Agenda

- Time Series - patterns
- Moving Average
- Exponential Smoothing
- Regression

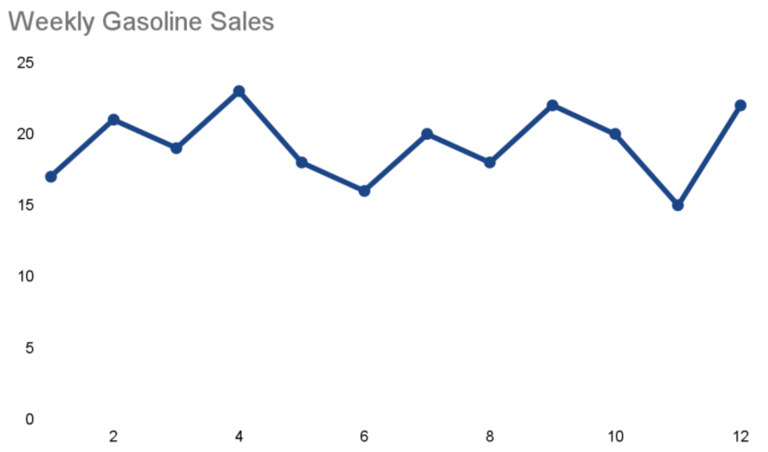
Time Series

- **Time series:** A sequence of observations on a variable measured at successive points in time.
- The measurements may be taken every hour, day, week, month, year, or any other regular interval. The pattern of the data is important in understanding the series' past behavior.
- If the behavior of the times series data of the past is expected to continue in the future, it can be used as a guide in selecting an appropriate forecasting method.

Time Series: Horizontal pattern

- Exists when the data fluctuate randomly around a constant mean over time.

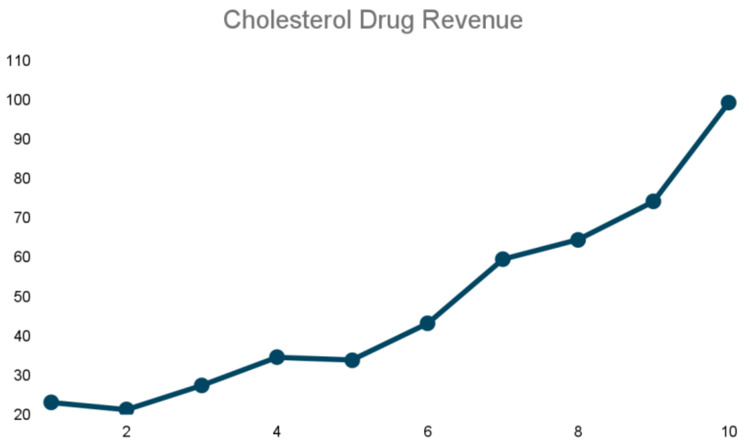
- **Stationary time series**: It denotes a time series whose statistical properties are independent of time:
  - The process generating the data has a **constant mean**.
  - The **variability** of the time series is **constant** over time.



Time Series: Trend pattern

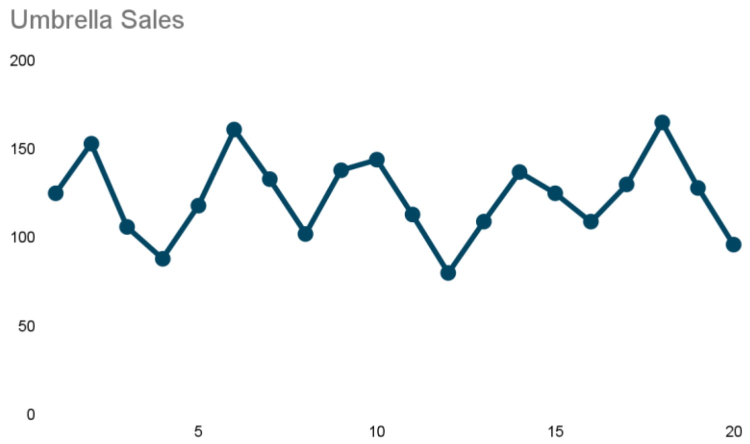
Trend Pattern:

- A **trend pattern** shows gradual shifts or movements to relatively higher or lower values over a longer period of time.
    - Population increases or decreases.
    - Shifting demographic characteristics of the population.
    - Improving technology.
    - Changes in the competitive landscape.
    - Changes in consumer preferences.



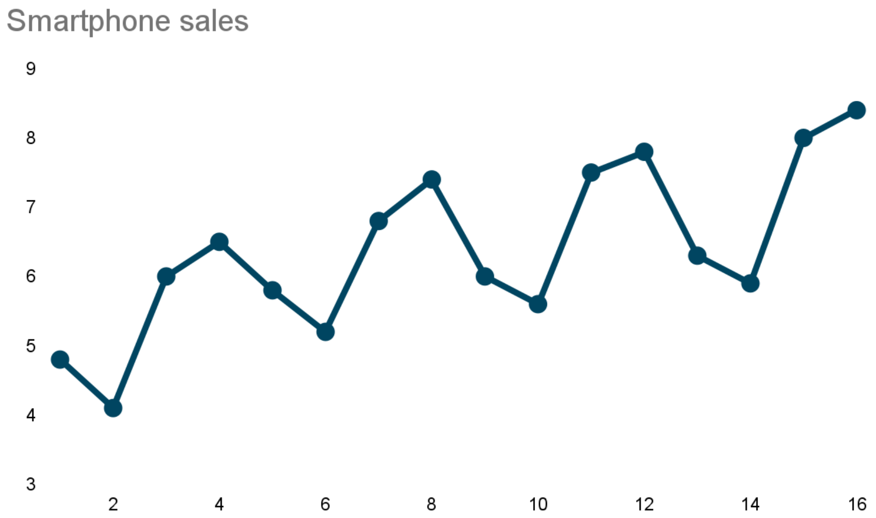
Time Series: Seasonal pattern

- **Seasonal patterns** are recurring patterns over successive periods of time.
  - Example: A retailer that sells bathing suits expects low sales activity in the fall and winter months, with peak sales in the spring and summer months to occur every year.
- The time series plot not only exhibits a seasonal pattern over a one-year period but also for less than one year in duration.
  - Example: daily traffic volume shows within-the-day “seasonal” behavior, with peak levels occurring during rush hour, moderate flow during the rest of the day, and light flow from midnight to early morning.



Time Series: Trend + Seasonality

Some time series include both a trend and a seasonal pattern



Time Series: Cyclic pattern

- A **cyclical pattern** exists if the time series plot shows an alternating sequence of points below and above the trendline that lasts for more than one year.
    - Example: Periods of moderate inflation followed by periods of rapid inflation can lead to a time series that alternates below and above a generally increasing trendline.
- Cyclical effects are often combined with long-term trend effects and referred to as trend-cycle effects.

Time Series pattern - discussion

- The underlying pattern in the time series is an important factor in selecting a forecasting method.
- A time series plot should be one of the first analytic tools.
- We need to use a forecasting method that is capable of handling the pattern exhibited by the time series effectively.

Introduction

- **Naïve forecasting method**: Using the most recent data to predict future data.
- The key concept associated with measuring forecast accuracy is **forecast error**.
- **Measures** to determine how well a particular forecasting method is able to reproduce the time series data that are already available.

FORECAST ERROR

$$e_t = y_t - \hat{y}_t$$

MEAN FORECAST ERROR (MFE)

$$\text{MFE} = \frac{\sum_{t=k+1}^{n} e_t}{n-k}$$

MEAN ABSOLUTE ERROR (MAE)

$$\text{MAE} = \frac{\sum_{t=k+1}^{n} |e_t|}{n-k}$$

MEAN SQUARED ERROR (MSE)

$$\text{MSE} = \frac{\sum_{t=k+1}^{n} e_t^2}{n-k}$$

MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

$$\text{MAPE} = \frac{\sum_{t=k+1}^{n} \left| \left(\frac{e_t}{y_t} \right) 100 \right|}{n-k}$$

Introduction

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21	17.00	4.00	4.00	16.00	19.05	19.05
3	19	19.00	0.00	0.00	0.00	0.00	0.00
4	23	19.00	4.00	4.00	16.00	17.39	17.39
5	18	20.00	−2.00	2.00	4.00	−11.11	11.11
6	16	19.60	−3.60	3.60	12.96	−22.50	22.50
7	20	19.00	1.00	1.00	1.00	5.00	5.00
8	18	19.14	−1.14	1.14	1.31	−6.35	6.35
9	22	19.00	3.00	3.00	9.00	13.64	13.64
10	20	19.33	0.67	0.67	0.44	3.33	3.33
11	15	19.40	−4.40	4.40	19.36	−29.33	29.33
12	22	19.00	3.00	3.00	9.00	13.64	13.64
		Totals	4.52	26.81	89.07	2.75	141.34

	Naïve Method	Average of Past Values
MAE	3.73	2.44
MSE	16.27	8.10
MAPE	19.24%	12.85%

The average of past values provides more accurate forecasts for the next period than using the most recent observation

Moving Average

Uses the average of the most recent k data values in the time series as the forecast for the next period.

MOVING AVERAGE FORECAST

$$\hat{y}_{t+1} = \frac{\Sigma(\text{most recent } k \text{ data values})}{k} = \frac{\sum_{i=t-k+1}^{t} y_i}{k} = \frac{y_{t-k+1} + \cdots + y_{t-1} + y_t}{k} \tag{8.6}$$

where

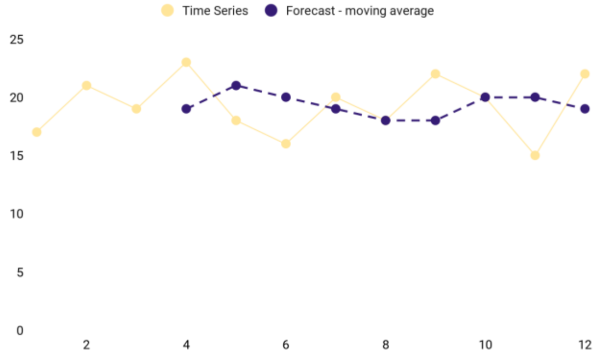
$\hat{y}_{t+1}$ = forecast of the time series for period $t+1$

y_t = actual value of the time series in period t

k = number of periods of time series data used to generate the forecast

Moving Average

Week	Time Series Value	Forecast	Forecast Error	Absolute Value of Forecast Error	Squared Forecast Error	Percentage Error	Absolute Value of Percentage Error
1	17						
2	21						
3	19						
4	23	19	4	4	16	17.39	17.39
5	18	21	–3	3	9	–16.67	16.67
6	16	20	–4	4	16	–25.00	25.00
7	20	19	1	1	1	5.00	5.00
8	18	18	0	0	0	0.00	0.00
9	22	18	4	4	16	18.18	18.18
10	20	20	0	0	0	0.00	0.00
11	15	20	–5	5	25	–33.33	33.33
12	22	19	3	3	9	13.64	13.64
		Totals	**0**	**24**	**92**	**–20.79**	**129.21**



$$MAE = \frac{\sum_{t=4}^{12} |e_t|}{n-3} = \frac{24}{9} = 2.67$$

$$MSE = \frac{\sum_{t=4}^{12} |e_t^2|}{n-3} = \frac{92}{9} = 10.22$$

$$MAPE = \frac{\sum_{t=4}^{12} \left| \left(\frac{e_t}{y_t} \right) 100 \right|}{n-3} = \frac{129.21}{9} = 14.36\%$$

Exponential Smoothing

Uses a weighted average of past time series values as a forecast.

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

Smoothing constant (α)- weight given to the actual period in time, t and weight given to the forecast is $1-\alpha$.

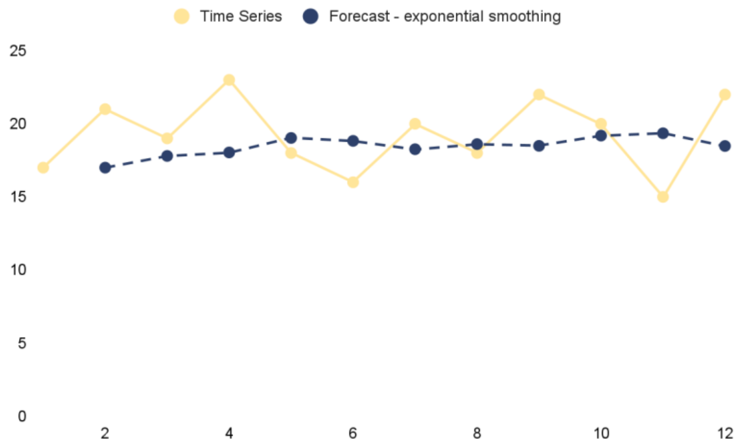
$$\hat{y}_2 = \alpha y_1 + (1-\alpha)\hat{y}_1$$

Forecast for period 2

Exponential Smoothing

$\alpha = 0.2$

Week	Time Series Value	Forecast	Forecast Error	Squared Forecast Error
1	17			
2	21	17.00	4.00	16.00
3	19	17.80	1.20	1.44
4	23	18.04	4.96	24.60
5	18	19.03	−1.03	1.06
6	16	18.83	−2.83	8.01
7	20	18.26	1.74	3.03
8	18	18.61	−0.61	0.37
9	22	18.49	3.51	12.32
10	20	19.19	0.81	0.66
11	15	19.35	−4.35	18.92
12	22	18.48	3.52	12.39
		Totals	**10.92**	**98.80**



Using regression for forecasting

- Regression → the linear relationship between the independent variable and the dependent variable that minimizes the MSE.
- Regression analysis can be used to forecast a time series with a linear trend.

Dependent variable → **Variable to be forecasted**

Independent variable → **Time period**

$$\hat{y}_t = b_0 + b_1 t$$

Trendline

$$\hat{y}_t = 20.4 + 1.1t$$

Substituting, t = 11 → 32,500 bicycles sale in the next time period

$$\hat{y}_{11} = 20.4 + 1.1(11) = 32.5$$

Using regression for forecasting - Complexity

Non-linear trend → $\hat{y}_t = b_0 + b_1 t + b_2 t^2 + b_3 t^3$

Autoregressive model → $\hat{y}_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + b_3 y_{t-3}$

Trendline

Independent variables → previous values of the time series

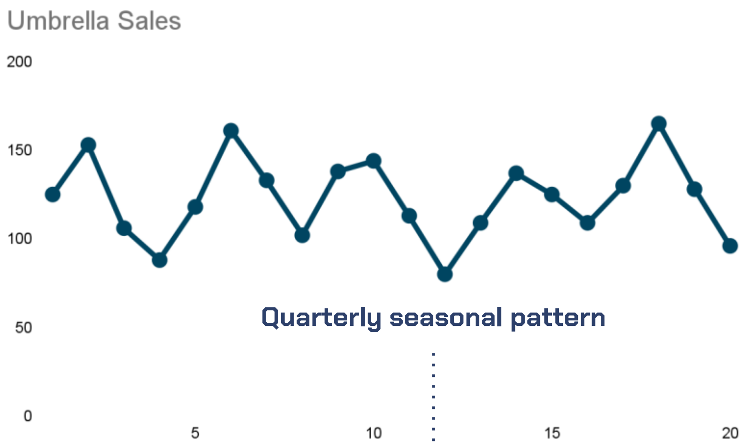
Using regression for forecasting - Seasonality

Seasonality without ***trend***

season → **dummy variable**

$\text{Qtr1}_t = 1$ if period t is quarter 1; 0 otherwise.
$\text{Qtr2}_t = 1$ if period t is quarter 2; 0 otherwise.
$\text{Qtr3}_t = 1$ if period t is quarter 3; 0 otherwise.

$$\hat{y}_t = b_0 + b_1\text{Qtr1}_t + b_2\text{Qtr2}_t + b_3\text{Qtr3}_t$$



Quarterly seasonal pattern

(k-1) dummy variables = 4-1 = ***3***

Using regression for forecasting - Seasonality

Period	Year	Quarter	Qtr1	Qtr2	Qtr3	Sales (1,000s)
1	1	1	1	0	0	4.8
2		2	0	1	0	4.1
3		3	0	0	1	6.0
4		4	0	0	0	6.5
5	2	1	1	0	0	5.8
6		2	0	1	0	5.2
7		3	0	0	1	6.8
8		4	0	0	0	7.4
9	3	1	1	0	0	6.0
10		2	0	1	0	5.6
11		3	0	0	1	7.5
12		4	0	0	0	7.8
13	4	1	1	0	0	6.3
14		2	0	1	0	5.9
15		3	0	0	1	8.0
16		4	0	0	0	8.4

Sales = 6.07 – 1.36 Qtr1 – 2.03 Qtr2 – 0.304 Qtr3 + 0.146 t

- *Quarter 1*: Sales = 4.71 + 0.146*t*
- *Quarter 2:* Sales = 4.04 + 0.146*t*
- *Quarter 3:* Sales = 5.77 + 0.146*t*
- *Quarter 4:* Sales = 6.07 + 0.146*t*

Using regression - Causal forecasting

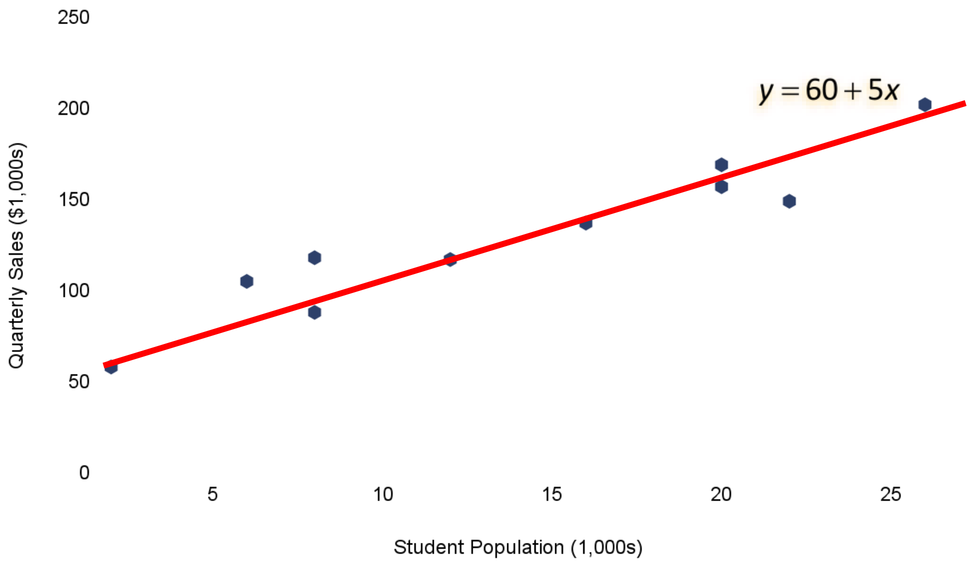
- The relationship of the variable to be forecast with other variables may also be used to develop a forecasting model.

  Examples:

  - Advertising expenditures when sales are to be forecast.
  - The mortgage rate when new housing construction is to be forecast.
  - Grade point average when starting salaries for recent college graduates are to be forecast.
  - The price of a product when the demand for the product is to be forecast.
  - The value of the Dow Jones Industrial Average when the value of an individual stock is to be forecast.
  - Daily high temperature when electricity usage is to be forecast.

Using regression - Causal forecasting

Restaurant	Student Population (1,000s)	Quarterly Sales ($1,000s)
1	2	58
2	6	105
3	8	88
4	8	118
5	12	117
6	16	137
7	20	157
8	20	169
9	22	149
10	26	202



Determining the best model for forecasting

- A visual inspection can indicate whether seasonality appears to be a factor and whether a linear or nonlinear trend seems to exist.
- For causal modeling, scatter charts can indicate whether strong linear or nonlinear relationships exist between the independent and dependent variables.
- If certain relationships appear totally random, this may lead you to exclude these variables from the model.
- While working with large data sets, it is recommended to divide your data into training and validation sets.
- Based on the errors produced by the different models for the validation set, you can pick the model that minimizes some forecast error measure, such as MAE, MSE or MAPE.
- There are software packages that will automatically select the best model to use.
- Ultimately, the user should decide which model to use based on the software output and his managerial knowledge.

Summary

Time Series - patterns

Moving Average

Exponential Smoothing

Regression