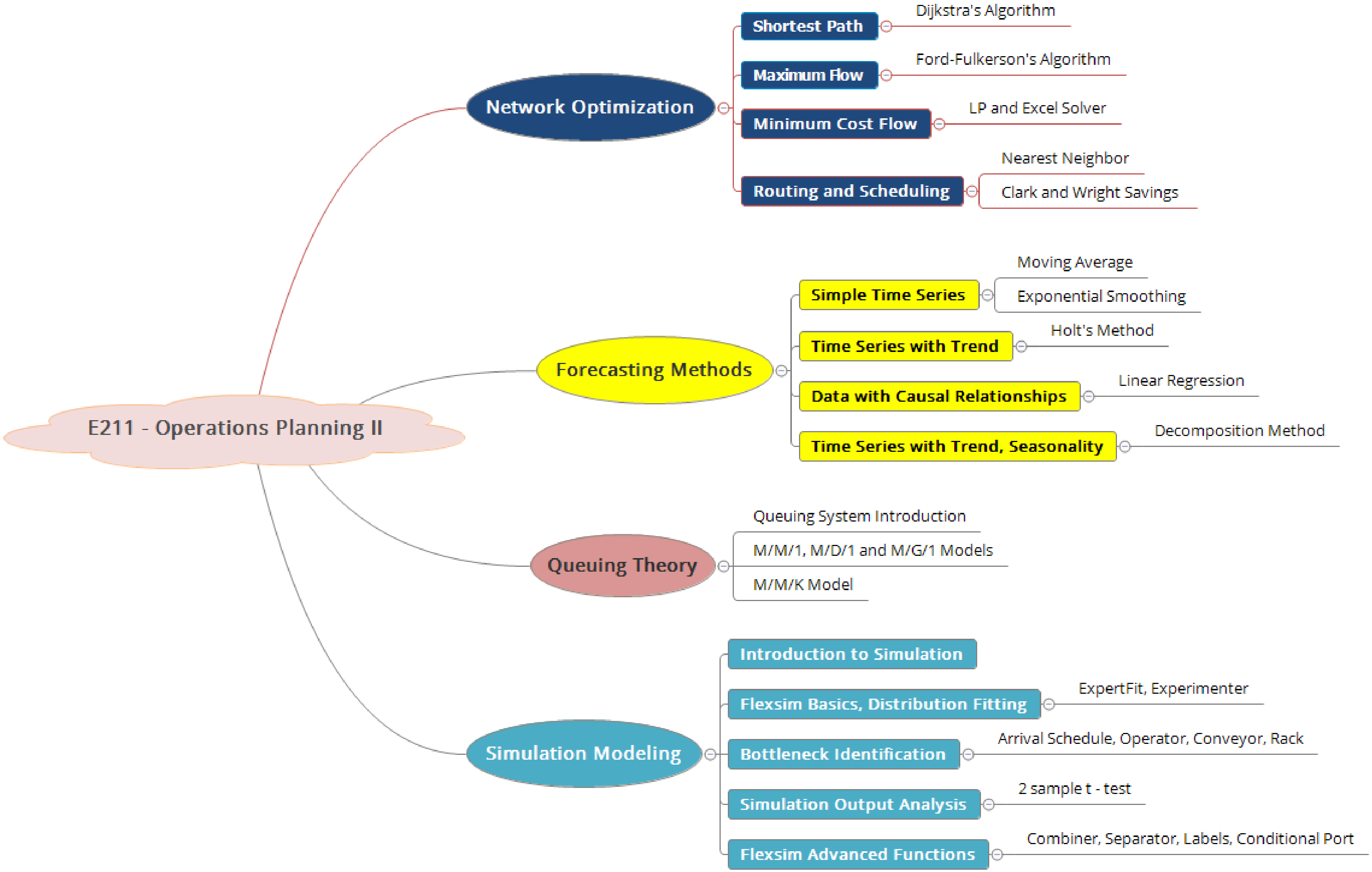
Problem 05

The Art of Forecasting (Part 2)

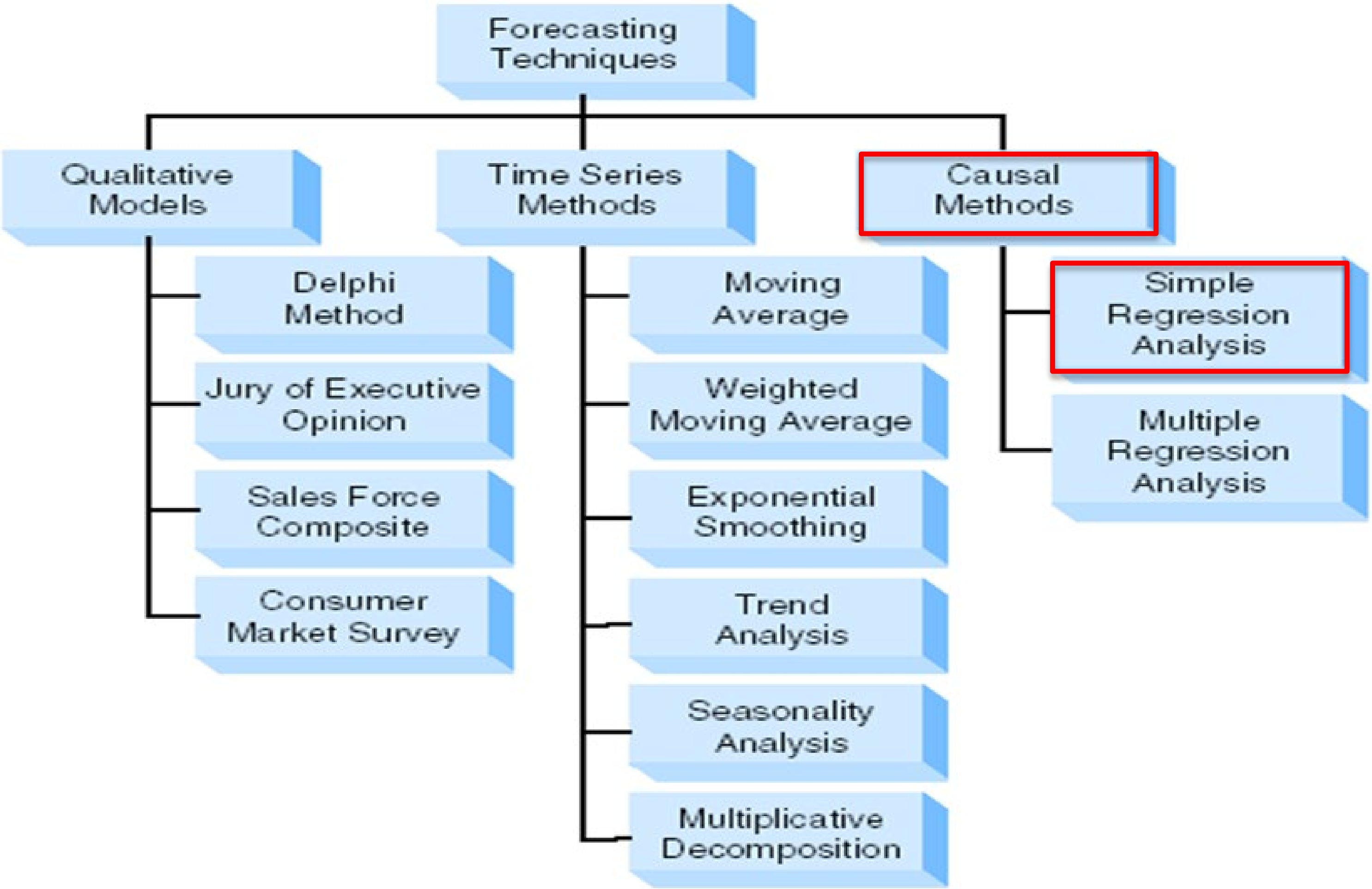
E211 – Operations Planning II

SCHOOL OF ENGINEERING

Module Coverage: E211 Topic Tree



Types of Forecasts



Simple Linear Regression

- The simple linear regression model assumes that the relationship between the dependent (response) variable, which is denoted y, and the independent (predictor) variable, denoted x, can be approximated by a straight line.

- Examples of simple linear regression model:
  - Forecasting price of house (y) from size of house (x)
  - Forecasting sales of product (y) from advertising spending (x)

Simple Linear Regression

- Relates dependent variable to one independent variable in the form of a linear equation

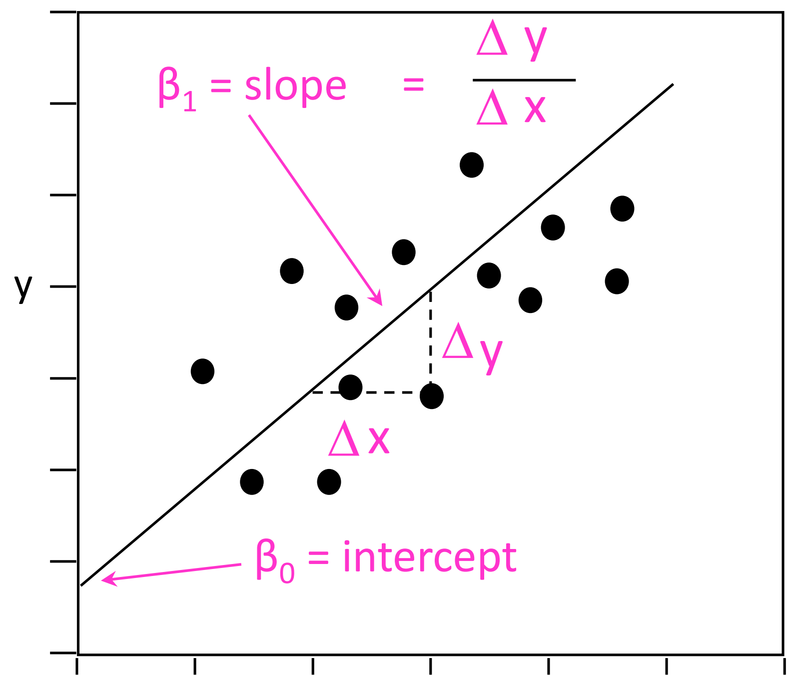
Dependent variable

Independent variable

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Unknown parameters to be estimated from data

Error variable



$\beta_1 > 0 \Rightarrow$ Positive Association
$\beta_1 < 0 \Rightarrow$ Negative Association
$\beta_1 = 0 \Rightarrow$ No Association

Note that β = Beta , ε = Epsilon

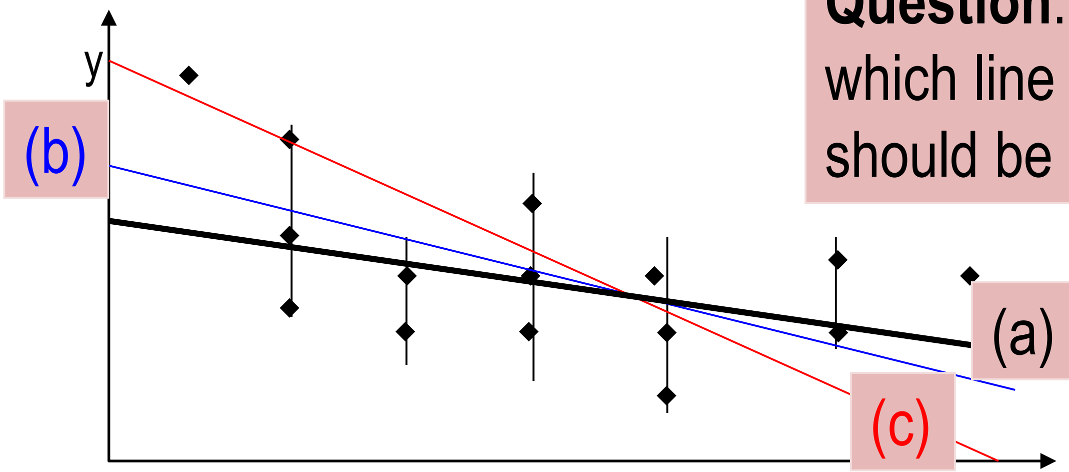
Model assumptions and ε

- At any given value of x, the population of ε is normally distributed with mean zero and a variance s^2 that does not depend on the value of x.

- The value of ε corresponding to an observed value of y is statistically independent of the value of ε corresponding to any other observed value of y.

Estimating the parameters

- We can draw a line through the data and obtain the equation of the line to represent the relationship.



Question: (Lines (a), (b) and (c), which line is better? Which line should be regarded as a good line?

Answer: A good line is the line that minimizes the sum of squared differences between the points and the line.

Least Squares Estimation of β_0, β_1

- Linear regression model:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

- Estimation of β_0 and β_1 using Least Squares Method

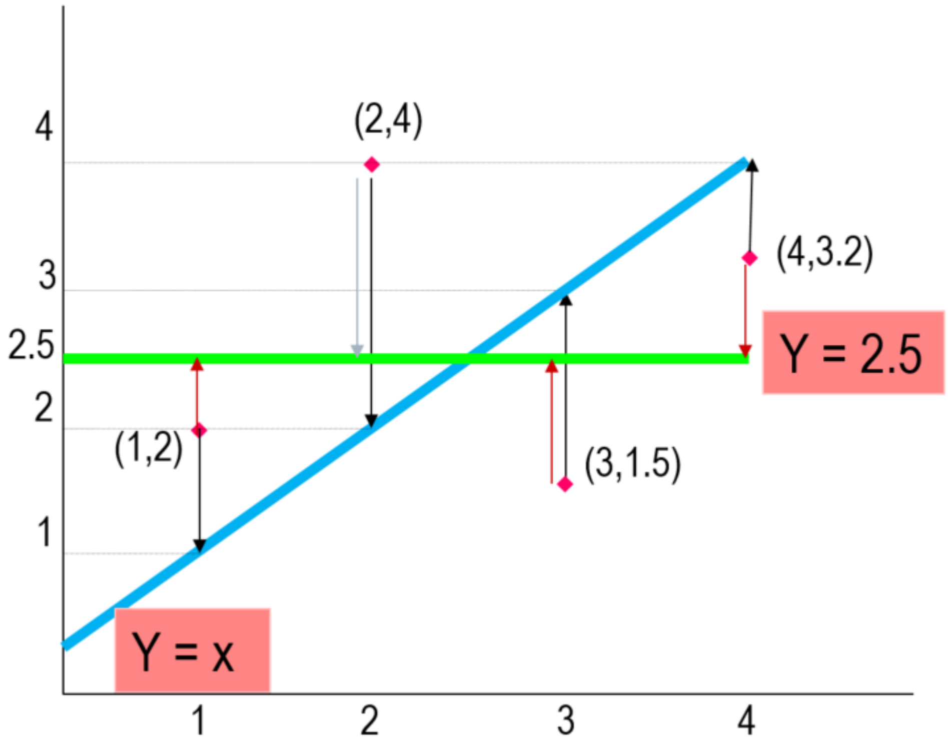
- **Goal**: Choose values (estimates) that minimize the sum of squared errors (differences) of observed data to the plotted straight-line

WORKSHEET Question: The Least Squares (Regression) Line

Sum of squared differences

Line y = x: $(2 - 1)^2 + (4 - 2)^2 + (1.5 - 3)^2 + (3.2 - 4)^2 = 7.89$

Line y = 2.5: $(2 - 2.5)^2 + (4 - 2.5)^2 + (1.5 - 2.5)^2 + (3.2 - 2.5)^2 = 3.99$

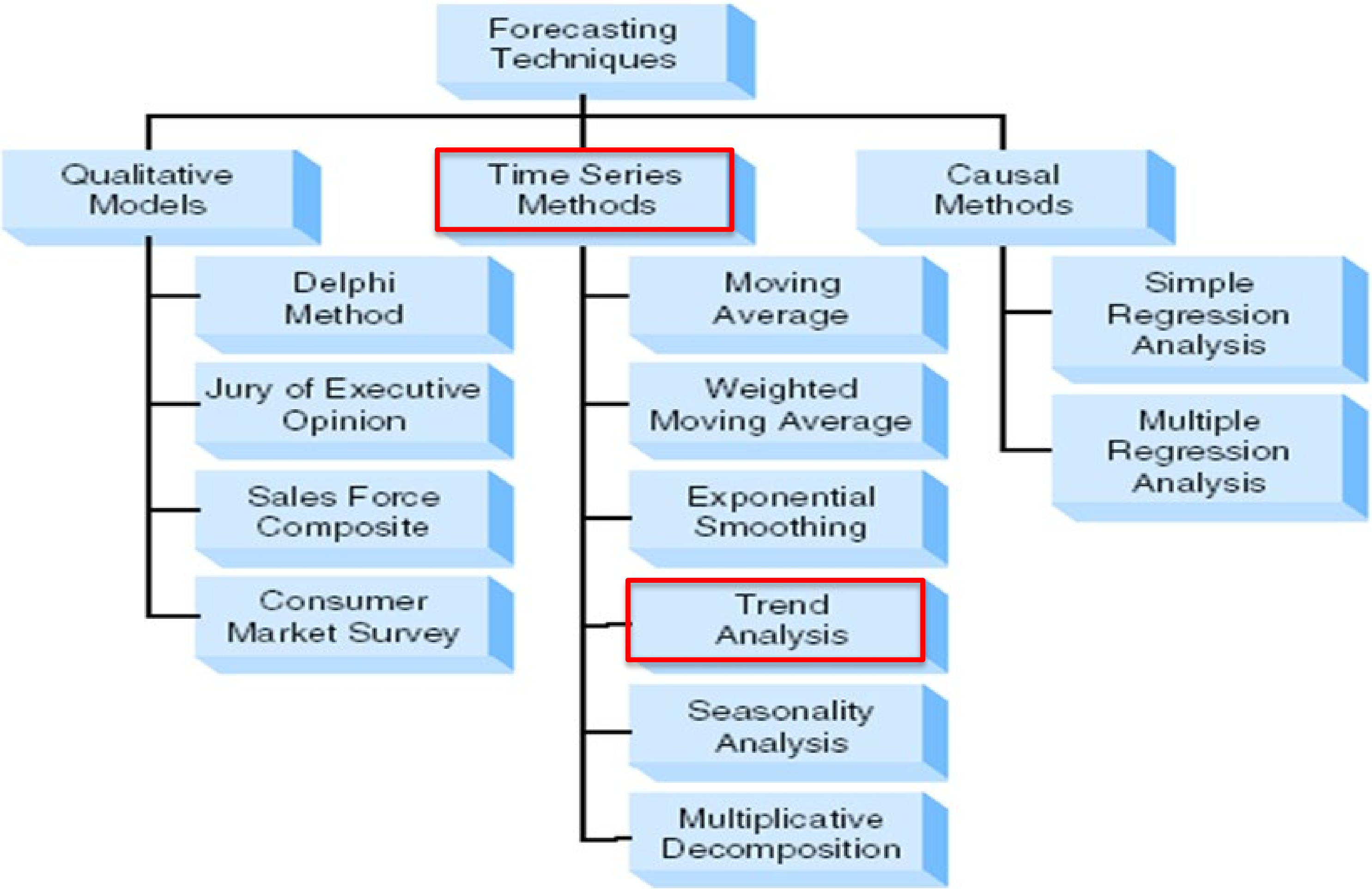


By comparing the two lines

The line Y = 2.5 is better

The smaller the sum of squared differences the better the fit of the line to the data.

Types of Forecasts



Holt’s Method

- Linear exponential smoothing technique
  - Trend adjusted exponential smoothing method.
  - Holt's two-parameter exponential smoothing model extends simple exponential smoothing to include a linear-trend component.
- Forecasts made for both
  - Time series level L_t which corresponds to constant in regression equation (β_0), and
  - Time series trend T_t corresponding to slope in regression equation (β_1).

Holt's Model

- The Holt's Model is built using 3 equations below

$$L_{t+1} = \alpha A_{t+1} + (1-\alpha)(L_t + T_t) \quad \text{------- a}$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1-\beta)T_t \quad \text{------- b}$$

$$H_{t+m} = L_t + mT_t \quad \text{------------------- c}$$

where, Equation *a, b and c* correspond to *the* *estimate of the time series' level*, *estimate of the time series' trend* and *the overall forecasting equation* respectively.

- What do these notations mean??

Notations of Holt's Model

L_{t+1} : estimate of time series' level for period t +1

A_{t+1} : actual value of time series at period t +1

L_t : estimate of time series' level for period t

T_t : estimate of time series' trend for period t

$\alpha(\text{Alpha})$: smoothing constant for the level estimate;

$\alpha \in [0, 1]$

$\beta(\text{Beta})$: smoothing constant for the trend estimate;

$\beta \in [0, 1]$

m : number of periods ahead to be forecasted

H_{t+m} : Holt's forecast value for period t + m

Discussions about Holt's Model

- In equation (a), the smoothing value, L_{t+1} , is predicted based on the current observation and the previous smoothed value (L_t) which is adjusted by adding a trend factor (T_t).
- The trend in equation (b) evolves by using the weighted average of the recent change of the smoothed value ($L_{t+1} - L_t$) and the previous trend T_t.
- Equation (c) is a forecast equation which is used to forecast *m* periods into the future.

Selection of Smoothing Constants

- When small values are used, less weight is given to historical data, effectively smoothing out random noise. This is appropriate under stable conditions.
- If the level and/or trend of the underlying data are changing appreciably over time, then higher values of the smoothing constants are suitable for quicker responses to level and/or trend changes.
- Because one can never really be sure if forecasting errors are a result of random noise or a real change, a compromise must be made.

P05 Suggested Solution

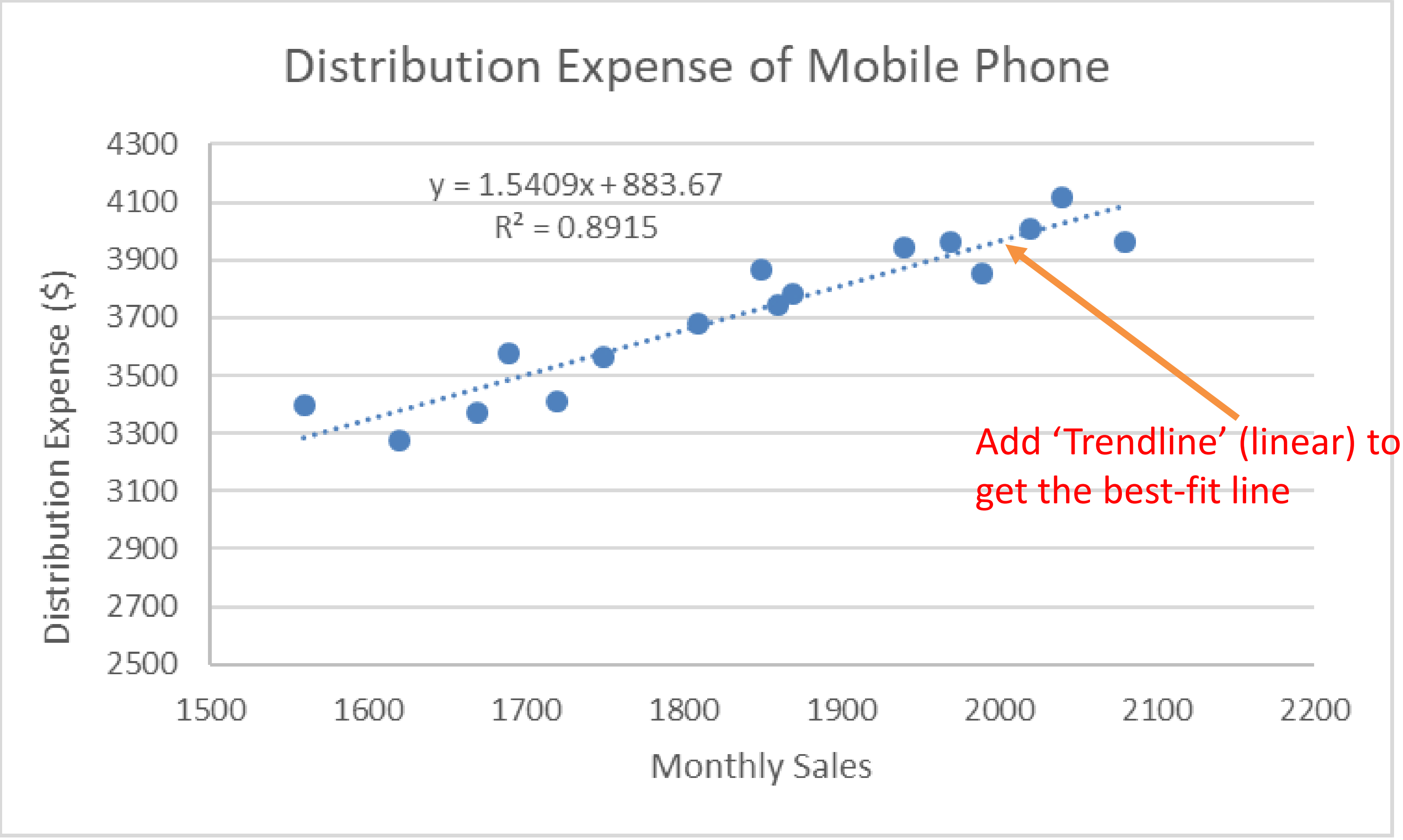
Causal Forecasting: Using Mobile Phone Monthly Sales (Sales Figure) to Forecast Distribution Expense of Mobile Phone

- Linear regression model:

  $y = \beta_0 + \beta_1 x$

  - y – Distribution Expense;
  - x – Mobile Phone Monthly Sales (Sales Figure);
  - β_0 - intercept value;
  - β_1, the change in Distribution Expense due to change in Mobile Phone Monthly Sales: we expect a positive correlation (β_1>0).
  - Use regression analysis (least squared error method) to determine β_0 and β_1

Scatter Plot for Distribution Expense of Mobile Phone Data



β_0 = 883.67
β_1 = 1.5409
R^2= 0.8915

Using Excel Regression Analysis – Summary Output

Regression Statistics	
Multiple R	0.94419
R Square	0.891495
Adjusted R Square	0.883745
Standard Error	88.84372
Observations	16

Correlation Coefficient (r)

Coefficient of Determination (R^2)

Estimate of the standard deviation of the points around the line

P-value for testing the significance of the independent variable

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	883.66789	265.2852797	3.33101	0.004947	314.687551	1452.64822	314.6875506	1452.64822
X Variable 1	1.5408734	0.143670567	10.72505	3.9E-08	1.23273071	1.84901615	1.232730712	1.84901615

Intercept and slope of the regression line equation

Interpretation of *r* and R^2 values

Correlation coefficient, r ('Multiple R') = 0.9442

- With r close to 1, it indicates that there is a strong relationship between the Distribution Expense and the Mobile Phone Monthly Sales.
- As β_1 (slope) is positive, it indicates a positive relationship between Distribution Expense and Mobile Phone Monthly Sales.

Coefficient of determination, R^2 ('R Square') = r*r =0.8915

- 89.15% of the variation in the Distribution Expense is explained by the variation in the Mobile Phone Monthly Sales.
- The rest of the variations may be due to other factors such as labour cost, fuel price, etc.

Testing the Significance of the Independent Variable

The p-value for each independent variable tests the null hypothesis that the variable has no correlation with the dependent variable.

- Testing the significance of the slope β_1:
  - P-value = 3.9E-08 < 0.05, we reject the null hypothesis → there is enough evidence to suggest that changes in x will affect y. In other words, there is correlation between x and y.
- Testing the significance of the y-intercept β_0:
  - P-value = 0.004947< 0.05, it is reasonable to include the y-intercept into the model.

P-value for testing the significance of the independent variable

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	883.66789	265.2852797	3.33101	0.004947	314.687551	1452.64822	314.6875506	1452.64822
X Variable 1	1.5408734	0.143670567	10.72505	3.9E-08	1.23273071	1.84901615	1.232730712	1.84901615

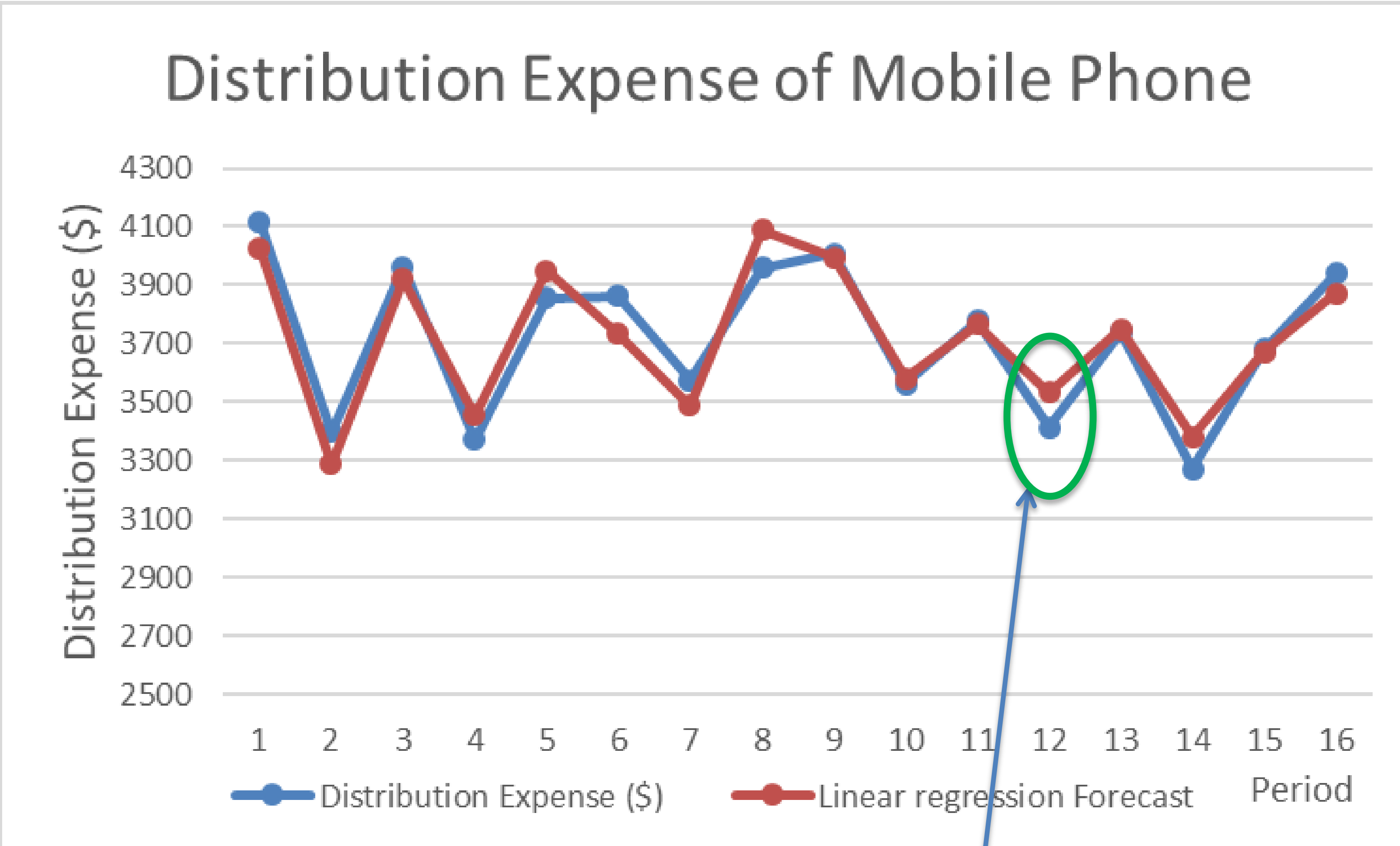
Using Regression Model to Predict y

Regression model: $y = 1.5409x + 883.6679$

- If Mobile Phone Monthly Sales = 1200,
  Predicted Distribution Expense:
  $y = 1.5409 (1200) + 883.6679 = \$2732.75$

- If Mobile Phone Monthly Sales = 2500,
  Predicted Distribution Expense:
  $y = 1.5409 (2500) + 883.6679 = \$4735.92$

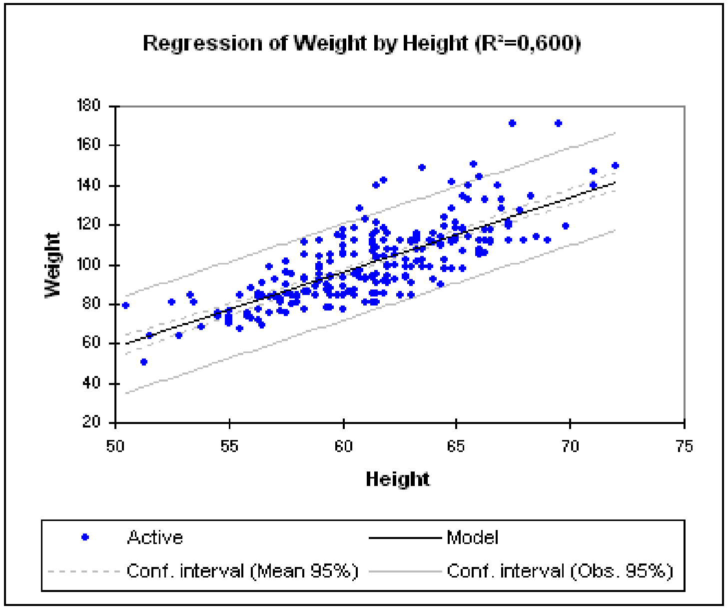
Using Linear Regression to Forecast Distribution Expense of Mobile Phone



- Other factors are affecting Distribution Expense in certain periods in addition to Mobile Phone Monthly Sales.

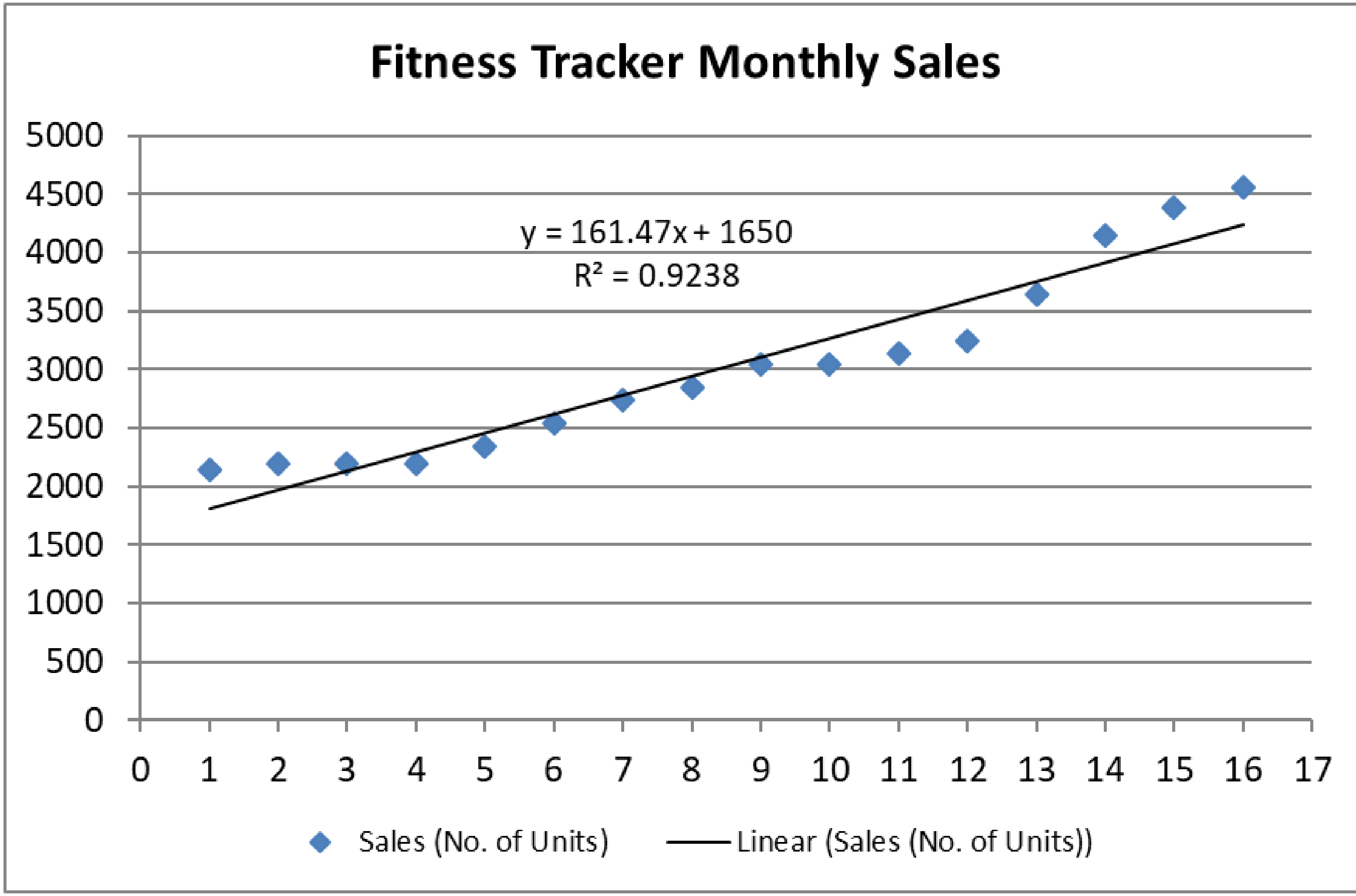
Using Linear Regression to Forecast Distribution Expense of Mobile Phone

- If we already have a good estimate of Mobile Phone Monthly Sales for the next month, linear regression may be a good method to forecast next month's Distribution Expense as there is a strong correlation between Distribution Expense and Mobile Phone Monthly Sales.
- Ways to improve linear regression forecast:
  - Generate confidence intervals of predicted values to provide range estimates (upper and lower bounds).



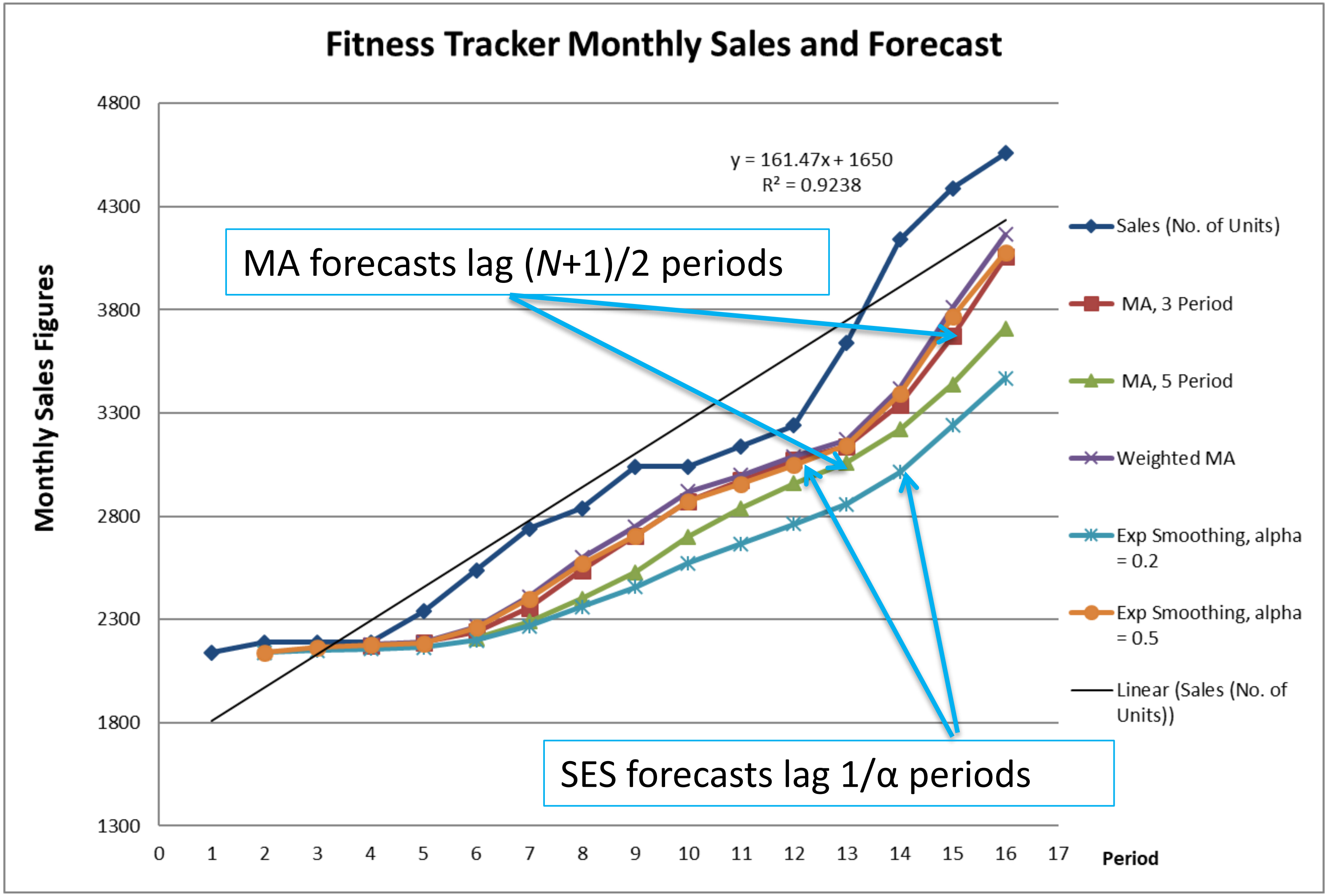
  - Use more predictors (independent variables) – multiple regression model.

Trend Analysis:
Forecasting for Fitness Tracker Monthly Sales



- There is a visible upward trend in the Fitness Tracker Monthly Sales data, will Linear Regression work well for the foresting?

Fitness Tracker Monthly Sales: Forecasting using MA, SES and Linear Regression Methods



- Which method gives a better forecast?
  - Moving Average method will lag by (N+1)/2 periods.
  - SES method will lag by 1/α periods.
  - Linear Regression method does not seem to follow closer to the actuals.

Fitness Tracker Monthly Sales: Forecasting Errors for MA, SES and Linear Regression Methods

Forecasting methods	Measures of forecasting error		
	MAD	MAPE	MSE
MA (3-Period)	346.41	10.07%	169317.09
MA (5-Period)	540.91	15.41%	350627.27
Weighted MA (3-Period)	297.69	8.69%	125926.92
Exp Smoothing, (alpha = 0.2)	515.83	14.91%	400384.09
Exp Smoothing, (alpha = 0.5)	290.53	8.55%	128689.60
Linear Regression	184.38	6.13%	45672.79

- There is a visible trend in the Fitness Tracker Monthly Sales data, Linear Regression method is the best among the 6 forecasting methods as it has the lowest forecasting errors.
- The concern is that Linear Regression treats all data points in time series equally. Therefore, it does not address the rate of change of the recent years well.
- What other forecasting method can handle both level, L and trend, T?

- To begin the forecasting process, Peter can use the intercept as their estimate for L_0 (= 1650) and the slope for T_0 (= 161.4706) --- The results from regression analysis.
- Choice of smoothing constants plays significant roles in determining accuracy of forecasts.
- Can select different sets of smoothing constants and see which set is better.

E.g. choose α (Alpha) = 0.8 and β (Beta) = 0.1.

Fitness Tracker Monthly Sales: Forecasting using Holt's Method ($\alpha = 0.8$, $\beta = 0.1$)

		L_t	T_t	Holt's H_t
Period	A_t (Sales)	1650.0000	161.4706	
1	2140	2074.2941	187.7529	1811.4706
2	2190	2204.4094	181.9892	2262.0471
3	2190	2229.2797	166.2773	2386.3986
4	2190	2231.1114	149.8327	2395.5570
5	2340	2348.1888	146.5572	2380.9441
6	2540	2530.9492	150.1775	2494.7460
7	2740	2728.2253	154.8874	2681.1267
8	2840	2848.6225	151.4384	2883.1127
9	3040	3032.0122	154.6335	3000.0609
10	3040	3069.3291	142.9018	3186.6457
11	3140	3154.4462	137.1234	3212.2310
12	3240	3250.3139	132.9978	3291.5696
13	3640	3588.6623	153.5329	3383.3117
14	4140	4060.4390	185.3572	3742.1952
15	4390	4361.1593	196.8935	4245.7963
16	4560	4559.6106	197.0493	4558.0528
17	???			4756.6599
18	???			4953.7092
19	???			5150.7585

T_0: Slope of linear trend line

L_0: intercept of linear trend line

With $\alpha = 0.8$, $\beta = 0.1$

Oct-18, $F_{17} = 4756.66$

Nov-18, $F_{18} = 4953.71$

Dec-18, $F_{19} = 5150.76$

$H_{17} = L_{16} + T_{16}$

$H_{19} = L_{16} + 3 * T_{16}$

$L_{12} = 0.8*A_{12} + (1-0.8)*(L_{11}+T_{11})$

$T_{12} = 0.1*(L_{12}-L_{11}) + (1-0.1)*T_{11}$

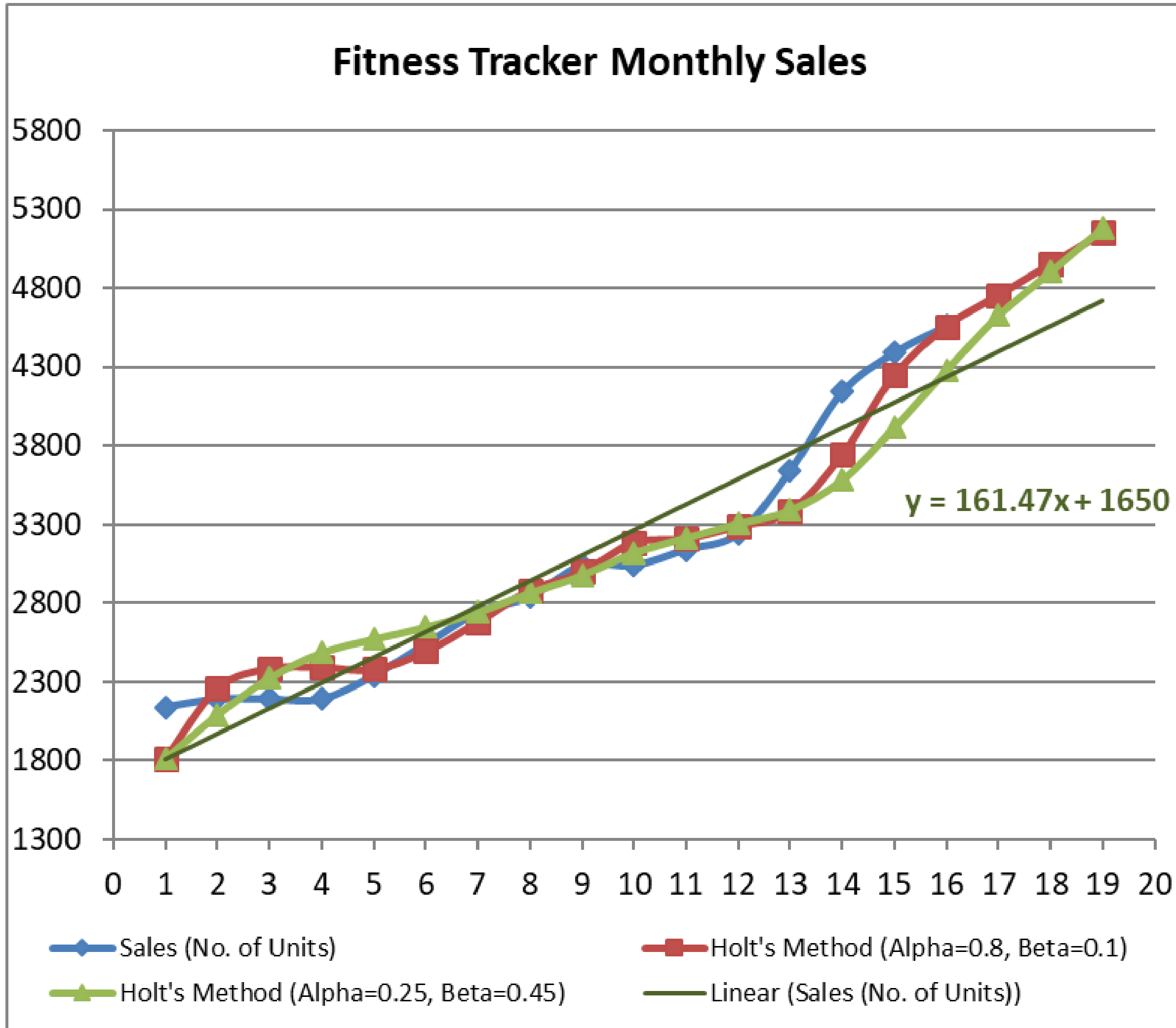
Fitness Tracker Monthly Sales: Forecasting using Holt's Method

- Comparison of different sets of smoothing constants:

  Set 1: α = 0.8 and β = 0.1; Set 2: α = 0.25 and β = 0.45.

The red line (set 1) can better capture the change in the historical data than the green line (set 2) which visibly lags behind the change in the data.

Choice of smoothing constants can significantly affect the performance of the forecasting method.



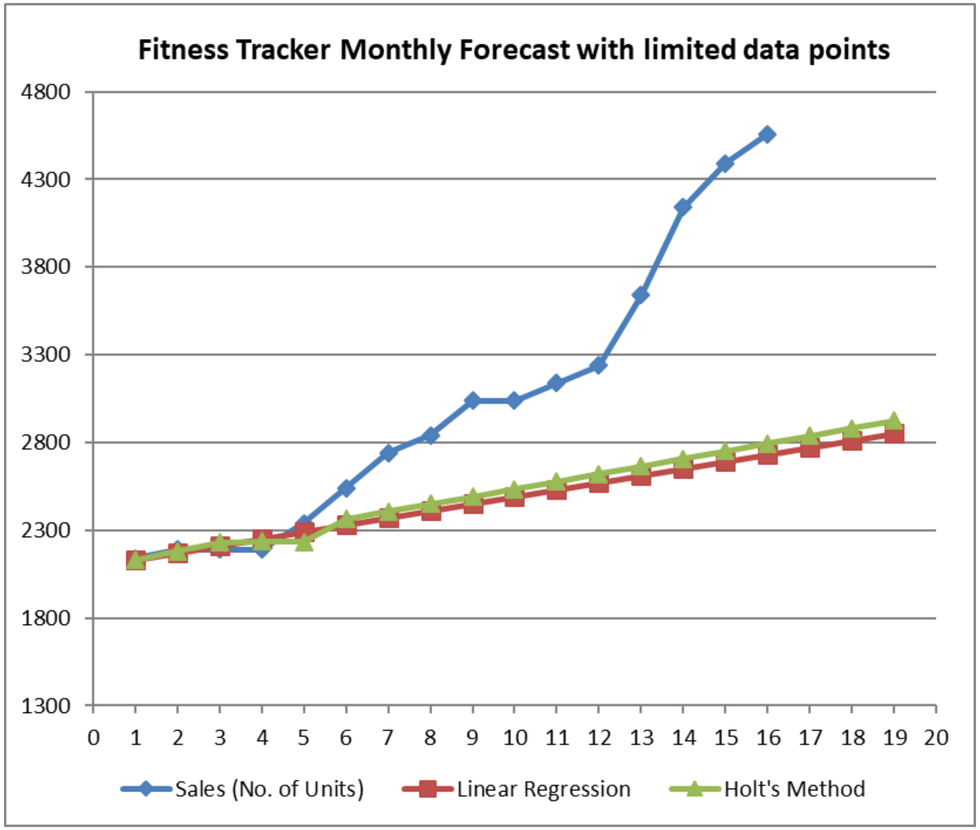
Fitness Tracker Monthly Sales: Which Forecasting Method is Better?

Forecasting methods	Measures of forecasting error		
	MAD	MAPE	MSE
MA (3-Period)	346.41	10.07%	169317.09
MA (5-Period)	540.91	15.41%	350627.27
Weighted MA (3-Period)	297.69	8.69%	125926.92
Exp Smoothing, (alpha = 0.2)	515.83	14.91%	400384.09
Exp Smoothing, (alpha = 0.5)	290.53	8.55%	128689.60
Linear Regression	184.38	6.13%	45672.79
Holt's (α = 0.8, β = 0.1)	131.36	4.64%	29931.90
Holt's (α = 0.25, β = 0.45)	192.53	6.34%	61954.59

The Holt's Method with α = 0.8 and β = 0.1 gives the best forecast in terms of the forecasting errors.

- The Moving Average and Simple Exponential Smoothing methods do not work well because the forecasted values tend to lag when there is a trend in the data set.
  - N-period Moving Average method will lag by (N+1)/2 periods
  - Simple Exponential Smoothing method with smoothing constant α will lag by 1/α periods.
- The better performance of Holt's Method over Linear Regression is due to:
  - The rate of increase in the historical data (Fitness Tracker Monthly Sales) is not constant over the time periods.
  - Linear Regression treats all data in the time series equally.
  - Holt's Method can put more weight on more recent data and therefore can better capture the change in the data.

Fitness Tracker Monthly Sales:
Forecasting with Limited Data Points - 5 Data Points (Jun-17 to Oct-17)



- In Oct-17, there were only 5 data points available.
- If just based on the 5 data points (Jun-17 to Oct-17), forecasts from both Linear Regression model and Holt's method cannot capture the change in the actual sales data well. Forecasting is not a one-off effort. Forecasts should be constantly updated as more data is made available.
- When more data is available, slope and intercept of both methods can better capture the overall trend and level; also they will be less susceptible to one or two points' randomness in data.

Selection of Forecasting Methods

1. Conduct linear regression and conclude whether the trend is significant based on the P value of slope.
2. If there is no trend, moving average and simple exponential smoothing can be used; if there is a trend, choose between linear regression and Holt's method.
3. If recent data's trend is significant, Holt's method may be better. Linear regression treats all data points in time series equally.
4. Forecasting is still only a method of predicting future events and thus, validation should be performed with/when actual data is available. Only when actual data is available, then we can compute the forecasting errors for the methods to decide which is better.

Learning Objectives

- Apply Linear Regression on historical data to determine whether there is any linear relationship between two factors and derive the forecast.
  - Interpret the values of Coefficient of Determination and Correlation coefficient.
  - Derive the forecast from the regression line.
- Perform forecasting on time series data with trend.
  - Describe how Holt’s Method works in forecasting.
  - Apply the Holt’s method to derive forecast for time series data with trend component.
  - Evaluate the impact of smoothing constants on forecasting accuracy of Holt’s Method.

Overview of E211 Operations Planning II Module