textit{cause} changes in these other variables, improving living standards might be expected to prolong lives and, incidentally, also increase the numbers of doctors and TVs.