Forecasting

- predict the future
- time series

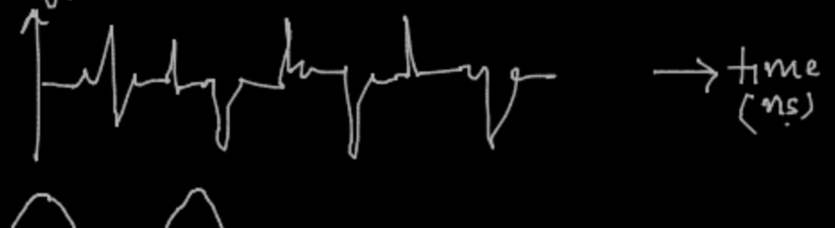
x y
x_1 x_2 x_3 y

Univariate Analysis

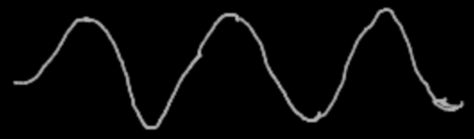
Time	(y) profits
Jan 1st	—
Feb 2nd	—
March 3rd	—

Monthly, daily, Weekly, Yearly, Eva
ms, ns, hours, mins

Global Warming



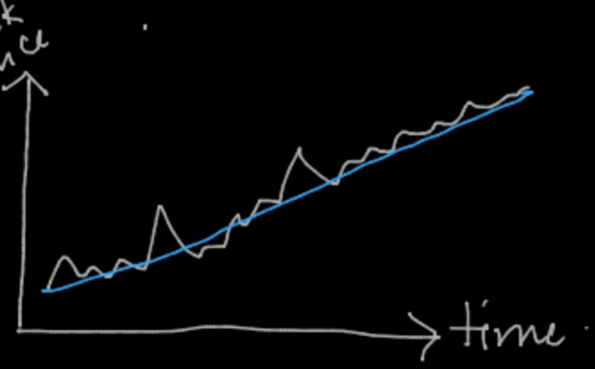
Traffic



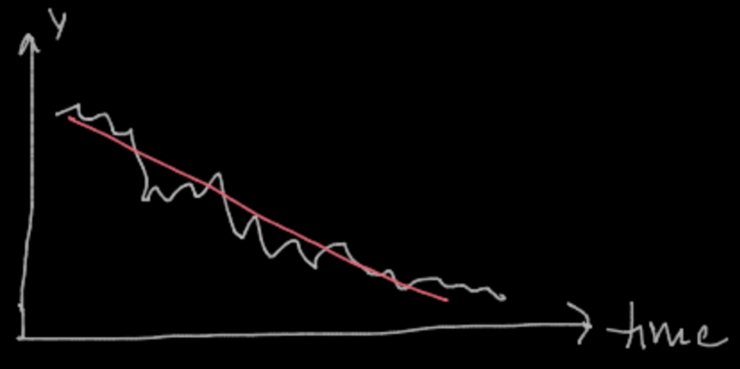
Time Series Data

- Systematic
  - Level
  - Trend
  - Seasonality
  - cyclicity
- Non Systematic
  - Errors
  - one time

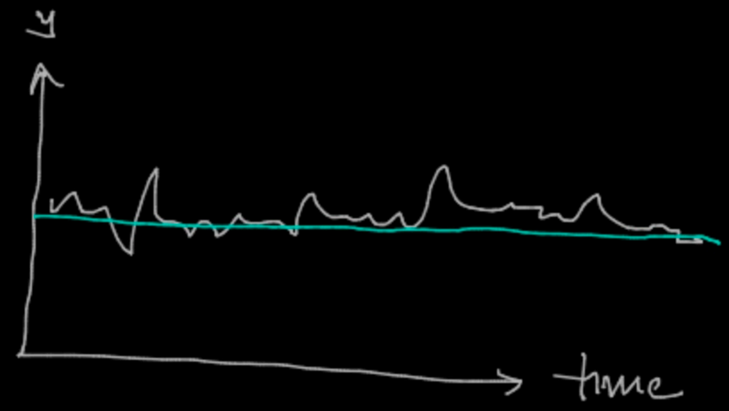
Trend



Upward trend

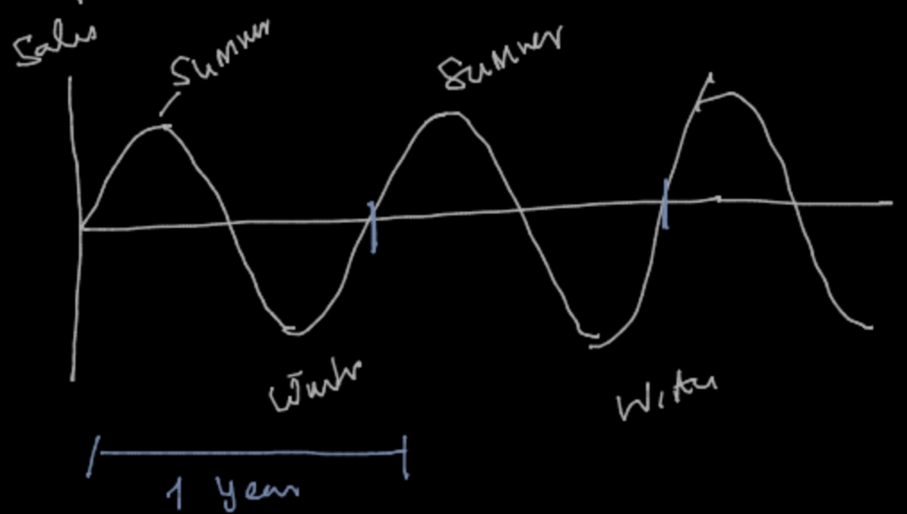


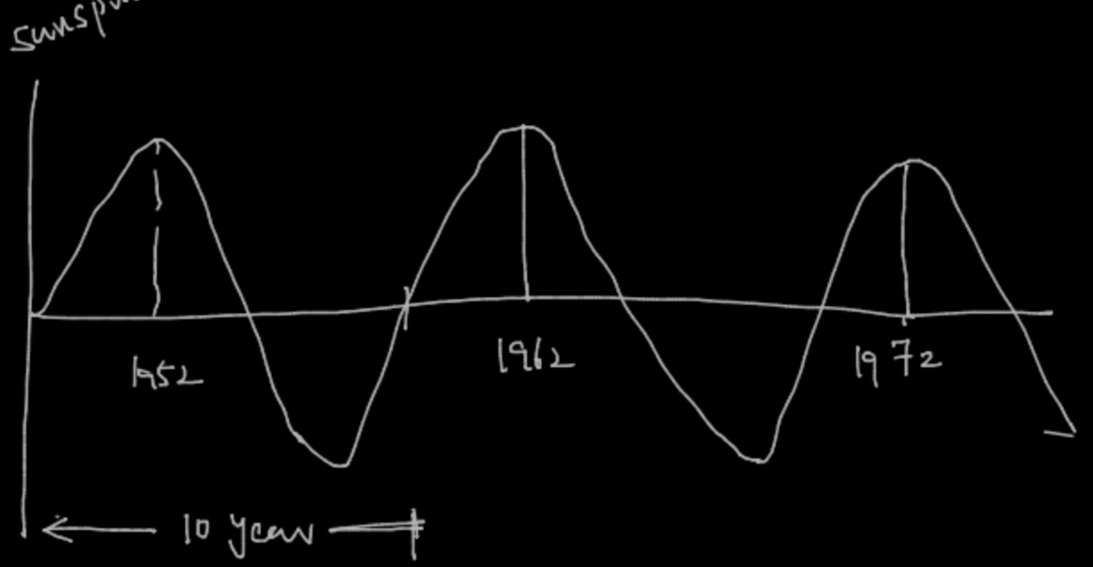
Downward



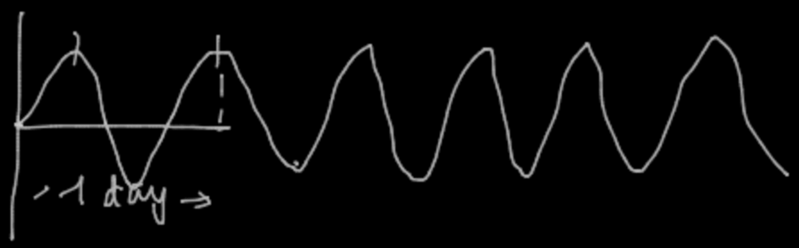
No / Constant trend.

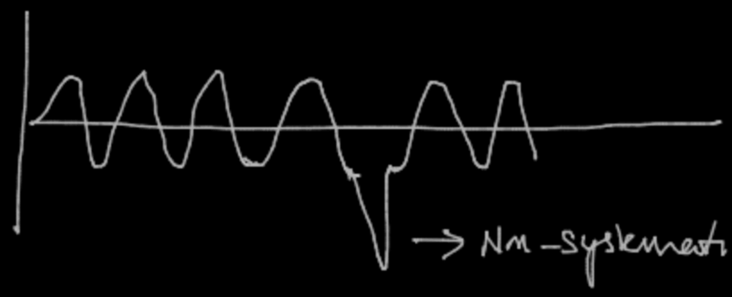
Seasonality





Cyclicity





PowerPoint File Edit View Insert Format Arrange Tools Slide Show Window Help zoom 80% Tue 16 Jan 6:25 PM

You are screen sharing Stop Share

Ffill & bfill

2000 Jan

Increase no. of records

Daily → upsampling

Monthly → Daily

May

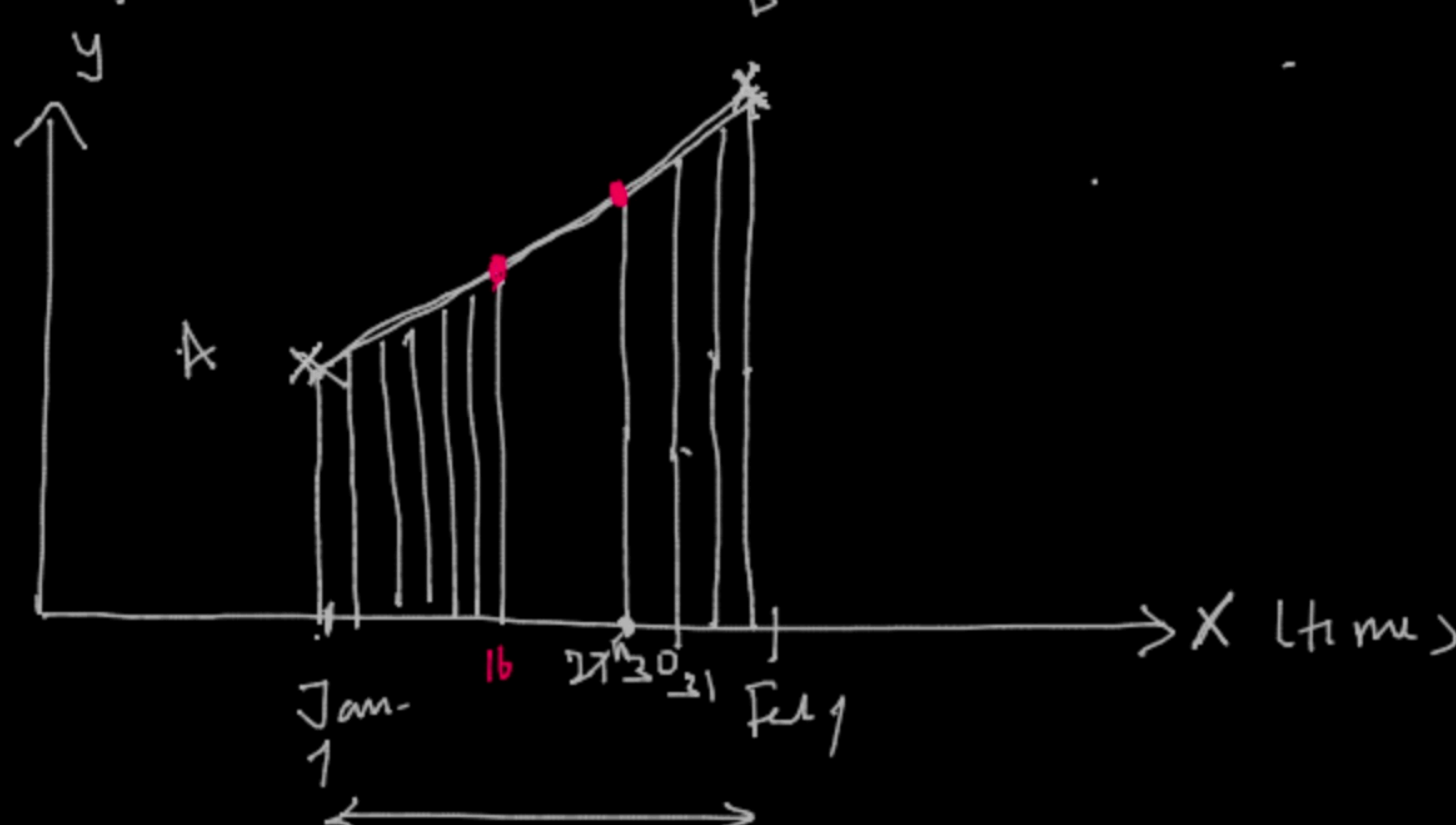
2012 June Null.

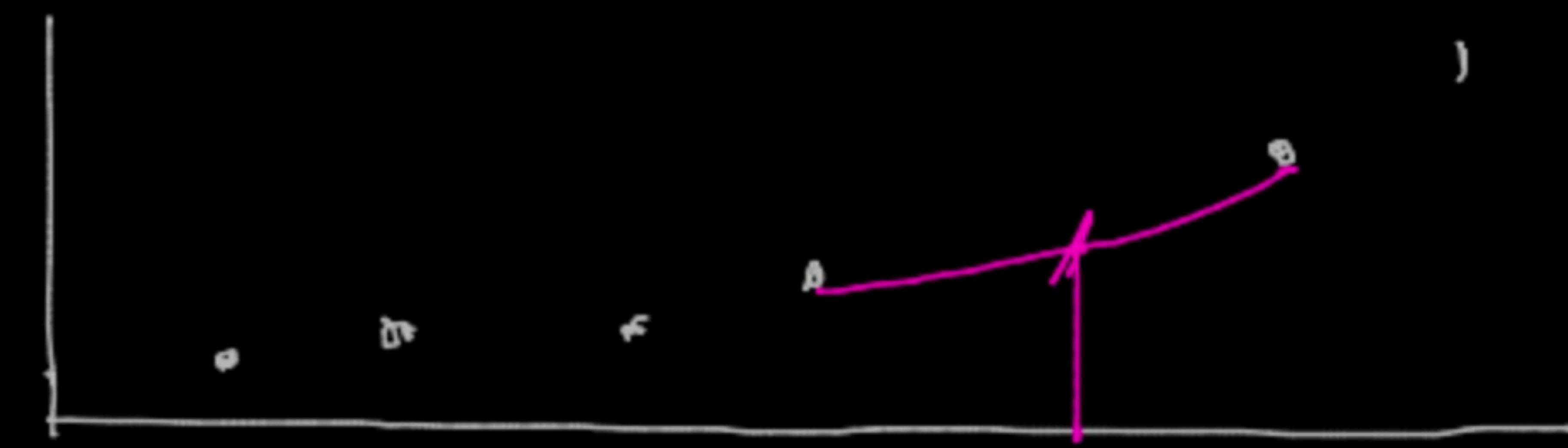
2012 July

2023 Dec

(1) Interpolation

Jan 27th





Interpolation

2007	Month No	Value
Jan	1	
Feb	2	
Mar	3	
	.	
	.	
	.	
2023 Dec	n	

Q1 →

Q2 →

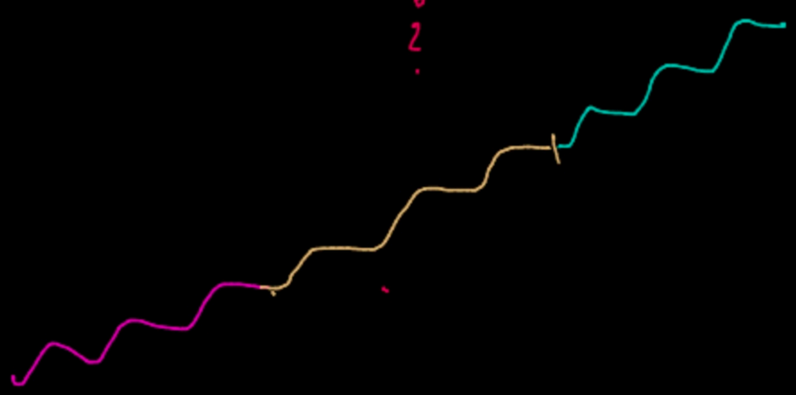
3 / 3

Splitting the data →

	Time	y
Jan 2001	1	–
Feb 2001	2	–
Mar 2001	3	–
	.	.
	.	.
	.	.
	.	.
	.	.
Dec 2012	144	–
	.	.
	2	
	0	
	2	
	.	

2/3

1/3

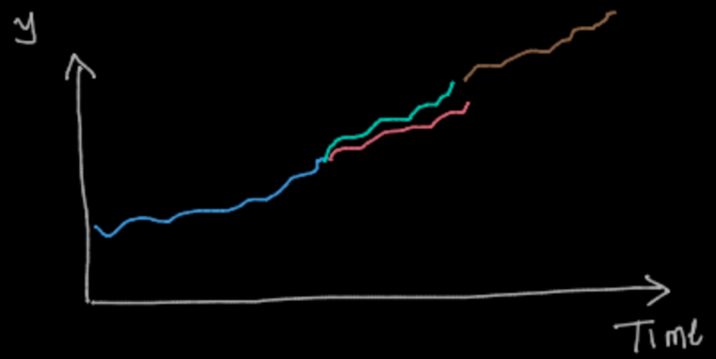


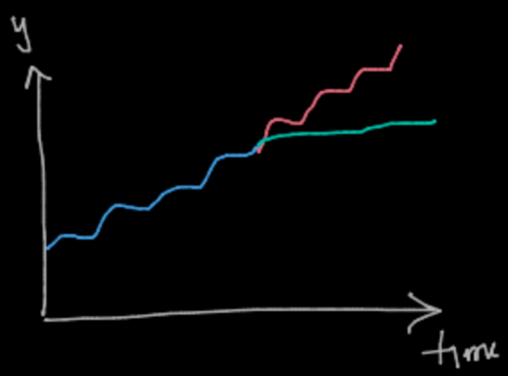
x_1	x_2	x_3	y
–	–	–	
.	.	.	.
.	.	.	.
.	.	.	.

80% .

20% .

Shuffled and randomly selected





Forecasting Models

1. Naive Model.
2. Moving Averages
3. Exponential smoothing →
4. Auto regression Models

1. $\hat{y}_t = y_{t-1}$

	y	$\hat{y}$	$y - \hat{y}$	$(y-\hat{y})^2$
Dec	10	Null	.	
Jan	9	10	.	
Feb.	10.5	9	.	
Mar		10.5	.	

MSE ✓

$\hat{y}_t = y_{t-k}$

	Time (t)	Stock y_t
Jan 23	1	y_{t-12}
	2	.
	3	.
	.	.
	.	.
	.	.
Oct	.	y_{t-3}
Nov 23	.	y_{t-2}
Dec 23	n	y_{t-1}
Jan 24		y_t
Feb 24		y_{t+1}

y_t y_{t-k}.

$k = 6$

-lags.

Assumptions for residual

- Mean = 0
- Normally dist
- Constant Variance → Homoscedasticity
- No auto correlation

2. Moving Averages

Window Size = 3

- Trailing Window
- Centred Window

6 monthly
12 monthly

Centred Window	Time (t)	y_t	Trailing y_t
Null	Jan	—	Null
—	Feb	—	Null
—	Mar	—	—
—	Apr	—	—
—	May	—	—
	Jun	—	—
	July	—	
	Aug	—