Forecasting

Ms Swapnil Shrivastava
swapnil@cdac.in

Content

- Forecasting Overview
- Time Series Forecasting
  - Smoothing methods: Moving Average, Exponential Smoothing
  - Trend projection
  - Seasonal Forecasting
- Causal Forecasting
  - Simple Linear Regression
- Qualitative Methods

Forecasting

- Forecasting is used to predict future aspects of a business operation.
- The success of an organization depends on how well management is able to anticipate the future and develop appropriate strategies.
- Forecasting methods can be classified as
  - Quantitative
  - Qualitative

Quantitative Methods

- Quantitative Forecasting methods can be used when
  - Past information about the variable being forecast is available
  - The information can be quantified
  - A reasonable assumption is that the pattern of the past will continue into the future
- Quantitative forecasting techniques includes
  - Time Series Methods
  - Causal Forecasting Methods

Time Series Method

- A time series is a set of observations of a variable measured at successive points in time or over successive periods of time.
  - Historical sales data
- The time series methods discover a pattern in the historical data and then extrapolate this pattern into the future values of the time series.
- Smoothing Methods, Trend Projection, Seasonal Forecasting

Causal Forecasting Methods

- The assumption is that the variable we want to forecast exhibits a cause-effect relationship with one or more other variables.
- Regression Analysis can be used as causal forecasting method
  - The sales volume of many products is influenced by advertising expenditures, so regression analysis may be used to develop an equation showing how these two variables are related. Then, once the advertising budget has been set for the next period, we can substitute this value into the equation to develop a prediction or forecast of the sales volume for that period.

Qualitative Methods

- Generally involve the use of expert judgement to develop forecasts.
  - A panel of experts might develop a consensus forecast of the prime rate for a year from now.
- They can be applied when the information on the variable being forecast cannot be quantified and when historical data either are not applicable or available.
  - New product launch

Components of Time Series

- Trend Component: it is the gradual shifts or movements to relatively higher or lower values over a longer period of time.
- Cyclical Component: any recurring sequence of points above and below the trend line lasting more than one year.
- Seasonal Component: regular patterns over one year period. E.g. manufacturers of snow removal equipment and heavy clothes have peak sales during winter.
- Irregular Component: includes deviations of actual time series values from those expected given the effects of the trend, cyclical and seasonal components.

Smoothing Methods

- To “ smooth out” random fluctuations caused by the irregular components of the time series.
- Appropriate for a stable time series, that is, one that exhibits no significant trend, cyclical or seasonal effects.
- Moving averages, weighted moving averages and exponential smoothing

Moving Averages

- It uses the average of the most recent n data values in the time series as the forecast for the next period.

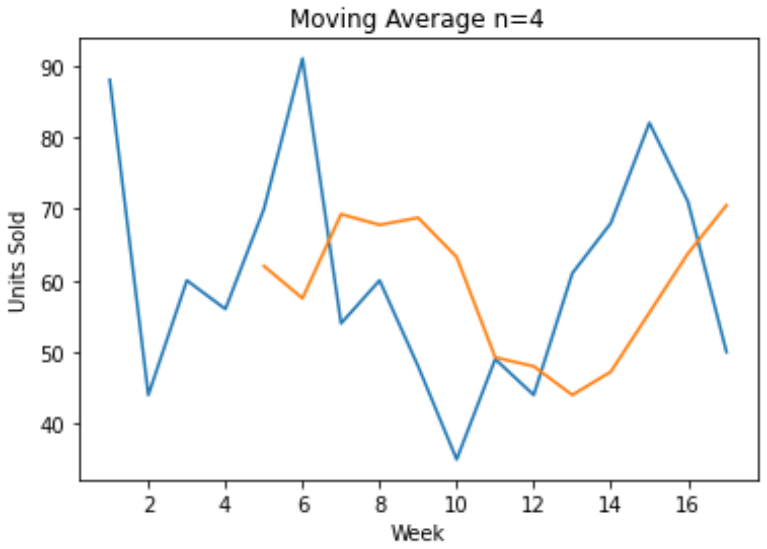
$$\text{Moving average} = \frac{\Sigma\,(\text{most recent n data values})}{n}$$

- The term moving indicates that as a new observation becomes available for the time series, it replaces the oldest observation and a new average is calculated.
- Select the number of data values to be included in the moving average.

Tablet Computer Sales Moving Average

```
import pandas as pd
import matplotlib.pyplot as plt

df=pd.read_excel("F:\Business
Analytics\Data_Files\Tablet Computer
Sales.xlsx", skiprows=2)
print(df)
df["ma_4"]=df["Units
Sold"].rolling(window=4).mean().shift(1)
print(df)
plt.plot(df[df.columns[0]],df[df.columns[1]])
plt.plot(df[df.columns[0]],df[df.columns[2]])
plt.xlabel("Week")
plt.ylabel("Units Sold")
plt.title("Moving Average n=4")
plt.show()
```



Forecast Accuracy

- The quality of a forecast depends on how accurate it is in predicting future values of a time series.
- Different values of n will produce different forecasts. Determine the best value of k that provides high accuracy.
- The error or residual in a forecast is the difference between the forecast and the actual value of the time series (once it is known).
- It is simply the vertical distance between the forecast and the data for the same time period.
- Error metrics are used to determine effectiveness of different forecasting models.

Error Metrics

- Mean Absolute Deviation (MAD): is the absolute difference between the actual value and the forecast, averaged over a range of forecasted values.

$$\text{MAD} = \frac{\sum_{t=1}^{n} |At - Ft|}{n}$$

- Mean Square Error(MSE): squares the difference between the actual value and the forecast, averaged over a range of forecasted values.

$$\text{MSE} = \frac{\sum_{t=1}^{n} (A_t - Ft)^2}{n}$$

Error Metrics

- Root Mean Square Error(RMSE) is square root of MSE.

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n}(A_t - Ft)^2}{n}}$$

- Mean Absolute Percentage Error (MAPE): is the average of absolute error divided by actual observation values.

$$\text{MAPE} = \frac{\sum_{t=1}^{n} \left|\frac{(A_t - Ft)}{A_t}\right|}{n} \text{x } 100$$

Tablet Computer Sales Error Metrics

```
import pandas as pd
from sklearn.metrics import mean_squared_error
from sklearn.metrics import mean_absolute_error
from sklearn.metrics import mean_absolute_percentage_error


df=pd.read_excel("F:\Business Analytics\Data_Files\Tablet Computer Sales.xlsx", skiprows=2)
print(df)
df["ma_4"]=df["Units Sold"].rolling(window=4).mean().shift(1)
print(df)
df=df.dropna(how='any')
print(mean_absolute_percentage_error(df["Units Sold"], df["ma_4"])) #MAPE
print(mean_squared_error(df["Units Sold"], df["ma_4"], squared=False)) #RMSE
print(mean_squared_error(df["Units Sold"], df["ma_4"], squared=True)) #MSE
print(mean_absolute_error(df["Units Sold"], df["ma_4"])) #MAD
```

Exponential Smoothing

- It uses a weighted average of past time series values as the forecast.

$$F_{t+1} = \alpha Y_t + (1- \alpha)F_t$$

where

F_{t+1} – forecast for the times series for period t+1

Y_t – actual value of the time series in period t

F_t – forecast of the time series for period t

α – smoothing constant ($0 <= \alpha <= 1$)

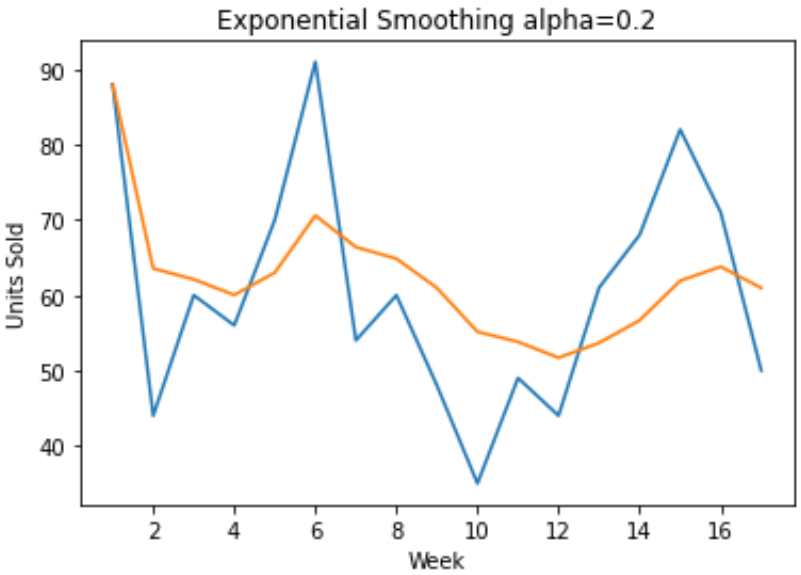
Forecast Accuracy

- If substantial random variability, a small value of α is preferred.
- For relatively little random variable, larger values of α are chosen.
- Choose the value of α that improves forecasting accuracy or reduces forecasting errors.
- Tablet Computer Sales example using Excel's Exponential Smoothing Tool

Tablet Computer Sales Exponential Smoothing

```
import pandas as pd
import matplotlib.pyplot as plt

df=pd.read_excel("F:\Business
Analytics\Data_Files\Tablet Computer
Sales.xlsx", skiprows=2)
print(df)
df["es_0p2"]=df["Units
Sold"].ewm(alpha=0.2).mean()
print(df)
plt.plot(df[df.columns[0]],df[df.columns[1]])
plt.plot(df[df.columns[0]],df[df.columns[2]])
plt.xlabel("Week")
plt.ylabel("Units Sold")
plt.title("Exponential Smoothing alpha=0.2")
plt.show()
```

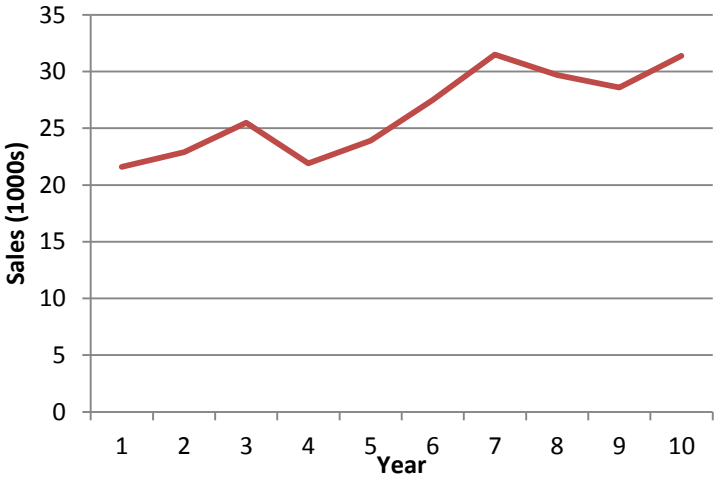


Trend Projection

- To forecast the values of a time series that exhibits a long-term linear trend.
- The type of time series for which the trend projection method is applicable shows a consistent increase or decrease over time.
- Consider the time series for bicycle sales of a particular manufacturer over the past 10 years.

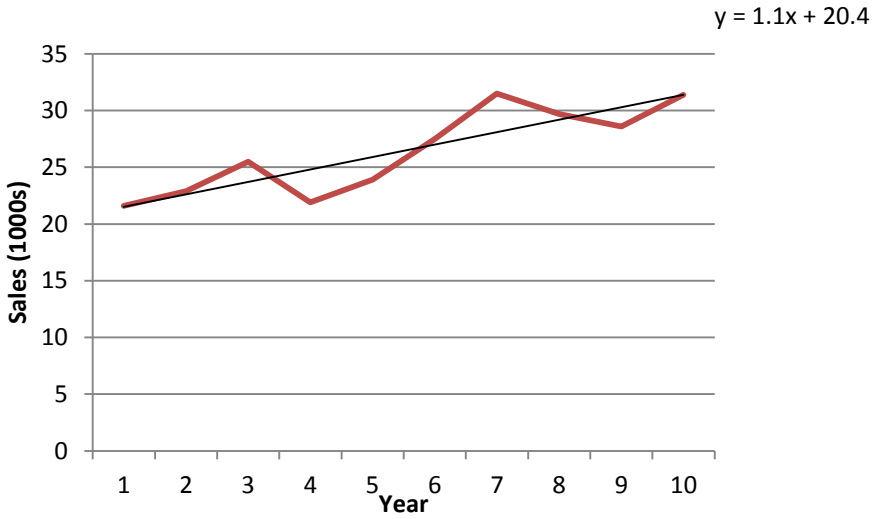
Bicycle Sales Time Series

Year (t)	Sales (1000s) (Yt)
1	21.6
2	22.9
3	25.5
4	21.9
5	23.9
6	27.5
7	31.5
8	29.7
9	28.6
10	31.4



Moving Average
Linear Regression

Trend Represented By A Linear Function for Bicycle Sales



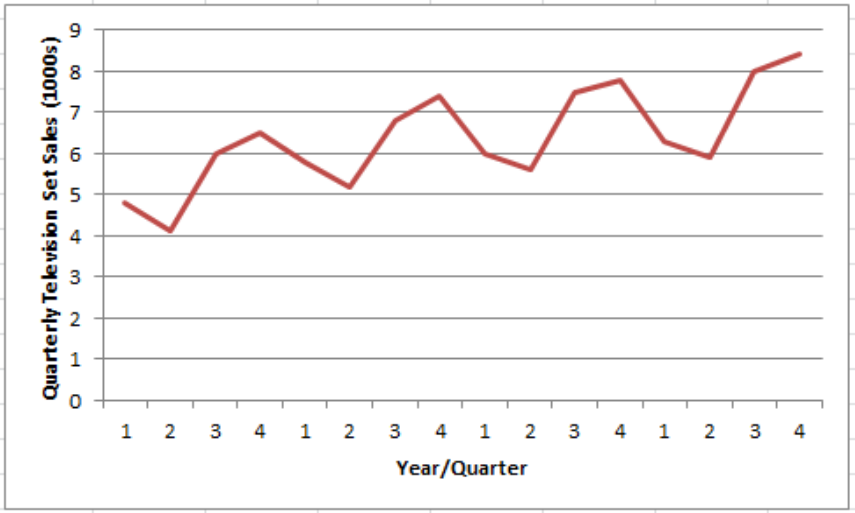
For next year T11 = 20.4 + 1.1 (11) = 32.5
For additional 2 and 3 years in future, T12=33.6 and T13=34.7

Seasonal Forecasting Example: Television Set Sales

To illustrate the use of multiplicative model on quarterly data with trend, seasonal and irregular components. These data show television set sales (in thousands of units) for a particular manufacturer over the past four years.

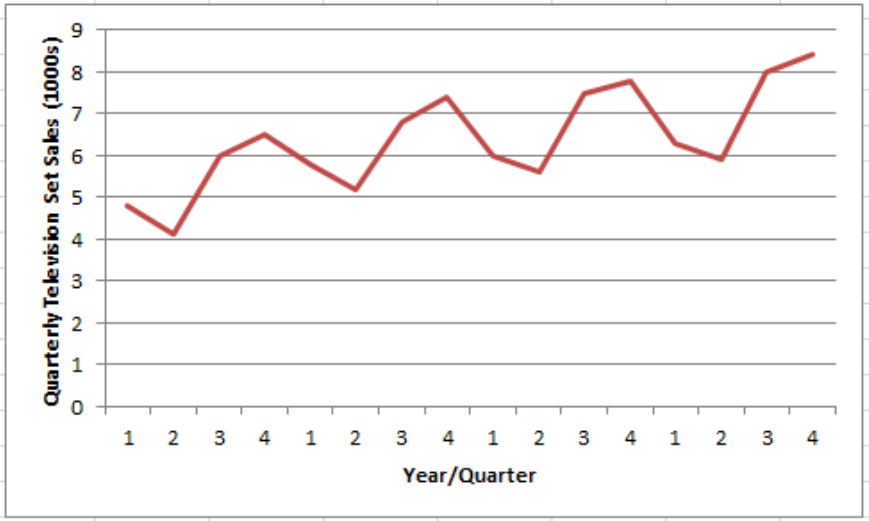
Example: Television Set Sales

Year	Quarter	Sales(1000s)
1	1	4.8
	2	4.1
	3	6
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8
	4	8.4



Example: Step1: Calculate Sample Mean

Year	Quarter	Sales(1000s)
1	1	4.8
	2	4.1
	3	6
	4	6.5
2	1	5.8
	2	5.2
	3	6.8
	4	7.4
3	1	6
	2	5.6
	3	7.5
	4	7.8
4	1	6.3
	2	5.9
	3	8
	4	8.4
	Mean	6.4



Example :Step 2 – Calculate Season Averages

			Year				
		1	2	3	4		
Quarter	1	4.8	5.8	6	6.3		5.7
	2	4.1	5.2	5.6	5.9		5.2
	3	6	6.8	7.5	8		7.1
	4	6.5	7.4	7.8	8.4		7.5
		Mean	6.4				

Example: Step 3 – Calculate Seasonal Factors

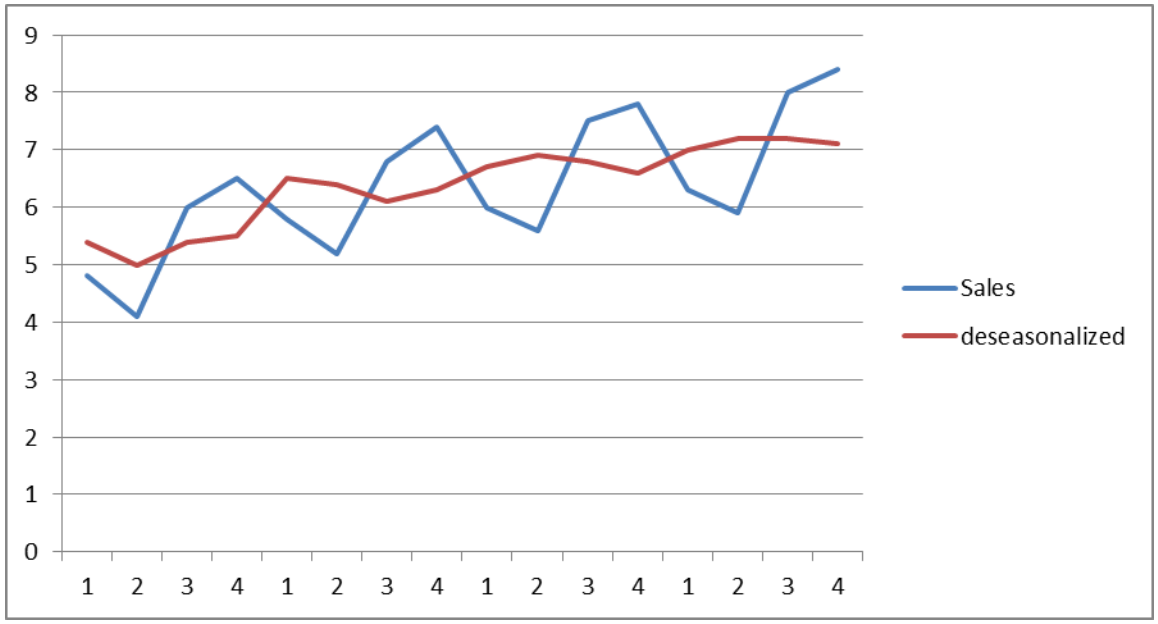
			Year					
		1	2	3	4			**SI**
Quarter	1	4.8	5.8	6	6.3		5.7	0.9
	2	4.1	5.2	5.6	5.9		5.2	0.8
	3	6	6.8	7.5	8		7.1	1.1
	4	6.5	7.4	7.8	8.4		7.5	1.2
		Mean	6.4					

Example: Step4 – De-Seasonalized Data

			Year					
		1	2	3	4			**SI**
Quarter	1	4.8	5.8	6	6.3		5.7	0.9
	2	4.1	5.2	5.6	5.9		5.2	0.8
	3	6	6.8	7.5	8		7.1	1.1
	4	6.5	7.4	7.8	8.4		7.5	1.2
		Mean	6.4					
Deseasonalized Data								
			Year					
		1	2	3	4			
Quarter	1	5.4	6.5	6.7	7.0			
	2	5.0	6.4	6.9	7.2			
	3	5.4	6.1	6.8	7.2			
	4	5.5	6.3	6.6	7.1			

Deseasonalized Data

Deseasonalized Data					
			Year		
		1	2	3	4
Quarter	1	5.4	6.5	6.7	7.0
	2	5.0	6.4	6.9	7.2
	3	5.4	6.1	6.8	7.2
	4	5.5	6.3	6.6	7.1



Forecasting with Seasonality

Deseasonalized Data							
			Year				
		1	2	3	4	**5**	
Quarter	1	5.4	6.5	6.7	7.0	**6.4**	
	2	5.0	6.4	6.9	7.2	**5.8**	
	3	5.4	6.1	6.8	7.2	**7.9**	
	4	5.5	6.3	6.6	7.1	**8.4**	
				Deseasonal forecast=			7.2

- Forecasting for year 5 using moving average of 4 periods.
- Using MA(4). The average of last 4 data points from year 4: 7.2
  - De-seasonal forecast = 7.2
- Forecast for Q3: De-seasonal Forecast * Q3 Seasonal Factor
  - Final Forecast = 7.2 *1.1 = 7.9

Methods for Estimating Seasonal Factors

- Step 1: Sample Mean. Compute the sample mean of the entire data set.
- Step 2: Seasonal Averages. Average the observations for the N like periods in the data.
  - For example, average all summers, winters, etc.
- Step 3: Seasonal Factors. Divide the averages from Step 2 by the sample mean.
  - The resulting N numbers will exactly add to N and correspond to the N seasonal factors.
- Step 4: De-seasonalization: To remove seasonality from a series, simply divide each observation in the data by the appropriate seasonal factor.
  - The resulting series will have no seasonality and is called a de-seasonalized series.

Regression Analysis

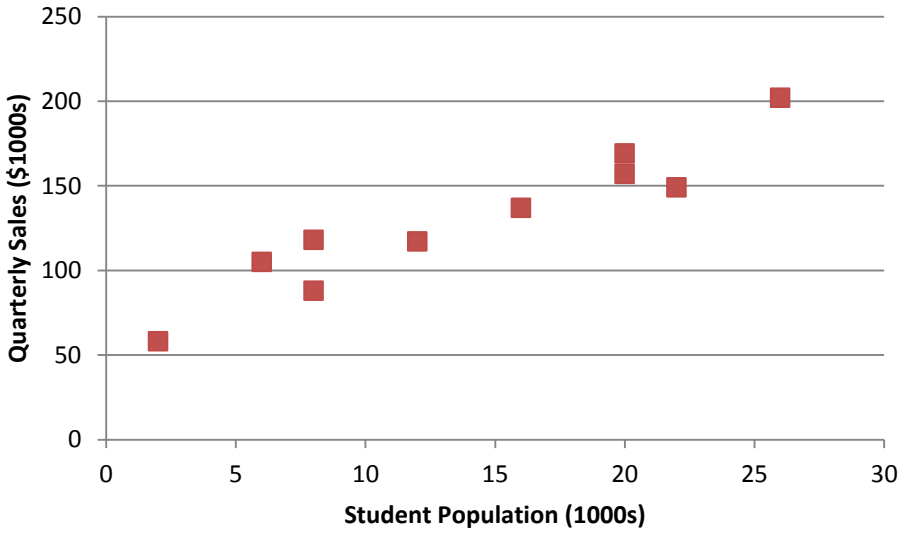
- It is a statistical technique that can be used to develop a mathematical equation showing how variables are related.
  - Dependent or response variable: the variable being predicted
  - Independent or predictor variable: variable or variables being used to predict the value of the dependent variable.
- Simple linear regression: Regression analysis involving one independent variable and one dependent variable for which the relationship between the variables is approximated by a straight line.

Using Regression Analysis as a Causal Forecasting Method

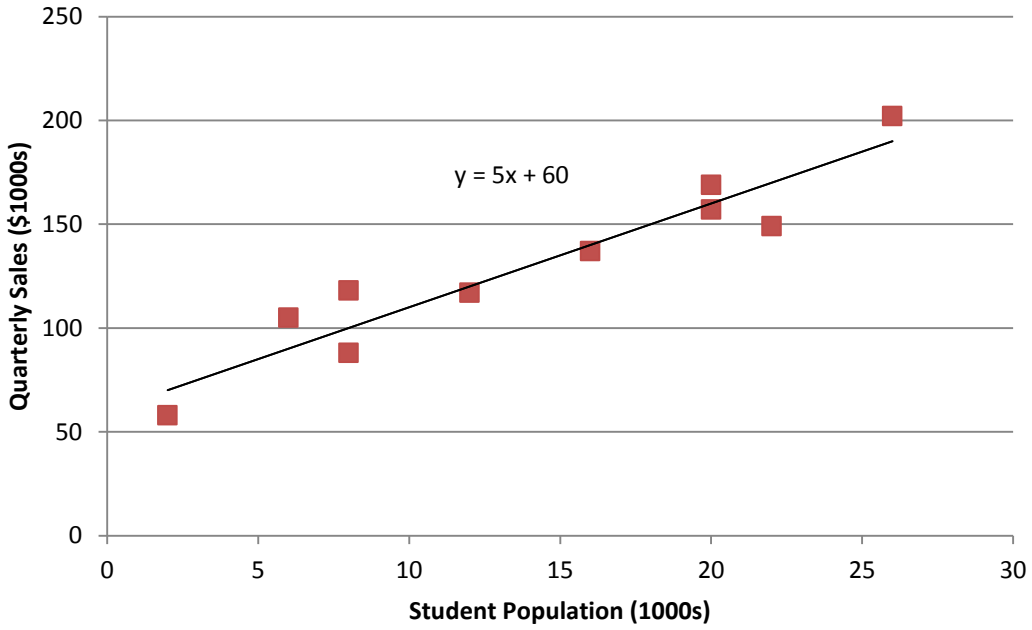
Armand's Pizza Parlors, a chain of Italian restaurants doing business in a five-state area. The most successful locations have been near college campuses. The managers believe that quarterly sales for these restaurants (denoted by y) are related positively to the size of the student population (denoted by x); that is, restaurants near campuses with a large population tend to generate more sales than those located near the campuses with a small population. Using regression model analysis we can develop an equation showing how the dependent variable y is related to the independent variable x. This equation can then be used to forecast quarterly sales for restaurants located near college campuses given the size of the student population.

Data on Quarterly Sales and Student Population for 10 Restaurants

Restaurant	Y= Quarterly Sales ($1000s)	X= Student Population (1000s)
1	58	2
2	105	6
3	88	8
4	118	8
5	117	12
6	137	16
7	157	20
8	169	20
9	149	22
10	202	26



The estimated Regression Equation for Armand's Pizza Parlors



The quarterly sales for a new restaurant to be located near a campus with 16,000 students is forecasted as $140,000

New Product Problem

- In the case of new products or new designs in the market, there is limited demand data.
- How to Forecast in such cases?

Qualitative Approaches

- Delphi Method: a method that obtains forecasts through group consensus.
- Expert Judgement: based on the judgement of a single expert or represent the consensus of a group of experts.
- Scenario Writing: a method that consists of developing a conceptual scenario of the future based on a well-defined set of assumptions.
- Intuitive Approaches: are based on the ability of the human mind to process information that, in most cases, is difficulty to quantify.

Summary

- Forecasting techniques help managers develop appropriate strategies for future.
- Quantitative forecasting methods include time series methods and causal methods.
- Qualitative forecasting methods may be used when little or no historical data are available.
- Application in production schedules, raw material purchasing plan, inventory policies and sales quota.