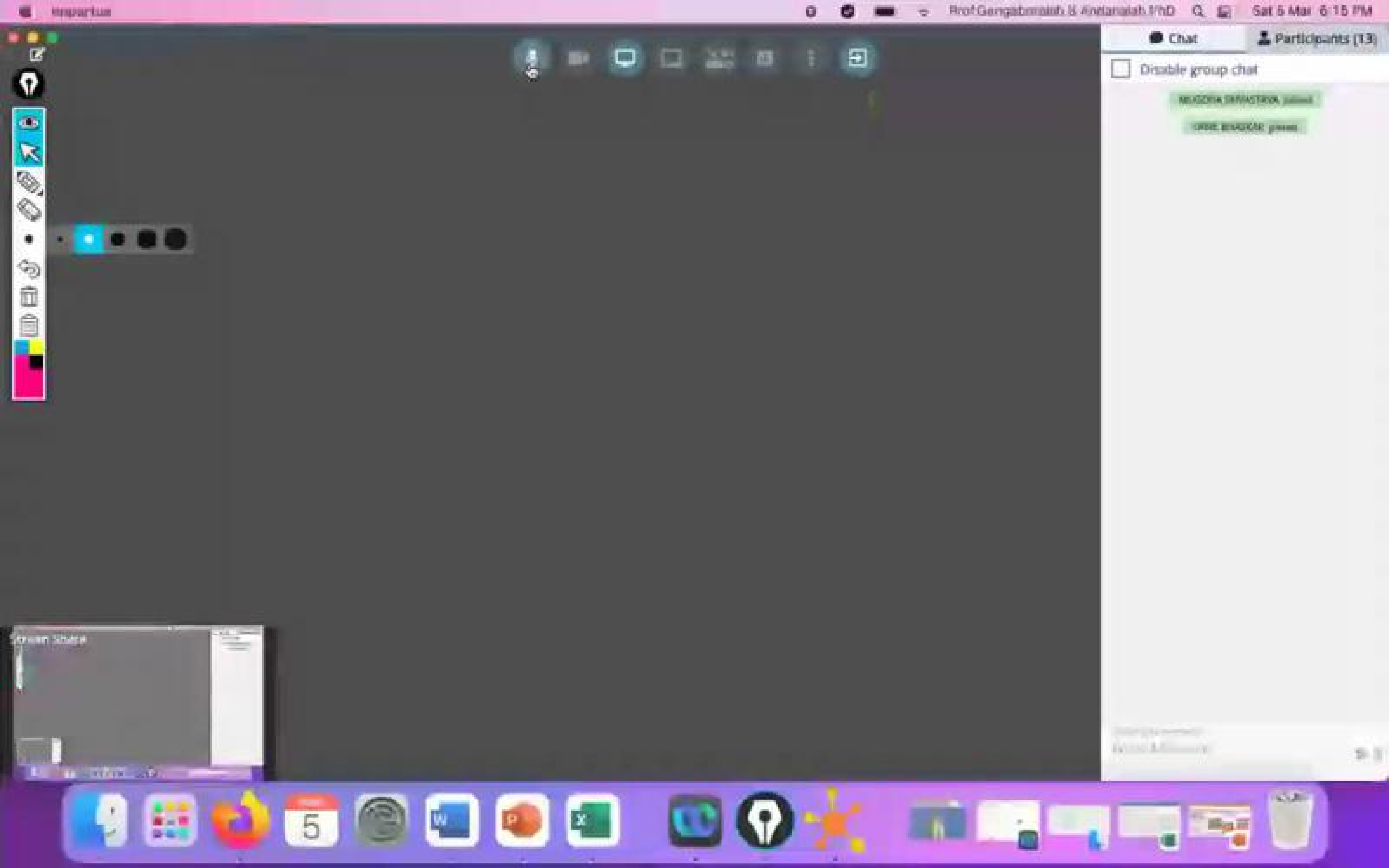
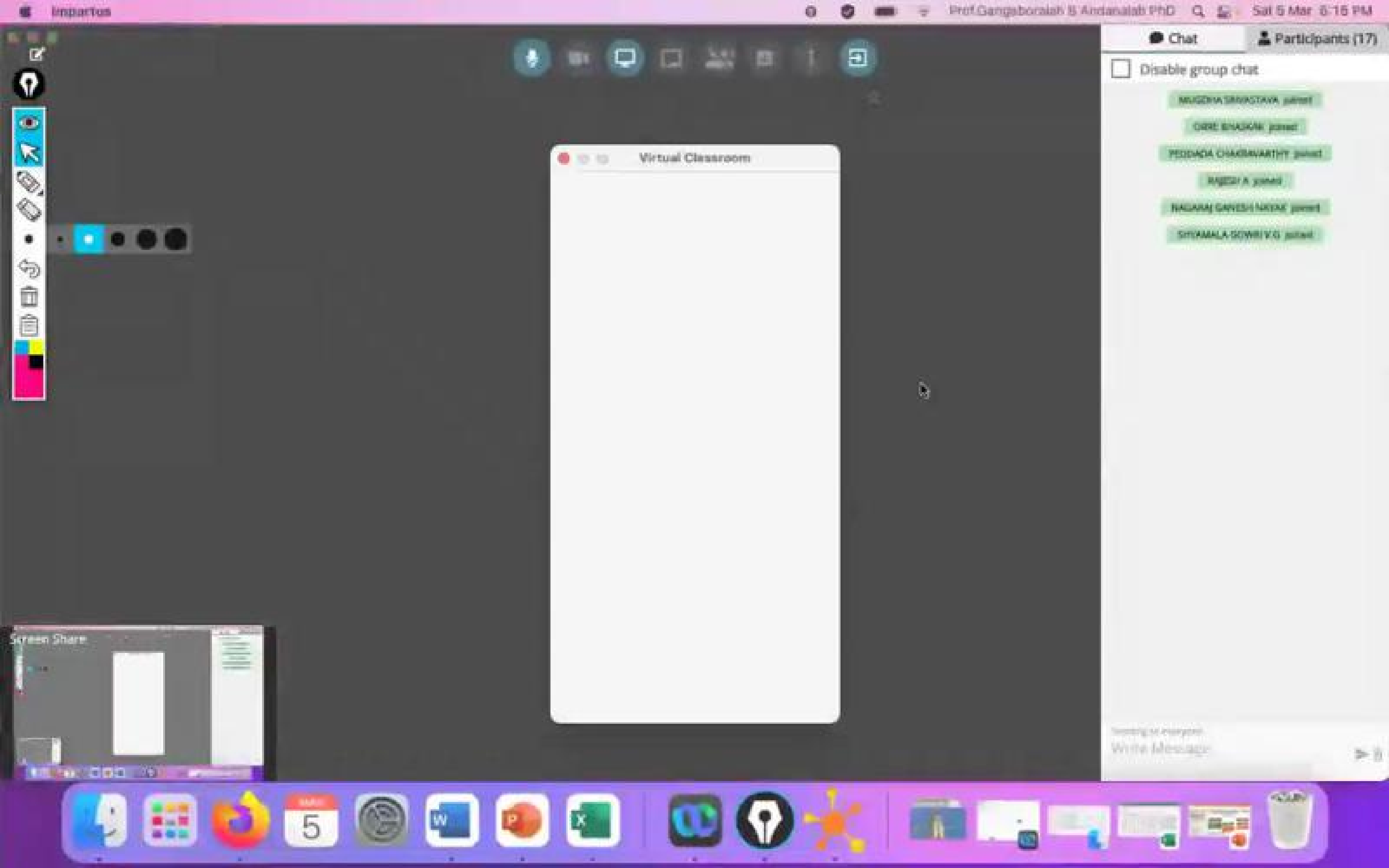
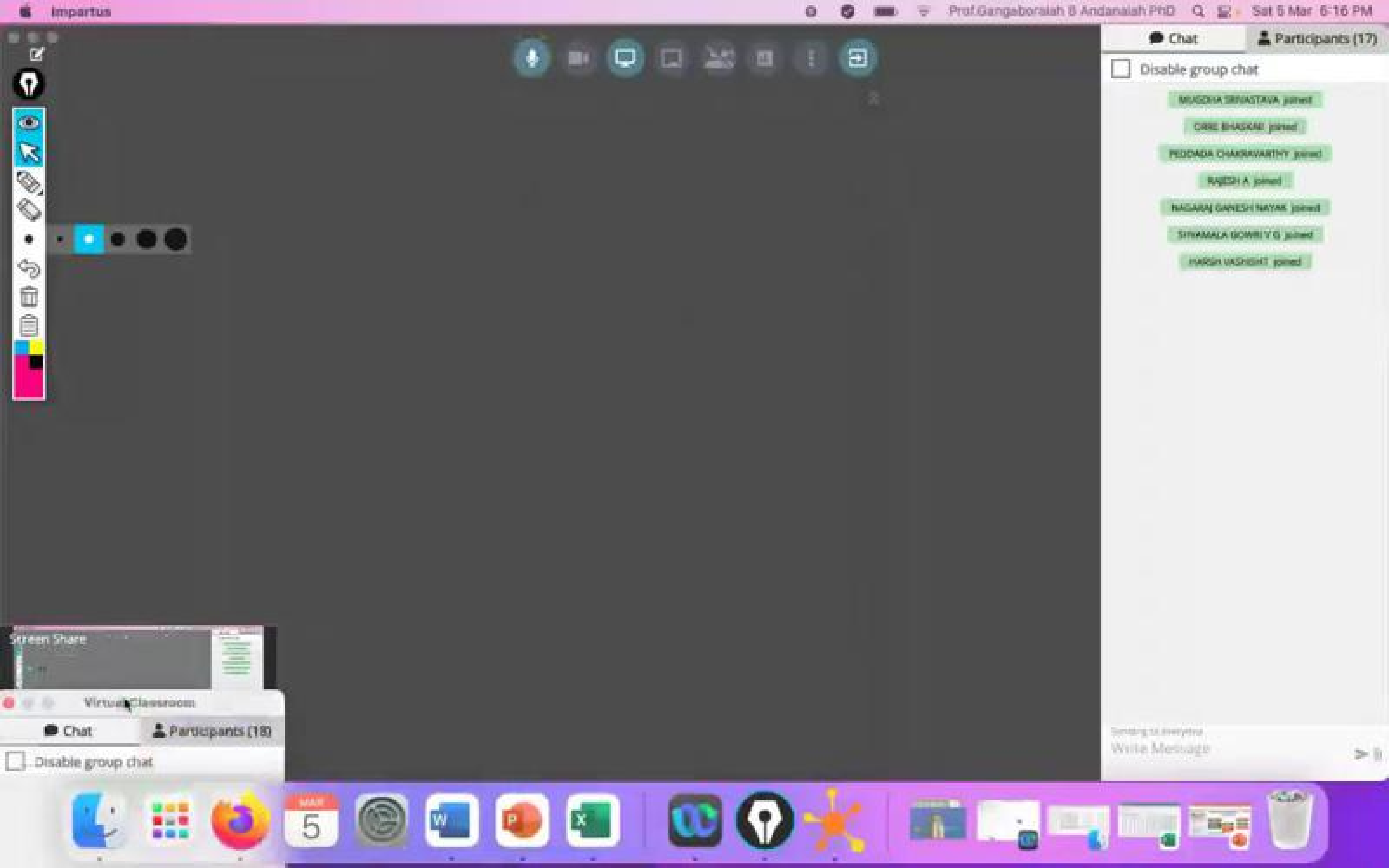

Sat 5 Mar 6:15 PM
Chat
Participants (13)
Disable group chat
Screen Share


Sat 5 Mar 6:16 PM
Chat
Participants (17)
Disable group chat
Virtual Classroom
Screen Share
Write Message



Impartus
Prof.Gangaboraiah B Andanaiah PhD
Sat 5 Mar 6:16 PM
Chat
Participants (17)
Disable group chat
MUGDHA SRIVASTAVA joined
PEDDADA CHAKRAVARTHY joined
RAJESH A joined
NAGARAJ GANESH NAYAK joined
HARSH VASHISHT joined
Write Message
Screen Share
Virtual Classroom
Chat
Participants (18)
Disable group chat

Statistical Methods for Data Science

Time Series ➡ Types of forecasting methods

Forecasting Methods

Qualitative
- Personal Opinion or Judgment
- Panel consensus
- Delphi Method
- Market Research

Quantitative

Time Series
- Smoothing methods
- Exponential smoothing
- Trend projection method

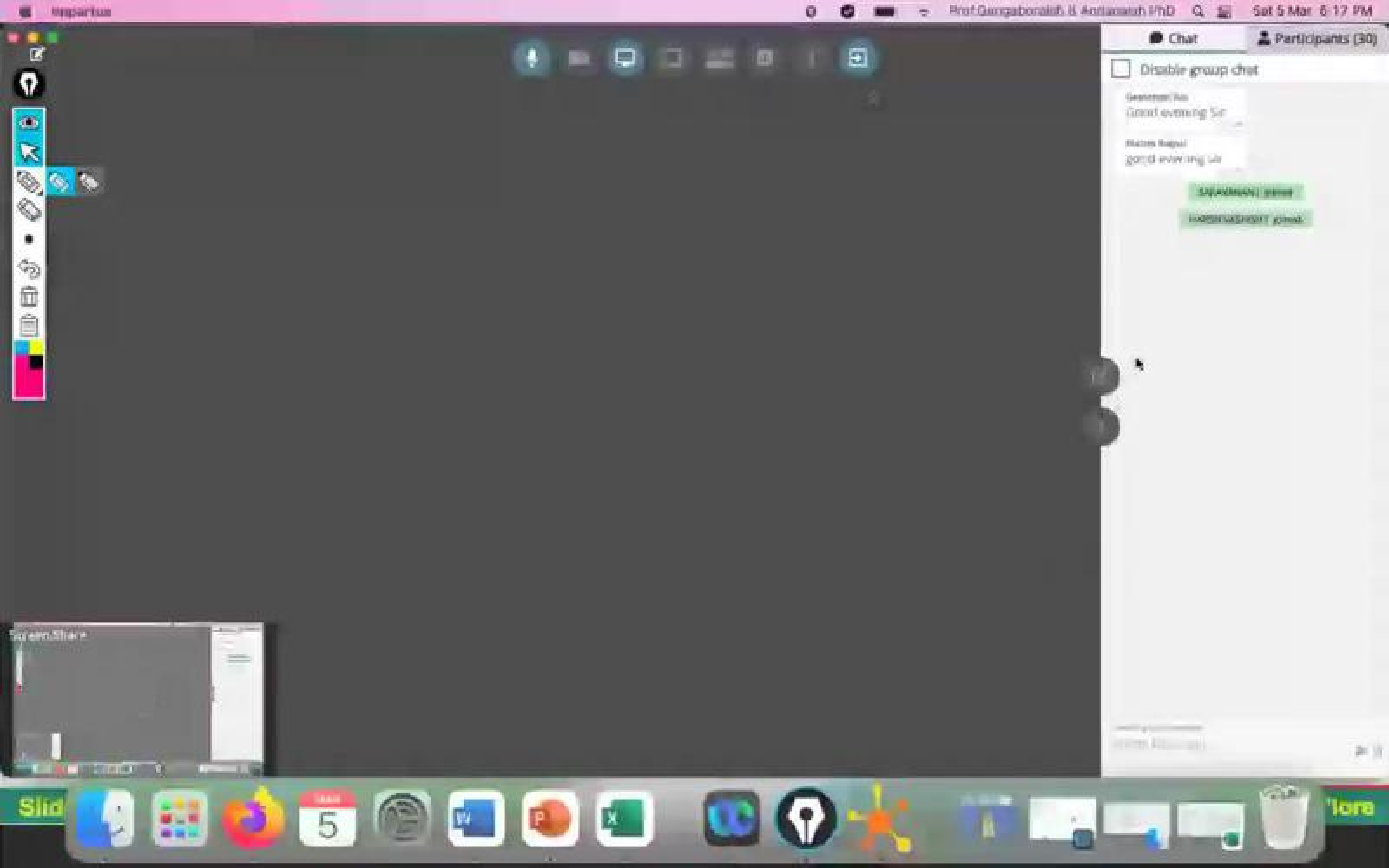
Causal
- Regression Trend Analysis

Time Series ➡ Types of forecasting methods

Virtual Classroom

Chat | Participants (23)

Disable group chat


Sat 5 Mar 6:17 PM
Chat
Participants (30)
Disable group chat
Good evening Sir
Screen Share

Time Series ➡ Qualitative forecasting

- **Personal Opinion**
  - Individuals forecasts future based on their own judgment or opinion without any formal model
  - Such assessments are relatively reliable and accurate

Delphi Method

The aim of the Delphi method is to construct consensus forecasts from a group of experts in a structured iterative manner. A facilitator is appointed in order to implement and manage the process. The Delphi method generally involves the following stages:

Delphi Method

- A panel of experts is assembled.
- Forecasting tasks/challenges are set and distributed to the experts.
- Experts return initial forecasts and justifications. These are compiled and summarised in order to provide feedback.

Delphi Method

- Feedback is provided to the experts, who now review their forecasts in light of the feedback. This step may be iterated until a satisfactory level of consensus is reached.
- Final forecasts are constructed by aggregating the experts' forecasts.

Statistical Methods for Data Science

Time Series ➡ Qualitative forecasting

ethod

Chat
Participants (36)
Disable group chat

aboraiah PhD (Stats)
Former Professor of Statistics | KIMS, B'lore

Delphi Method

- Each stage of the Delphi method comes with its own challenges. In what follows, we provide some suggestions and discussions about each one of these.

Statistical Methods for Data Science

Time Series ➡ Definition

A time series is defined as a set of observations on a variable generated sequentially in time. The measurement of the variable may be made continuously or at discrete (equally spaced) intervals.

Often a variable continuous in time is measured at regular intervals and this produces a discrete series of data.

Time Series ➡ Definition

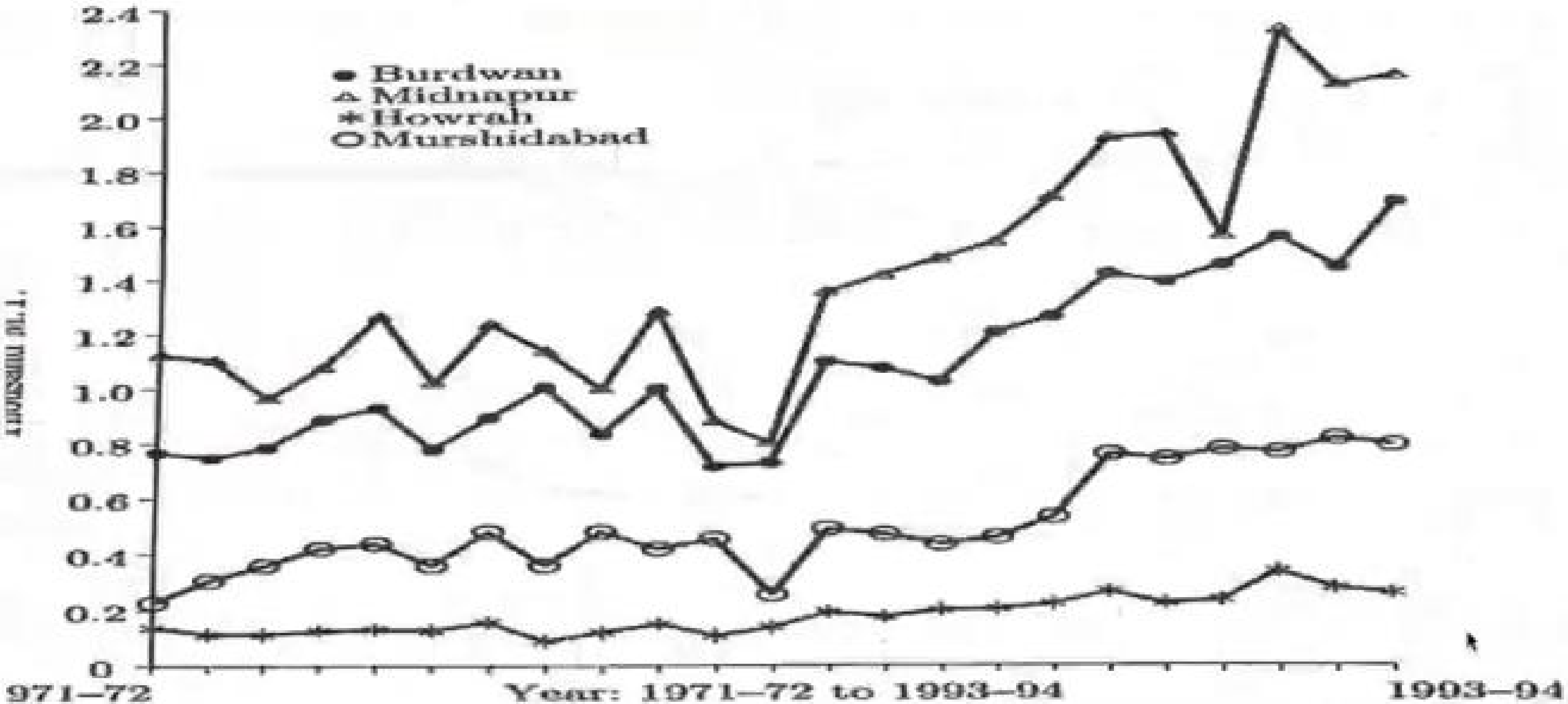
Thus a time series may be represented as a set of data $(y_1, y_2, y_3, \ldots, y_n)$, y_t denoting the value of the variable y at time t.

Example of continuous variable:

-Temperature on a chemical reactor

- Level of tide at a particular site

- Amplitude of an electrical signal



'ig 14.2 District-wise production of rice for four districts in West Bengal.

Series ➡ Definition

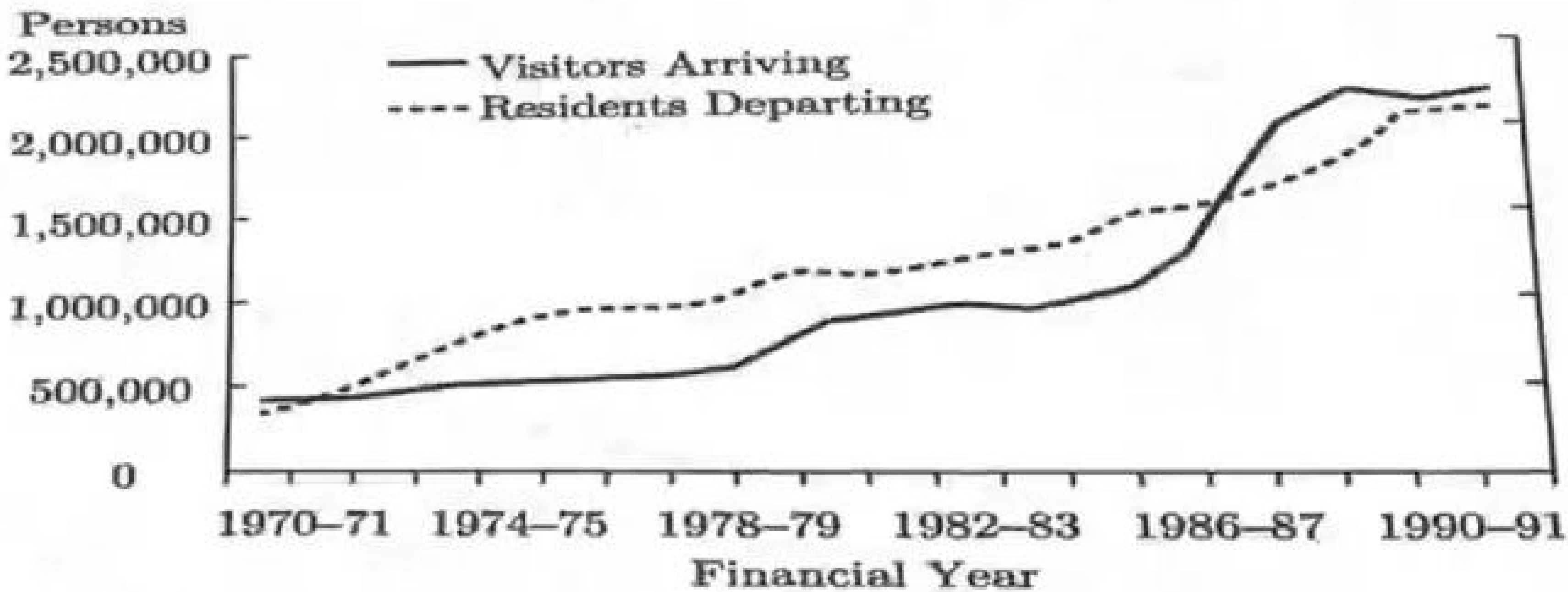


Fig 14.3 Short-term movement to and from Australia.
[*Source*: Immigration update, December, Quarter 1991, Bureau of Immigration Research, Statistics Section, Australia.]

Statistical Methods for Data Science

Time Series ➡ Definition

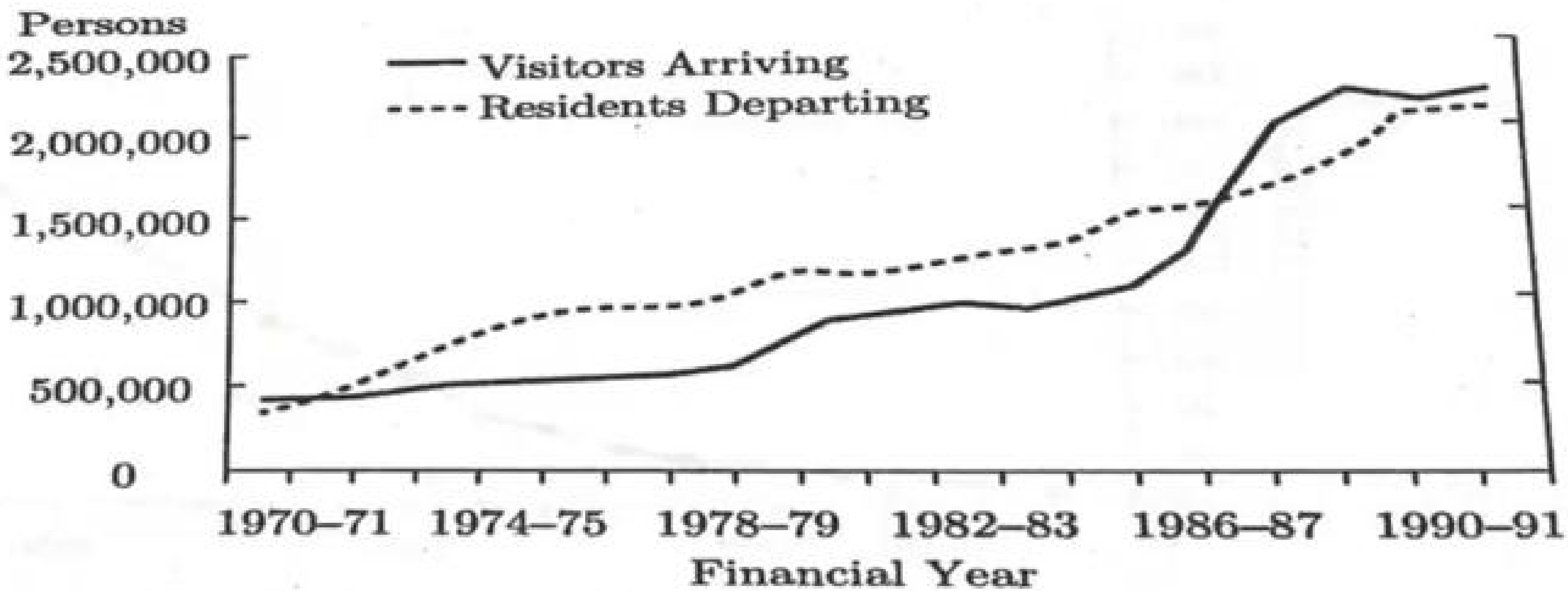


Fig 14.3 Short-term movement to and from Australia.
[*Source*: Immigration update, December, Quarter 1991, Bureau of Immigration Research, Statistics Section, Australia.]

980

5).]

Time Series ➡ Decomposition Models

This is an importan
of time series analy
adjustment. It seeks
observed time serie
series (that could be
original by addition
each of these has a
of behaviour. For ex
usually decompose

Time Series ➡ Decomposition Models

This is an important technique for all types of time series analysis, especially for seasonal adjustment. It seeks to construct, from an observed time series, a number of component series (that could be used to reconstruct the original by additions or multiplications) where each of these has a certain characteristic or type of behaviour. For example, time series are usually decomposed into:

Time Series ➡ Decomposition Models

S_t: the seasonal component at time t, reflecting seasonality (seasonal variation). A seasonal pattern exists when a time series is influenced by seasonal factors. Seasonality occurs over a fixed and known period (e.g., the quarter of the year, the month, or day of the week)

Time Series ➡ Decomposition Models

S_t: the seasonal component at time t, reflecting seasonality (seasonal variation). A seasonal pattern exists when a time series is influenced by seasonal factors. Seasonality occurs over a fixed and known period (e.g., the quarter of the year, the month, or day of the week)

Time Series ➡ Decomposition Models

I_t: the irregular component (or "noise") at time t, which describes random, irregular influences. It represents the residuals or remainder of the time series after the other components have been removed.

Time Series ➡ Decomposition Models

Hence a time series using an additive model can be thought of as:

$$y_t = T_t + C_t + S_t + I_t$$

Where as the multiplicative model can be thought of as:

Time Series ➡ Decomposition Models

An additive model would be used when the variations around the trend do not vary with the level of the time series whereas a multiplicative model would be appropriate if the trend is proportional to the level of the time series.

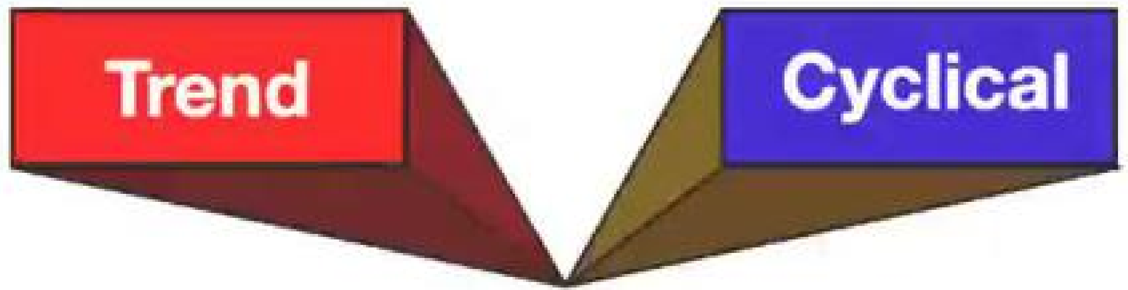
Statistical Methods for Data Science

Time Series ➡ Decomposition Models

Sometimes the trend and cyclical components are grouped into one, called the trend-cycle component. The trend-cycle component can just be referred to as the "trend" component, even though it may contain cyclical behaviour. For example, a seasonal decomposition of time series plot decomposes a time series into seasonal, trend and irregular components using loess and plots the components separately, whereby the cyclical component (if present in the data) is included in the "trend" component plot.

Statistical Methods for Data Science

Time Series ➡ Time Series Components



Time Series ➡ Time Series Components

Trend:

Response

Mo., Qtr., Yr.

Time Series ➡ Time Series Components

Trend:

- Persistent, overall upward or downward pattern

Response

Mo., Qtr., Yr.

Time Series ➡ Time Series Components

Cyclical:

Repeating up & down movements above trend line

Time Series ➡ Time Series Components

Cyclical:

Repeating up & down movements above trend line

- Due to interactions of factors influencing economy

Virtual Classroom
Chat
Participants (47)
Disable group chat
DEEPAK PANWAR JOINED

Time Series ➡ Smoothing Techniques – Moving Averages

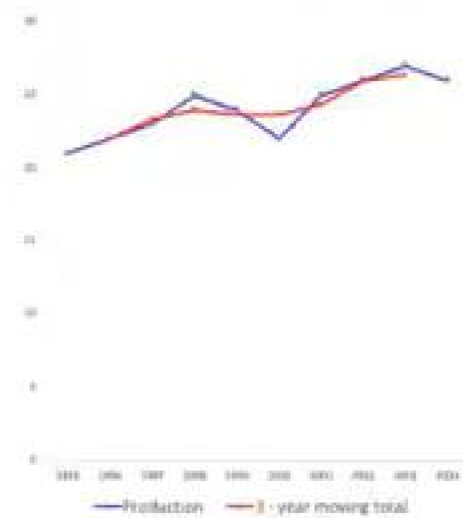
- Appropriate for a time series with a horizontal pattern ie., the data that are stationary.
  - Moving Average (the average of the most recent k data values forms the forecast for the next period)

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33

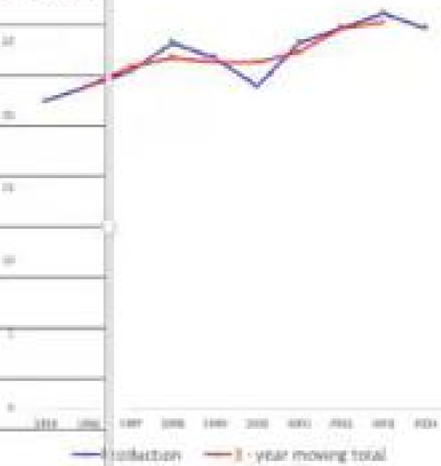


Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33

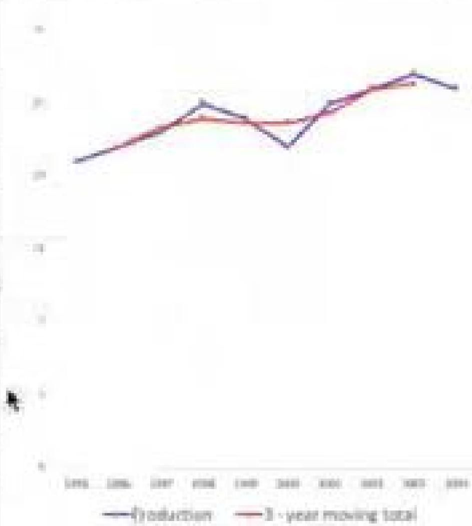


Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33



Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

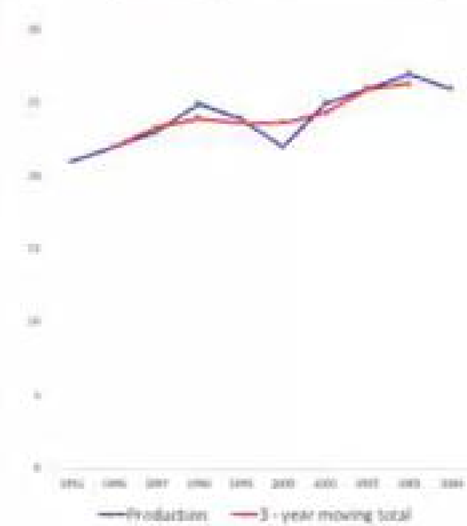
Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33



Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

Following data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33

Time Series ➡ Smoothing Techniques – Moving Averages

…ollowing data shows production volume (in '000 tones). Compute 3-year moving average for all available years

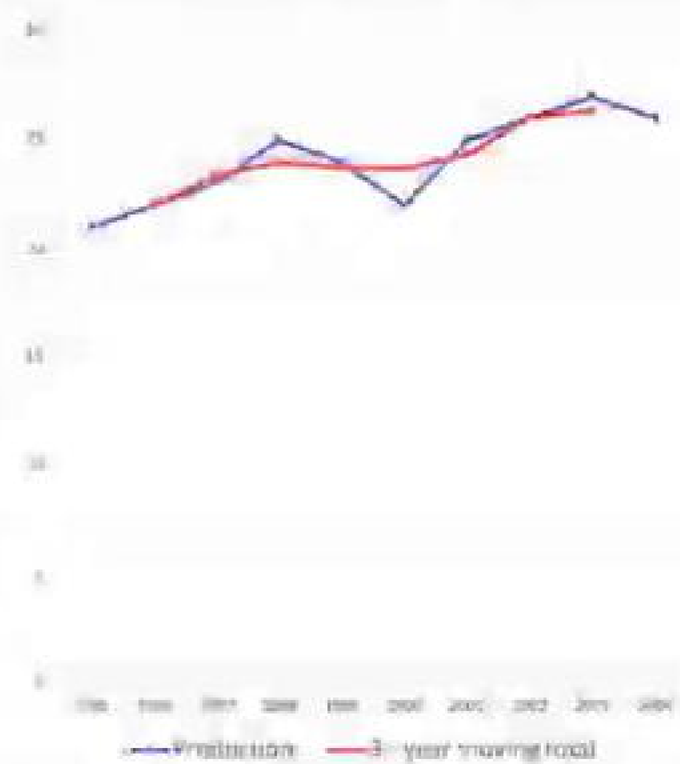
Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33

atistical Methods for Data Science

Series ➡ Smoothing Techniques – Moving Averages

i data shows production volume (in '000 tones). Compute 3-year moving average for all available years

r	roduction	3 - year moving total	3 - year moving average
5	21		
6	22		
7	23	66	22.00
8	25	70	23.33
9	24	72	24.00
0	22	71	23.67
1	25	71	23.67
2	26	73	24.33
3	27	78	26.00
4	26	79	26.33

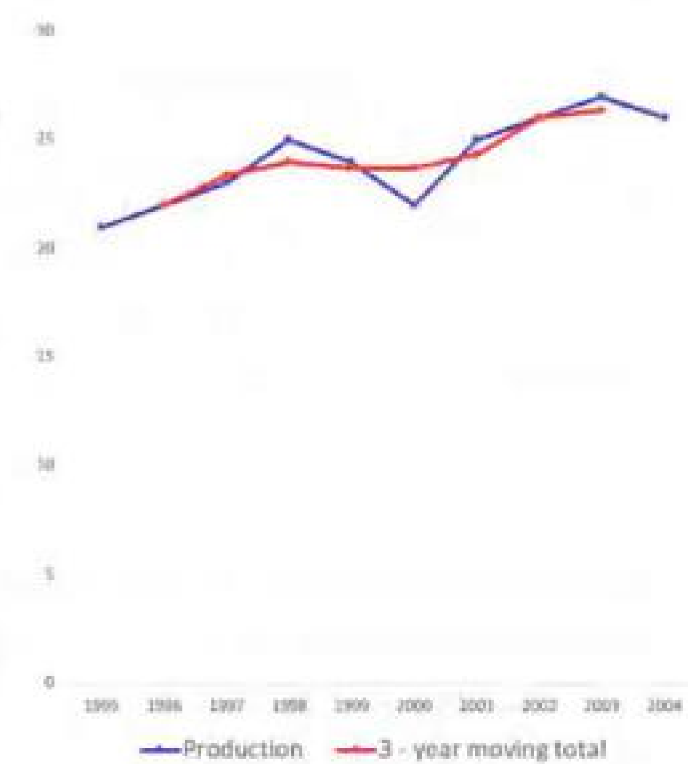


Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Moving Averages

…ollowing data shows production volume (in '000 tones). Compute 3-year moving average for all available years

Year	Production	3 - year moving total	3 - year moving average
1995	21		
1996	22		
1997	23	66	22.00
1998	25	70	23.33
1999	24	72	24.00
2000	22	71	23.67
2001	25	71	23.67
2002	26	73	24.33
2003	27	78	26.00
2004	26	79	26.33



Time Series ➡ Smoothing Techniques – Exponential

- Exponential (Weighted average of all the past time series values with exponentially decreasing importance in the forecast)

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

- α=1⇒ Forecast value is just the previous actual value. A naïve forecast.
- α=0⇒ Current Forecast value is just the previous forecast value.
- How to choose value of k and α?

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Exponential (Weighted average of all the past time series values with exponentially decreasing importance in the forecast)

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

- α=1⇒ Forecast value is just the previous actual value. A naïve forecast.
- α=0⇒ Current Forecast value is just the previous forecast value.
- How to choose value of k and α?

r s

e

e. A

cast

KIMS, B'lore

Virtual Classroom

Chat

Participants (44)

Disable group chat

WASIM AKRAM JOINED

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

Virtual Classroom

Chat | Participants (44)

Disable group chat

Chinnam Gari Chara...
82.26

Geetanjali Rai
82.26

Alok
82.26

Balajee R
sir...we are using the that year actual to come up with prediciton ?

Geetanjali Rai
@Balajee Yt is from previous year

Alok
should we not take 79

79*0.2

ZEENATH BIBI JOINED

S L S N V V S Sudhakar
I think we need to take 79*0.2+.8*82.26

Sending to everyone
Write Message

Do ... the pr... 3 dif...

- Fo... r
- Th... othing do... ificant ini... Moving Av... values are
- Fo... anged as

$$F_{t+1} \ldots -\alpha)F_t$$

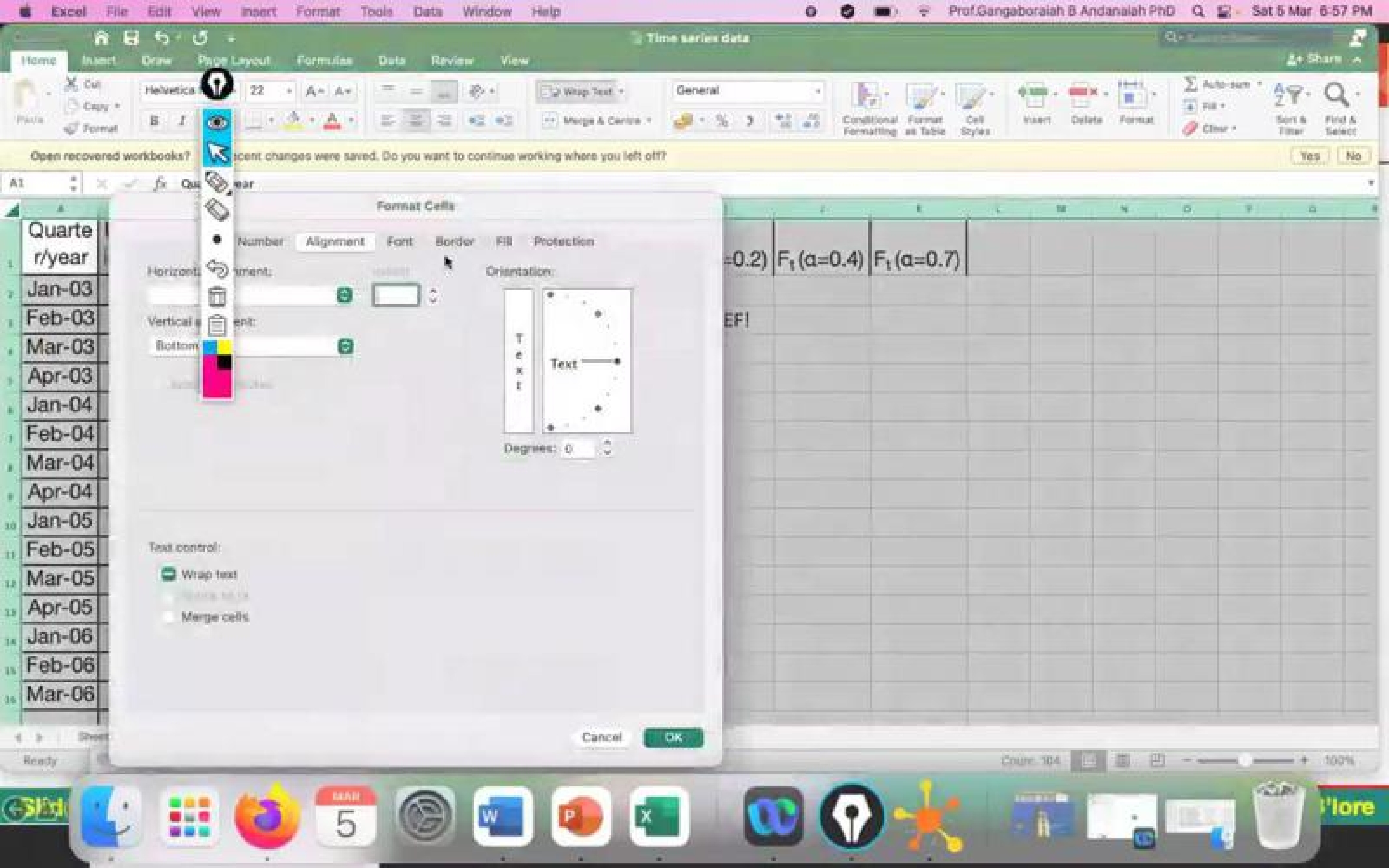
16.26
65.81
82.07

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.5	82.5	82.5
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
	MSE	2.70	1.04	0.26

• …owing data gives the cargo exports y_t (in units of 10^3 metric tons) at the principal port in India for a period of 20 years

…ocovariance, autocorrelation and auto regression

	lag 1 (y_1	lag 2 (y_2 =	lag3 (y_1 =	lag 1 (y_4 =	lag 5 (y_1 =	yt-m	y1-m	y2-m	y3-m	y4-m	y5-m	(yt-m)('	(yt-m)('	(yt-m)(y	(yt-m)(y	(yt-m)(y5-m)
1230	1345	1382	1416	1593	1802	-1010.2	-895	-858	-824	-647	-438.2	9E+05	9E+05	832607	653801	442670
1345	1382	1416	1593	1802	1817	-895.2	-858	-824	-647	-438	-423.2	8E+05	7E+05	579373	392277	378849
1382	1416	1593	1802	1817	1995	-858.2	-824	-647	-438	-423	-245.2	7E+05	6E+05	376063	363190	210431
1416	1593	1802	1817	1995	2212	-824.2	-647	-438	-423	-245	-28.2	5E+05	4E+05	348801	202094	23242
1593	1802	1817	1995	2212	2607	-647.2	-438	-423	-245	-28	366.8	3E+05	3E+05	158693	18251	-2E+05
1802	1817	1995	2212	2607	2698	-438.2	-423	-245	-28	367	457.8	2E+05	1E+05	12357	-2E+05	-2E+05
1817	1995	2212	2607	2698	2703	-423.2	-245	-28	367	458	462.8	1E+05	11934	-2E+05	-2E+05	-2E+05
1995	2212	2607	2698	2703	2738	-245.2	-28	367	458	463	497.8	6915	-89939	-1E+05	-1E+05	-1E+05
2212	2607	2698	2703	2738	2810	-28.2	367	458	463	498	569.8	-10344	-12910	-13051	-14038	-16068
2607	2698	2703	2738	2810	2632	366.8	458	463	498	570	391.8	2E+05	2E+05	182593	209003	143712
2698	2703	2738	2810	2632	2548	457.8	463	498	570	392	307.8	2E+05	2E+05	260854	179366	140911
2703	2738	2810	2632	2548	2881	462.8	498	570	392	308	640.8	2E+05	3E+05	181325	142450	296562
2738	2810	2632	2548	2881	2725	497.8	570	392	308	641	484.8	3E+05	2E+05	153223	318990	241333
2810	2632	2548	2881	2725	2911	569.8	392	308	641	485	670.8	2E+05	2E+05	365128	276239	382222
2632	2548	2881	2725	2911	2759	391.8	308	641	485	671	518.8	1E+05	3E+05	189945	262819	203266


Excel File Edit View Insert Format Tools Data Window Help
Prof.Gangaboraiah B Andanaiah PhD
Sat 5 Mar 6:57 PM
Time series data
Home Insert Draw Page Layout Formulas Data Review View
Share
Helvetica
22
Wrap Text
Merge & Centre
General
Conditional Formatting
Format as Table
Cell Styles
Insert
Delete
Format
Auto-sum
Fill
Clear
Sort & Filter
Find & Select
Open recovered workbooks?
cent changes were saved. Do you want to continue working where you left off?
Yes
No
A1
Quarte r/year
Jan-03
Feb-03
Mar-03
Apr-03
Jan-04
Feb-04
Mar-04
Apr-04
Jan-05
Feb-05
Mar-05
Apr-05
Jan-06
Feb-06
Mar-06
=0.2)
F1 (α=0.4)
F1 (α=0.7)
EF!
Format Cells
Number
Alignment
Font
Border
Fill
Protection
Vertical
Bottom
Orientation:
Text
Degrees: 0
Text control:
Wrap text
Merge cells
Cancel
OK
Ready
Count: 104
100%

Time series data

J4 fx =(0.4*H3)+(0.6*J3)

A	B	C	D	E	F	G	H	I	J	K
Jan-03	82.5	NA	NA	NA		1	82.5	NA	NA	NA
Feb-03	81.3	82.5	82.5	82.5		2	81.3	82.5	82.5	82.5
Mar-03	81.3	82.26	82.02	81.66		3	81.3	82.26	82.02	81.66
Apr-03	79	81.61	80.81	79.8		4	79	82.068	81.732	81.408
Jan-04	76.6	80.61	79.13	77.56		5	76.6	81.4544	80.6392	79.7224
Feb-04	78	80.09	78.68	77.87		6	78	80.48352	79.02352	77.53672
Mar-04	78.4	79.75	78.57	78.24		7	78.4	79.98682	78.61411	77.86102
Apr-04	78	79.4	78.34	78.07		8	78	79.66945	78.52847	78.2383
Jan-05	78.8	79.28	78.52	78.58		9	78.8	79.33556	78.31708	78.07149
Feb-05	78.7	79.16	78.59	78.66		10	78.7	79.22845	78.51025	78.58145
Mar-05	78.4	79.01	78.52	78.48		11	78.4	79.12276	78.58615	78.66443
Apr-05	80	79.21	79.11	79.54		12	80	78.97821	78.51169	78.47933
Jan-06	80.7	79.51	79.75	80.35		13	80.7	79.18257	79.10701	79.5438
Feb-06	80.7	79.75	80.13	80.6		14	80.7	79.48605	79.74421	80.35314
Mar-06	80.8	79.96	80.4	80.74		15	80.8	79.72884	80.12652	80.59594
3/2006						16		79.94307	80.39591	80.73878

Sheet1 Sheet2

...e Series ➡ Smoothing Techniques – Exponential

Done the analysis on the previous problem for 3 different values of α

For α=0.7, it is better

The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed

Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.5	82.5	82.5
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
	MSE	2.70	1.04	0.26

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.5	82.5	82.5
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006				

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	82.07	81.73	81.41
1/2004	76.60	81.45	80.64	79.72
2/2004	78.00	80.48	79.02	77.54
3/2004	78.40	79.99	78.61	77.86
4/2004	78.00	79.67	78.53	78.24
1/2005	78.80	79.34	78.32	78.07
2/2005	78.70	79.23	78.51	78.58
3/2005	78.40	79.12	78.59	78.66
4/2005	80.00	78.98	78.51	78.48
1/2006	80.70	79.18	79.11	79.54
2/2006	80.70	79.49	79.74	80.35
3/2006	80.80	79.73	80.13	80.60
3/2006				

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.5	82.5	82.5
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74

I17 f_x =(0.2*H16)+(0.8*I16)

Jan-03	82.5	1	82.5	NA	NA	NA
Feb-03	81.3	2	81.3	82.50	82.50	82.50
Mar-03	81.3	3	81.3	82.26	82.02	81.66
Apr-03	79	4	79	82.07	81.73	81.41
Jan-04	76.6	5	76.6	81.45	80.64	79.72
Feb-04	78	6	78	80.48	79.02	77.54
Mar-04	78.4	7	78.4	79.99	78.61	77.86
Apr-04	78	8	78	79.67	78.53	78.24
Jan-05	78.8	9	78.8	79.34	78.32	78.07
Feb-05	78.7	10	78.7	79.23	78.51	78.58
Mar-05	78.4	11	78.4	79.12	78.59	78.66
Apr-05	80	12	80	78.98	78.51	78.48
Jan-06	80.7	13	80.7	79.18	79.11	79.54
Feb-06	80.7	14	80.7	79.49	79.74	80.35
Mar-06	80.8	15	80.8	79.73	80.13	80.60
3/2006		16		79.94	80.40	80.74

Chat | Participants (50)

Disable group chat

yt is the actual value ?

Sai Kumar K C: Sir, Y value is % Utilization, Right?

without F1...what will we do for F2

ahh...oke

SRIVATSAN K R JOINED

got it

thanks

PRASHANT SINGH JOINED

MALAVIKA SHAMESH JOINED

Sunil Kumar Pradhan: Sir - could you please share the column H values here

BHATT NAYAN SURESHKUMAR JOINED

B C HEMANTH KUMAR JOINED

Statistical Methods for Data Science

Time Series ⇨ Smoothing Techniques – Exponential

Done the analysis on the previous problem for 3 different values of α

For α=0.7, it is better

The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed

Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74

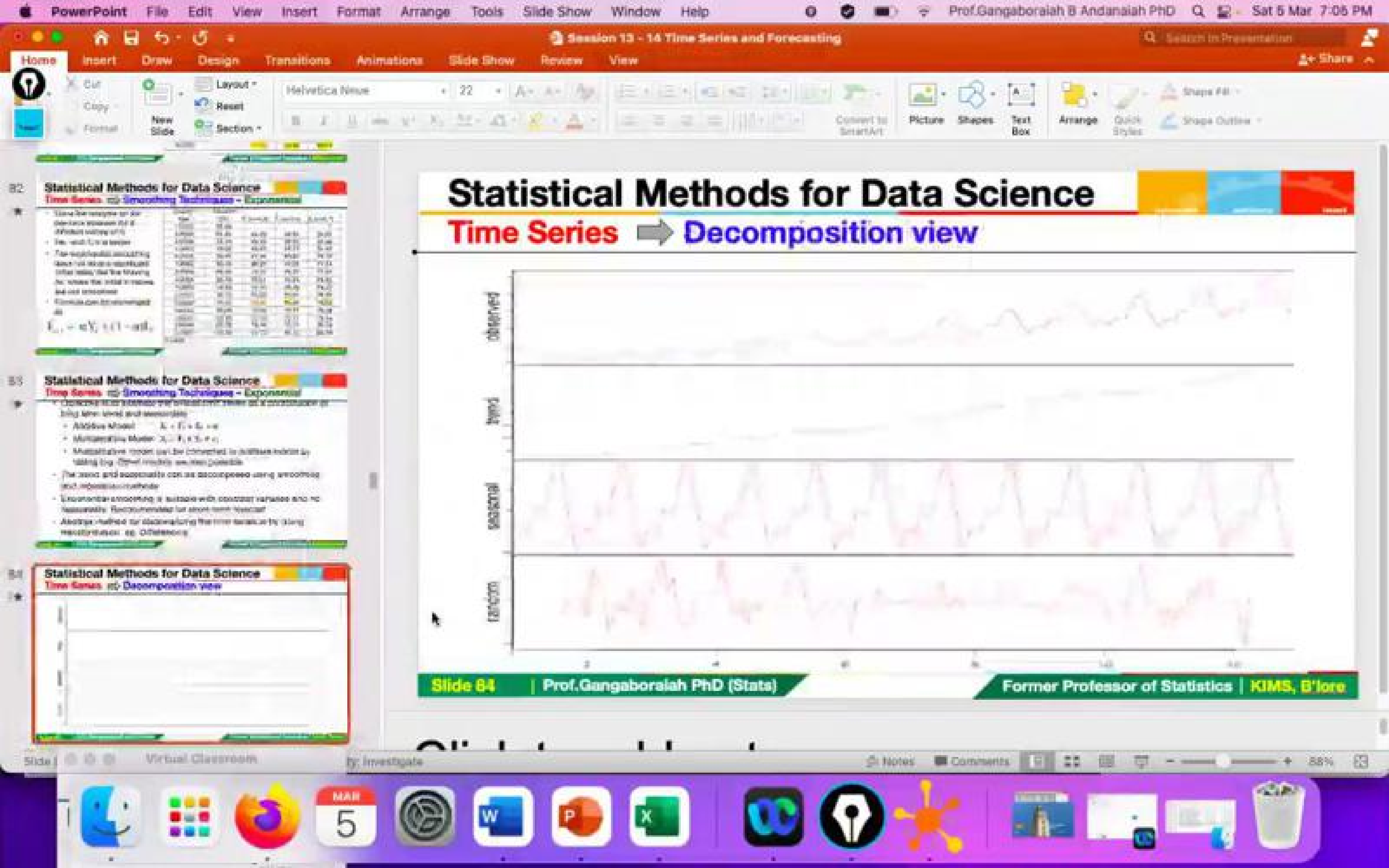
Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74


Statistical Methods for Data Science
Time Series
Decomposition view
observed
trend
seasonal
random
Slide 84
Prof.Gangaboraiah PhD (Stats)
Former Professor of Statistics
KIMS, B'lore
Session 13 - 14 Time Series and Forecasting

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
3/2006	-	79.94	80.40	80.74

Time Series ➡ Smoothing Techniques – Exponential

- Objective is to estimate the overall time series as a combination of long term trend and seasonality
  - Additive Model: $X_t = F_t + S_t + \varepsilon_t$
  - Multiplicative Model: $X_t = F_t \times S_t \times \varepsilon_t$
  - Multiplicative model can be converted to additive model by taking log. Other models are also possible
- The trend and seasonality can be decomposed using smoothing and regression methods
- Exponential smoothing is suitable with constant variance and no seasonality. Recommended for short-term forecast.
- Another method for stationarizing the time series is by doing transformation .eg. Differencing

Time Series ➡ Which performance should be used?

- Choice of the modeler
- **Advantages**
  - MSE gives greater weight to larger deviations (which could result from outliers)
  - MAD gives less weight to larger deviations
  - MAPE gives less overall weight to a large deviation if the time series value is large
  - LAD tells us if all deviations fall below some threshold value
- Although we illustrate all 4 techniques, in general focus will be on MSE and MAD.

Statistical Methods for Data Science

Time Series ⇨ Smoothing Techniques – Moving Averages

Quarte	Utilization	2-Qtr	Abs(Yt -Ft)	$(Y_t - F_t)^2$	3-Qtr	Abs(Yt -	$(Y_t - F_t)^2$	4-Qtr	Abs(	$(Y_t - F_t)^2$
Jan-03	82.5									
Feb-03	81.3	81.9	0.6	0.36	81.7					
Mar-03	81.3	81.3	0	0.00	80.5	0.4	0.16			
Apr-03	79	80.2	1.15	1.32	79	1.53	2.34	81	2.03	4.12
Jan-04	76.6	77.8	1.2	1.44	77.9	2.37	5.62	79.6	2.95	8.70
Feb-04	78	77.3	0.7	0.49	77.7	0.13	0.02	78.7	0.73	0.53
Mar-04	78.4	78.2	0.2	0.04	78.1	0.73	0.53	78	0.4	0.16
Apr-04	78	78.2	0.2	0.04	78.4	0.13	0.02	77.8	0.25	0.06
Jan-05	78.8	78.4	0.4	0.16	78.5	0.4	0.16	78.3	0.5	0.25
Feb-05	78.7	78.8	0.05	0.00	78.6	0.2	0.04	78.5	0.22	0.05
Mar-05	78.4	78.6	0.15	0.02	79	0.23	0.05	78.5	0.08	0.01
Apr-05	80	79.2	0.8	0.64	79.7	0.97	0.94	79	1.02	1.04
Jan-06	80.7	80.4	0.35	0.12	80.5	1	1.00	79.5	1.25	1.56
Feb-06	80.7	80.7	0	0.00	80.7	0.23	0.05	80	0.75	0.56
Mar-06	80.8	80.8	0.05	0.00		0.07	0.00	80.6	0.25	0.06
	MSE	0.33	5.85	4.64	0.84	8.39	10.94	1.42	10.4	17.11
	MSE			0.33			0.84			1.43
	MAD		0.42			0.65			0.87	

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
4/2006	-	79.94	80.40	80.74

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
4/2006	-	79.94	80.40	80.74

Virtual Classroom

Chat | Participants (50)

Disable group chat

Time Series ➡ Smoothing Techniques – Exponential

Time Series ➡ Smoothing Techniques – Exponential

Objective is to estimate the overall time series as a combination of long term trend and seasonality

- Additive Model: $X_t = F_t + S_t + \varepsilon_t$
- Multiplicative Model: $X_t = F_t \times S_t \times \varepsilon_t$
- Multiplicative model can be converted to additive model by taking log. Other models are also possible

Excel File Edit View Insert Format Tools Data Window Help

Prof.Gangaboraiah B Andanaiah PhD Sat 5 Mar 7:10 PM

Time series data

Home Insert Draw Page Layout Formulas Data Review View

...en recovered workbooks? Your recent changes were saved. Do you want to continue working where you left off? Yes No

uarte /year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
an-03	82.5	NA	NA	NA
eb-03	81.3	82.5	82.5	82.5
ar-03	81.3	82.26	82.02	81.66
or-03	79	81.61	80.81	79.8
Jan-04	76.6	80.61	79.13	77.56
Feb-04	78	80.09	78.68	77.87
Ma...				
Ap...				
Ja...				
Fe...				
M...				
Ap...				
Ja...				
Fe...				
M...				

Exponential Smoothing

Input

Input Range: $B$2:$B$16

Damping factor: 0.2

Labels

Output options

Output Range

New Worksheet Ply

New Workbook

Chart Output

Standard Errors

OK

Cancel

Quarter/year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1	82.5	NA	NA	NA
2	81.3	82.50	82.50	82.50
3	81.3	82.26	82.02	81.66
4	79	82.07	81.73	81.41
5	76.6	81.45	80.64	79.72
6	78	80.48	79.02	77.54
7	78.4	79.99	78.61	77.86
8	78	79.67	78.53	78.24
9	78.8	79.34	78.32	78.07
10	78.7	79.23	78.51	78.58
11	78.4	79.12	78.59	78.66
12	80	78.98	78.51	78.48
13	80.7	79.18	79.11	79.54
14	80.7	79.49	79.74	80.35
15	80.8	79.73	80.13	80.60

Time series data

Open ... red workbooks? Your recent changes were saved. Do you want to continue working where you left off?

M5 f_x =B2

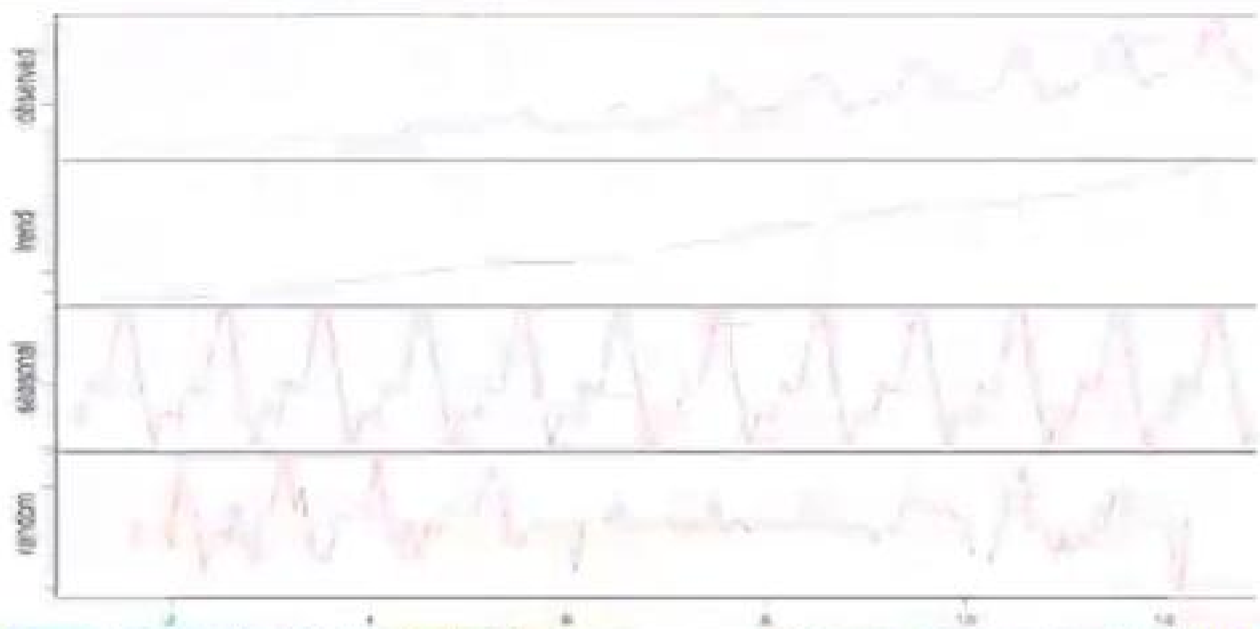
Qua r/ye	Utilizat ion (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
Jan-	82.5	NA	NA	NA
Feb-	81.3	82.5	82.5	82.5
Mar-	81.3	82.26	82.02	81.66
Apr-	79	81.61	80.81	79.8
Jan-04	76.6	80.61	79.13	77.56
Feb-04	78	80.09	78.68	77.87
Mar-04	78.4	79.75	78.57	78.24
Apr-04	78	79.4	78.34	78.07
Jan-05	78.8	79.28	78.52	78.58
Feb-05	78.7	79.16	78.59	78.66
Mar-05	78.4	79.01	78.52	78.48
Apr-05	80	79.21	79.11	79.54
Jan-06	80.7	79.51	79.75	80.35
Feb-06	80.7	79.75	80.13	80.6
Mar-06	80.8	79.96	80.4	80.74

Quarter/year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)		
1	82.5	NA	NA	NA		
2	81.3	82.50	82.50	82.50		
3	81.3	82.26	82.02	81.66	#N/A	#N/A
4	79	82.07	81.73	81.41	82.5	#N/A
5	76.6	81.45	80.64	79.72	81.54	#N/A
6	78	80.48	79.02	77.54	81.348	#N/A
7	78.4	79.99	78.61	77.86	79.4696	1.52869259
8	78	79.67	78.53	78.24	77.17392	2.14517351
9	78.8	79.34	78.32	78.07	77.834784	2.13318006
10	78.7	79.23	78.51	78.58	78.2868568	1.7546587
11	78.4	79.12	78.59	78.66	78.0573914	0.60117149
12	80	78.98	78.51	78.48	78.6514783	0.56970232
13	80.7	79.18	79.11	79.54	78.8902957	0.40049473
14	80.7	79.49	79.74	80.35	78.4580591	0.46119176
15	80.8	79.73	80.13	80.60	79.0014118	0.9063126

Sheet1 Sheet2

Statistical Methods for Data Science

Time Series ➡ Decomposition view



Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
4/2006	-	79.94	80.40	80.74

In stochastic time series,

$$Y_t,\ t \in Z = \{\pm 1, \pm 2, \ldots\}$$

is a family of random variables, Y_t denoting the value of the characteristic of interest at time t.

Thus, $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ is seen as a realized value of the random vector $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$ with joint probability density function $f_{\mathbf{Y}}(y)$.

Index set:

- The index set **T** is a collection of all time functions that can result from random experiment, usually the index **T** denote time. X_t are independent and identically distributed (iid) random variables with mean μ_t and variance σ_t^2. Each realization of this process gives an ensemble or a data set.

Statistical Methods for Data Science

Time Series ➡ Definition of Stochastic Process

The joint distribution function of a finite random variables

$$\{Y_{t_1}, \ldots Y_{t_n}\}, t_1 < t_2 < \ldots < t_n$$

from the collection $\{Y_t, t \in T\}$ is

$$F_{Y_{t_1}, \ldots Y_{t_n}}(y_1, \ldots, y_n) = P[Y_{t_1} \le y_{t_1}, \ldots, P[Y_{t_n} \le y_{t_n}],$$

$$(y_{t_1}, \ldots, y_{t_n}) \in R^n$$

Static

Virtual Classroom

Chat

Participants (46)

Disable group chat

ltp sir please ignore i copend it

Virtual Classroom

Chat | Participants (46)

Disable group chat

Np sir please ignore I copied it

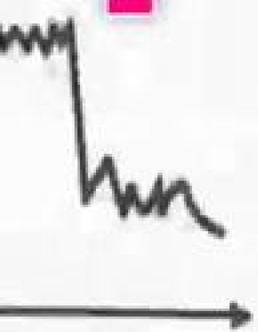
Statistical Methods for Data Science

Time Series ➡ Stationary Stochastic Process in Time series

A special kind of stochastic process is based on the assumption that the process is in a particular state of equilibrium. This type of assumption is called stationarity.

A stochastic process is strictly stationary if its properties are unaffected by a change of origin.

ne series



me series.

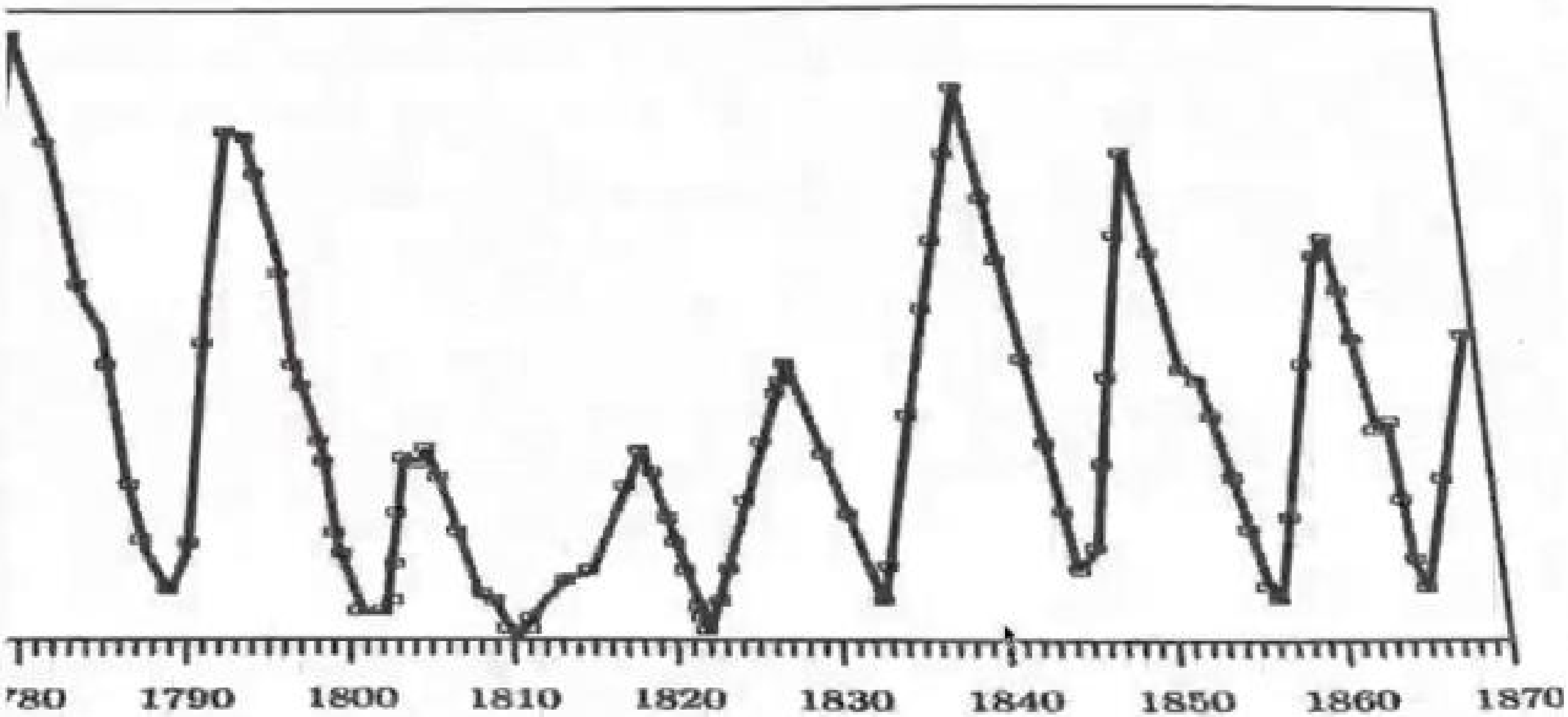


Fig 15.3 The Wolfer sunspot numbers, 1770-1869.
(Source: Box and Jenkins, 1976)

However, many non-stationary time series can be so modified that the reduced to time series obeys the original series.

The two main component which cause lack of stationarity are trend and seasonality. In fitting the stationary time series model, we therefore, assume that the trend and seasonality have been [illegible] from the original series.

- Stationarity implies that the joint distribution of pair of observations y_{t_1}, y_{t_2} viz.,

$$f(y_{t_1}, y_{t_2}) = f(y_{t_1+h}, y_{t_2+h})$$

For any integer h. That is, the joint distribution of any pair of observations on time points which differ by a constant quantity is the same for such pairs.

Science

es

Autocovarianc

Chat

Participants (48)

Disable group chat

...nce and Autocorrelation

t... diagrams for lag k ($k = 1, 2$).

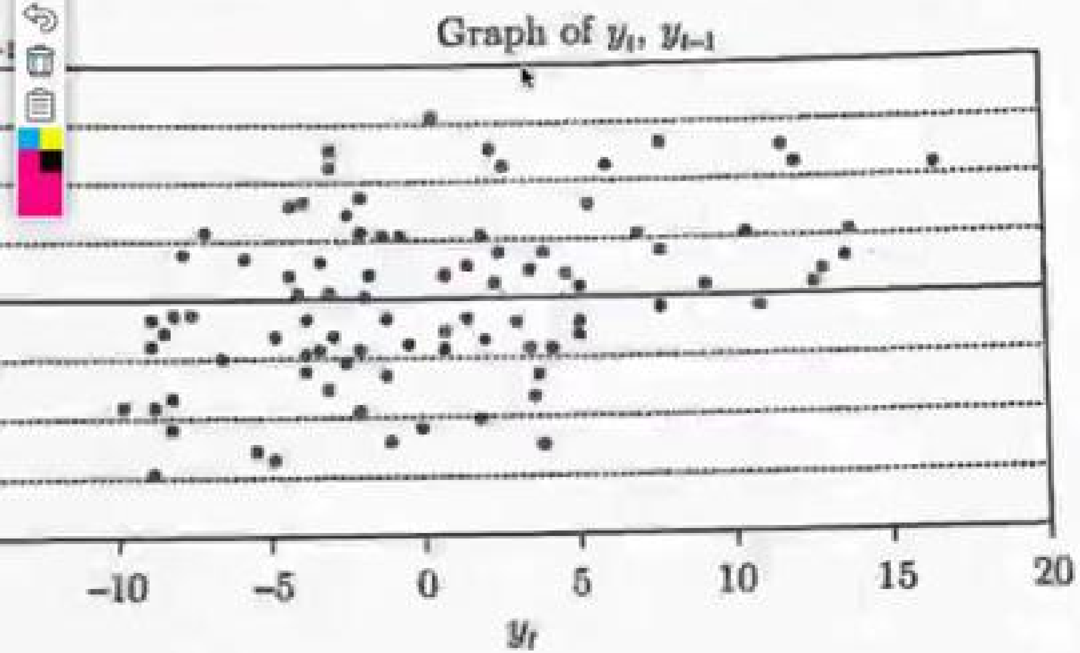


Fig 15.4(b) $Y_t = 0.6\varepsilon_{t-1} + \varepsilon_t$

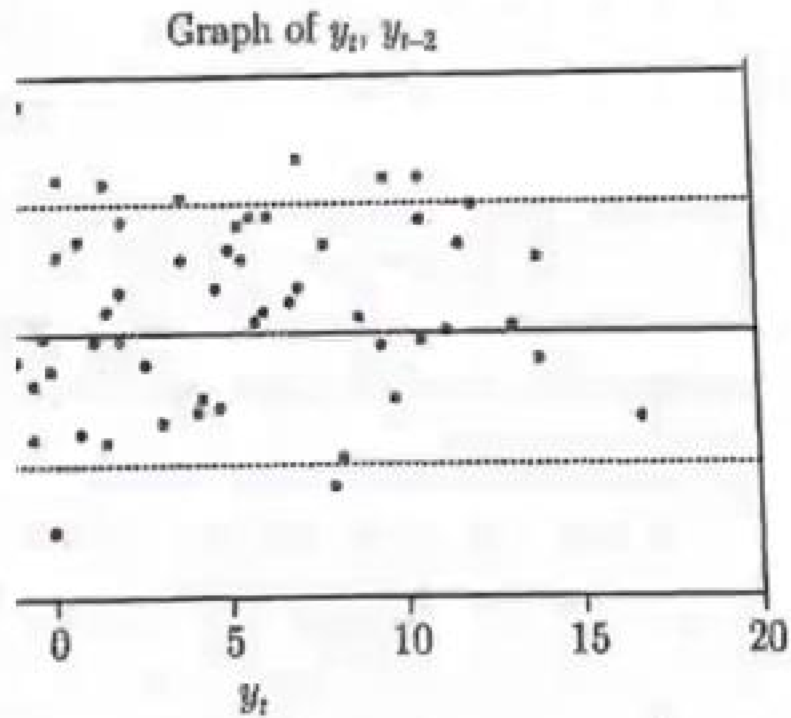


Fig 15.5(b) $Y_t = 0.7\varepsilon_t + \varepsilon_{t-1}$

Time Series ➡ Autocovariance and Autocorrelation

- Let $\{Y_t\}$ be a Time Series with $E\{Y^2\} < \infty$. The mean function of $\{Y_t\}$ is $\mu_y(t) = E(Y_t)$.
- The covariance function of $\{Y_t, Y_{t+h}\}$ is
- $\gamma_h = \text{Cov}(Y_t, Y_{t+h}) = E[(Y_t - \mu)(Y_{t+h} - \mu)]$

Statistical Methods for Data Science

Time Series ➡ Autocovariance and Autocorrelation

- Let $\{Y_t\}$ be a Time Series with $E\{Y^2\} < \infty$. The mean function of $\{Y_t\}$ is $\mu_y(t) = E(Y_t)$.
- The covariance function of $\{Y_t, Y_{t+h}\}$ is
- $\gamma_h = Cov(Y_t, Y_{t+h}) = E[(Y_t-\mu)(Y_{t+h}-\mu)]$

Statistical Methods for Data Science

Time Series ➡ Autocovariance and Autocorrelation

- Clearly,

$$|\gamma_h| \le \gamma_0, \forall\, h = 1, 2, \ldots$$

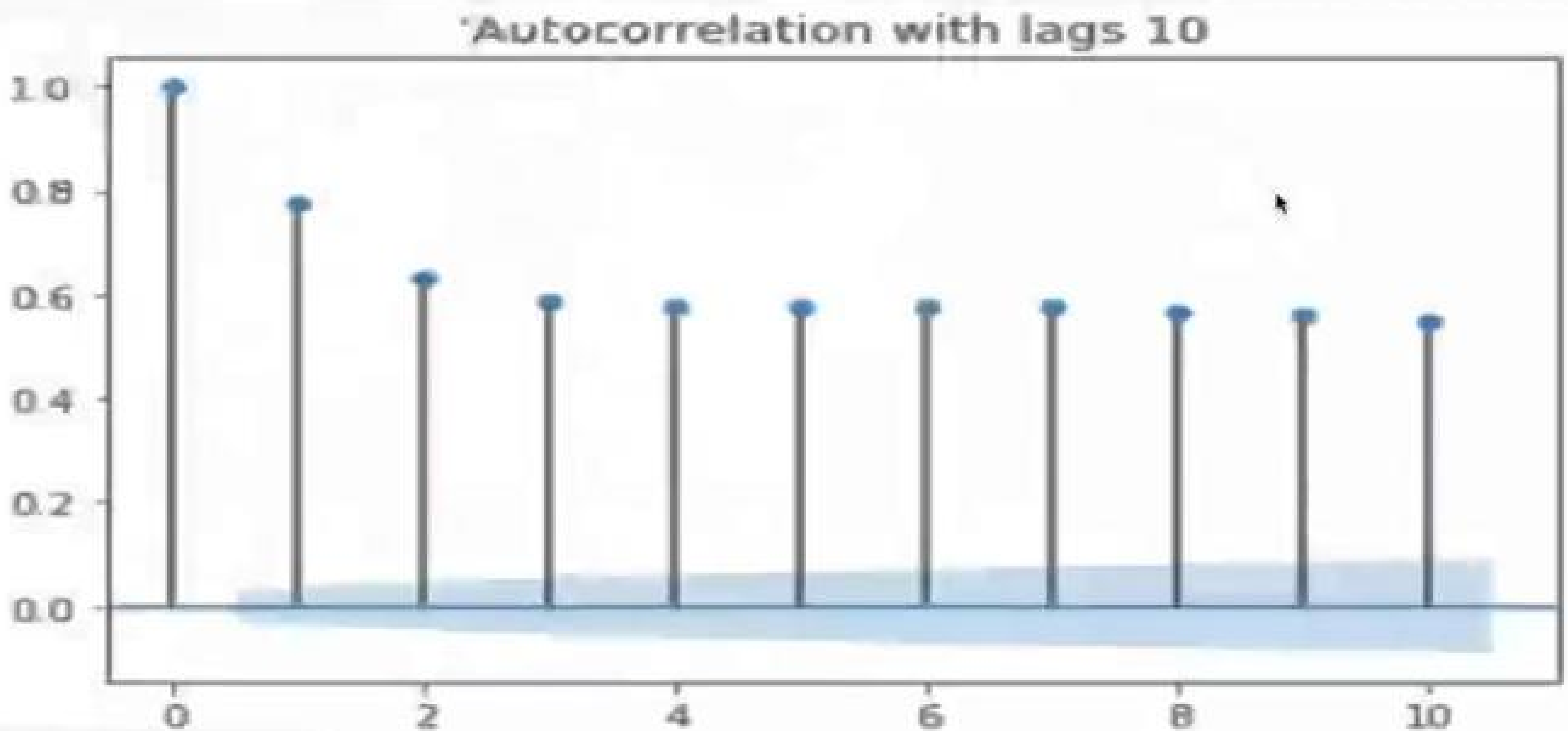
Chat

Participants (48)

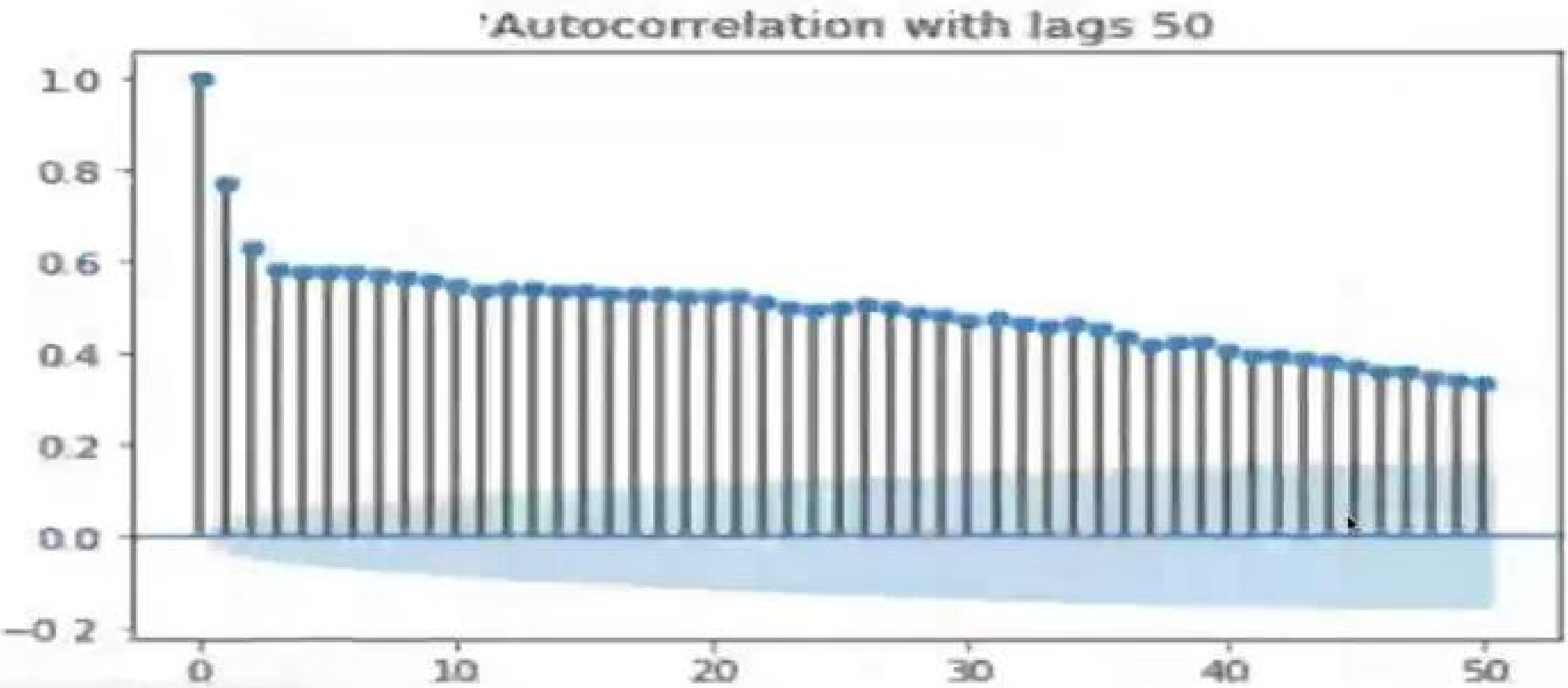
Disable group chat

is it due to rounding error?

me Series → Autocorrelation



Virtual Classroom
Chat
Participants (48)
Disable group chat



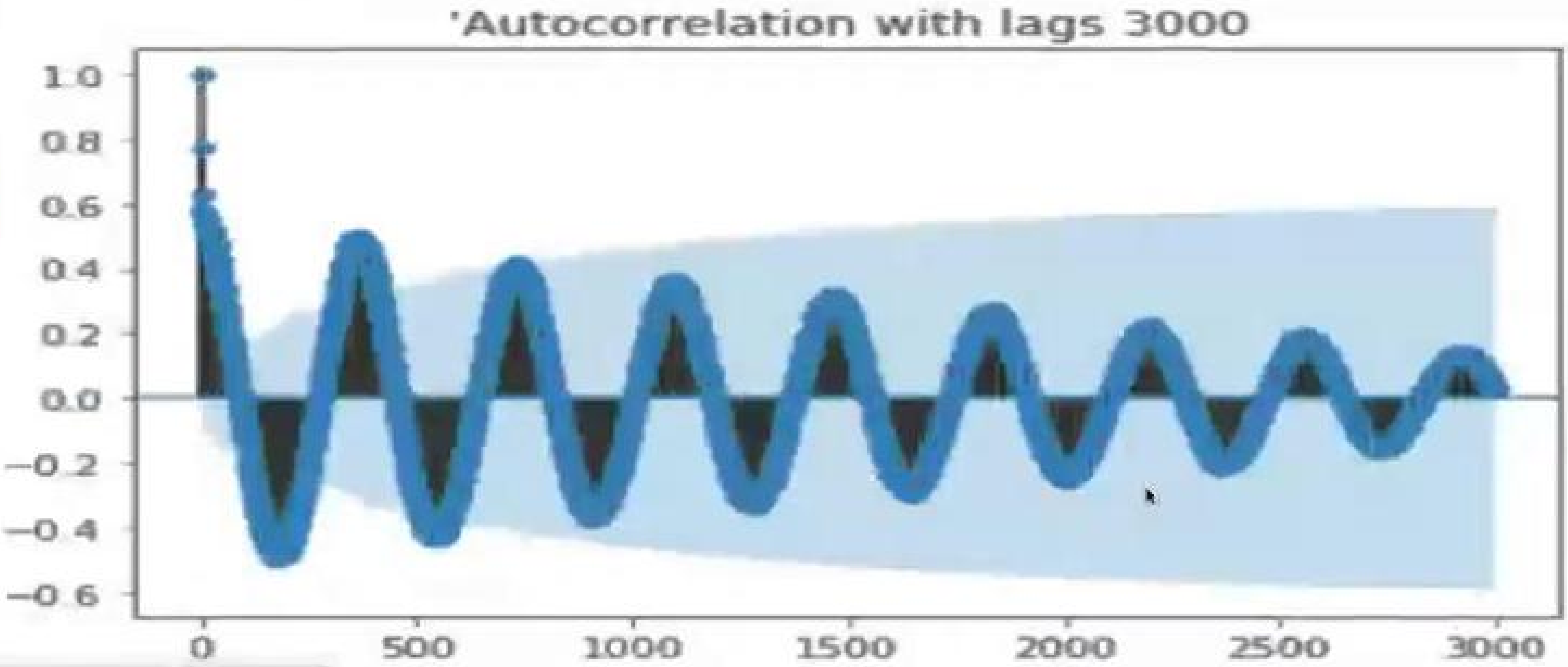
Virtual Classroom

Chat

Participants (48)

Disable group chat

months based on current data .. How



Virtual Classroom

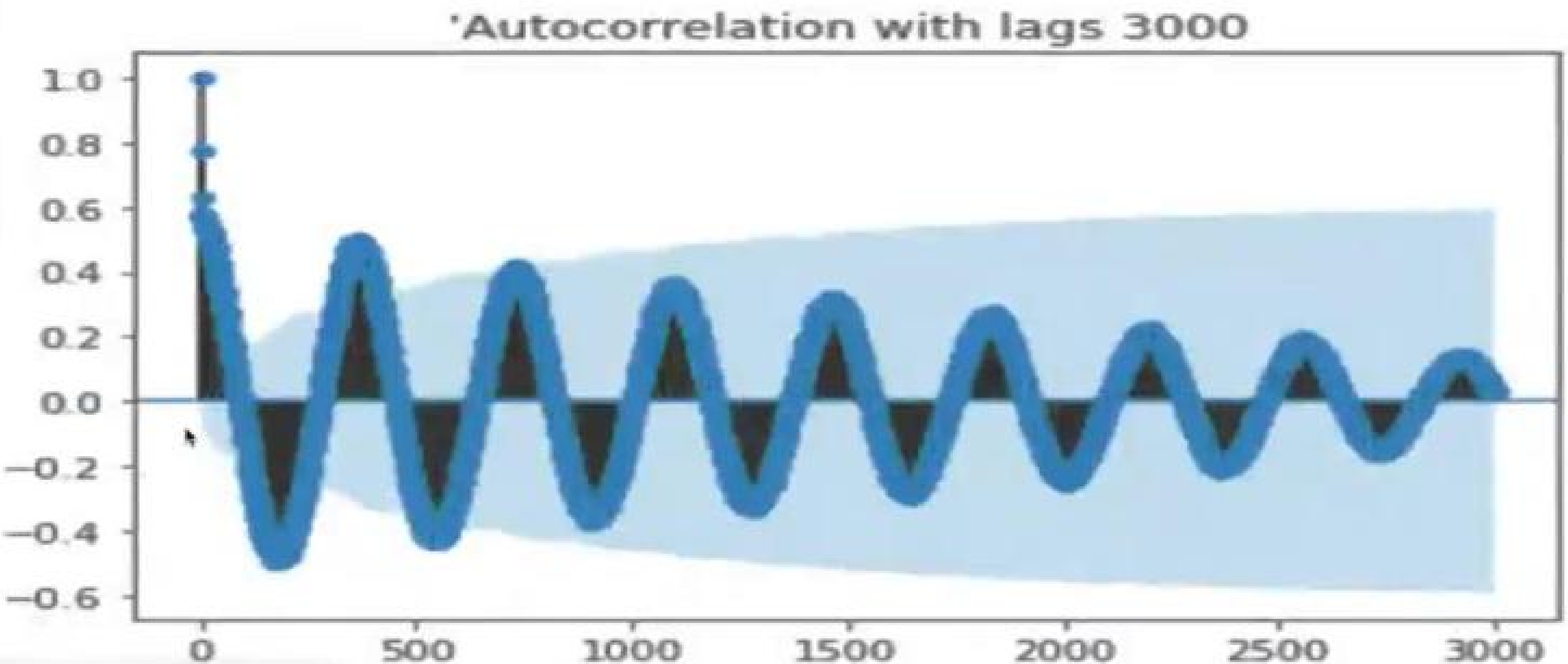
Chat | Participants (48)

Disable group chat

sir, if we need to forecast for next 3 months based on current data .. How

Statistical Methods for Data Science

Time Series ➡ Autocorrelation



Chat
Participants (48)
Disable group chat

Time Series ➡ Partial – Autocorrelation Function (PACF)

- The partial correlation between two variables is a conditional correlation taking into account their dependence on all other remaining variables
  - Eg … n is

$$\frac{\mathrm{Cov}(X_t, X_{t-3} \mid X_{t-1}, X_{t-2})}{\sqrt{\mathrm{Var}(X_t \mid X_{t-1}, X_{t-2})\,\mathrm{Var}(X_{t-3} \mid X_{t-1}, X_{t-2})}}$$

Virtual Classroom
Chat
Participants (48)
Disable group chat
sir, if we need to forecast for next 3 months based on current data .. How

Statistical Methods for Data Science

Time Series ➡ Partial – Autocorrelation Function (PACF)

- The partial correlation between two variables is a conditional correlation taking into account their dependence on all other remaining variables
  - Eg [illegible] is

$$\frac{\operatorname{Cov}(X_t, X_{t-3} \mid X_{t-1}, X_{t-2})}{\sqrt{\operatorname{Var}(X_t \mid X_{t-1}, X_{t-2})\operatorname{Var}(X_{t-3} \mid X_{t-1}, X_{t-2})}}$$

Statistical Methods for Data Science

Time Series ➡ Measurement of forecast error

- To try to determine which one of these forecasting methods gives the "best" forecast, we evaluate the "success" of each method using a *performance measure*

Statistical Methods for Data Science

Time Series ➡ Measurement of forecast error

- To try to determine which one of these forecasting methods gives the “best” forecast, we evaluate the “success” of each method using a *performance measure*

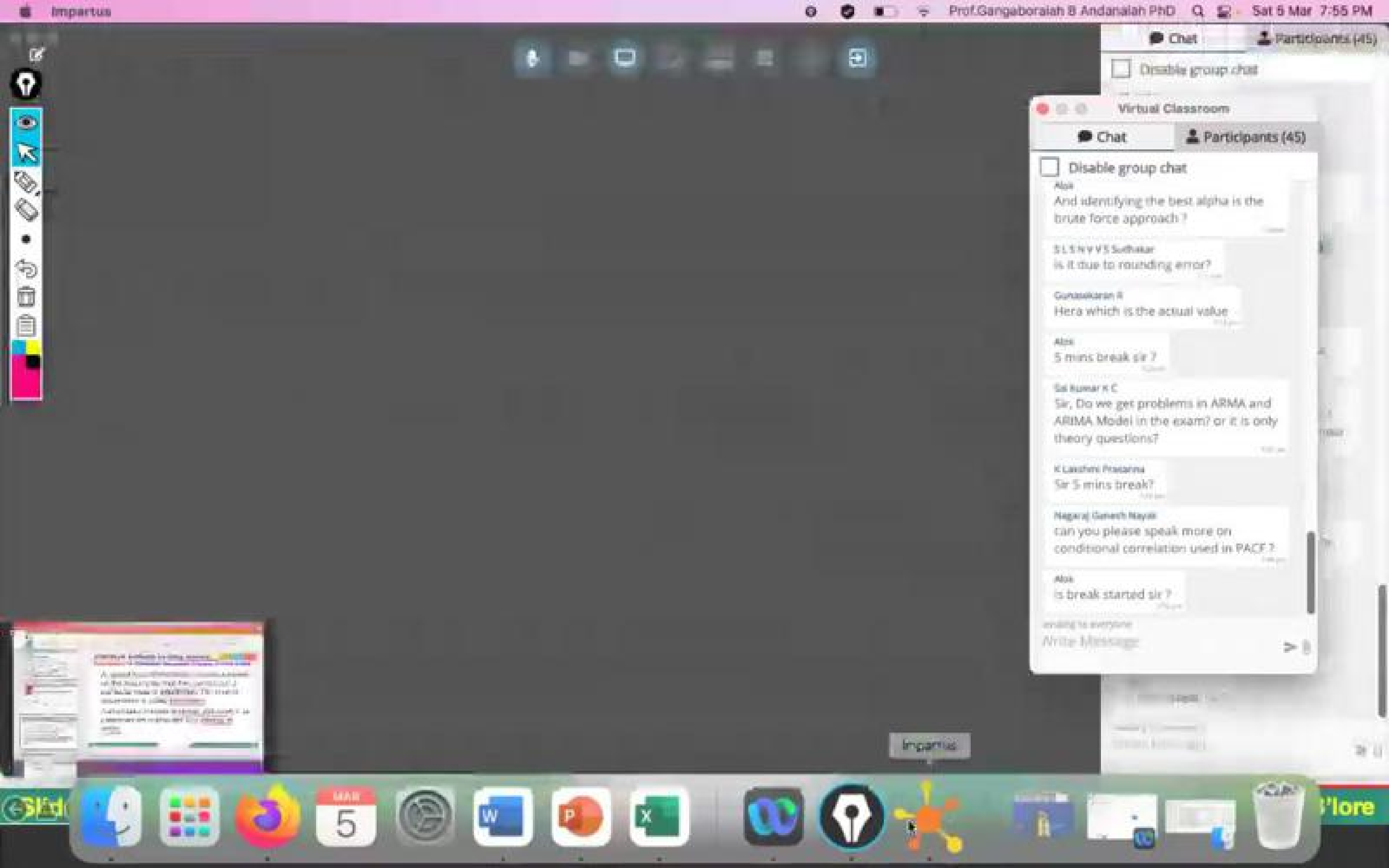
Statistical Methods for Data Science

Time Series ➡ Measurement of forecast error

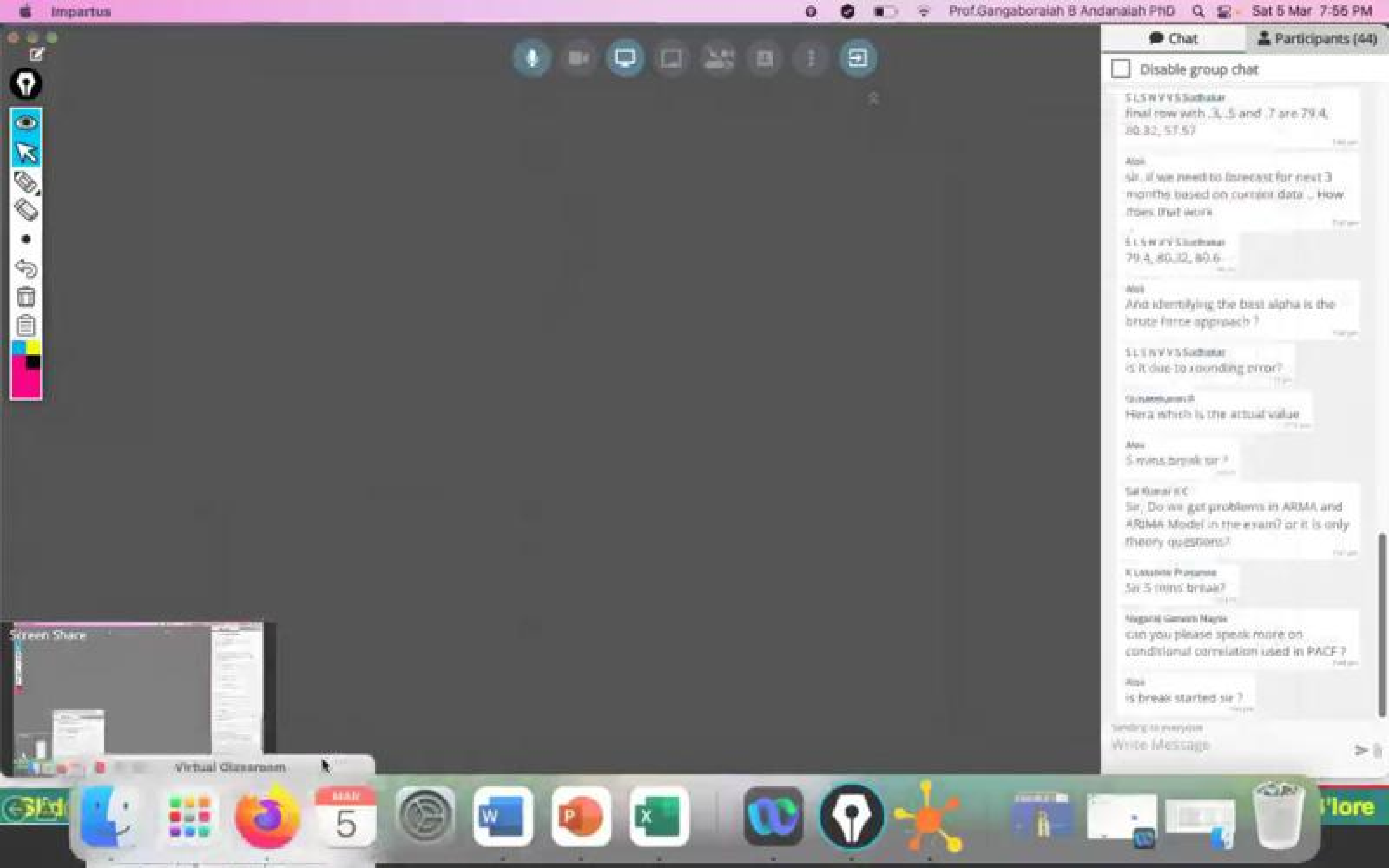
- To try to determine which one of these forecasting methods gives the "best" forecast, we evaluate the "success" of each method using a ***performance measure***

Time Series ➡ Which performance should be used?

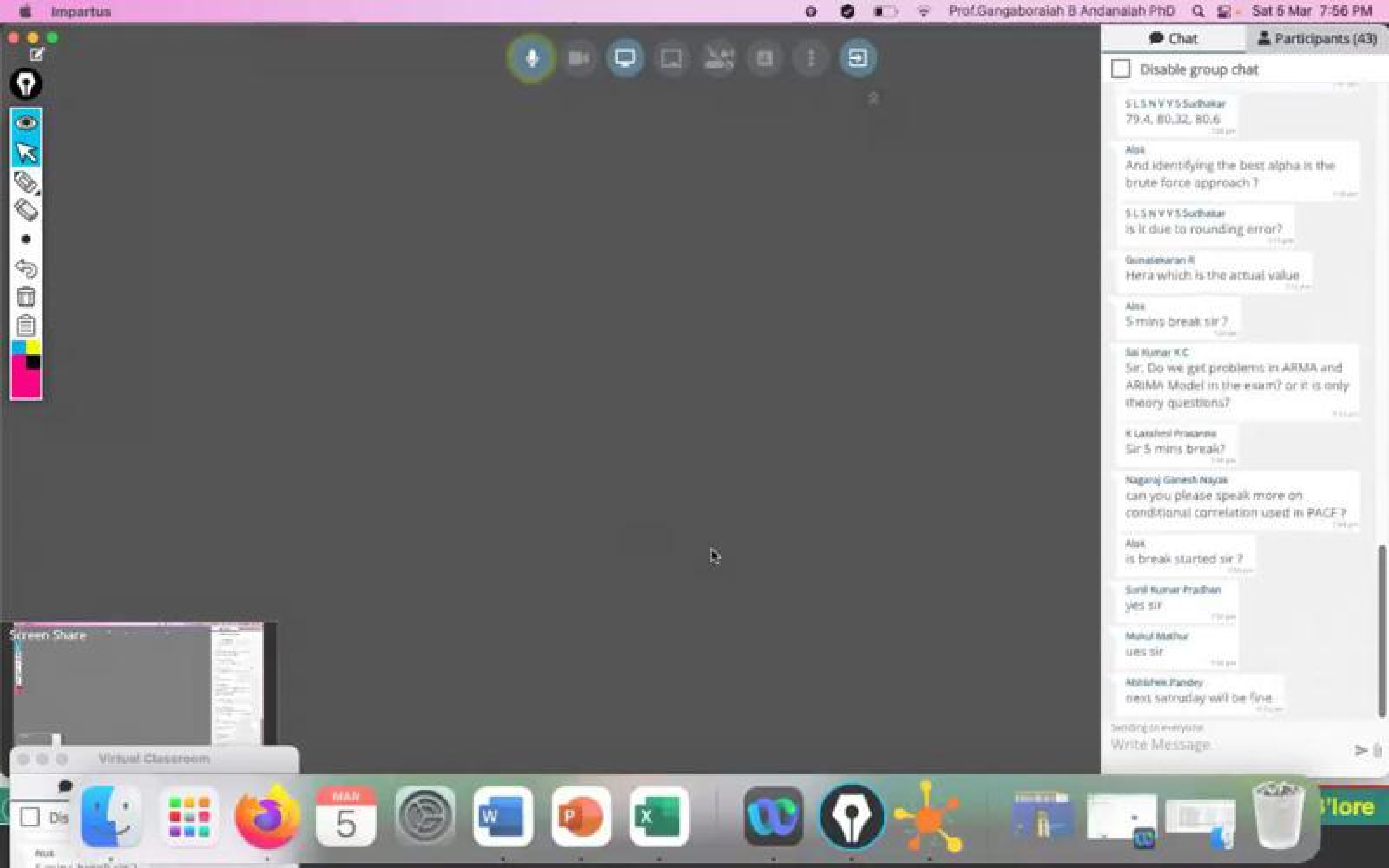
- Choice of the modeler
- **Advantages**
  - MSE gives greater weight to larger deviations (which could result from outliers)



Impartus
Prof.Gangaboraiah B Andanaiah PhD
Sat 5 Mar 7:55 PM
Virtual Classroom
Chat
Participants (45)
Disable group chat
Alok
And identifying the best alpha is the brute force approach ?
S L S N V V S Sudhakar
is it due to rounding error?
Hera which is the actual value
Alok
5 mins break sir ?
Sai Kumar K C
Sir, Do we get problems in ARMA and ARIMA Model in the exam? or it is only theory questions?
K Lakshmi Prasanna
Sir 5 mins break?
Nagaraj Ganesh Nayak
can you please speak more on conditional correlation used in PACF ?
Alok
is break started sir ?
Write Message
MAR
5



Impartus
Prof.Gangaboraiah B Andanaiah PhD
Sat 5 Mar 7:56 PM
Chat
Participants (44)
Disable group chat
S L S N V V S Sudhakar
final row with .3, .5 and .7 are 79.4, 80.32, 57.57
sir, if we need to forecast for next 3 months based on current data .. How does that work
S L S N V V S Sudhakar
79.4, 80.32, 80.6
And identifying the best alpha is the brute force approach ?
S L S N V V S Sudhakar
is it due to rounding error?
Here which is the actual value
5 mins break sir ?
Sir, Do we get problems in ARMA and ARIMA Model in the exam? or it is only theory questions?
Sir 5 mins break?
can you please speak more on conditional correlation used in PACF ?
is break started sir ?
Write Message
Screen Share
Virtual Classroom
MAR
5


Impartus
Prof.Gangaboraiah B Andanaiah PhD
Sat 6 Mar 7:56 PM
Chat
Participants (43)
Disable group chat
S L S N V V S Sudhakar
79.4, 80.32, 80.6
Alok
And identifying the best alpha is the brute force approach ?
S L S N V V S Sudhakar
is it due to rounding error?
Gunasekaran R
Hera which is the actual value
Alok
5 mins break sir ?
Sai Kumar K C
Sir, Do we get problems in ARMA and ARIMA Model in the exam? or it is only theory questions?
K Lakshmi Prasanna
Sir 5 mins break?
Nagaraj Ganesh Nayak
can you please speak more on conditional correlation used in PACF ?
Alok
is break started sir ?
Sunil Kumar Pradhan
yes sir
Mukul Mathur
ues sir
Abhishek Pandey
next satruday will be fine
Sending to everyone
Write Message
Screen Share
Virtual Classroom
MAR
5

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α

Quarter/ Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.56
2/2004	78.00	80.09	78.68	77.87
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.59	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
MSE				

Virtual Classroom
Chat | Participants (44)
Disable group chat
ARIMA Model in the exam? or is it only theory questions?

Statistical Methods for Data Science

Time Series ➡ Smoothing Techniques – Exponential

- Done the analysis on the previous problem for 3 different values of α
- For α=0.7, it is better
- The exponential smoothing does not have a significant initial delay like the Moving Av. where the initial k values are not smoothed
- Formula can be rearranged as

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

Quarter Year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
1/2003	82.50			
2/2003	81.30	82.50	82.50	82.50
3/2003	81.30	82.26	82.02	81.66
4/2003	79.00	81.61	80.81	79.80
1/2004	76.60	80.61	79.13	77.58
2/2004	78.00	80.09	78.68	77.67
3/2004	78.40	79.75	78.57	78.24
4/2004	78.00	79.40	78.34	78.07
1/2005	78.80	79.28	78.52	78.58
2/2005	78.70	79.16	78.50	78.66
3/2005	78.40	79.01	78.52	78.48
4/2005	80.00	79.21	79.11	79.54
1/2006	80.70	79.51	79.75	80.35
2/2006	80.70	79.75	80.13	80.60
3/2006	80.80	79.96	80.40	80.74
MSE				

79	81.61	80.81	79.8	4	79	82.07	81.73	81.41	9.41	7.46	5.80
76.6	80.61	79.13	77.56	5	76.6	81.45	80.64	79.72	23.57	16.32	9.75
78	80.09	78.68	77.87	6	78	80.48	79.02	77.54	6.17	1.05	0.21
78.4	79.75	78.57	78.24	7	78.4	79.99	78.61	77.86	2.52	0.05	0.29
78	79.4	78.34	78.07	8	78	79.67	78.53	78.24	2.79	0.28	0.06
78.8	79.28	78.52	78.58	9	78.8	79.34	78.32	78.07	0.29	0.23	0.53
78.7	79.16	78.59	78.66	10	78.7	79.23	78.51	78.58	0.28	0.04	0.01
78.4	79.01	78.52	78.48	11	78.4	79.12	78.59	78.66	0.52	0.03	0.07
80	79.21	79.11	79.54	12	80	78.98	78.51	78.48	1.04	2.22	2.31
80.7	79.51	79.75	80.35	13	80.7	79.18	79.11	79.54	2.30	2.54	1.34
80.7	79.75	80.13	80.6	14	80.7	79.49	79.74	80.35	1.47	0.91	0.12
80.8	79.96	80.4	80.74	15	80.8	79.73	80.13	80.60	1.15	0.45	0.04
				16		79.94	80.40	80.74			

Time series data

=(H3-I3)^2

tilizat n (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)
82.5	NA	NA	NA
81.3	82.5	82.5	82.5
81.3	82.26	82.02	81.66
79	81.61	80.81	79.8
76.6	80.61	79.13	77.56
78	80.09	78.68	77.87
78.4	79.75	78.57	78.24
78	79.4	78.34	78.07
78.8	79.28	78.52	78.58
78.7	79.16	78.59	78.66
78.4	79.01	78.52	78.48
80	79.21	79.11	79.54
80.7	79.51	79.75	80.35
80.7	79.75	80.13	80.6
80.8	79.96	80.4	80.74

Quarter/year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)	MSE (0.2)	MSE (0.4)	MSE (0.7)
1	82.5	NA	NA	NA	NA	NA	NA
2	81.3	82.50	82.50	82.50	1.44	1.44	1.44
3	81.3	82.26	82.02	81.66	0.92	0.52	0.13
4	79	82.07	81.73	81.41	9.41	7.46	5.80
5	76.6	81.45	80.64	79.72	23.57	16.32	9.75
6	78	80.48	79.02	77.54	6.17	1.05	0.21
7	78.4	79.99	78.61	77.86	2.52	0.05	0.29
8	78	79.67	78.53	78.24	2.79	0.28	0.06
9	78.8	79.34	78.32	78.07	0.29	0.23	0.53
10	78.7	79.23	78.51	78.58	0.28	0.04	0.01
11	78.4	79.12	78.59	78.66	0.52	0.03	0.07
12	80	78.98	78.51	78.48	1.04	2.22	2.31
13	80.7	79.18	79.11	79.54	2.30	2.54	1.34
14	80.7	79.49	79.74	80.35	1.47	0.91	0.12
15	80.8	79.73	80.13	80.60	1.15	0.45	0.04

Sheet1 Sheet2

Time series data

...en recovered workbooks? Your recent changes were saved. Do you want to continue working where you left off? Yes No

fx =SUM(M3:M16)

B	C	D	E	F	G	H	I	J	K	L	M	N	O
76.6	80.61	79.13	77.56		5	76.6	81.45	80.64	79.72		23.57	16.32	9.75
78	80.09	78.68	77.87		6	78	80.48	79.02	77.54		6.17	1.05	0.21
78.4	79.75	78.57	78.24		7	78.4	79.99	78.61	77.86		2.52	0.05	0.29
78	79.4	78.34	78.07		8	78	79.67	78.53	78.24		2.79	0.28	0.06
78.8	79.28	78.52	78.58		9	78.8	79.34	78.32	78.07		0.29	0.23	0.53
78.7	79.16	78.59	78.66		10	78.7	79.23	78.51	78.58		0.28	0.04	0.01
78.4	79.01	78.52	78.48		11	78.4	79.12	78.59	78.66		0.52	0.03	0.07
80	79.21	79.11	79.54		12	80	78.98	78.51	78.48		1.04	2.22	2.31
80.7	79.51	79.75	80.35		13	80.7	79.18	79.11	79.54		2.30	2.54	1.34
80.7	79.75	80.13	80.6		14	80.7	79.49	79.74	80.35		1.47	0.91	0.12
80.8	79.96	80.4	80.74		15	80.8	79.73	80.13	80.60		1.15	0.45	0.04
					16		79.94	80.40	80.74		[illegible]		

Sheet1 Sheet2

B	C	D	E	F	G	H	I	J	K	L	M	N	O
76.6	80.61	79.13	77.56		5	76.6	81.45	80.64	79.72		23.57	16.32	9.75
78	80.09	78.68	77.87		6	78	80.48	79.02	77.54		6.17	1.05	0.21
78.4	79.75	78.57	78.24		7	78.4	79.99	78.61	77.86		2.52	0.05	0.29
78	79.4	78.34	78.07		8	78	79.67	78.53	78.24		2.79	0.28	0.06
78.8	79.28	78.52	78.58		9	78.8	79.34	78.32	78.07		0.29	0.23	0.53
78.7	79.16	78.59	78.66		10	78.7	79.23	78.51	78.58		0.28	0.04	0.01
78.4	79.01	78.52	78.48		11	78.4	79.12	78.59	78.66		0.52	0.03	0.07
80	79.21	79.11	79.54		12	80	78.98	78.51	78.48		1.04	2.22	2.31
80.7	79.51	79.75	80.35		13	80.7	79.18	79.11	79.54		2.30	2.54	1.34
80.7	79.75	80.13	80.6		14	80.7	79.49	79.74	80.35		1.47	0.91	0.12
80.8	79.96	80.4	80.74		15	80.8	79.73	80.13	80.60		1.15	0.45	0.04
					16		79.94	80.40	80.74		53.87	33.53	22.11

=O17/15

1	82.5	NA	NA	NA		NA	NA	NA
2	81.3	82.50	82.50	82.50		1.44	1.44	1.44
3	81.3	82.26	82.02	81.66		0.92	0.52	0.13
4	79	82.07	81.73	81.41		9.41	7.46	5.80
5	76.6	81.45	80.64	79.72		23.57	16.32	9.75
6	78	80.48	79.02	77.54		6.17	1.05	0.21
7	78.4	79.99	78.61	77.86		2.52	0.05	0.29
8	78	79.67	78.53	78.24		2.79	0.28	0.06
9	78.8	79.34	78.32	78.07		0.29	0.23	0.53
10	78.7	79.23	78.51	78.58		0.28	0.04	0.01
11	78.4	79.12	78.59	78.66		0.52	0.03	0.07
12	80	78.98	78.51	78.48		1.04	2.22	2.31
13	80.7	79.18	79.11	79.54		2.30	2.54	1.34
14	80.7	79.49	79.74	80.35		1.47	0.91	0.12
15	80.8	79.73	80.13	80.60		1.15	0.45	0.04
16		79.94	80.40	80.74		53.87	33.53	22.11
					MSE	3.591235366	2.2355817	1.473684

Time Series ➡ Autoregressive (AR) models

- An **autoregressive (AR) model** is a representation of some type of random process; used to describe certain time-varying processes. The autoregressive model specifies that the output variable depends linearly on its own previous values and on a stochastic (random) error term; thus the model is in the form of a stochastic difference equation.

$$y_t = c_t + \emptyset_1 y_{t-1} + \emptyset_2 y_{t-2} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t$$

=M17/15

1	82.5	NA	NA	NA	NA	NA	NA
2	81.3	82.50	82.50	82.50	1.44	1.44	1.44
3	81.3	82.26	82.02	81.66	0.92	0.52	0.13
4	79	82.07	81.73	81.41	9.41	7.46	5.80
5	76.6	81.45	80.64	79.72	23.57	16.32	9.75
6	78	80.48	79.02	77.54	6.17	1.05	0.21
7	78.4	79.99	78.61	77.86	2.52	0.05	0.29
8	78	79.67	78.53	78.24	2.79	0.28	0.06
9	78.8	79.34	78.32	78.07	0.29	0.23	0.53
10	78.7	79.23	78.51	78.58	0.28	0.04	0.01
11	78.4	79.12	78.59	78.66	0.52	0.03	0.07
12	80	78.98	78.51	78.48	1.04	2.22	2.31
13	80.7	79.18	79.11	79.54	2.30	2.54	1.34
14	80.7	79.49	79.74	80.35	1.47	0.91	0.12
15	80.8	79.73	80.13	80.60	1.15	0.45	0.04
16		79.94	80.40	80.74	53.87	33.53	22.11
				MSE	3.59	2.24	1.47

Microsoft Excel

Quarter/year	Utilization (%)	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)	MSE (0.2)	MSE (0.4)	MSE
1	82.5	NA	NA	NA	NA	NA	NA
2	81.3	82.50	82.50	82.50	1.44	1.44	
3	81.3	82.26	82.02	81.66	0.92	0.52	
4	79	82.07	81.73	81.41	9.41	7.46	
5	76.6	81.45	80.64	79.72	23.57	16.32	
6	78	80.48	79.02	77.54	6.17	1.05	
7	78.4	79.99	78.61	77.86	2.52	0.05	
8	78	79.67	78.53	78.24	2.79	0.28	
9	78.8	79.34	78.32	78.07	0.29	0.23	
10	78.7	79.23	78.51	78.58	0.28	0.04	
11	78.4	79.12	78.59	78.66	0.52	0.03	
12	80	78.98	78.51	78.48	1.04	2.22	
13	80.7	79.18	79.11	79.54	2.30	2.54	
14	80.7	79.49	79.74	80.35	1.47	0.91	
15	80.8	79.73	80.13	80.60	1.15	0.45	

Time series data

=VARP(H17)

G	H	I	P	Q	R	S
1	82.5	81.3	NA	NA	NA	NA
2	81.3	79	82.50	82.50	82.50	
3	81.3	76.6	82.26	82.02	81.66	
4	79	78	82.07	81.73	81.41	
5	76.6	78.4	81.45	80.64	79.72	
6	78	78	80.48	79.02	77.54	
7	78.4	78.8	79.99	78.61	77.86	
8	78	78.7	79.67	78.53	78.24	
9	78.8	78.4	79.34	78.32	78.07	
10	78.7	80	79.23	78.51	78.58	
11	78.4	80.7	79.12	78.59	78.66	
12	80	80.7	78.98	78.51	78.48	
13	80.7	80.8	79.18	79.11	79.54	
14	80.7		79.49	79.74	80.35	
15	80.8		79.73	80.13	80.60	
Mean	79.55	=VARP(H17)	79.94	80.40	80.74	
						MSE

Formula Builder

VARP

Number1 = 79.54666667

H17

Number2 = number

Result: 0

Done

VARP

This function is available for compatibility with Excel 2007 and earlier.

Calculates variance based on the entire population (ignores logical values and text in the population)

More help on this function

Sheet1 Sheet2

Time series data

=VARP(H2:H16)

G	H	I	P	Q	R
4	79	78	82.07	81.73	81.41
5	76.6	78.4	81.45	80.64	79.72
6	78	78	80.48	79.02	77.54
7	78.4	78.8	79.99	78.61	77.86
8	78	78.7	79.67	78.53	78.24
9	78.8	78.4	79.34	78.32	78.07
10	78.7	80	79.23	78.51	78.58
11	78.4	80.7	79.12	78.59	78.66
12	80	80.7	78.98	78.51	78.48
13	80.7	80.8	79.18	79.11	79.54
14	80.7		79.49	79.74	80.35
15	80.8		79.73	80.13	80.60
Mean	79.55	2.47	79.94	80.40	80.74

MSE

Formula Builder

VARP

Number1 = {82.5;81.3;81.3;79;76.6;78;78.4;...

H2:H16

Result: 2.465155556

VARP

This function is available for compatibility with Excel 2007 and earlier.

Calculates variance based on the entire population (ignores logical values and text in the population).

Quarter/ year	Utilization (%)	lag 2	y1=yt-mean	F_t (α=0.2)	F_t (α=0.4)	F_t (α=0.7)	MSE (0.2)	MSE (0.4)	M
1	82.5	81.3	2.95	NA	NA	NA	NA	NA	NA
2	81.3	79	1.75	82.50	82.50	82.50	1.44	1.44	
3	81.3	76.6	1.75	82.26	82.02	81.66	0.92	0.52	
4	79	78	-0.55	82.07	81.73	81.41	9.41	7.46	
5	76.6	78.4	-2.95	81.45	80.64	79.72	23.57	16.32	
6	78	78	-1.55	80.48	79.02	77.54	6.17	1.05	
7	78.4	78.8	-1.15	79.99	78.61	77.86	2.52	0.05	
8	78	78.7	-1.55	79.67	78.53	78.24	2.79	0.28	
9	78.8	78.4	-0.75	79.34	78.32	78.07	0.29	0.23	
10	78.7	80	-0.85	79.23	78.51	78.58	0.28	0.04	
11	78.4	80.7	-1.15	79.12	78.59	78.66	0.52	0.03	
12	80	80.7	0.45	78.98	78.51	78.48	1.04	2.22	
13	80.7	80.8	1.15	79.18	79.11	79.54	2.30	2.54	
14	80.7			79.49	79.74	80.35	1.47	0.91	
15	80.8			79.73	80.13	80.60	1.15	0.45	

Time series data

G	H	I	J	K	L	P	Q	R	S	T
8	78	78.7	-1.55	-0.85	1.3095	79.67	78.53	78.24		2.79
9	78.8	78.4	-0.75	-1.15	0.8562	79.34	78.32	78.07		0.29
10	78.7	80	-0.85	0.45	-0.384	79.23	78.51	78.58		0.28
11	78.4	80.7	-1.15	1.15	-1.322	79.12	78.59	78.66		0.52
12	80	80.7	0.45	1.15	0.5228	78.98	78.51	78.48		1.04
13	80.7	80.8	1.15	1.25	1.4455	79.18	79.11	79.54		2.30
14	80.7				8.9536	79.49	79.74	80.35		1.47
15	80.8			Autocovariance	0.6887	79.73	80.13	80.60		1.15
Mean	79.55			Autocorrelation		79.94	80.40	80.74		53.87
Variance	2.47								MSE	3.59

The following data gives the cargo exports y_t (in units of 10^3 metric tons) at the principal port in India for a period of 20 years

Find autocovariance, autocorrelation and auto regression

Year (t)	y_t	lag 1 ($y_1 = y_{t-1}$)	lag 2 ($y_2 = y_{t-2}$)	lag3 ($y_1 = y_{t-3}$)	lag 1 ($y_4 = y_{t-4}$)	lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m	y4-m	y5-m	(yt-m)(	(yt-m
1	1230	1345	1382	1416	1593	1802	-1010.2	-895	-858	-824	-647	-438.2	9E+05	9E+
2	1345	1382	1416	1593	1802	1817	-895.2	-858	-824	-647	-438	-423.2	8E+05	7E+
3	1382	1416	1593	1802	1817	1995	-858.2	-824	-647	-438	-423	-245.2	7E+05	6E+
4	1416	1593	1802	1817	1995	2212	-824.2	-647	-438	-423	-245	-28.2	5E+05	4E+
5	1593	1802	1817	1995	2212	2607	-647.2	-438	-423	-245	-28	366.8	3E+05	3E+
6	1802	1817	1995	2212	2607	2698	-438.2	-423	-245	-28	367	457.8	2E+05	1E+
7	1817	1995	2212	2607	2698	2703	-423.2	-245	-28	367	458	462.8	1E+05	119
8	1995	2212	2607	2698	2703	2738	-245.2	-28	367	458	463	497.8	6915	-899
9	2212	2607	2698	2703	2738	2810	-28.2	367	458	463	498	569.8	-10344	-129
10	2607	2698	2703	2738	2810	2632	366.8	458	463	498	570	391.8	2E+05	2E+
11	2698	2703	2738	2810	2632	2548	457.8	463	498	570	392	307.8	2E+05	2E+
12	2703	2738	2810	2632	2548	2881	462.8	498	570	392	308	640.8	2E+05	3E+
13	2738	2810	2632	2548	2881	2725	497.8	570	392	308	641	484.8	3E+05	2E+
14	2810	2632	2548	2881	2725	2911	569.8	392	308	641	485	670.8	2E+05	2E+
15	2632	2548	2881	2725	2911	2759	391.8	308	641	485	671	518.8	1E+05	3E+

rincipal port in India for a period of 20 years

lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m	y4-m	y5-m	(yt-m)('	(yt-m)('	(yt-m)(y	(yt-m)(y	(yt-m)(y5-m)
1802	-1010.2	-895	-858	-824	-647	-438.2	9E+05	9E+05	832607	653801	442670
1817	-895.2	-858	-824	-647	-438	-423.2	8E+05	7E+05	579373	392277	378849
1995	-858.2	-824	-647	-438	-423	-245.2	7E+05	6E+05	376063	363190	210431
2212	-824.2	-647	-438	-423	-245	-28.2	5E+05	4E+05	348801	202094	23242
2607	-647.2	-438	-423	-245	-28	366.8	3E+05	3E+05	158693	18251	-2E+05
2698	-438.2	-423	-245	-28	367	457.8	2E+05	1E+05	12357	-2E+05	-2E+05
2703	-423.2	-245	-28	367	458	462.8	1E+05	11934	-2E+05	-2E+05	-2E+05
2738	-245.2	-28	367	458	463	497.8	6915	-89939	-1E+05	-1E+05	-1E+05
2810	-28.2	367	458	463	498	569.8	-10344	-12910	-13051	-14038	-16068
2632	366.8	458	463	498	570	391.8	2E+05	2E+05	182593	209003	143712
2548	457.8	463	498	570	392	307.8	2E+05	2E+05	260854	179366	140911
2881	462.8	498	570	392	308	640.8	2E+05	3E+05	181325	142450	296562
2725	497.8	570	392	308	641	484.8	3E+05	2E+05	153223	318990	241333
2911	569.8	392	308	641	485	670.8	2E+05	2E+05	365128	276239	382222
2759	391.8	308	641	485	671	518.8	1E+05	3E+05	189945	262819	203266

principal port in India for a period of 20 years

lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m	y4-m	y5-m	(yt-m)(	(yt-m)(	(yt-m)(y	(yt-m)(y	(yt-m)(y5-m)
1802	-1010.2	-895	-858	-824	-647	-438.2	9E+05	9E+05	832607	653801	442670
1817	-895.2	-858	-824	-647	-438	-423.2	8E+05	7E+05	579373	392277	378849
1995	-858.2	-824	-647	-438	-423	-245.2	7E+05	6E+05	376063	363190	210431
2212	-824.2	-647	-438	-423	-245	-28.2	5E+05	4E+05	348801	202094	23242
2607	-647.2	-438	-423	-245	-28	366.8	3E+05	3E+05	158693	18251	-2E+05
2698	-438.2	-423	-245	-28	367	457.8	2E+05	1E+05	12357	-2E+05	-2E+05
2703	-423.2	-245	-28	367	458	462.8	1E+05	11934	-2E+05	-2E+05	-2E+05
2738	-245.2	-28	367	458	463	497.8	6915	-89939	-1E+05	-1E+05	-1E+05
2810	-28.2	367	458	463	498	569.8	-10344	-12910	-13051	-14038	-16068
2632	366.8	458	463	498	570	391.8	2E+05	2E+05	182593	209003	143712
2548	457.8	463	498	570	392	307.8	2E+05	2E+05	260854	179366	140911
2881	462.8	498	570	392	308	640.8	2E+05	3E+05	181325	142450	296562
2725	497.8	570	392	308	641	484.8	3E+05	2E+05	153223	318990	241333
2911	569.8	392	308	641	485	670.8	2E+05	2E+05	365128	276239	382222
2759	391.8	308	641	485	671	518.8	1E+05	3E+05	189945	262819	203266

principal port in India for a period of 20 years

lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m	y4-m	y5-m	(yt-m)(y1-m)	(yt-m)(y2-m)
1802	-1010.2	-895.2	-858.2	-824.2	-647.2	-438.2	904331.04	866953.64
1817	-895.2	-858.2	-824.2	-647.2	-438.2	-423.2	768260.64	737823.84
1995	-858.2	-824.2	-647.2	-438.2	-423.2	-245.2	707328.44	555427.04
2212	-824.2	-647.2	-438.2	-423.2	-245.2	-28.2	533422.24	361164.44
2607	-647.2	-438.2	-423.2	-245.2	-28.2	366.8	283603.04	273895.04
2698	-438.2	-423.2	-245.2	-28.2	366.8	457.8	185446.24	107446.64
2703	-423.2	-245.2	-28.2	366.8	457.8	462.8	103768.64	11934.24
2738	-245.2	-28.2	366.8	457.8	462.8	497.8	6914.64	-89939.36
2810	-28.2	366.8	457.8	462.8	497.8	569.8	-10343.76	-12909.96
2632	366.8	457.8	462.8	497.8	569.8	391.8	167921.04	169755.04
2548	457.8	462.8	497.8	569.8	391.8	307.8	211869.84	227892.84
2881	462.8	497.8	569.8	391.8	307.8	640.8	230381.84	263703.44
2725	497.8	569.8	391.8	307.8	640.8	484.8	283646.44	195038.04
2911	569.8	391.8	307.8	640.8	484.8	670.8	223247.64	175384.44
2759	391.8	307.8	640.8	484.8	670.8	518.8	120596.04	251065.44

Time series data

Open recovered workbooks? Your recent changes were saved. Do you want to continue working where you left off?

I1 =SUM(N4:N23)

	L	M	N	O	P	Q	R
18	670.8	518.8	120596.04	251065.44	189944.64	262819.44	203265.84
19	518.8		197238.24	149221.44	206472.24	159686.64	
20			310659.84	429848.64	332447.04		
21			325203.84	251514.24			
22			348011.04				
23							
24							
25		Count	19	18	17	16	15
26		Sumproduct	590150[illegible].96	4925219.12	3899349.48	2696177.84	1691211
27		autocovariance	310605.6295	273623.2844	229373.4988	168511.115	112747.4
28		Autocorrelation	0.909211544	0.80095602	0.671427087	0.493269395	0.330036638

Sheet1 Sheet2

Find autocovariance, autocorrelation and auto regression

Year (t)	y_t	lag 1 ($y_1 = y_{t-1}$)	lag 2 ($y_2 = y_{t-2}$)	lag3 ($y_1 = y_{t-3}$)	lag 1 ($y_4 = y_{t-4}$)	lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m
1	1230	1345	1382	1416	1593	1802	-1010.2	-895.2	-858.2	
2	1345	1382	1416	1593	1802	1817	-895.2	-858.2	-824.2	
3	1382	1416	1593	1802	1817	1995	-858.2	-824.2	-647.2	
4	1416	1593	1802	1817	1995	2212	-824.2	-647.2	-438.2	
5	1593	1802	1817	1995	2212	2607	-647.2	-438.2	-423.2	
6	1802	1817	1995	2212	2607	2698	-438.2	-423.2	-245.2	
7	1817	1995	2212	2607	2698	2703	-423.2	-245.2	-28.2	
8	1995	2212	2607	2698	2703	2738	-245.2	-28.2	366.8	
9	2212	2607	2698	2703	2738	2810	-28.2	366.8	457.8	
10	2607	2698	2703	2738	2810	2632	366.8	457.8	462.8	
11	2698	2703	2738	2810	2632	2548	457.8	462.8	497.8	
12	2703	2738	2810	2632	2548	2881	462.8	497.8	569.8	
13	2738	2810	2632	2548	2881	2725	497.8	569.8	391.8	
14	2810	2632	2548	2881	2725	2911	569.8	391.8	307.8	
15	2632	2548	2881	2725	2911	2759	391.8	307.8	640.8	
16	2548	2881	2725	2911	2759		307.8	640.8	484.8	

Excel File Edit View Insert Format Tools Data Window Help

Time series data

Home Insert Draw Page Layout Formulas Data Review View

Open recovered workbooks? Your recent changes were saved. Do you want to continue working where you left off? Yes No

fx 1230

A	B	C	D	E	F	G	H	I	J
3	1382	1416	1593	1802	1817	1995	-858.2	-824.2	-647.2
4	1416	1593	1802	1817	1995	2212	-824.2	-647.2	-438.2
5	1593	1802	1817	1995	2212	2607	-647.2	-438.2	-423.2
6	1802	1817	1995	2212	2607	2698	-438.2	-423.2	-245.2
7	1817	1995	2212	2607	2698	2703	-423.2	-245.2	-28.2
8	1995	2212	2607	2698	2703	2738	-245.2	-28.2	366.8
9	2212	2607	2698	2703	2738	2810	-28.2	366.8	457.8
10	2607	2698	2703	2738	2810	2632	366.8	457.8	462.8
11	2698	2703	2738	2810	2632	2548	457.8	462.8	497.8
12	2703	2738	2810	2632	2548	2881	462.8	497.8	569.8
13	2738	2810	2632	2548	2881	2725	497.8	569.8	391.8
14	2810	2632	2548	2881	2725	2911	569.8	391.8	307.8
15	2632	2548	2881	2725	2911	2759	391.8	307.8	640.8
16	2548	2881	2725	2911	2759		307.8	640.8	484.8
17	2881	2725	2911	2759			640.8	484.8	670.8
18	2725	2911	2759				484.8	670.8	518.8
19	2911	2759					670.8	518.8	
20	2759								

Sheet1 Sheet2

The following data gives the cargo exports y_t (in units of 10^3 metric tons) at the principal port in India for a period of 20 years

Find autocovariance, autocorrelation and auto regression

Year (t)	y_t	lag 1 ($y_1 = y_{t-1}$)	lag 2 ($y_2 = y_{t-2}$)	lag3 ($y_1 = y_{t-3}$)	lag 1 ($y_4 = y_{t-4}$)	lag 5 ($y_1 = y_{t-5}$)	yt-m	y1-m	y2-m	y3-m
1	1230	1845	1382	1416	1593	1807	-1010.2	-895.2	-858.2	
2	1309	1987	1416	1593	1807	1817	-893.2	-858.2	-824.2	
3	1982	1416	1593	1807	1617	1993	-838.2	-824.2	-647.2	
4	1415	1593	1807	1617	1993	2212	-826.2	-647.2	-436.2	
5	1593	1807	1817	1903	2712	2602	-647.2	-436.2	-423.2	
6	1607	1617	1995	2212	2607	2698	-456.2	-423.2	-245.2	
7	1817	1903	2712	2607	2698	2703	-423.2	-245.2	-28.2	
8	1903	2212	2607	2698	2703	2738	-243.2	-28.2	566.8	
9	2212	2607	2698	2703	2738	2810	28.2	366.8	497.8	
10	2807	2698	2703	2738	2810	2632	366.8	457.8	467.8	
11	2698	2703	2738	2810	2632	2549	457.8	467.8	497.8	
12	2703	2738	2810	2632	2548	2881	467.8	497.8	569.8	
13	2738	2810	2632	2548	2881	2725	497.8	569.8	391.8	
14	2810	2611	2549	2881	2725	2911	588.8	391.8	307.8	
15	[illegible]	[illegible]	[illegible]	2775	2911	2759	391.8	307.8	640.8	