Session 3: Forecasting & Game Theory

Introduction to Forecasting

Agenda

- Definition & Demand Patterns
- Simple Forecasting Methods
- Forecast Error

Definition

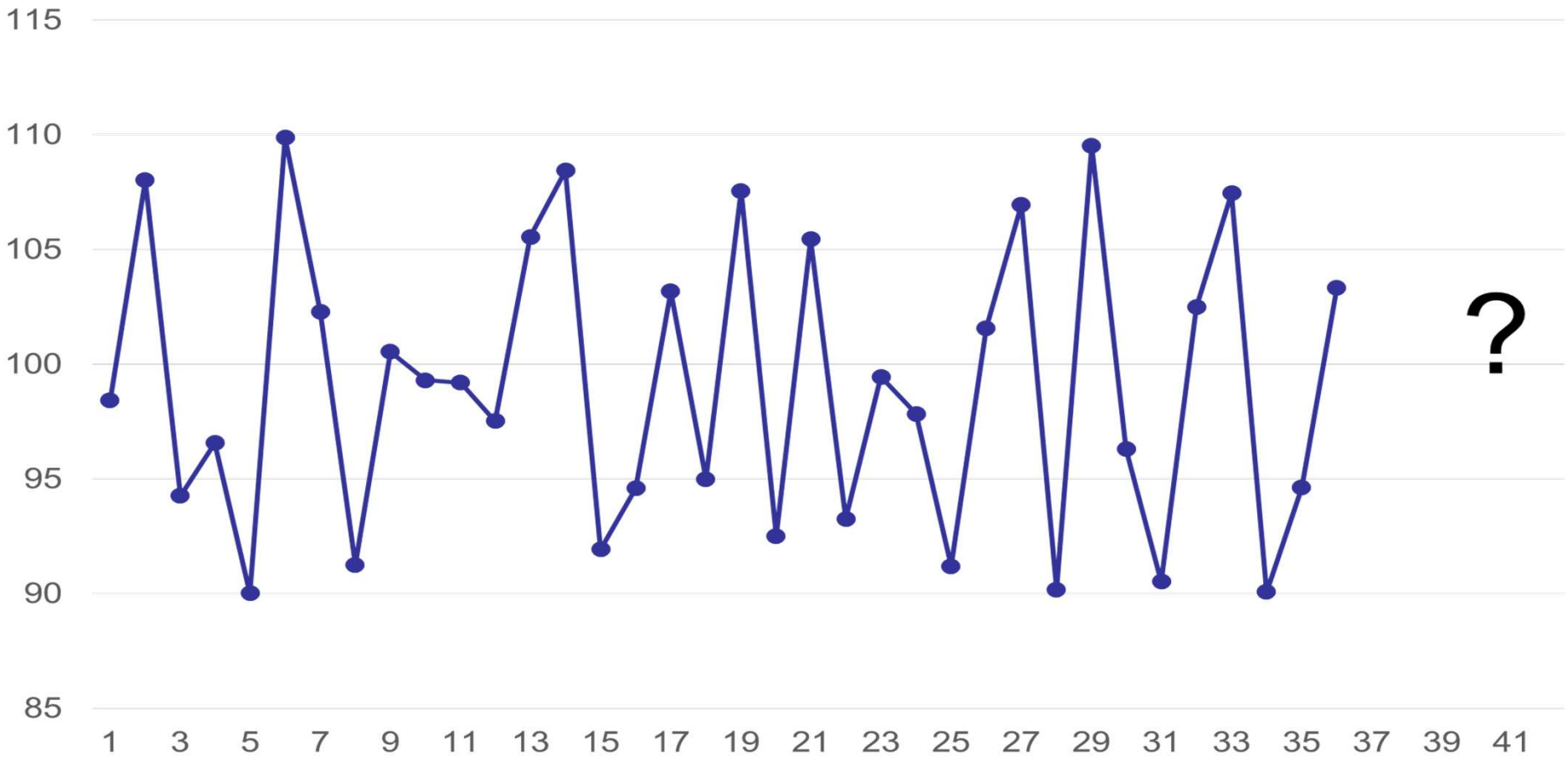
What is Forecasting?

- It is the process of creating statements about outcomes of variables that presently are uncertain and will only be realized in the future.

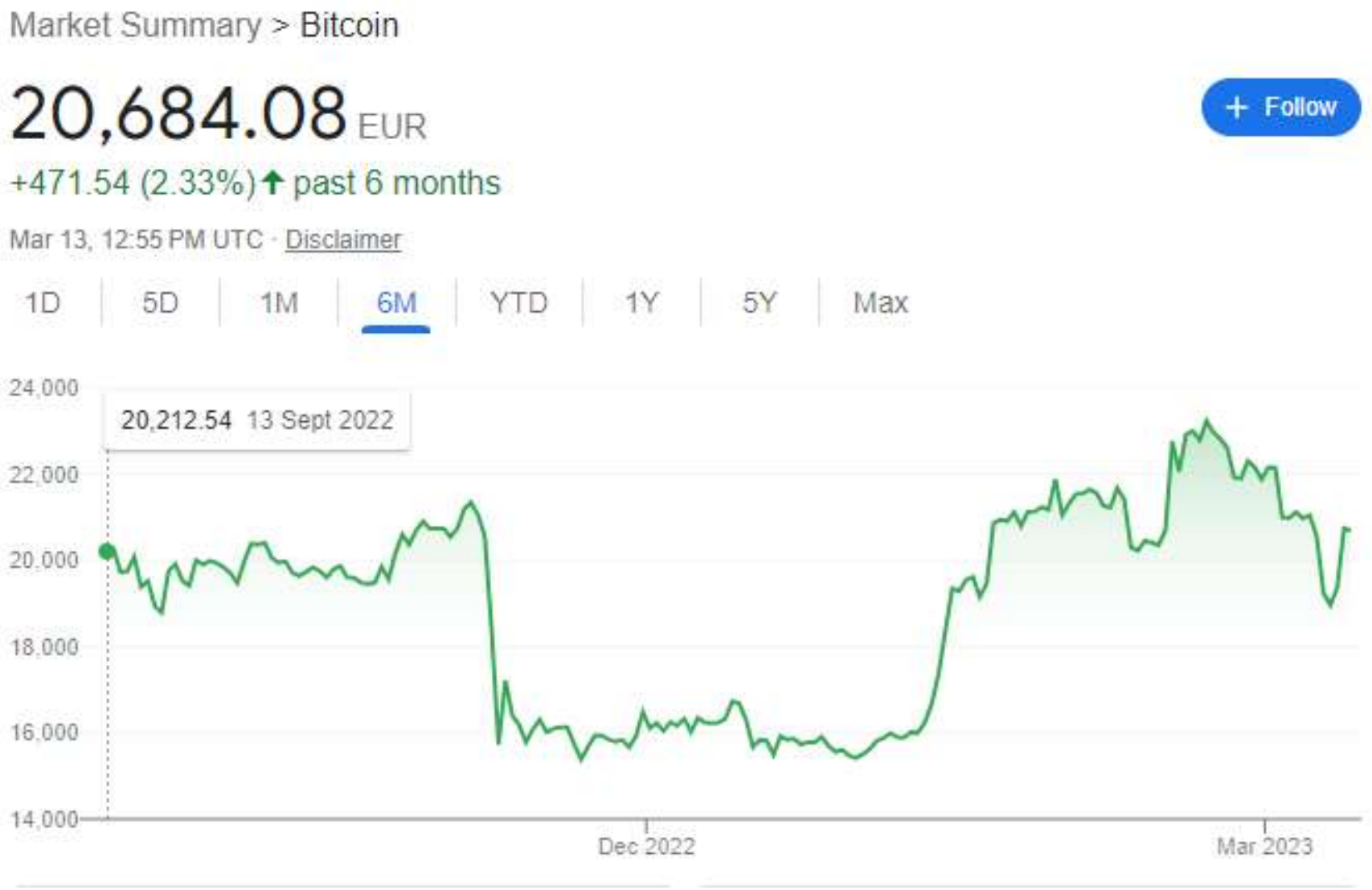
How?

- Mainly based on past and present data and analysis of trends

What is next?


115
110
105
100
95
90
85
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
?

What is next?


Market Summary > Bitcoin
20,684.08 EUR
+ Follow
+471.54 (2.33%) past 6 months
Mar 13, 12:55 PM UTC · Disclaimer
1D
5D
1M
6M
YTD
1Y
5Y
Max
24,000
22,000
20,000
18,000
16,000
14,000
20,212.54 13 Sept 2022
Dec 2022
Mar 2023

What is next?


Market Summary > Bitcoin
61,918.15 EUR
+ Follow
+37,734.91 (156.04%) ↑ past 6 months
Mar 8, 12:04 PM UTC · Disclaimer
1D 5D 1M 6M YTD 1Y 5Y Max
70,000
60,000
50,000
40,000
30,000
20,000
24,183.24 8 Sept 2023
Dec 2023
Mar 2024

What is next?


Market Summary > Bitcoin
92,868.47 EUR
+36,213.14 (63.92%) ↑ past year
+ Follow
May 13, 1:51 PM UTC · Disclaimer
1D
5D
1M
6M
YTD
1Y
5Y
Max
120,000
100,000
80,000
60,000
40,000
Oct 2024
Mar 2025

Definition

Advantages

- Reduce uncertainty and anticipate change(s) in the market

Disadvantages

- No one can be absolutely sure what the future holds
- *Silent assumption*: the future will be like the past; there is no room for human intelligence

Example #1

- Thomas Watson (CEO of IBM), forecasted the demand for computers.
- He predicted that the world market demand for computers would be 5. Yes, you read correctly, not 5 million, just 5 computers. In his defense, he made this forecast in the 1950s.

Example #2

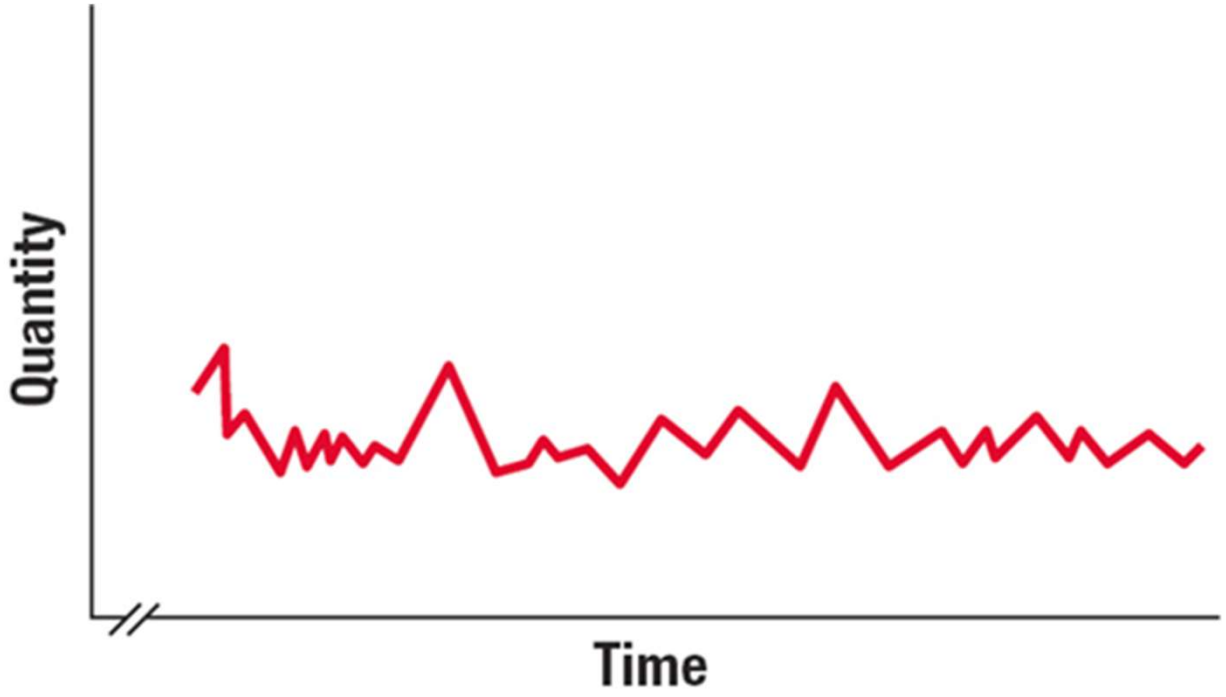
- In the 1960s, the managers at Decca Recording were offered the opportunity to publish the music of a Liverpool guitar band. Decca's forecast for sales of this band was pessimistic "guitar groups are on the way out" was the management consensus.
- Unfortunately for Decca, the band that they rejected was *the Beatles*, a band that subsequently went on to become one of the most successful music bands in history.

Patterns

- **Time series**
  - The repeated observations of data points in their order of occurrence (time order)
- **There are five basic time series patterns**
  1. Horizontal
  2. Trend
  3. Seasonal
  4. Cyclical
  5. Random

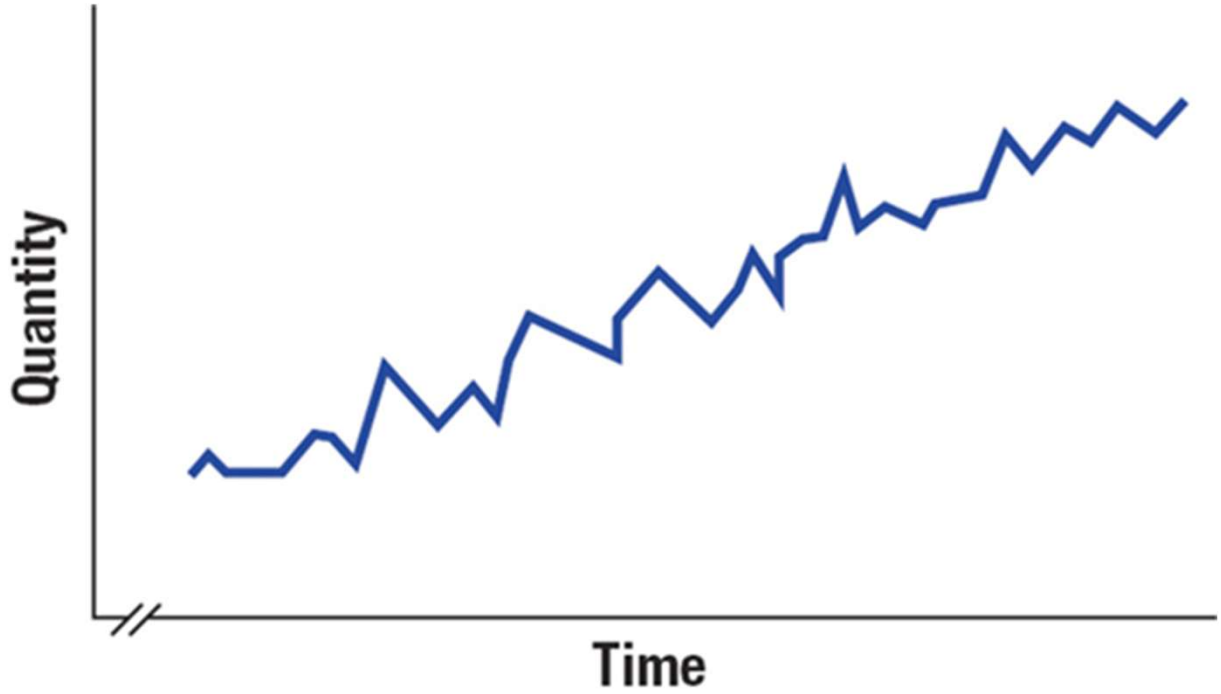
1. Horizontal

A horizontal pattern exists when data values fluctuate around a constant mean



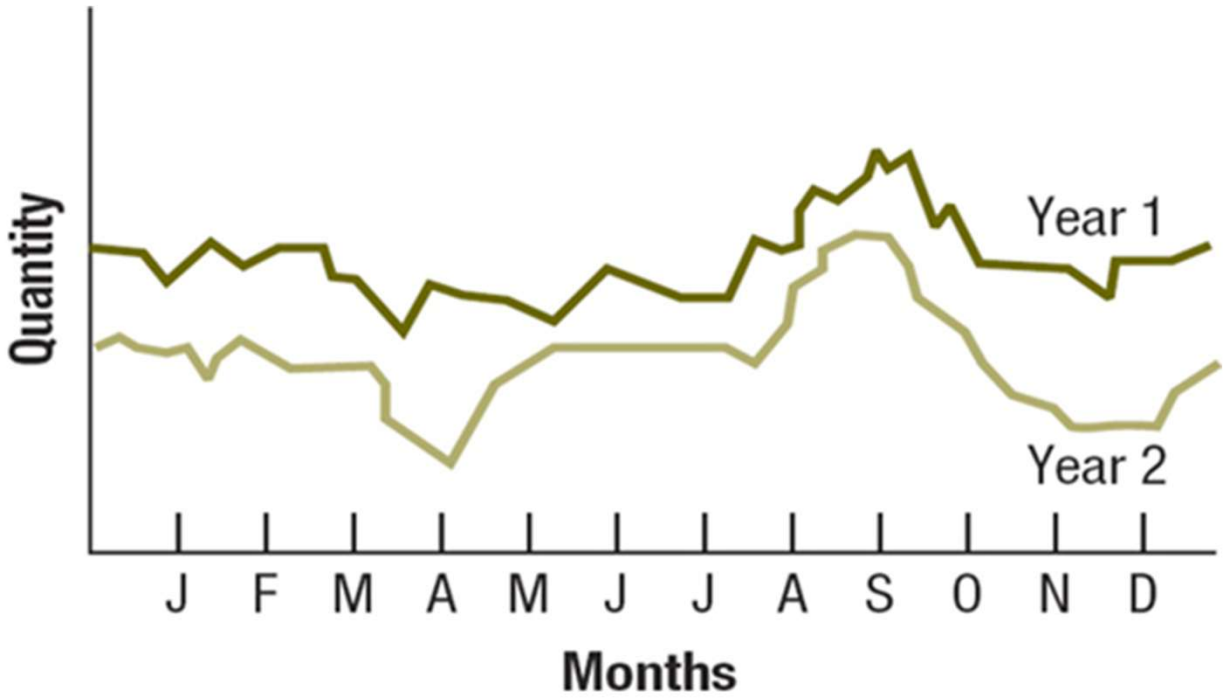
2. Trend

A trend exists when there is a long-term increase or decrease in the data. It does not have to be linear. Sometimes we will refer to a trend “changing direction” when it might go from an increasing trend to a decreasing trend



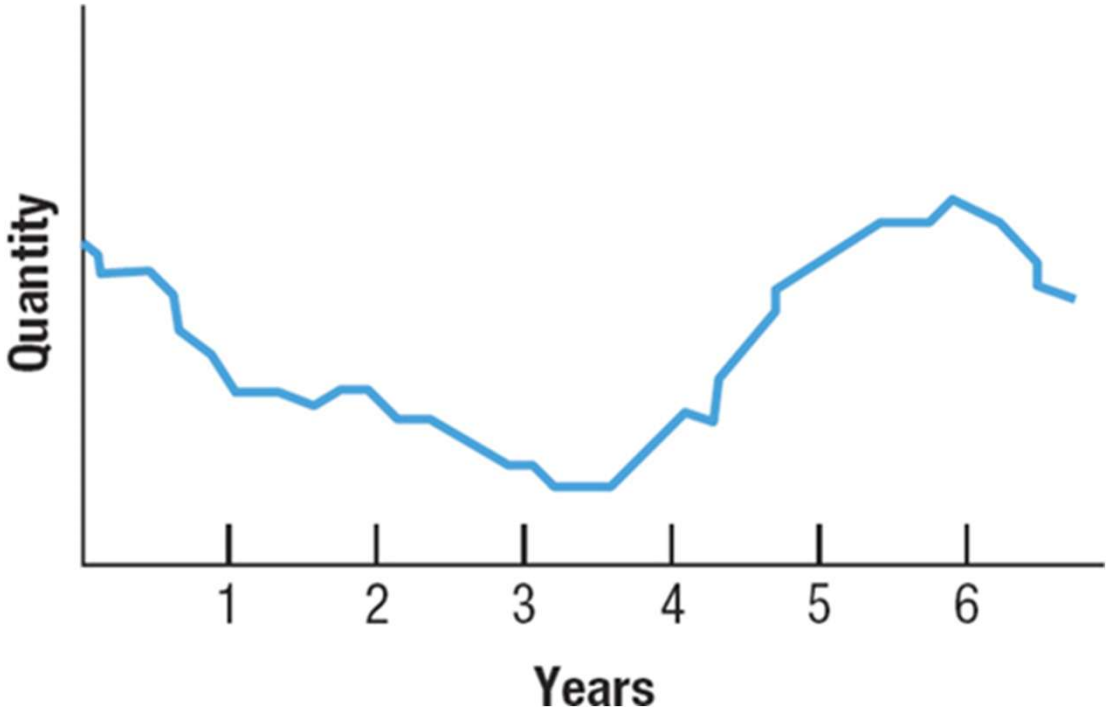
3. Seasonal

A seasonal pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week). Seasonality is always of a fixed and known period



4. Cyclical

A cyclic pattern exists when data exhibit rises and falls that are not of fixed period. The duration of these fluctuations is usually of at least 2 years



5. Random

No specific pattern can be identified

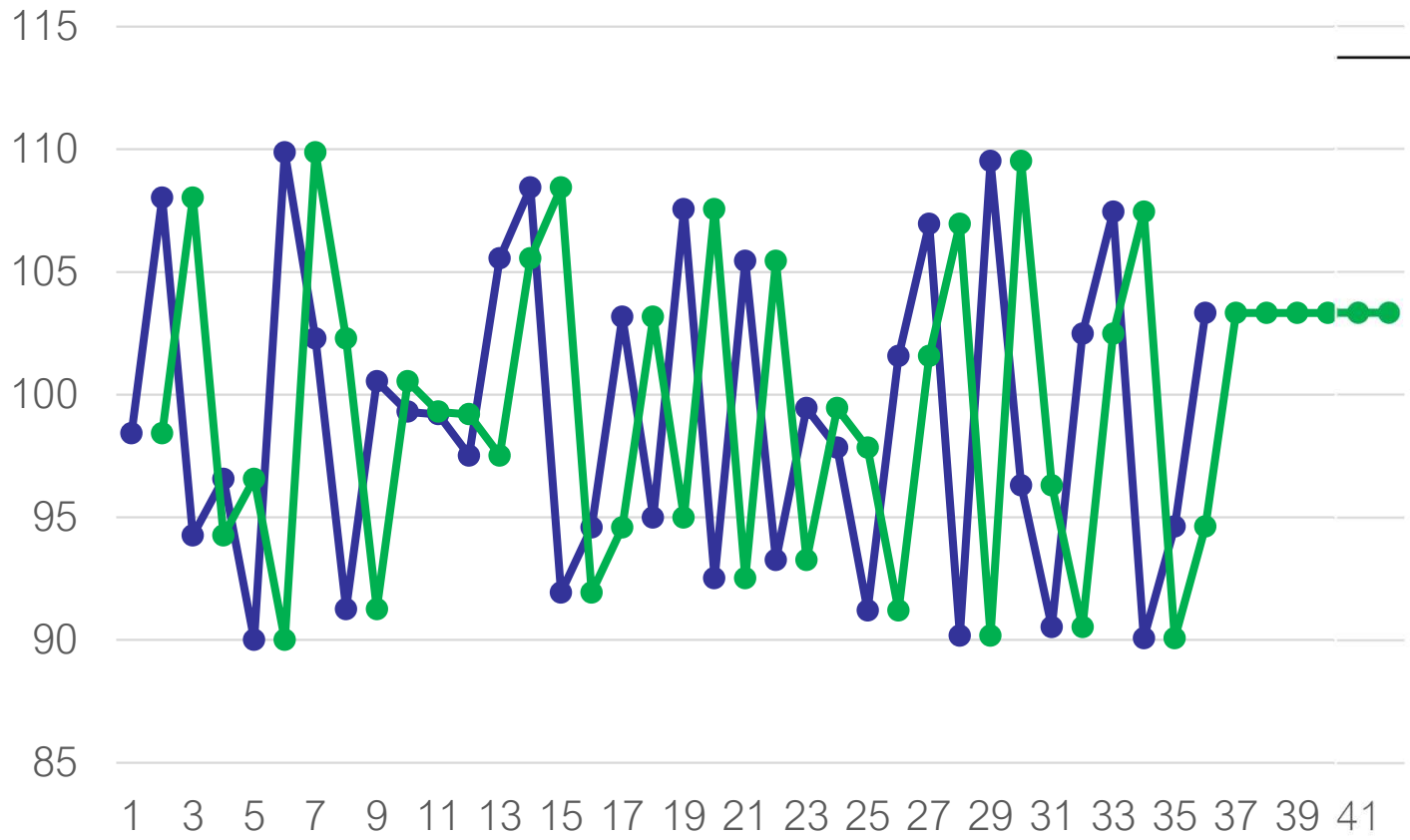
Forecasting Methods

- Naïve method
- Global Average method
- Moving Average method
- Weighted Moving Average method

Naïve Method

- The forecast for the next period equals the demand for the current period $\hat{y}_{t+1} = y_t$
- Comments:
  - Only the very last observation is considered
  - No smoothing or averaging is performed
  - Shortest memory possible

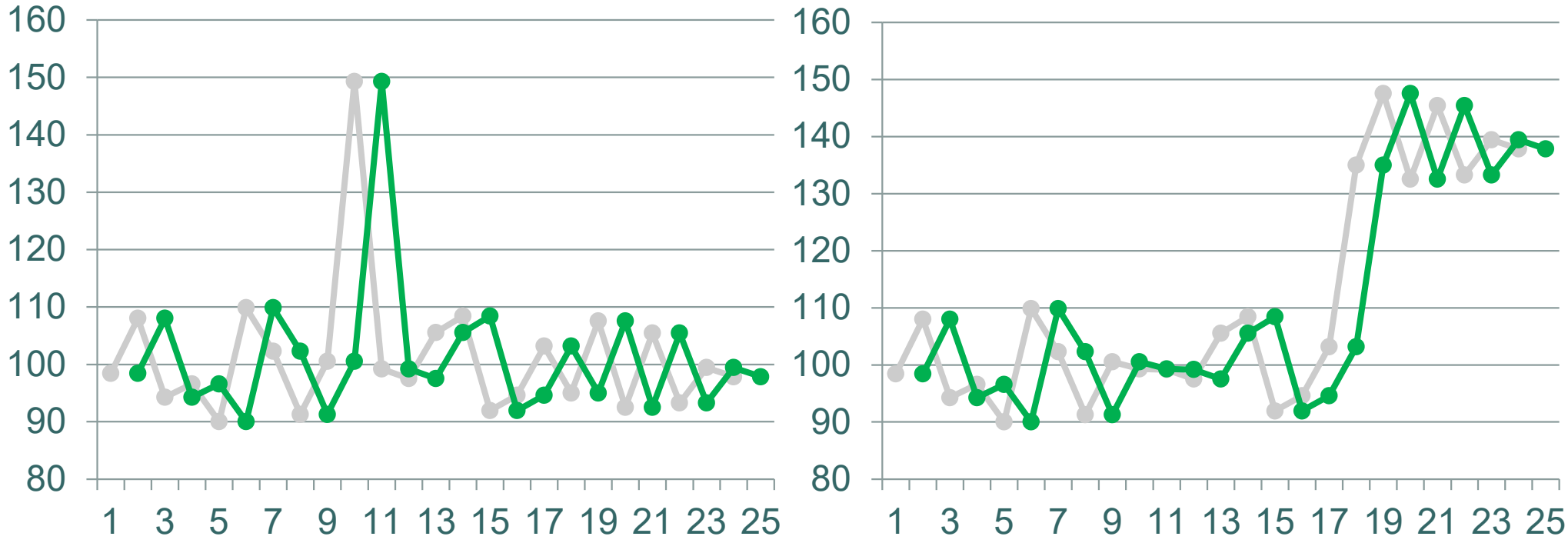
Naïve Method



t	Y	Naïve
1	98.43	
2	108.02	98.43
3	94.26	108.02
4	96.56	94.26
…	…	…
34	90.07	107.45
35	94.62	90.07
36	103.33	94.62
37		103.33
38		103.33
39		103.33

$$\hat{y}_{t+h} = y_t$$

Naïve Method: pros and cons

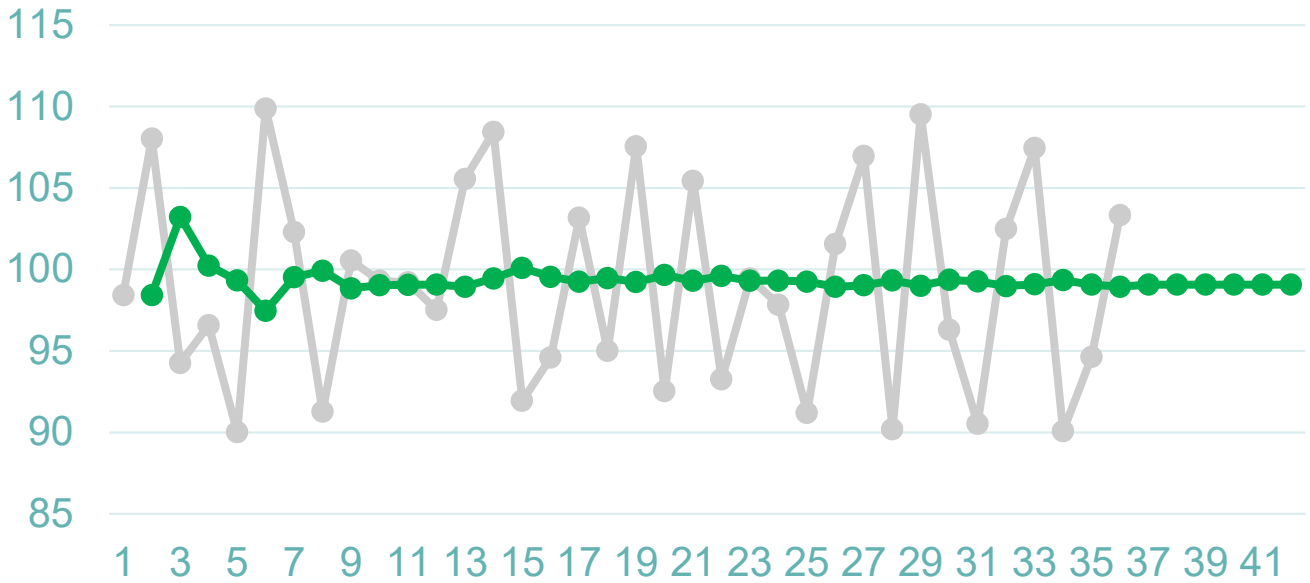


- Outliers are copied forward
- No smoothing
- Fast reaction to level changes

Global Average Method

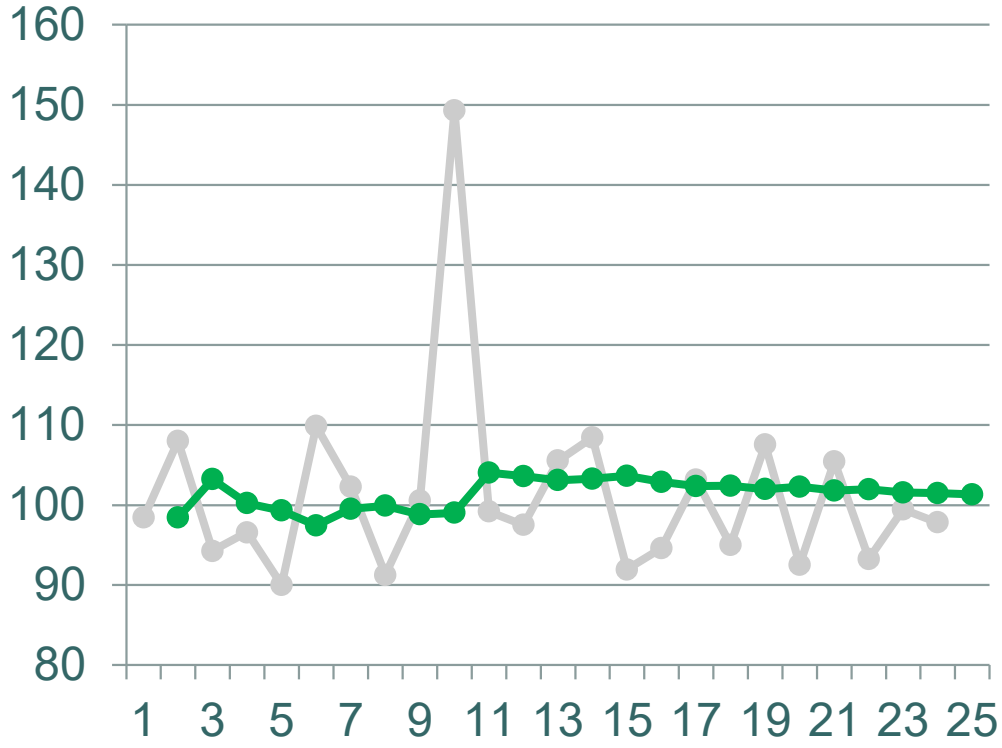
- No parameters, all values contribute equally
- The forecast is set as the mean of the observed sample
- Longest memory possible

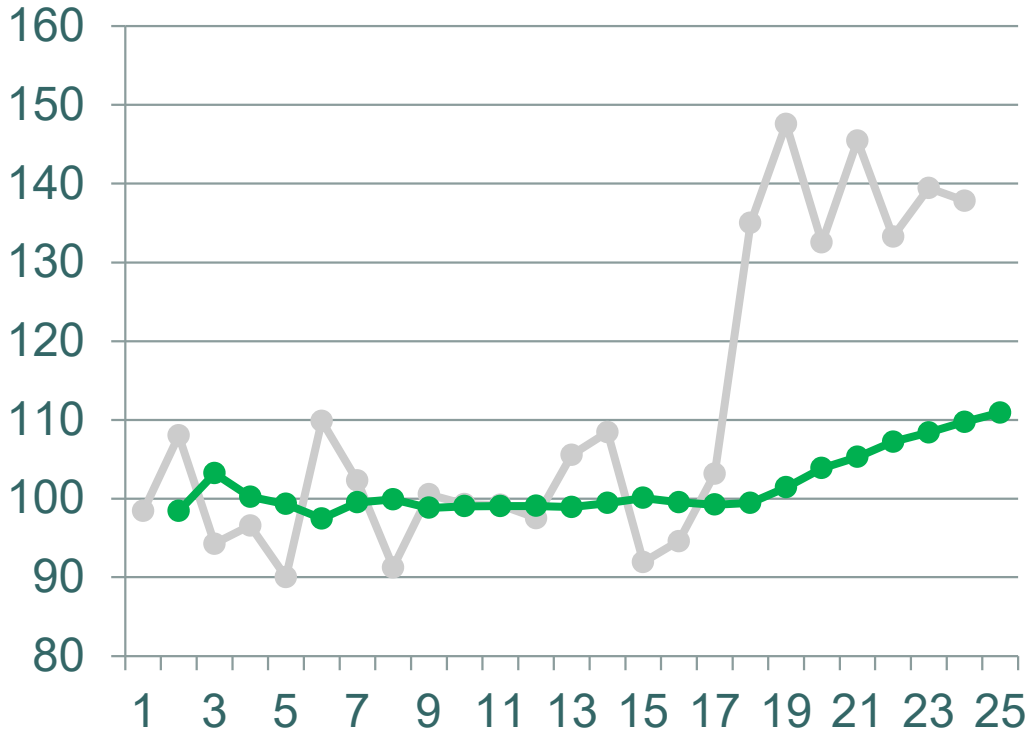
$$\hat{y}_{t+h} = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$



t	Y	Average
1	98.43	
2	108.02	98.43
3	94.26	103.22
4	96.56	100.23
…	…	…
34	90.07	99.34
35	94.62	99.07
36	103.33	98.94
37		99.06
38		99.06
39		99.06

Global Average Method: pros and cons





- Outliers are smoothed out
- Slow reaction to level changes

Moving Average Method

- Moving average: one parameter (k)
- The forecast is set as the average of the k-most-recent observations
- The value of k corresponds to the degree of smoothing and the length of the memory
- Selecting an appropriate value for k renders the model robust against outliers (high k values) or level changes (low k values)

$$\hat{y}_{\mathrm{t+h}} = \frac{1}{k}\sum_{i=0}^{k-1} y_{t-i}$$

Moving Average Method: Examples

t	Y	MA(3)	MA(7)
1	98.43		
2	108.02		
3	94.26		
4	96.56	100.23	
5	90.01	99.61	
6	109.87	93.61	
7	102.29	98.81	
8	91.26	100.72	99.92
9	100.54	101.14	98.90
10	99.31	98.03	97.83
…	…	…	…
34	90.07	100.15	100.49
35	94.62	100.00	98.08
36	103.33	97.38	98.71
37		96.01	97.83
38		96.01	97.83
39		96.01	97.83

MA(3):

$$\hat{y}_4 = \frac{1}{3}(y_1 + y_2 + y_3)$$

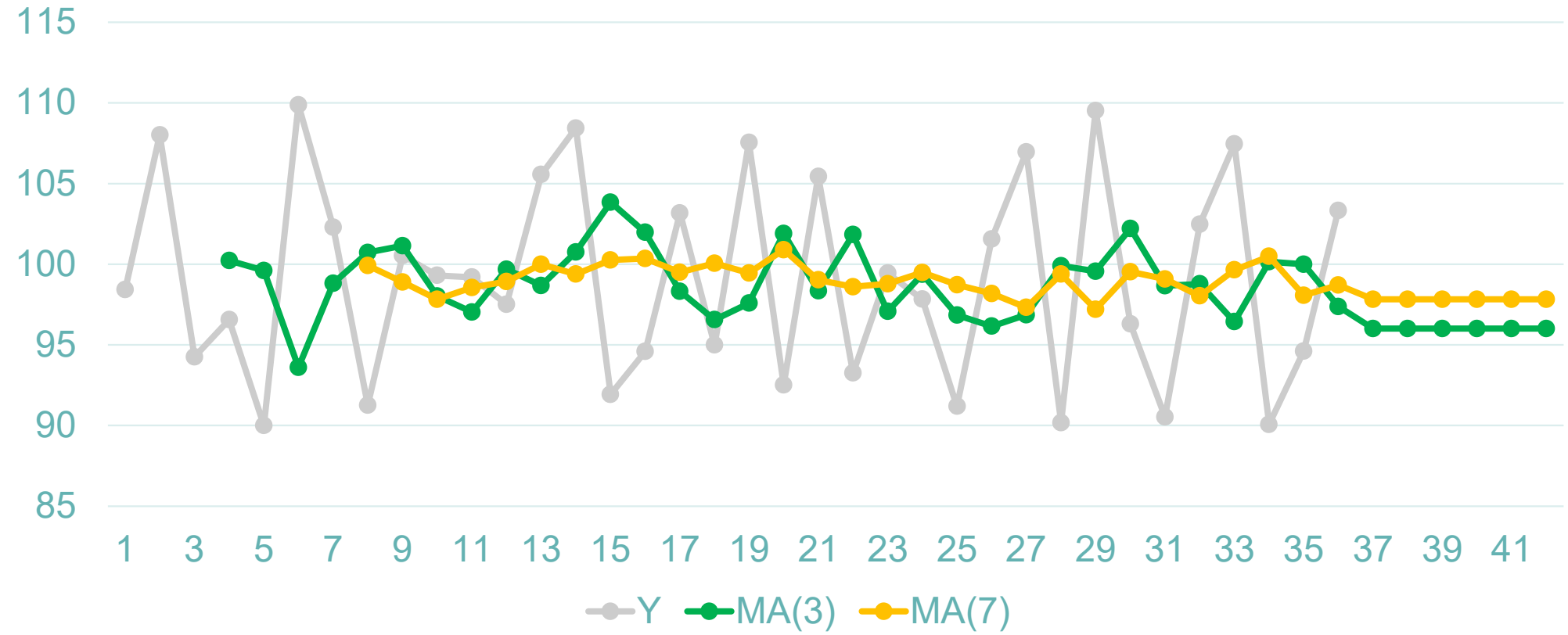
$$\hat{y}_{36} = \frac{1}{3}(y_{33} + y_{34} + y_{35})$$

MA(7):

$$\hat{y}_8 = \frac{1}{7}(y_1 + y_2 + \cdots + y_7)$$

$$\hat{y}_{36} = \frac{1}{7}(y_{29} + y_{30} + \cdots + y_{35})$$

Moving Average Method: Examples



- What is the best value for k?
- Does it make sense to give equal weight to all k recent observations?

Moving Average Method: Examples

- Compute a three-week moving average forecast for the number of arrivals at Piraeus port in week 4. The numbers of arrivals for the past three weeks were as follows:

Week	Arrivals
1	400
2	380
3	411

- If the actual number of arrivals in week 4 is 415, what is the forecast error for week 4?
- What is the forecast for week 5?

Moving Average Method: Examples

The moving average forecast at the end of week 3 is:

$$\hat{y}_4 = \frac{1}{3}(y_1 + y_2 + y_3) = \frac{1}{3}(400+380+411) = 397$$

The forecast error for week 4 is:

$$y_4 - \hat{y}_4 = 415 - 397 = 18$$

The forecast for week 5 requires the actual arrivals from weeks 2 through 4, the three most recent weeks of data

$$\hat{y}_5 = \frac{1}{3}(y_2 + y_3 + y_4) = \frac{1}{3}(380+411+415) = 402$$

Weighted Moving Average Method

- In the weighted moving average method, each historical data point in the average can have its **own weight,** provided that the sum of the weights equals one
- The average is obtained by multiplying the weight of each period by the actual data point for that period, and then adding the products together

Forecasts

- There are four forecasts:

	#1	#2	#3	#4
Week 1	70	50	29	43
Week 2	55	32	52	44
Week 3	40	48	62	54
Week 4	80	60	47	49

- Which forecast is the best one?
- This cannot be answered **before** you observe the true data

Forecast Error

- We define the **forecast error (FE)** for period t as the difference between the forecast for period t and the actual value for period t:

  FE at t = Actual value at t – Forecasted value at t

	#1	#2	#3	#4	Actual
Week 1	70	50	29	43	38
Week 2	55	32	52	44	49
Week 3	40	48	62	54	59
Week 4	80	60	47	49	44

- Calculate the forecast error (FE) for #1 and #2

Forecast Error Calculations

- The FE for #1 is:

  Week 1: 38 – 70 = -32

  Week 2: 49 – 55 = -6

  Week 3: 59 – 40 = 19

  Week 4: 44 – 80 = -36

	#1	True	Error
Week 1	70	38	**-32**
Week 2	55	49	**-6**
Week 3	40	59	**19**
Week 4	80	44	**-36**

Cumulative Forecast Error

- The *Cumulative Forecast Error (CFE)* takes the sum of the forecast errors

$$CFE = \sum_{t=1}^{N} FE_t$$

- Then, we can calculate the *Average Forecast Error* as:

$$\frac{CFE}{N}$$

where *N* is the number of observations

Forecast Error Calculations

- The Cumulative Sum of Forecast Error (CFE) for #1 is:

  CFE = (-32) + (-6) + 19 + (-36) = -55

  Average FE = -55/4 = -13.75

	Error
Week 1	**-32**
Week 2	**-6**
Week 3	**19**
Week 4	**-36**

- What can we comment on the (average) forecast error?
- The FE for #2 is: -12, 17, 11, -16; i.e., the average forecast error for #2 equals with 0
- What can we comment on the #2?

Quality of a Forecast

- A first way to measure the quality of a forecast is to see if it is right, on average; i.e., having the average forecast error be 0

- **Calculate** the average forecast error for #3
- The average forecast error for #3 = 0

	#3	True
Week 1	29	38
Week 2	52	49
Week 3	62	59
Week 4	47	44

- Which forecast is the best one and why?
- Both are right, on average. Does this mean that they are equally good at forecasting?

Mean Squared Error

- The ***Mean Squared Error (MSE)*** takes the average of the squared forecast errors

$$MSE = \frac{\sum_{t=1}^{N} FE_t^2}{N}$$

- The idea behind this is to first square the errors and then average them. A negative forecast error and a positive forecast error combined will no longer cancel each other out

Mean Absolute Error

- An easier way is just to simply take the absolute values of the forecast errors and average those out
- We define the ***Mean Absolute Error (MAE)*** as the average of the absolute values of the forecast errors

$$MAE = \frac{\sum_{t=1}^{N} |FE_t|}{N}$$

- This is also referred to as the ***Mean Absolute Deviation (MAD)***

Mean Absolute Percentage Error

- The ***Mean Absolute Percentage Error (MAPE)*** is achieved by dividing the forecast errors by the actual value y_t

$$MAPE = \frac{\sum_{t=1}^{N} \left| \frac{FE_t}{y_t} \right| * 100}{N}$$

- This measure does not look at the forecast errors in absolute terms, but in relative terms

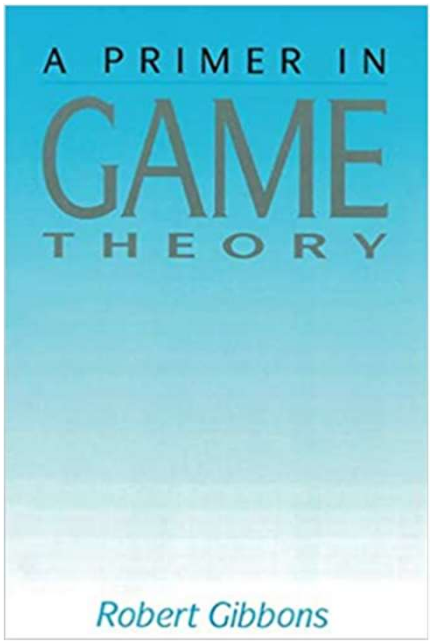
Key Takeaways

- Which measure of forecast quality you use is up to you
- Be careful when you report the forecast error and when you receive the forecast error
- Keep data from your old forecasts and assess the quality of those

Game Theory

Agenda

- Game Theory
- Nash equilibrium
- Public Goods Game



R. Gibbons. A Primer in Game Theory Paperback. Harlow: Prentice Hall, 1992

Video

Game Theory Explained in One Minute

https://www.youtube.com/watch?v=YueJukoFBMU

Movie: A Beautiful Mind

Classification of Problems

	A single decision maker	More than two decision makers
Deterministic	Optimisation	**Game Theory**
Stochastic	Decision Analysis	

Deterministic

Stochastic

Many Decision Makers

- Game Theory: is the study of multiperson decision problems (Gibbons 1992)
- Eleven game-theorists have now won the economics Nobel Prize (2014, Jean Tirole)
- Categories:
  - Cooperative / Non-Cooperative games
  - Complete / Incomplete information
  - Time order: Simultaneous / Sequential
- Main Elements: Players, Strategies, Payoffs

Solution Concept

- In Game Theory, it is used the term *solution concept*
- Non-Cooperative games: Nash equilibrium
- Cooperative games: Core, Shapley value
- Nash equilibrium: If each player has chosen a strategy and no player can benefit by changing strategies while the other players keep theirs unchanged
- Rational players

Prisoner's Dilemma

Prisoner's Dilemma

- Two friends had robbed a bank. The Police has arrested them for speeding without having not enough evidence to arrest them for the robbery
- In both friends are offered a deal by the Police and they have to decide what to do independently
- The deal is:
  - If both deny the robbery, they will be in prison for 1 year each
  - If both confess the robbery, they will be in prison for 6 years
  - If one confesses the robbery, then this guy will be free while the other will be in prison for 9 years

Prisoner's Dilemma: Main Elements

- Players: P1, P2
- Strategies: Deny, Confess
- Payoffs: a pair (x, y)

P1 \ P2	Deny	Confess
Deny	(1,1)	(9,0)
Confess	(0,9)	(6,6)

Nash equilibrium

- Is Deny vs Deny (1,1) Nash equilibrium?
- Given that P2 remains on ‘Deny’

P1 \ P2	Deny	Confess
Deny	(1,1)	(9,0)
Confess	(0,9)	(6,6)

Banned Cigarette Advertisements

- On April 1, 1970, President Richard Nixon signed legislation officially banning cigarette ads on television and radio
- The ban took effect at midnight on Jan. 2, 1971
- What will be the impact of restriction rate in profitability of tobacco companies?
- All four major tobacco companies registered higher profits than before

Food for Thought

- Magic box that doubles the money
- Two players
- $20 each
- How much should I put in the box?

Public Goods Game

- Four players (with the “same power”)
- Simultaneous game
- Box doubles the token/money
- “Public good" payoff is evenly divided among players
- Each subject also keeps the money they do not contribute
- Players are rational (i.e., to max their token/money)

How much should I put in the box?