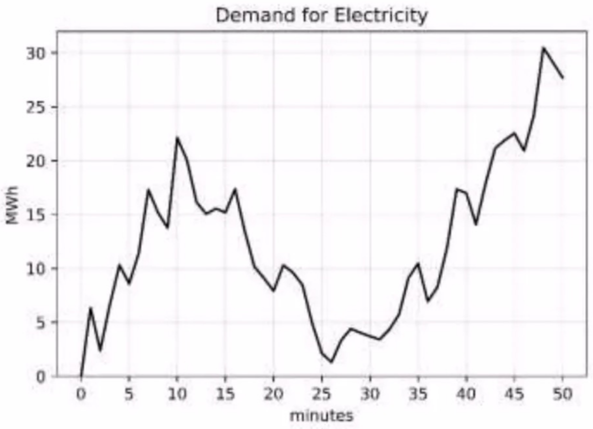
Time Series Forecasting

What Is a Time Series?

A sequence of data points organized in time order.

- The sequence captures data at equally spaced points in time.
- Data collected irregularly is not considered a time series.



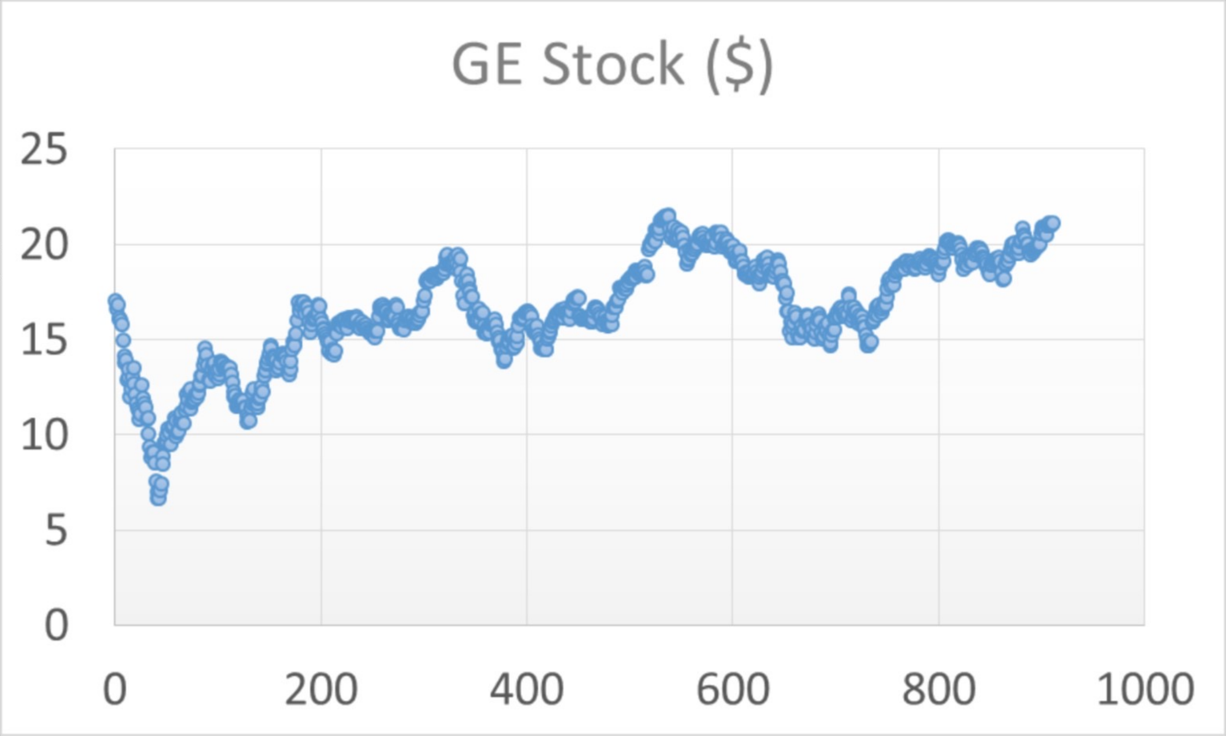
Motivations for Using Time Series

Time-series methods are used to do the following:

- Understand the generative process underlying the observed data
- Fit a model in order to monitor or forecast a process

Components of Time Series

- Trend
- Seasonality
- Random component

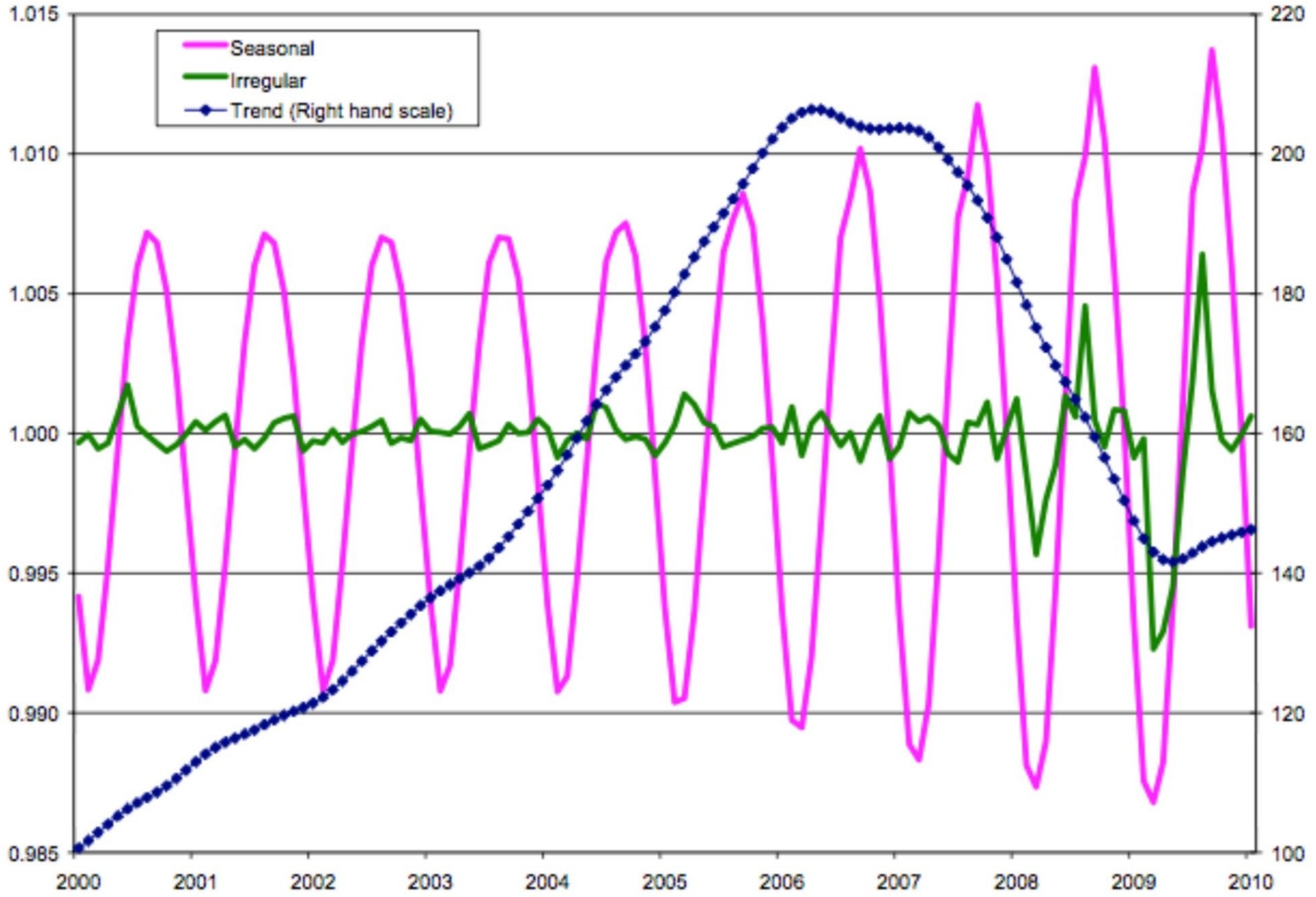


Time-Series Components

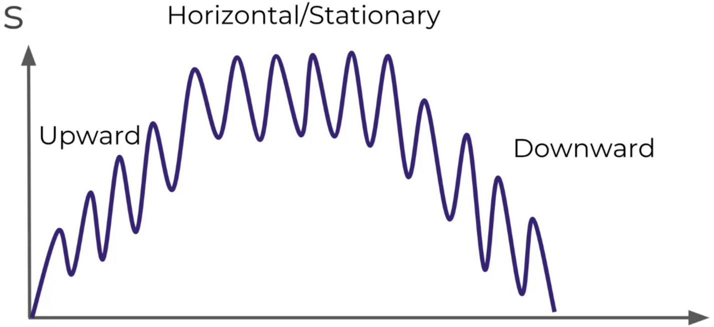
A time series has three components:

- **Trend** – long-term direction
- **Seasonality** – periodic behavior
- **Residual** – irregular fluctuations

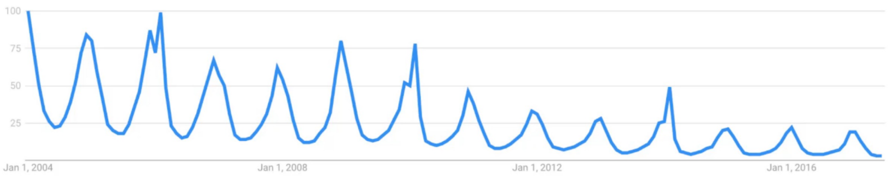
Trend, Seasonality and Randomness



Trends



Seasonality - Repeating trends

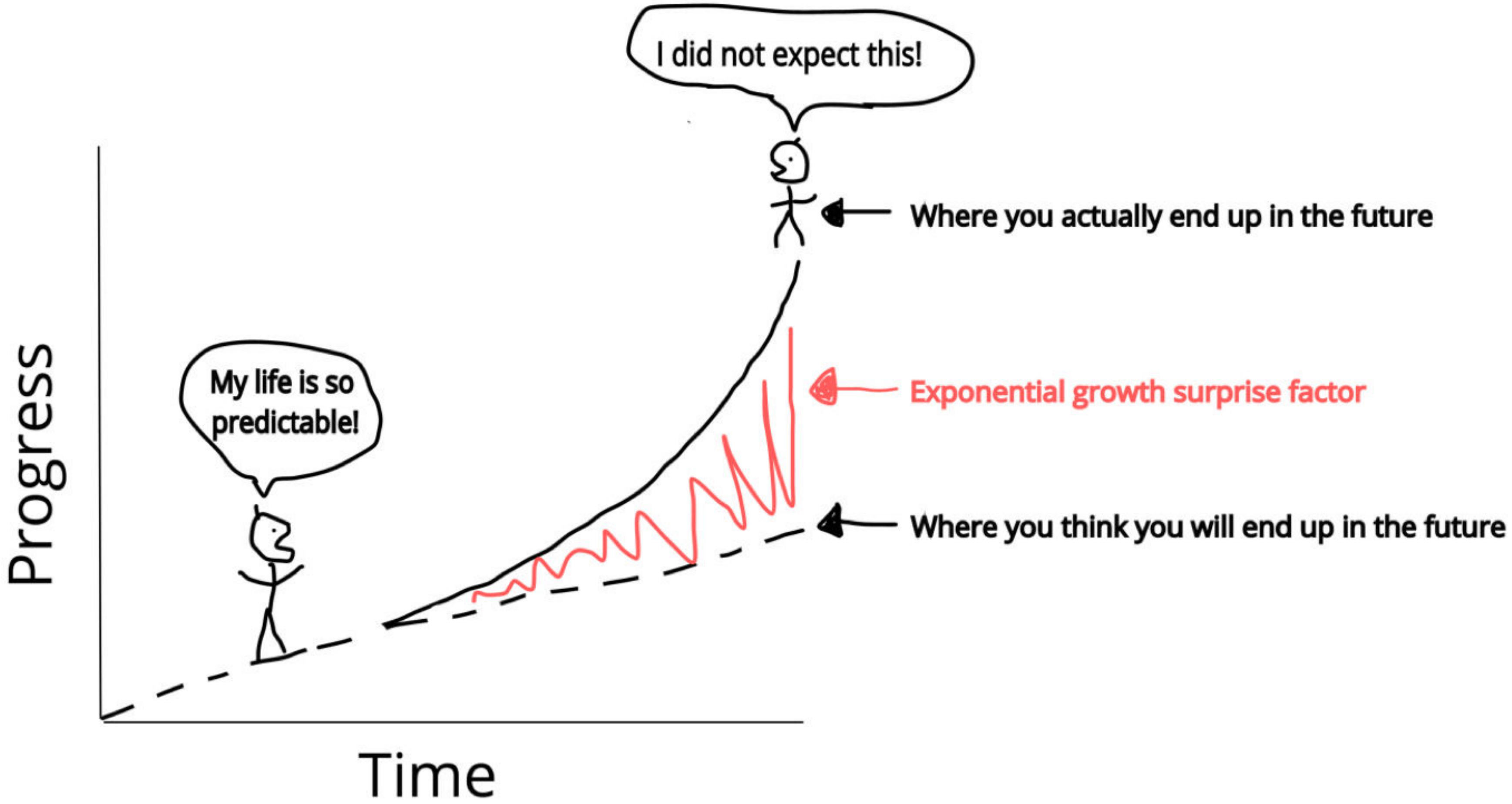


Google Trends - “Snowboarding”

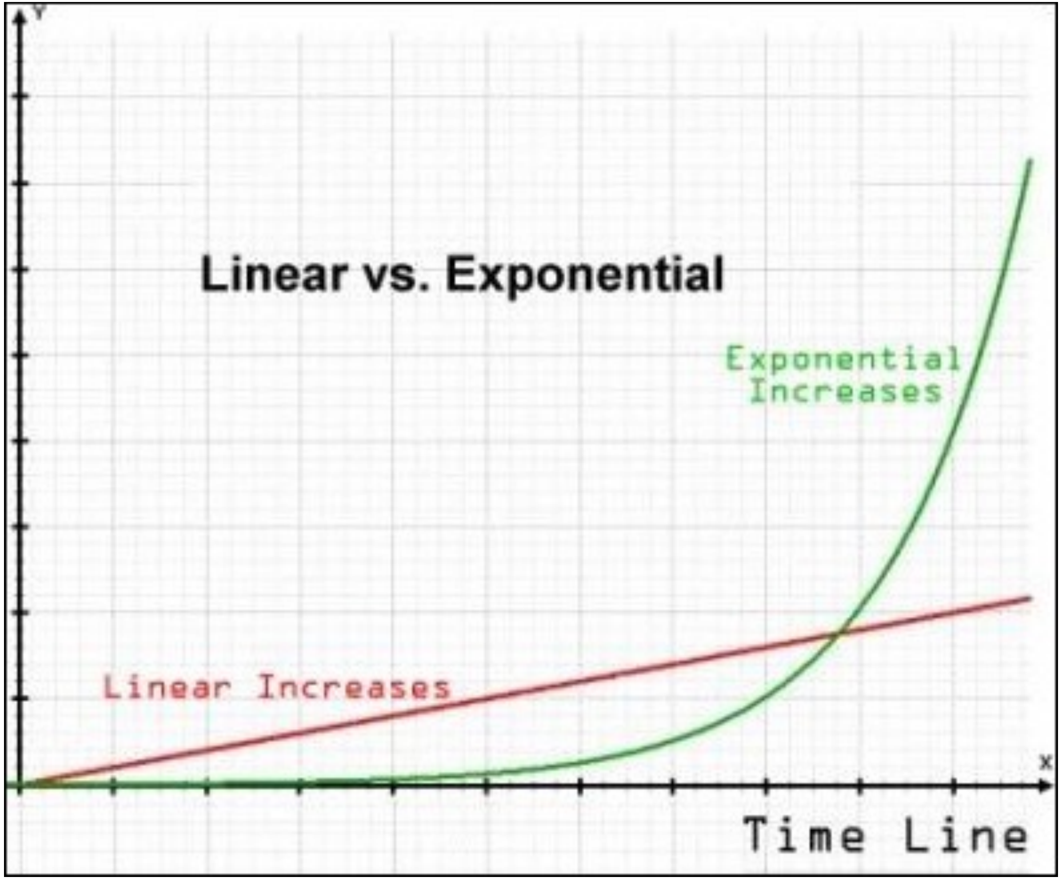
Cyclical - Trends with no set repetition.

SP500

ADDITIVE and MULTIPLICATIVE MODEL



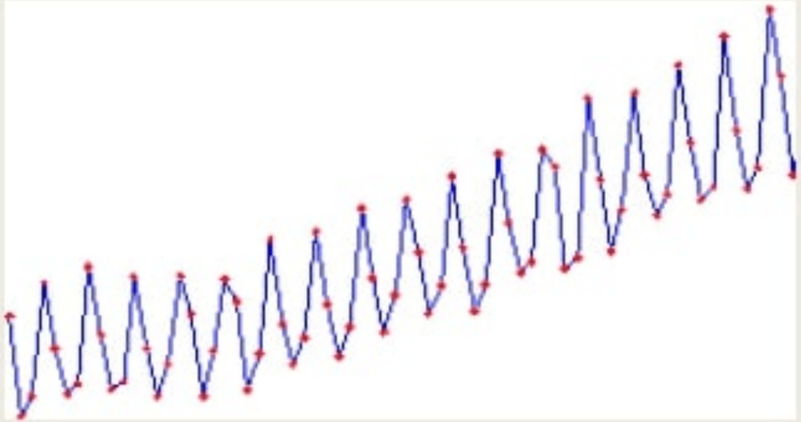
ADDITIVE and MULTIPLICATIVE MODEL



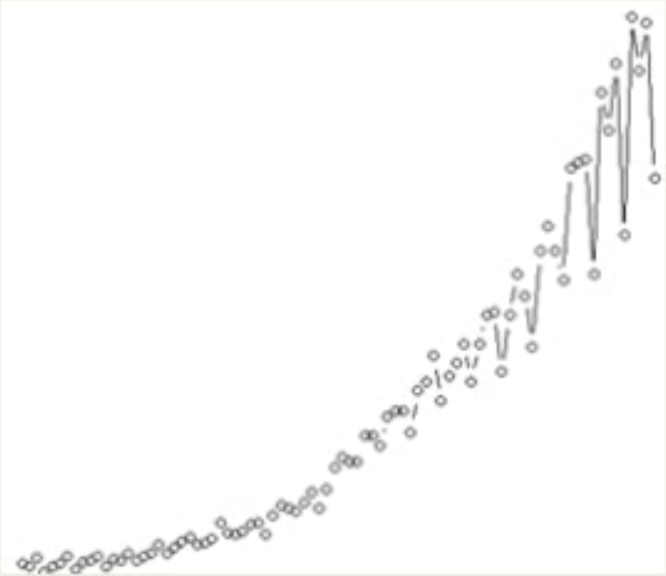
1. We use an Additive model for Error-Trend-Seasonality Decomposition when the trend is linear and Seasonality and trend seems to be constant over time

2. Multiplicative model when trend and seasonality increase in a non linear rate

Additive



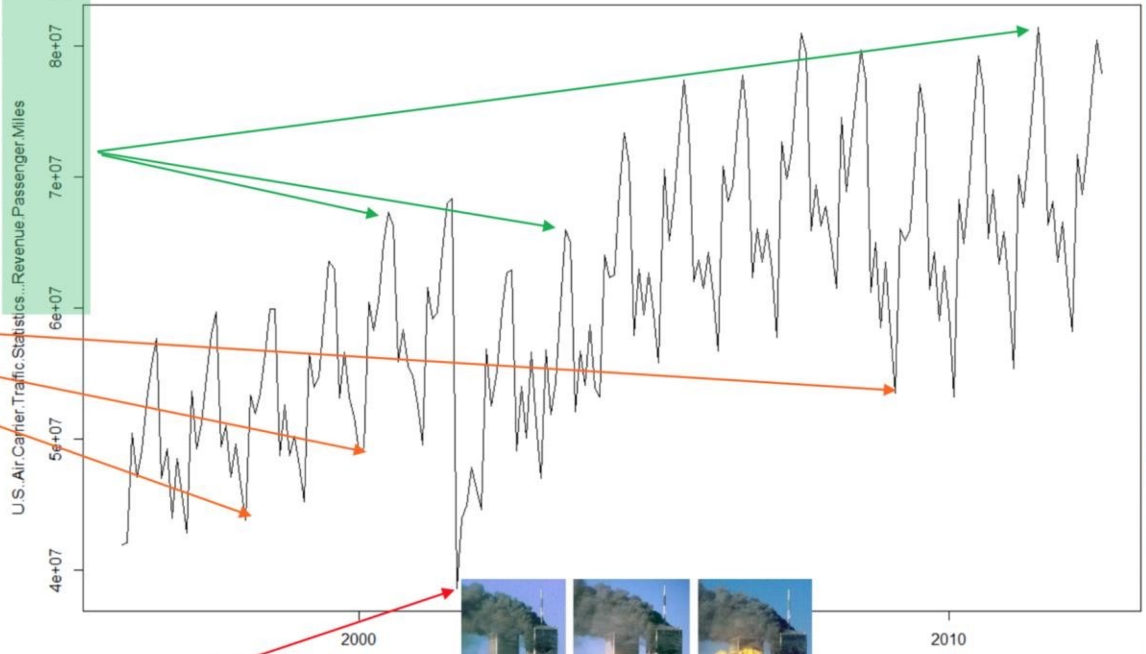
Multiplicative



US Air Carrier Traffic – Revenue Passenger Miles (‘000)

RPM

```
> milestimeseries <- ts(miles, frequency = 12, start = c(1996,1))
> milestimeseries
          Jan      Feb      Mar      Apr      May      Jun      Jul
1996 41972194 42054796 50443045 47112397 49118248 52880510
1997 45850623 42838949 53620994 49282817 51191842 54707221
1998 46514139 43769273 53361926 51968480 53515798 56460422
1999 47988560 45241211 56555731 53920855 54674958 59213000
2000 49045412 49306303 60443541 58286680 60533783 64903295
2001 52634354 49532578 61575055 59151645 59662416 64353323
2002 46224031 44615129 56897729 52542164 55116060 59745343
2003 51197175 47040806 56766580 51857453 54335598 60272900
2004 53979786 53179693 64035864 62340117 62530704 68866398
2005 59629608 55795165 70595861 65145552 68268899 72952959
2006 61035027 56729212 70799794 68120559 69352606 74099239
2007 63016013 57793832 72700241 69836156 71933109 76926452
2008 64667106 61504426 74575531 68906882 72725750 76162105
2009 58373786 53506580 66027341 65166300 65868254 71350227
2010 59651061 53240066 68307090 64953250 68850904 74474550
2011 61630362 55391206 70158268 67683558 71711448 76057910
2012 61940180 58243763 71696039 68669228 71887523 76760759
          Aug      Sep      Oct      Nov      Dec
1996 57723208 47035464 49263120 43937074 48539606
1997 59715433 49418190 51058879 47056048 49654209
```



Data sources:
http://www.bts.gov/xml/air_traffic/src/index.xml
and https://datamarket.com/data/set/281x/us-air-carrier-traffic-statistics-revenue-passenger-miles
Last accessed: 31-Mar-2016

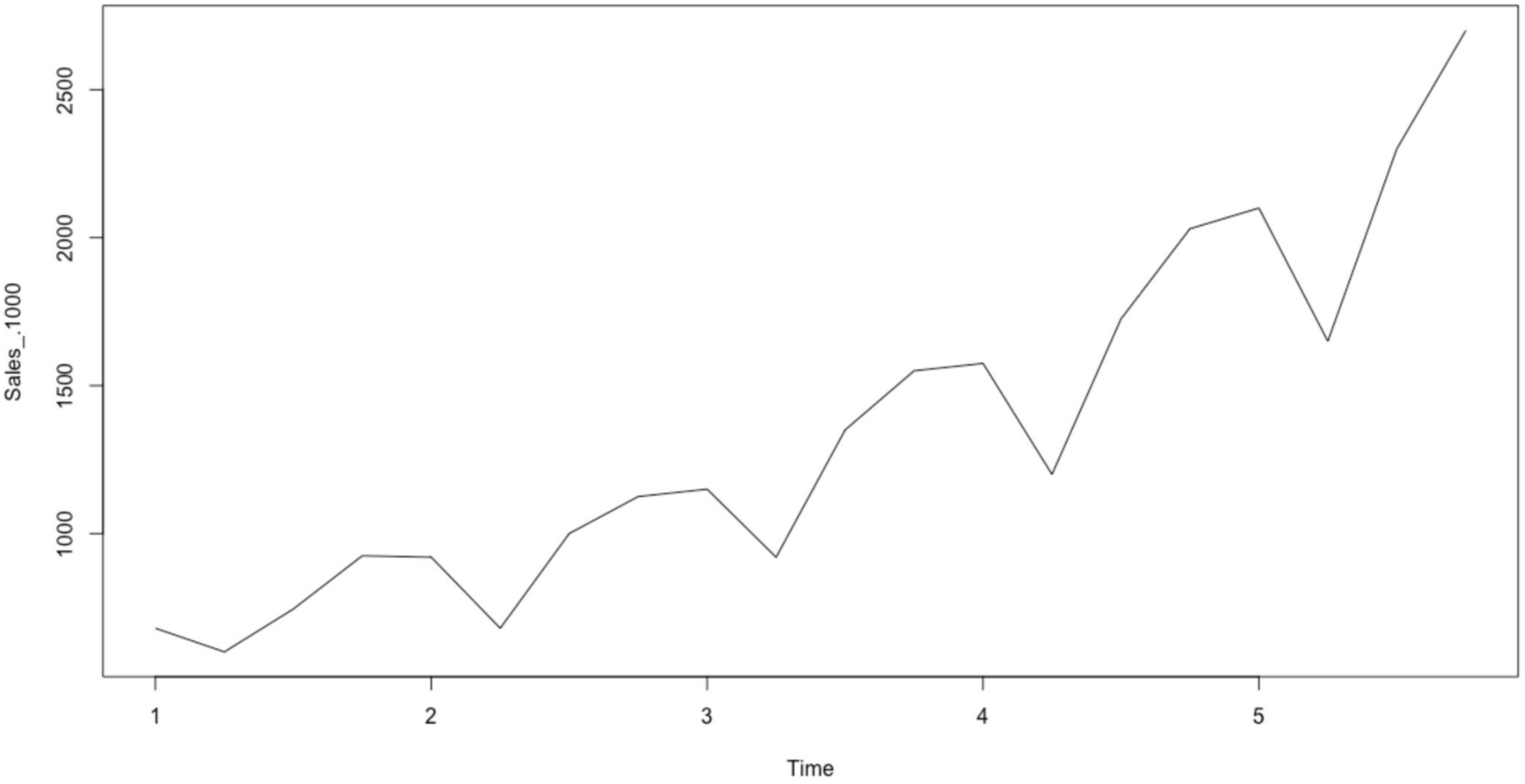
```
          Aug      Sep      Oct      Nov      Dec
1996 57723208 47035464 49263120 43937074 48539606
1997 59715433 49418190 51058879 47056048 49654209
1998 59927214 48751280 52578217 48734375 50208641
1999 63003663 53131972 56653901 53215500 51746821
2000 66256804 55900504 58373996 55590325 54822970
2001 68377080 38601868 43964788 44915764 47836501
2002 62944816 49096035 54019748 50106814 56656594
2003 64989766 52121480 56724551 54128776 58739845
2004 70961522 57881042 63021142 59453943 62680310
```

CURVE FITTING / REGRESSION ON TIME METHODS

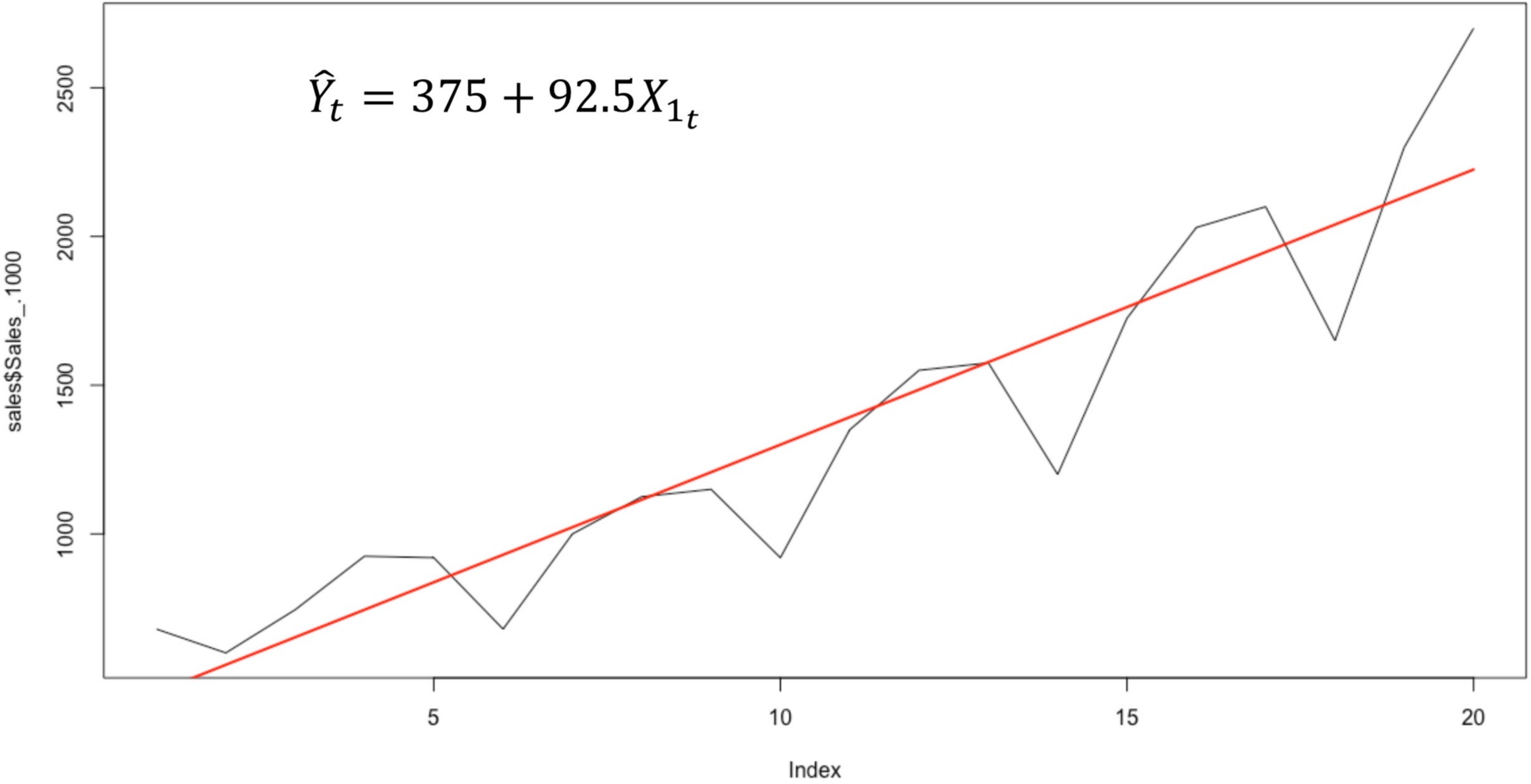
Regression on Time

- Use when trend is the most pronounced
- ACF decays exponentially and PACF has very few spikes

Quarter	Sales_$1000
1	680
2	600
3	745
4	925
5	920
6	680
7	1000
8	1125
9	1150
10	920
11	1350
12	1550
13	1575
14	1200
15	1725
16	2030
17	2100
18	1650



Regression Analysis



Incorporating Seasonality in Trend Analysis

- Take the trend prediction and actual prediction.
- Depending on additive or multiplicative model compute the deviation and map it as seasonality effect for each prediction.
- Take averages of the seasonality value. Use this to make future predictions.

Year	Quarter	Time variable (this is created)	Revenues (in $M)
2008	I	1	10.2
	II	2	12.4
	III	3	14.8
	IV	4	15
2009	I	5	11.2
	II	6	14.3
	III	7	18.4
	IV	8	18

```
Call:
lm(formula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-3.5595 -0.9384  0.4405  1.3265  1.9286

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  10.0393     1.5531   6.464  0.00065 ***
x             0.9440     0.3076   3.069  0.02196 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.993 on 6 degrees of freedom
Multiple R-squared: 0.6109,     Adjusted R-squared: 0.5461
F-statistic: 9.422 on 1 and 6 DF,  p-value: 0.02196
```

What is the Regression equation?

$$y = 10.0393 + 0.9440x$$

Seasonality: Multiplicative

Time	Observed values TSI* (assuming no impact of cyclicality)	Predicted values (per the regression) T*	SI* = TSI/T
1	10.2	10.983	0.929
2	12.4	11.927	1.040
3	14.8	12.871	1.150
4	15.0	13.815	1.086
5	11.2	14.759	0.759
6	14.3	15.703	0.911
7	18.4	16.647	1.105
8	18.0	17.591	1.023

* T: Trend; S: Seasonal; I: Irregular

Quarterly Seasonality

Time	Average seasonality factor
Q1	$0.844 \left(= \frac{0.929+0.759}{2}\right)$
Q2	0.975
Q3	1.127
Q4	1.054

Time	Observed values TSI* (assuming no impact of cyclicality)	Predicted values (per the regression) T*	SI* = TSI/T
1	10.2	10.983	0.929
2	12.4	11.927	1.040
3	14.8	12.871	1.150
4	15.0	13.815	1.086
5	11.2	14.759	0.759
6	14.3	15.703	0.911
7	18.4	16.647	1.105
8	18.0	17.591	1.023

Computations

- Trend $Y_9 = 10.039 + 0.944(9) = 18.535$

- Corrected for seasonality and randomness: $18.535 * 0.844 = 15.643$

Issues with Regressing on Time

- If there is no trend or if seasonality and fluctuations are more important than trend, then the coefficients behave weirdly

Goodness of Fit

- MAE (Mean absolute error) $\frac{\sum|y_i - \hat{y}_i|}{n}$
- MSE (Mean square error) $\frac{\sum(y_i - \hat{y}_i)^2}{n}$
- RMSE (Root mean square error) $\sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n}}$
- MAPE (Mean absolute percent error) $\frac{1}{n}\left(\frac{\sum|y_i - \hat{y}_i|}{y_i}\right) * 100$

SIMPLE FORECASTING / BENCHMARKING METHODS

Average Method

$$\hat{y}_{t+h} = \bar{y} = \frac{y_1 + \cdots + y_t}{t}$$

- Forecasts of all future values are equal to the **mean of the historical data**

Naïve Method

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

- Forecasts of all future values are equal to the last **observed** value
- A very useful method in many economic and financial time series

Seasonal Naïve Method

$$\hat{y}_{t+h} = y_{t+h-km}$$

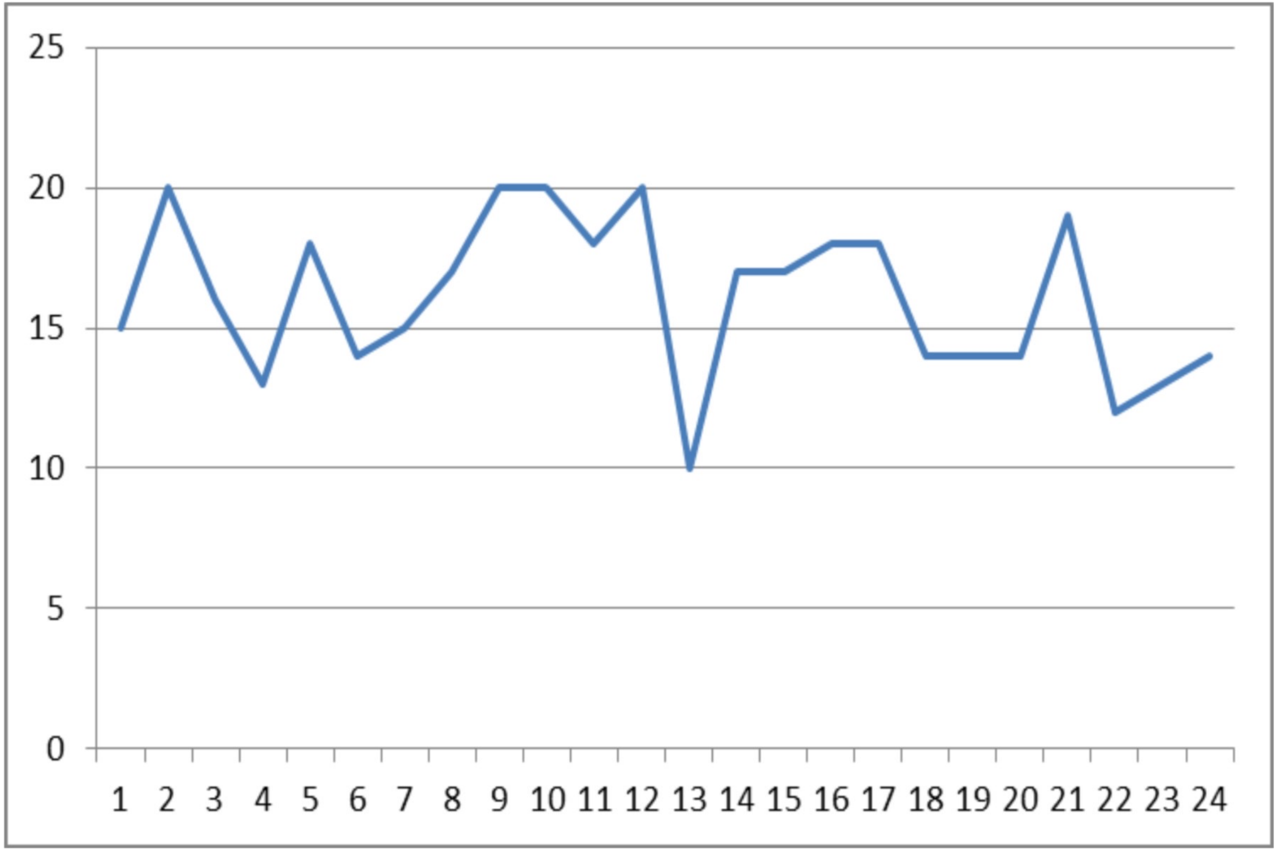
Integer part

where m is the seasonal period and $k = \left\lfloor \frac{h-1}{m} \right\rfloor + 1$

- Forecasts of all future values for a particular period (e.g., July) are equal to the last **observed** value for the same period (e.g., last July)

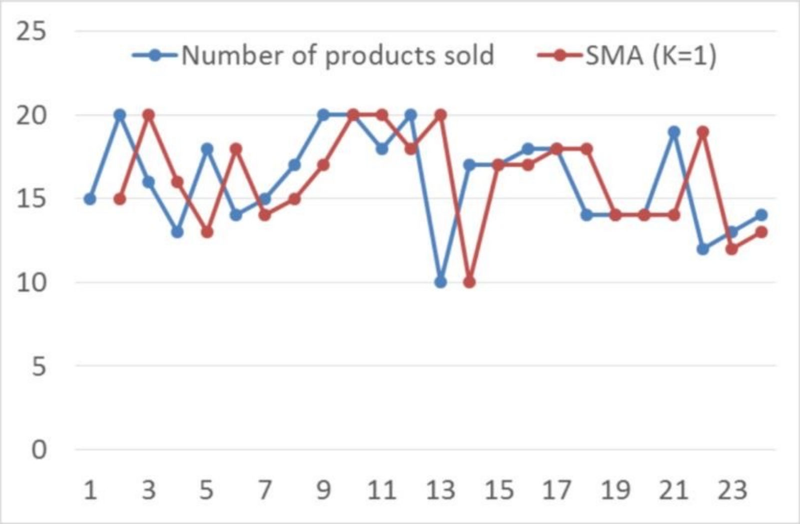
SMOOTHING METHODS

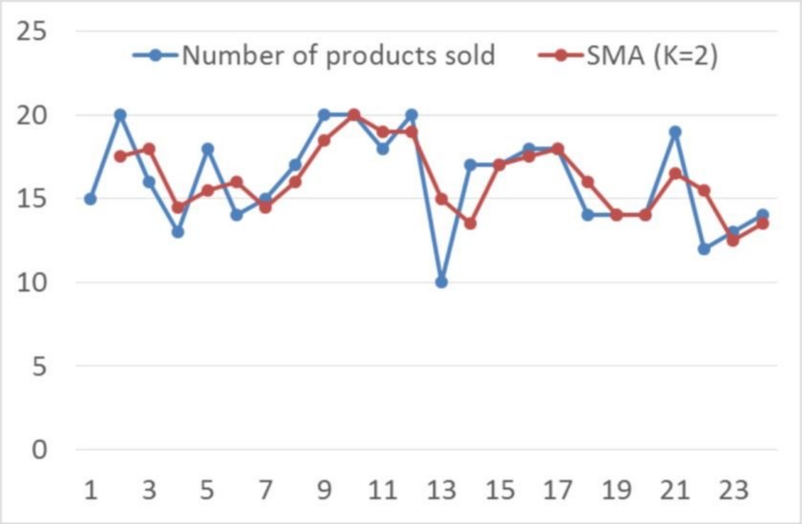
Stationary Model: Moving Averages

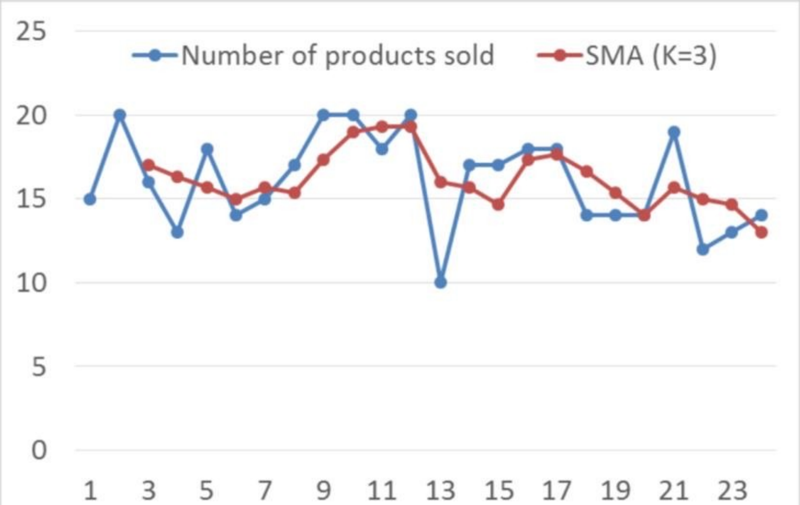


Stationary Model: Case 1 – Simple Moving Averages

1	Number of products sold	SMA (K=1)	Error	SMA (K=2)	Error	SMA (K=3)	Error
2	15						
3	20	15	5	17.5	2.5		
4	16	20	4	18	2	17	1
5	13	16	3	14.5	1.5	16.333333	3.33333
6	18	13	5	15.5	2.5	15.666667	2.33333
7	14	18	4	16	2	15	1
8	15	14	1	14.5	0.5	15.666667	0.66667
9	17	15	2	16	1	15.333333	1.66667
10	20	17	3	18.5	1.5	17.333333	2.66667
11	20	20	0	20	0	19	1
12	18	20	2	19	1	19.333333	1.33333
13	20	18	2	19	1	19.333333	0.66667
14	10	20	10	15	5	16	6
15	17	10	7	13.5	3.5	15.666667	1.33333
16	17	17	0	17	0	14.666667	2.33333
17	18	17	1	17.5	0.5	17.333333	0.66667
18	18	18	0	18	0	17.666667	0.33333
19	14	18	4	16	2	16.666667	2.66667
20	14	14	0	14	0	15.333333	1.33333
21	14	14	0	14	0	14	0
22	19	14	5	16.5	2.5	15.666667	3.33333
23	12	19	7	15.5	3.5	15	3
24	13	12	1	12.5	0.5	14.666667	1.66667
25	14	13	1	13.5	0.5	13	1
26			**2.91304**		**1.45652**		**1.78788**







Only decision point is K

Stationary Model: Case 2 – Weighted Moving Averages

$$\hat{Y}_{t+1} = w_1 Y_t + w_2 Y_{t-1} + \cdots + w_k Y_{t-k+1}$$

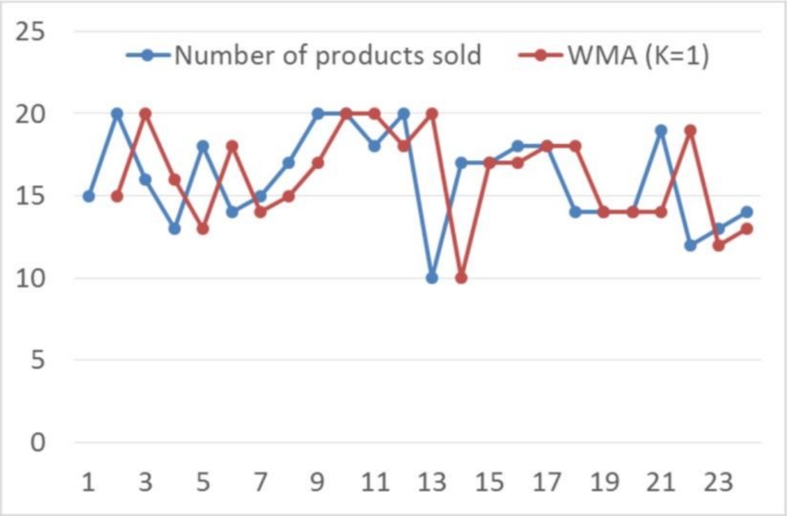
- Typically we choose a time period of moving average and weights are chosen such that the error is minimized

Stationary Model: Case 2 – Weighted Moving Averages

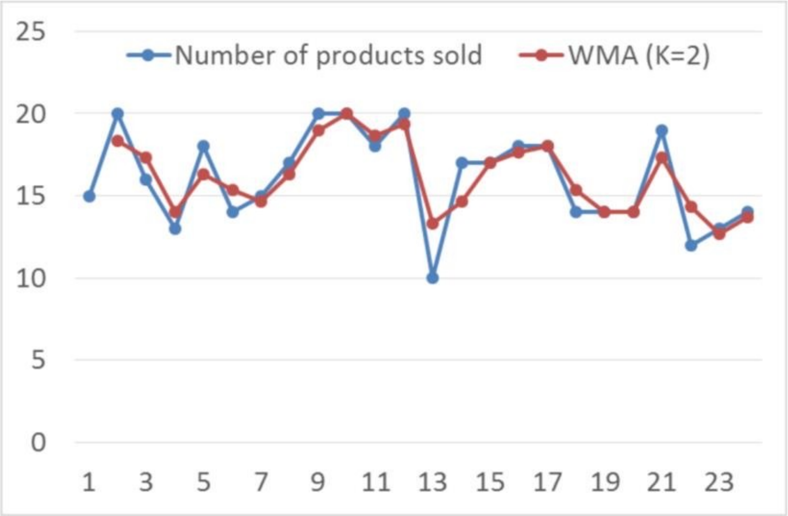
	A	B	C	D	E	F	G
1	Number of products sold	WMA (K=1)	Error	WMA (K=2)	Error	WMA (K=3)	Error
2	15						
3	20	=A2*1	=ABS(B3-A3)	=(A3*2+A2*1)/3	=ABS(A3-D3)		
4	16	=A3*1	=ABS(B4-A4)	=(A4*2+A3*1)/3	=ABS(A4-D4)	=(A4*3+A3*2+A2*1)/6	=ABS(A4-F4)
5	13	=A4*1	=ABS(B5-A5)	=(A5*2+A4*1)/3	=ABS(A5-D5)	=(A5*3+A4*2+A3*1)/6	=ABS(A5-F5)
6	18	=A5*1	=ABS(B6-A6)	=(A6*2+A5*1)/3	=ABS(A6-D6)	=(A6*3+A5*2+A4*1)/6	=ABS(A6-F6)
7	14	=A6*1	=ABS(B7-A7)	=(A7*2+A6*1)/3	=ABS(A7-D7)	=(A7*3+A6*2+A5*1)/6	=ABS(A7-F7)
8	15	=A7*1	=ABS(B8-A8)	=(A8*2+A7*1)/3	=ABS(A8-D8)	=(A8*3+A7*2+A6*1)/6	=ABS(A8-F8)
9	17	=A8*1	=ABS(B9-A9)	=(A9*2+A8*1)/3	=ABS(A9-D9)	=(A9*3+A8*2+A7*1)/6	=ABS(A9-F9)
10	20	=A9*1	=ABS(B10-A10)	=(A10*2+A9*1)/3	=ABS(A10-D10)	=(A10*3+A9*2+A8*1)/6	=ABS(A10-F10)
11	20	=A10*1	=ABS(B11-A11)	=(A11*2+A10*1)/3	=ABS(A11-D11)	=(A11*3+A10*2+A9*1)/6	=ABS(A11-F11)
12	18	=A11*1	=ABS(B12-A12)	=(A12*2+A11*1)/3	=ABS(A12-D12)	=(A12*3+A11*2+A10*1)/6	=ABS(A12-F12)
13	20	=A12*1	=ABS(B13-A13)	=(A13*2+A12*1)/3	=ABS(A13-D13)	=(A13*3+A12*2+A11*1)/6	=ABS(A13-F13)
14	10	=A13*1	=ABS(B14-A14)	=(A14*2+A13*1)/3	=ABS(A14-D14)	=(A14*3+A13*2+A12*1)/6	=ABS(A14-F14)
15	17	=A14*1	=ABS(B15-A15)	=(A15*2+A14*1)/3	=ABS(A15-D15)	=(A15*3+A14*2+A13*1)/6	=ABS(A15-F15)
16	17	=A15*1	=ABS(B16-A16)	=(A16*2+A15*1)/3	=ABS(A16-D16)	=(A16*3+A15*2+A14*1)/6	=ABS(A16-F16)
17	18	=A16*1	=ABS(B17-A17)	=(A17*2+A16*1)/3	=ABS(A17-D17)	=(A17*3+A16*2+A15*1)/6	=ABS(A17-F17)
18	18	=A17*1	=ABS(B18-A18)	=(A18*2+A17*1)/3	=ABS(A18-D18)	=(A18*3+A17*2+A16*1)/6	=ABS(A18-F18)
19	14	=A18*1	=ABS(B19-A19)	=(A19*2+A18*1)/3	=ABS(A19-D19)	=(A19*3+A18*2+A17*1)/6	=ABS(A19-F19)
20	14	=A19*1	=ABS(B20-A20)	=(A20*2+A19*1)/3	=ABS(A20-D20)	=(A20*3+A19*2+A18*1)/6	=ABS(A20-F20)
21	14	=A20*1	=ABS(B21-A21)	=(A21*2+A20*1)/3	=ABS(A21-D21)	=(A21*3+A20*2+A19*1)/6	=ABS(A21-F21)
22	19	=A21*1	=ABS(B22-A22)	=(A22*2+A21*1)/3	=ABS(A22-D22)	=(A22*3+A21*2+A20*1)/6	=ABS(A22-F22)
23	12	=A22*1	=ABS(B23-A23)	=(A23*2+A22*1)/3	=ABS(A23-D23)	=(A23*3+A22*2+A21*1)/6	=ABS(A23-F23)
24	13	=A23*1	=ABS(B24-A24)	=(A24*2+A23*1)/3	=ABS(A24-D24)	=(A24*3+A23*2+A22*1)/6	=ABS(A24-F24)
25	14	=A24*1	=ABS(B25-A25)	=(A25*2+A24*1)/3	=ABS(A25-D25)	=(A25*3+A24*2+A23*1)/6	=ABS(A25-F25)
26			**=AVERAGE(C3:C25)**		**=AVERAGE(E3:E25)**		**=AVERAGE(G3:G25)**

Stationary Model: Case 2 – Weighted Moving Averages

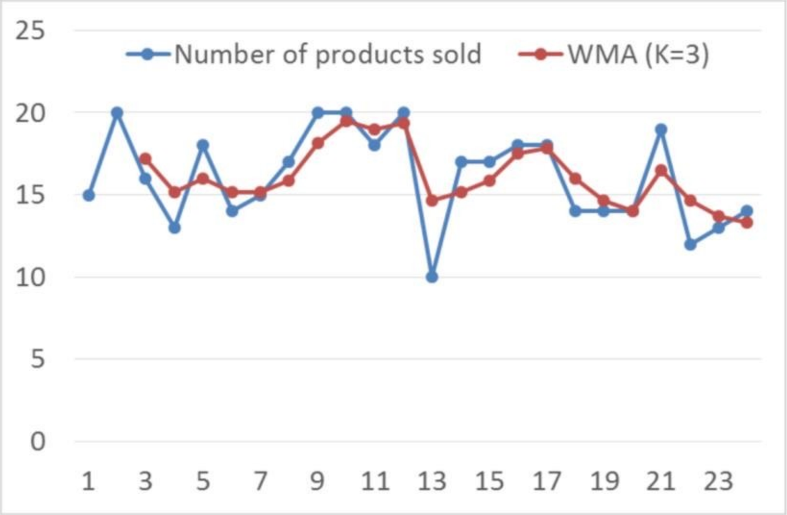
1	Number of products sold	WMA (K=1)	Error	WMA (K=2)	Error	WMA (K=3)	Error
2	15						
3	20	15	5	18.3333333	1.66666667		
4	16	20	4	17.3333333	1.33333333	17.1666667	1.16666667
5	13	16	3	14	1	15.1666667	2.16666667
6	18	13	5	16.3333333	1.66666667	16	2
7	14	18	4	15.3333333	1.33333333	15.1666667	1.16666667
8	15	14	1	14.6666667	0.33333333	15.1666667	0.16666667
9	17	15	2	16.3333333	0.66666667	15.8333333	1.16666667
10	20	17	3	19	1	18.1666667	1.83333333
11	20	20	0	20	0	19.5	0.5
12	18	20	2	18.6666667	0.66666667	19	1
13	20	18	2	19.3333333	0.66666667	19.3333333	0.66666667
14	10	20	10	13.3333333	3.33333333	14.6666667	4.66666667
15	17	10	7	14.6666667	2.33333333	15.1666667	1.83333333
16	17	17	0	17	0	15.8333333	1.16666667
17	18	17	1	17.6666667	0.33333333	17.5	0.5
18	18	18	0	18	0	17.8333333	0.16666667
19	14	18	4	15.3333333	1.33333333	16	2
20	14	14	0	14	0	14.6666667	0.66666667
21	14	14	0	14	0	14	0
22	19	14	5	17.3333333	1.66666667	16.5	2.5
23	12	19	7	14.3333333	2.33333333	14.6666667	2.66666667
24	13	12	1	12.6666667	0.33333333	13.6666667	0.66666667
25	14	13	1	13.6666667	0.33333333	13.3333333	0.66666667
26			**2.91304348**		**0.97101449**		**1.33333333**



Number of products sold
WMA (K=1)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23

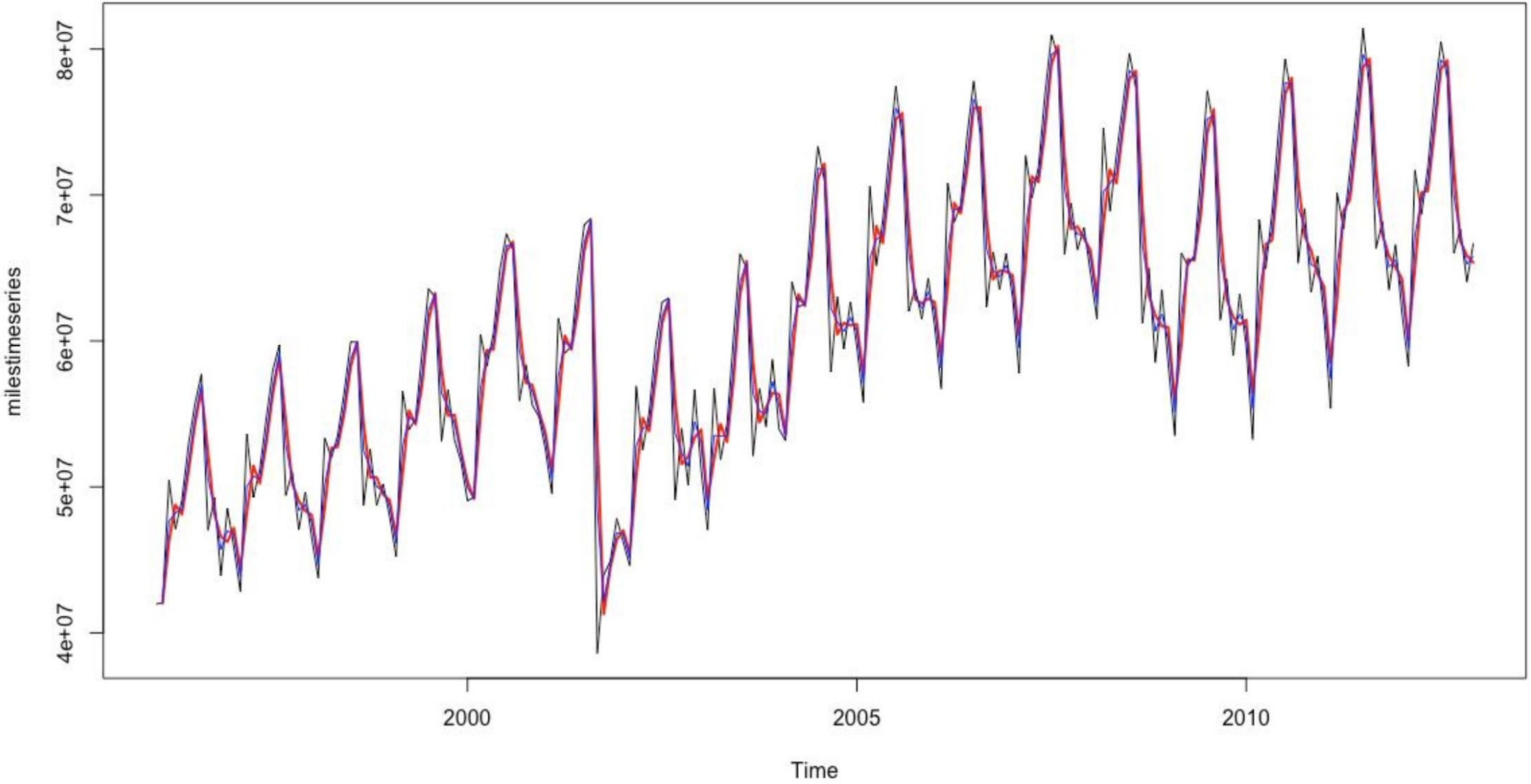


Number of products sold
WMA (K=2)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23



Number of products sold
WMA (K=3)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23

SMA and WMA – Revenue Passenger Miles

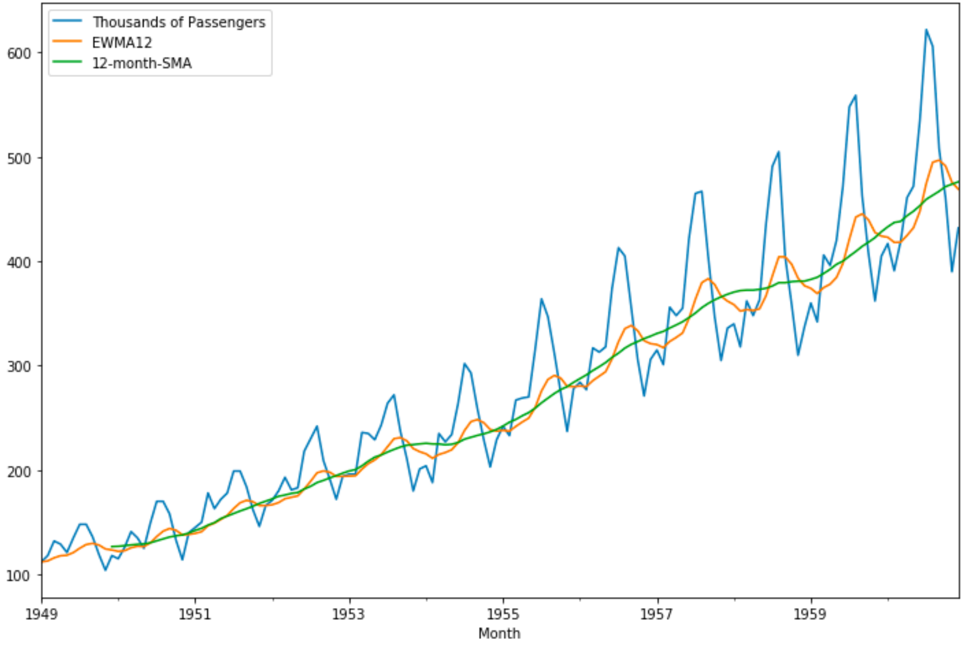



```
> MAPE-SMA 4.093731 > MAPE-WMA 2.729154
```

Exponential Weighted Moving Average

- We could theoretically attempt to use these simple moving averages to build a generalized model for the real world time series we're analyzing.
- Later on, we'll discover much more sophisticated models for this.

- EWMA- Exponentially Weighted Moving Averages
- Basic SMA has some "weaknesses".
    - It will never reach to full peak or valley of the data due to the averaging.



- EWMA will allow us to reduce the lag effect from SMA and it will put more weight on values that occurred more recently (by applying more weight to the more recent values, thus the name).

Exponential Smoothing

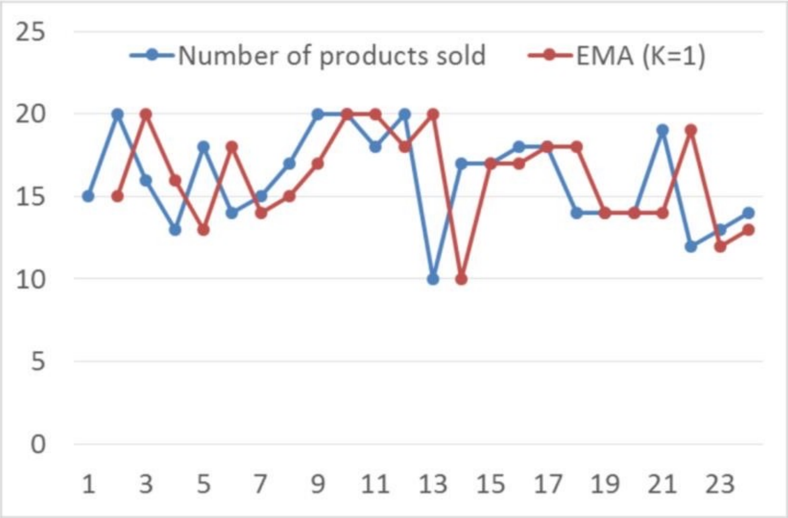
	A	B	C	D	E	F	G	H	I
1	Number of products sold	EMA (K=1)	Error	EMA (K=2)	Error	EMA (K=3)	Error	EMA (K=4)	Error
2	15								
3	20	=A2*1	=ABS(B3-A3)	=AVERAGE(A2:A3)					
4	16	=A3*K2+B3*L2	=ABS(B4-A4)	=A4*K3+D3*L3	=ABS(A4-D4)	=AVERAGE(A2:A4)	=ABS(A4-F4)		
5	13	=A4*K2+B4*L2	=ABS(B5-A5)	=A5*K3+D4*L3	=ABS(A5-D5)	=A5*K4+F4*L4	=ABS(A5-F5)	=AVERAGE(A2:A5)	=ABS(H5-A5)
6	18	=A5*K2+B5*L2	=ABS(B6-A6)	=A6*K3+D5*L3	=ABS(A6-D6)	=A6*K4+F5*L4	=ABS(A6-F6)	=A6*K5+H5*L5	=ABS(H6-A6)
7	14	=A6*K2+B6*L2	=ABS(B7-A7)	=A7*K3+D6*L3	=ABS(A7-D7)	=A7*K4+F6*L4	=ABS(A7-F7)	=A7*K5+H6*L5	=ABS(H7-A7)
8	15	=A7*K2+B7*L2	=ABS(B8-A8)	=A8*K3+D7*L3	=ABS(A8-D8)	=A8*K4+F7*L4	=ABS(A8-F8)	=A8*K5+H7*L5	=ABS(H8-A8)
9	17	=A8*K2+B8*L2	=ABS(B9-A9)	=A9*K3+D8*L3	=ABS(A9-D9)	=A9*K4+F8*L4	=ABS(A9-F9)	=A9*K5+H8*L5	=ABS(H9-A9)
10	20	=A9*K2+B9*L2	=ABS(B10-A10)	=A10*K3+D9*L3	=ABS(A10-D10)	=A10*K4+F9*L4	=ABS(A10-F10)	=A10*K5+H9*L5	=ABS(H10-A10)
11	20	=A10*K2+B10*L2	=ABS(B11-A11)	=A11*K3+D10*L3	=ABS(A11-D11)	=A11*K4+F10*L4	=ABS(A11-F11)	=A11*K5+H10*L5	=ABS(H11-A11)
12	18	=A11*K2+B11*L2	=ABS(B12-A12)	=A12*K3+D11*L3	=ABS(A12-D12)	=A12*K4+F11*L4	=ABS(A12-F12)	=A12*K5+H11*L5	=ABS(H12-A12)
13	20	=A12*K2+B12*L2	=ABS(B13-A13)	=A13*K3+D12*L3	=ABS(A13-D13)	=A13*K4+F12*L4	=ABS(A13-F13)	=A13*K5+H12*L5	=ABS(H13-A13)
14	10	=A13*K2+B13*L2	=ABS(B14-A14)	=A14*K3+D13*L3	=ABS(A14-D14)	=A14*K4+F13*L4	=ABS(A14-F14)	=A14*K5+H13*L5	=ABS(H14-A14)
15	17	=A14*K2+B14*L2	=ABS(B15-A15)	=A15*K3+D14*L3	=ABS(A15-D15)	=A15*K4+F14*L4	=ABS(A15-F15)	=A15*K5+H14*L5	=ABS(H15-A15)
16	17	=A15*K2+B15*L2	=ABS(B16-A16)	=A16*K3+D15*L3	=ABS(A16-D16)	=A16*K4+F15*L4	=ABS(A16-F16)	=A16*K5+H15*L5	=ABS(H16-A16)
17	18	=A16*K2+B16*L2	=ABS(B17-A17)	=A17*K3+D16*L3	=ABS(A17-D17)	=A17*K4+F16*L4	=ABS(A17-F17)	=A17*K5+H16*L5	=ABS(H17-A17)
18	18	=A17*K2+B17*L2	=ABS(B18-A18)	=A18*K3+D17*L3	=ABS(A18-D18)	=A18*K4+F17*L4	=ABS(A18-F18)	=A18*K5+H17*L5	=ABS(H18-A18)
19	14	=A18*K2+B18*L2	=ABS(B19-A19)	=A19*K3+D18*L3	=ABS(A19-D19)	=A19*K4+F18*L4	=ABS(A19-F19)	=A19*K5+H18*L5	=ABS(H19-A19)
20	14	=A19*K2+B19*L2	=ABS(B20-A20)	=A20*K3+D19*L3	=ABS(A20-D20)	=A20*K4+F19*L4	=ABS(A20-F20)	=A20*K5+H19*L5	=ABS(H20-A20)
21	14	=A20*K2+B20*L2	=ABS(B21-A21)	=A21*K3+D20*L3	=ABS(A21-D21)	=A21*K4+F20*L4	=ABS(A21-F21)	=A21*K5+H20*L5	=ABS(H21-A21)
22	19	=A21*K2+B21*L2	=ABS(B22-A22)	=A22*K3+D21*L3	=ABS(A22-D22)	=A22*K4+F21*L4	=ABS(A22-F22)	=A22*K5+H21*L5	=ABS(H22-A22)
23	12	=A22*K2+B22*L2	=ABS(B23-A23)	=A23*K3+D22*L3	=ABS(A23-D23)	=A23*K4+F22*L4	=ABS(A23-F23)	=A23*K5+H22*L5	=ABS(H23-A23)
24	13	=A23*K2+B23*L2	=ABS(B24-A24)	=A24*K3+D23*L3	=ABS(A24-D24)	=A24*K4+F23*L4	=ABS(A24-F24)	=A24*K5+H23*L5	=ABS(H24-A24)
25	14	=A24*K2+B24*L2	=ABS(B25-A25)	=A25*K3+D24*L3	=ABS(A25-D25)	=A25*K4+F24*L4	=ABS(A25-F25)	=A25*K5+H24*L5	=ABS(H25-A25)
26			**=AVERAGE(C3:C25)**		**=AVERAGE(E3:E25)**		**=AVERAGE(G3:G25)**		**=AVERAGE(I3:I25)**

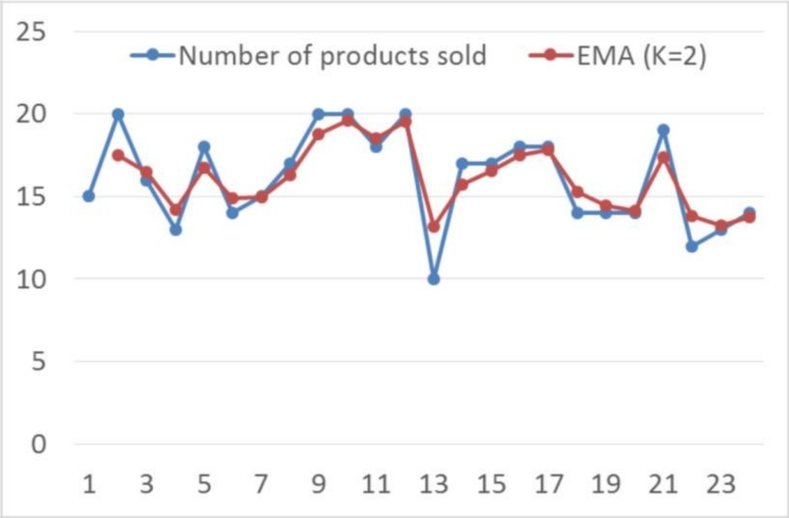
J	K	L
K	2/(K+1)	1-[2/(K+1)]
1	1	=1-K2
2	=2/(J3+1)	=1-K3
3	=2/(J4+1)	=1-K4
4	=2/(J5+1)	=1-K5

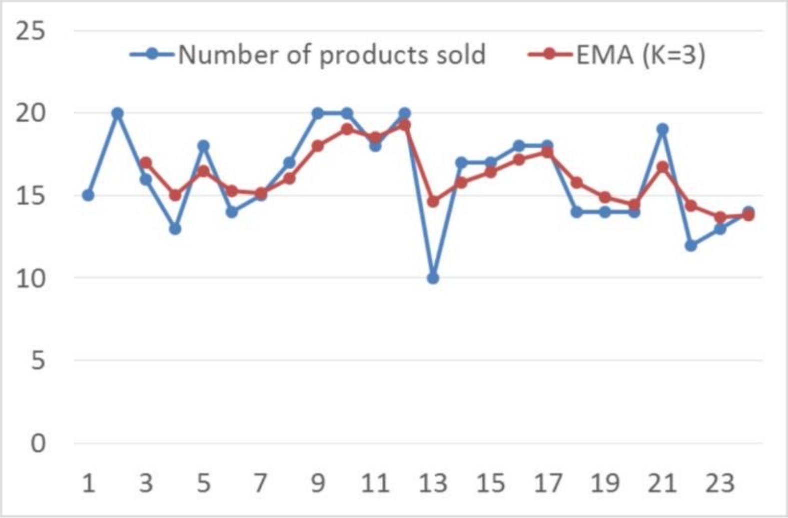
$$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$$

Exponential Smoothing

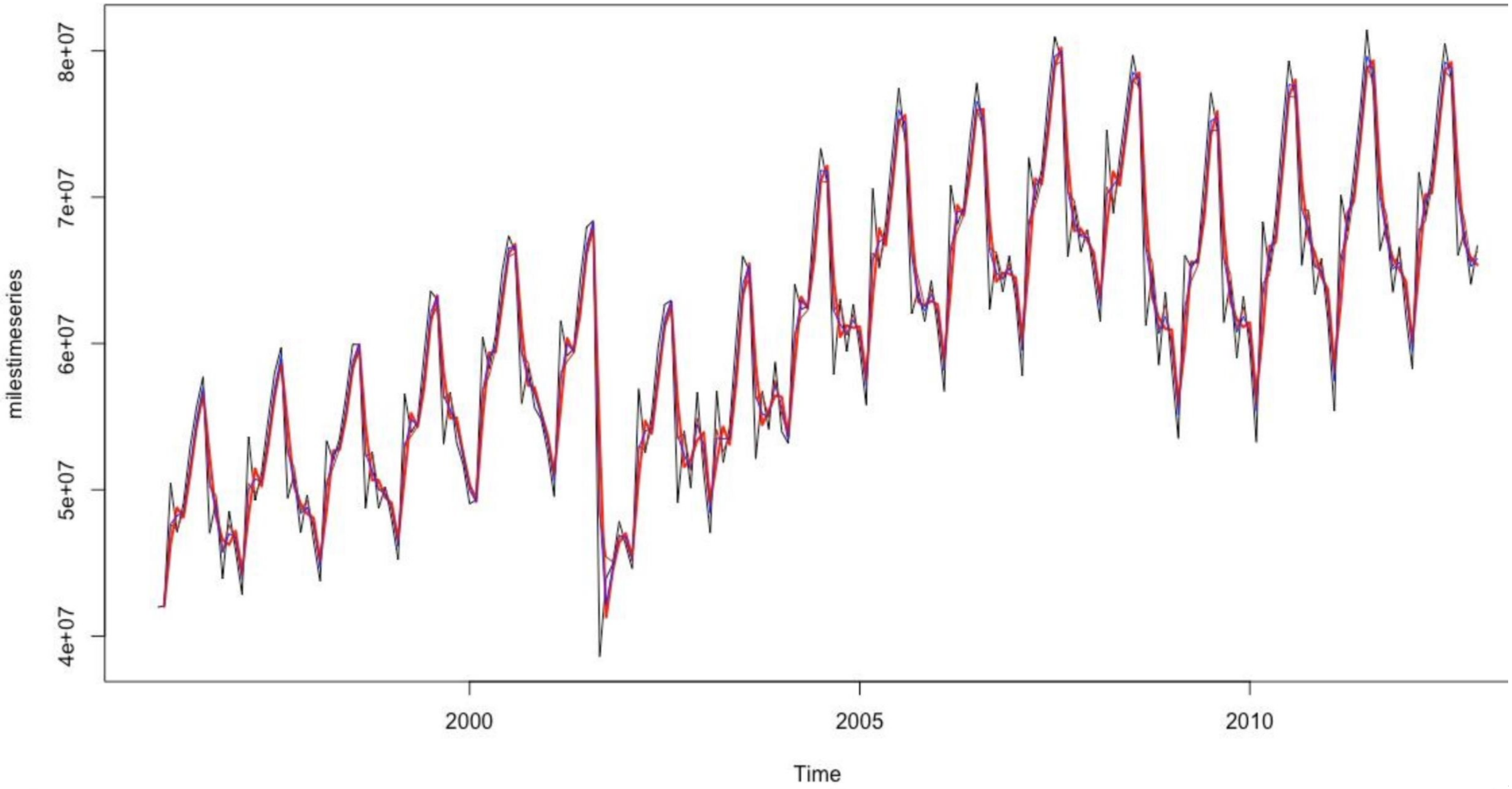
1	Number of products sold	EMA (K=1)	Error	EMA (K=2)	Error	EMA (K=3)	Error	EMA (K=4)	Error
2	15								
3	20	15	5	17.5					
4	16	20	4	16.5	0.5	17	1		
5	13	16	3	14.1666667	1.16666667	15	2	16	3
6	18	13	5	16.7222222	1.27777778	16.5	1.5	16.8	1.2
7	14	18	4	14.9074074	0.90740741	15.25	1.25	15.68	1.68
8	15	14	1	14.9691358	0.0308642	15.125	0.125	15.408	0.408
9	17	15	2	16.3230453	0.67695473	16.0625	0.9375	16.0448	0.9552
10	20	17	3	18.7743484	1.22565158	18.03125	1.96875	17.62688	2.37312
11	20	20	0	19.5914495	0.40855053	19.015625	0.984375	18.576128	1.423872
12	18	20	2	18.5304832	0.53048316	18.5078125	0.5078125	18.3456768	0.3456768
13	20	18	2	19.5101611	0.48983895	19.2539063	0.74609375	19.0074061	0.99259392
14	10	20	10	13.1700537	3.17005368	14.6269531	4.62695313	15.4044436	5.40444365
15	17	10	7	15.7233512	1.27664877	15.8134766	1.18652344	16.0426662	0.95733381
16	17	17	0	16.5744504	0.42554959	16.4067383	0.59326172	16.4255997	0.57440029
17	18	17	1	17.5248168	0.4751832	17.2033691	0.79663086	17.0553598	0.94464017
18	18	18	0	17.8416056	0.1583944	17.6016846	0.39831543	17.4332159	0.5667841
19	14	18	4	15.2805352	1.2805352	15.8008423	1.80084229	16.0599295	2.05992954
20	14	14	0	14.4268451	0.42684507	14.9004211	0.90042114	15.2359577	1.23595772
21	14	14	0	14.1422817	0.14228169	14.4502106	0.45021057	14.7415746	0.74157463
22	19	14	5	17.3807606	1.61923944	16.7251053	2.27489471	16.4449448	2.55505522
23	12	19	7	13.7935869	1.79358685	14.3625526	2.36255264	14.6669669	2.66696687
24	13	12	1	13.264529	0.26452895	13.6812763	0.68127632	14.0001801	1.00018012
25	14	13	1	13.754843	0.24515702	13.8406382	0.15936184	14.0001081	0.00010807
26			**2.91304348**		**0.84055449**		**1.23867161**		**1.48027795**


Number of products sold
EMA (K=1)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23


Number of products sold
EMA (K=2)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23

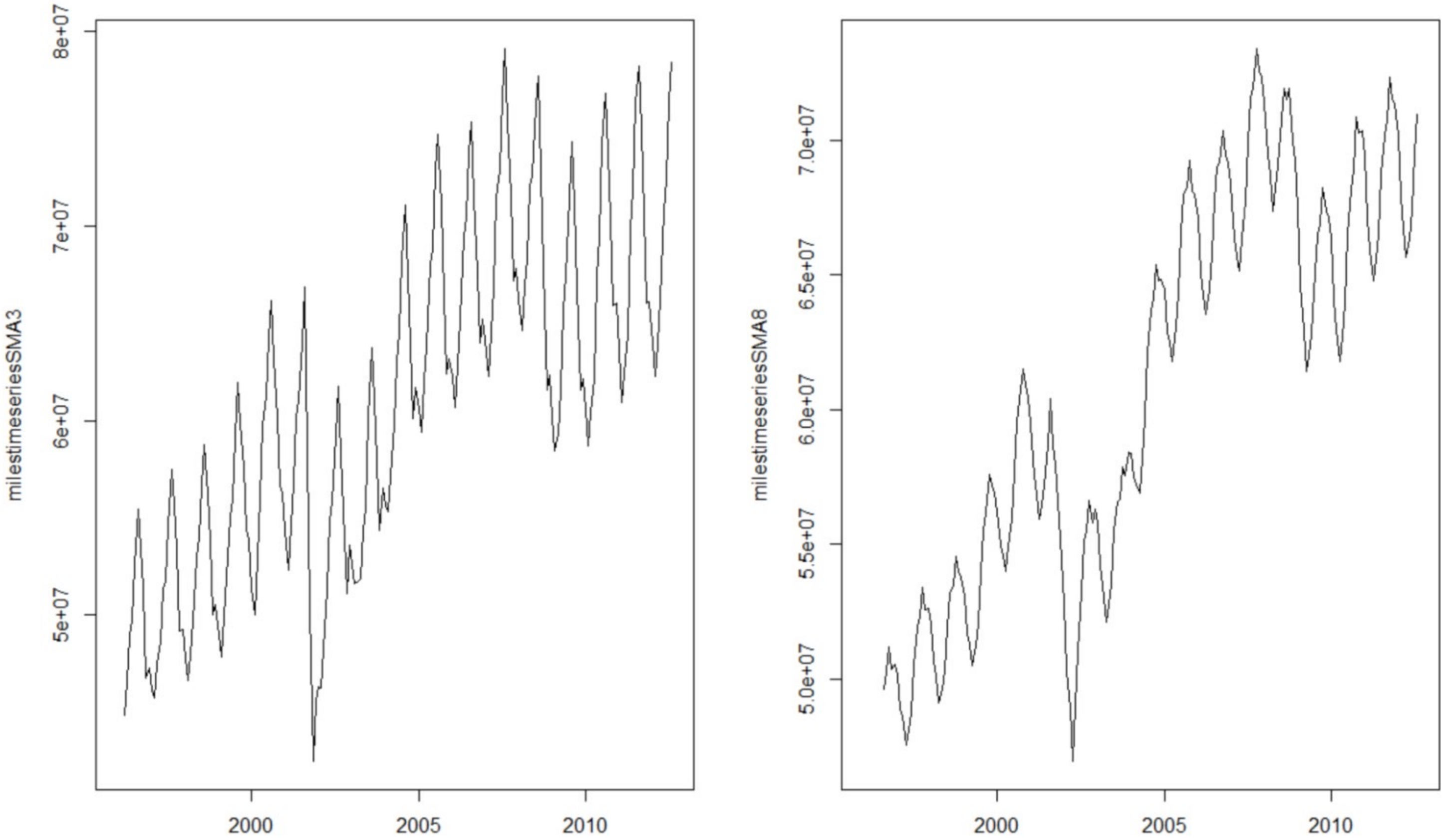

Number of products sold
EMA (K=3)
25
20
15
10
5
0
1 3 5 7 9 11 13 15 17 19 21 23

SMA, WMA and Exponential Smoothing – RPM



`> MAPE-SMA 4.087519 > MAPE-WMA 2.725013 > MAPE-EMA 2.542173`

Effect of *k* – Revenue Passenger Miles



SMA

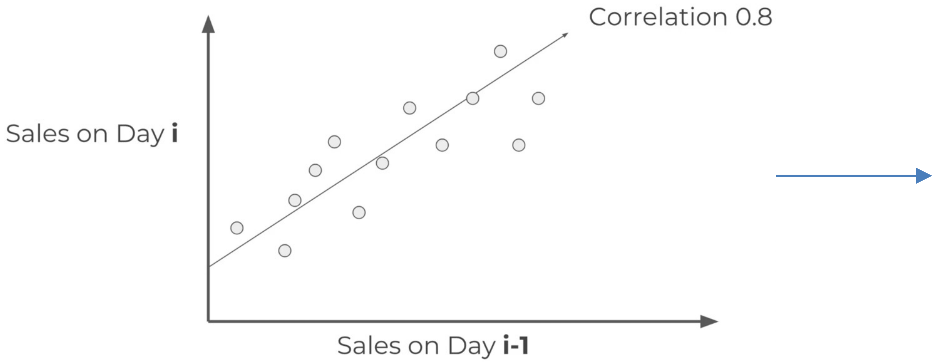
K=3 vs K=8

Advanced Time Series Methods
ARIMA

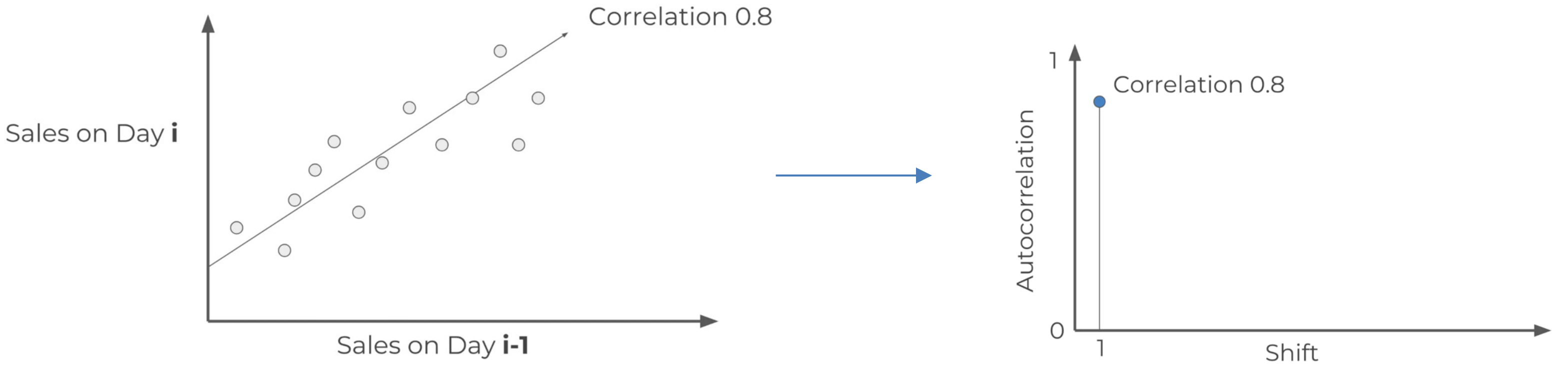
ACF

- An autocorrelation plot (also known as a Correlogram) shows the correlation of the series with itself, lagged by x time units.
- So the y axis is the correlation and the x axis is the number of time units of lag.

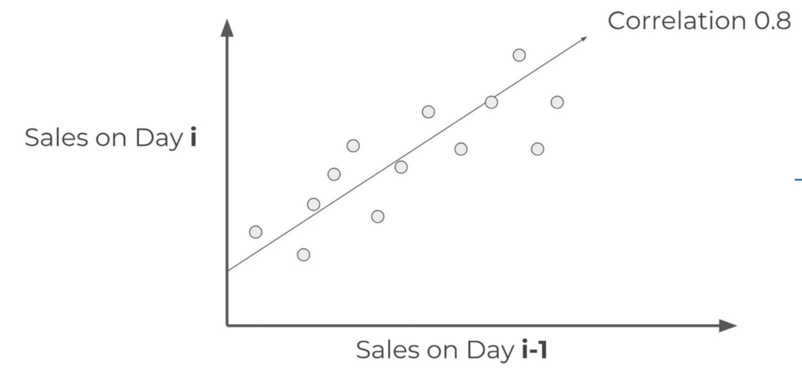
ACF

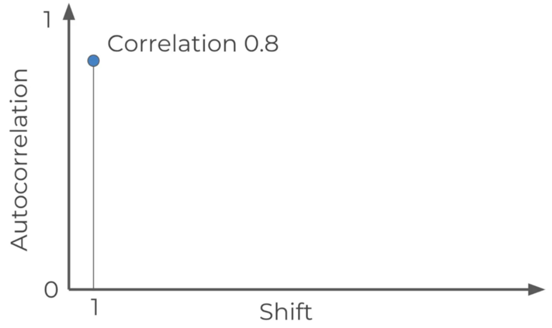


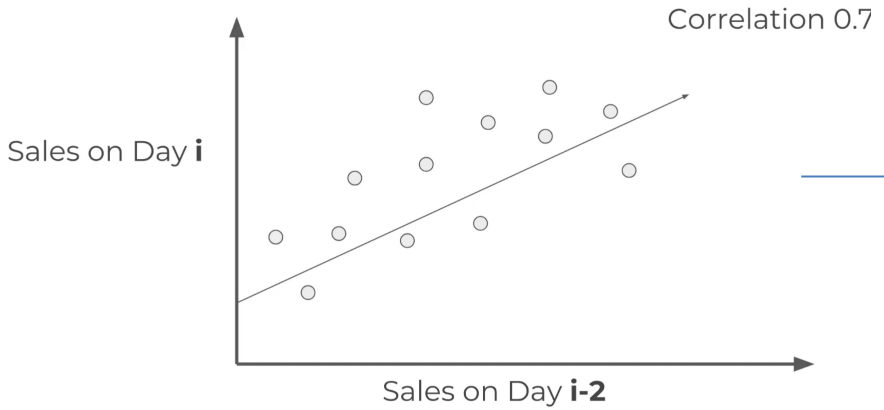
ACF


Correlation 0.8
Sales on Day **i**
Sales on Day **i-1**
1
Correlation 0.8
Autocorrelation
0
1
Shift

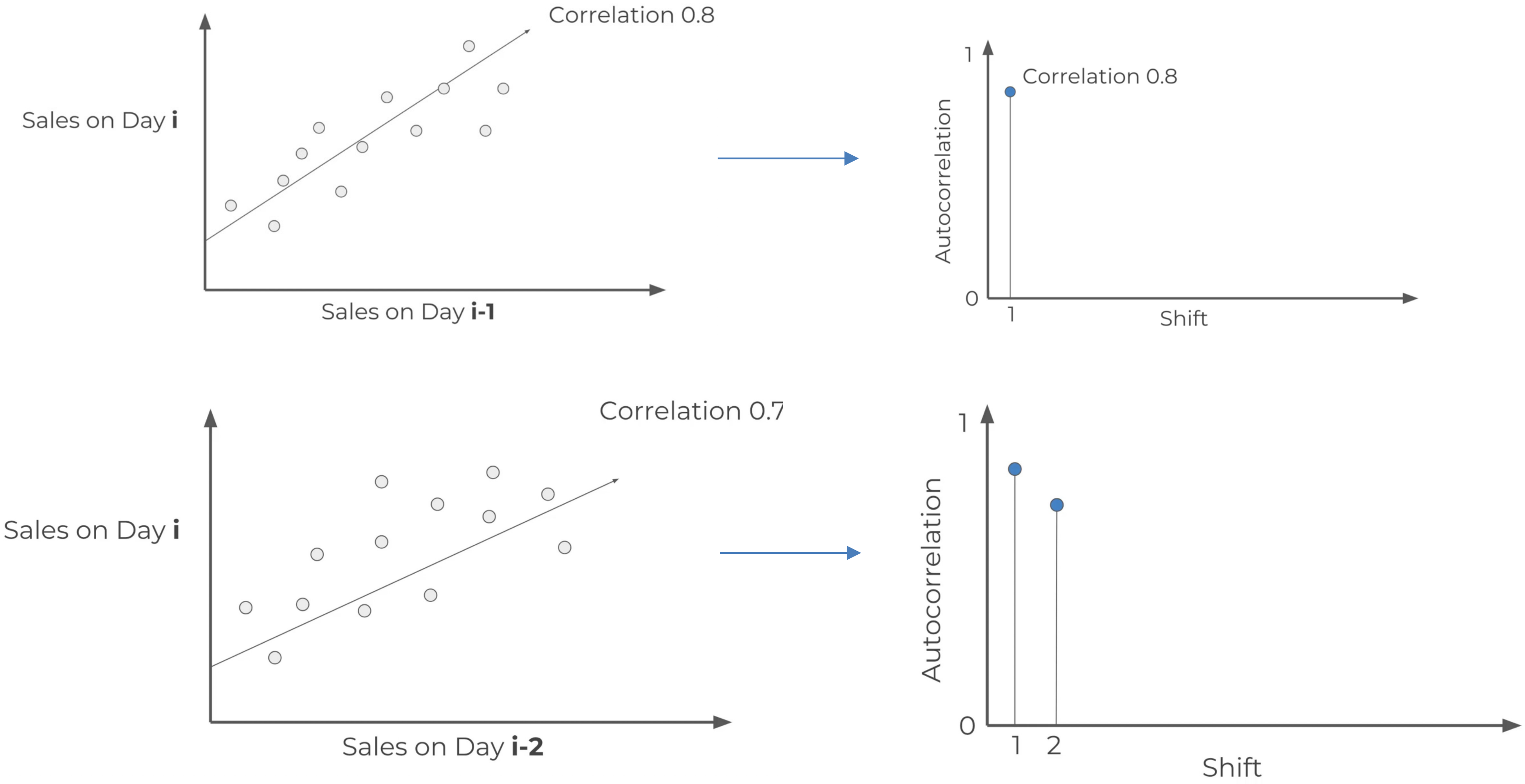
ACF


Correlation 0.8
Sales on Day i
Sales on Day i-1


1
Correlation 0.8
Autocorrelation
0
1
Shift

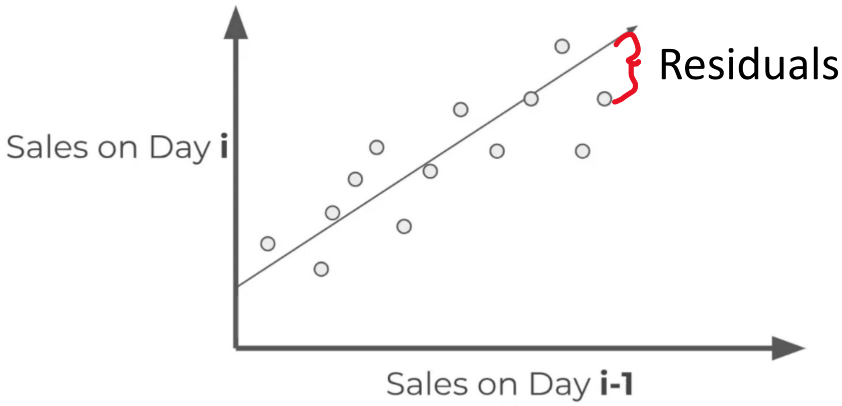

Correlation 0.7
Sales on Day i
Sales on Day i-2

ACF


Correlation 0.8
Sales on Day **i**
Sales on Day **i-1**
1
Correlation 0.8
Autocorrelation
0
1
Shift
Correlation 0.7
Sales on Day **i**
Sales on Day **i-2**
1
Autocorrelation
0
1
2
Shift

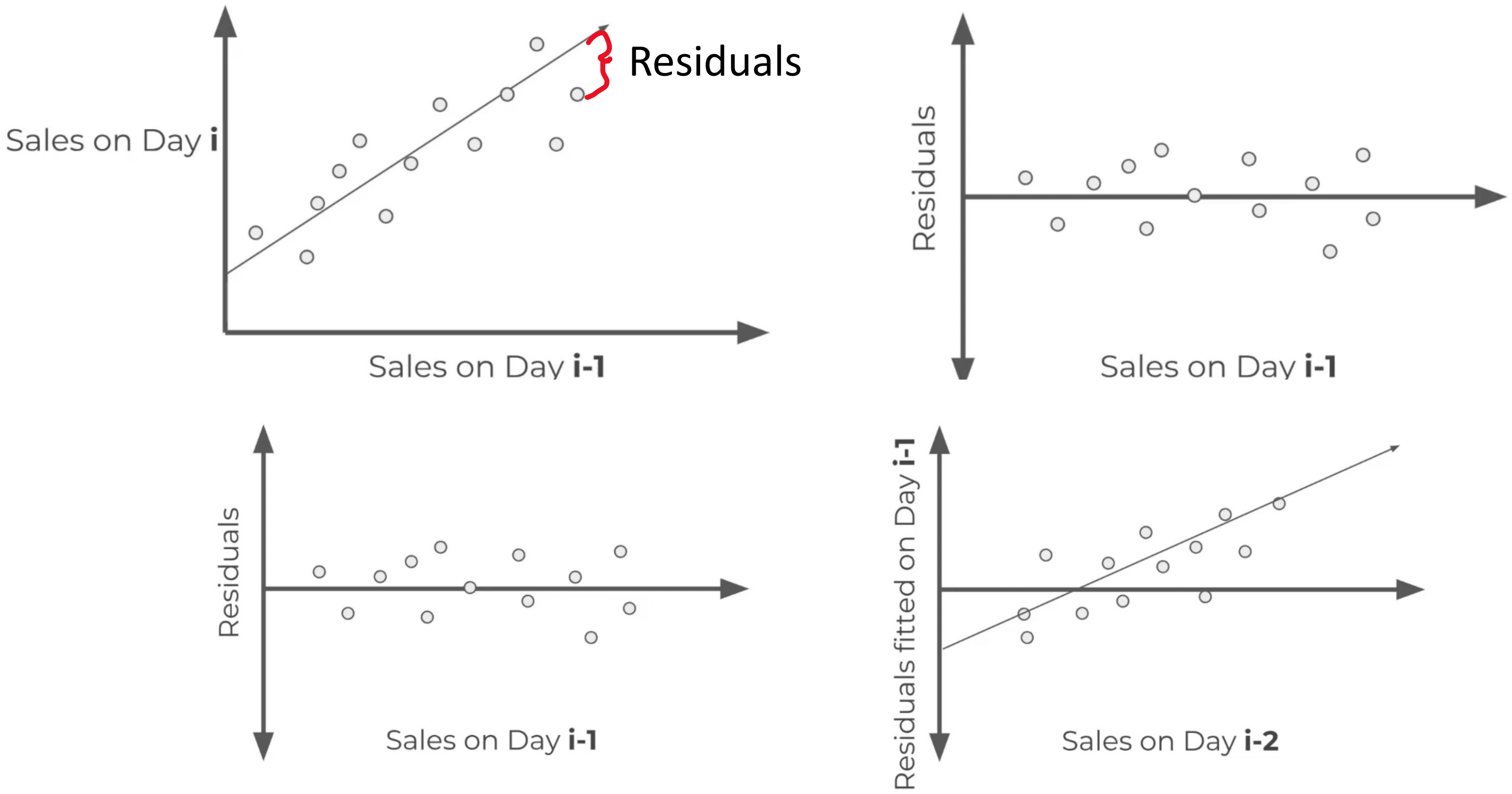
- An autocorrelation plot shows the correlation of the series with itself, lagged by x time units.
- You go on and do this for all possible time lags x and this defines the plot.
- Let’s see some typical examples!

PACF


Residuals
Sales on Day i
Sales on Day i-1


Residuals
Sales on Day i-1

PACF


Residuals
Sales on Day i
Sales on Day i-1
Residuals
Sales on Day i-1
Residuals
Sales on Day i-1
Residuals fitted on Day i-1
Sales on Day i-2

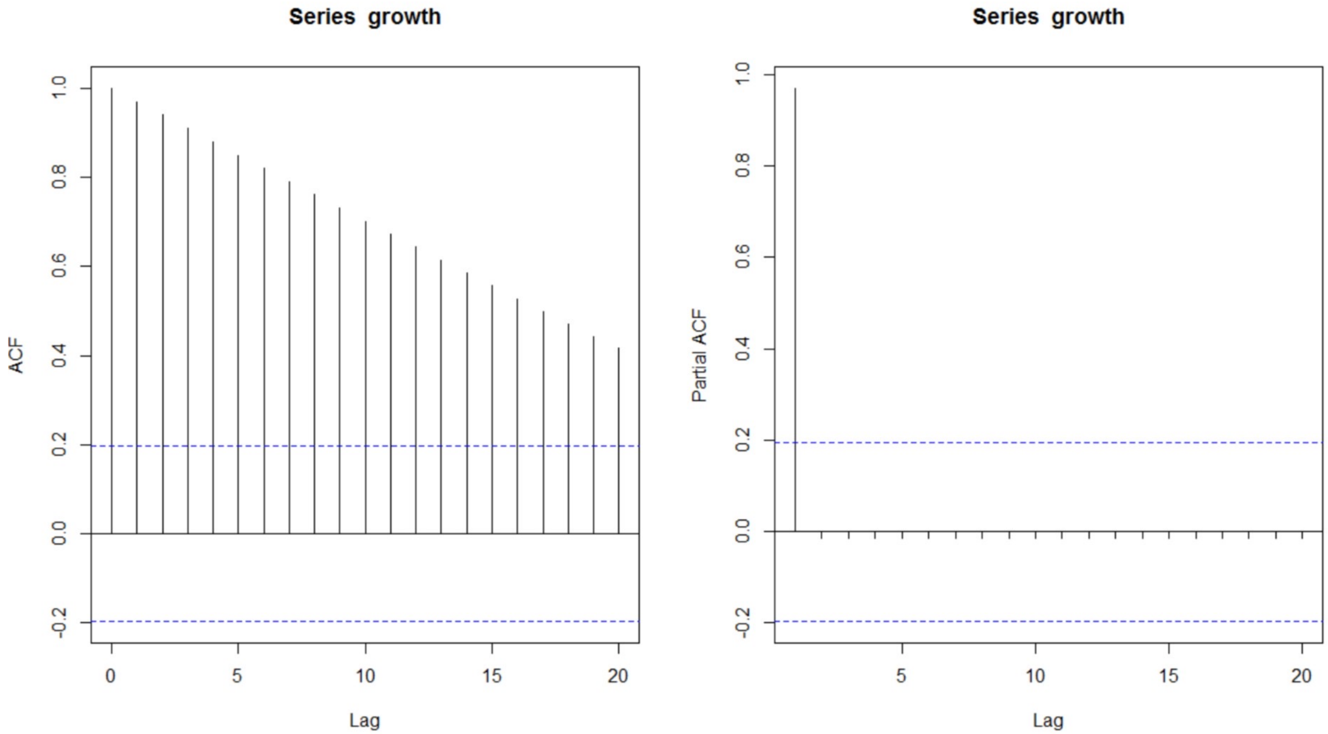
- We essentially plot out the relationship between the previous day’s residuals versus the real values of the current day.
- In general we expect the partial autocorrelation to drop off quite quickly.

ACF vs PACF

- The ACF describes the autocorrelation between an observation and another observation at a prior time step that includes direct and indirect dependence information.

- The PACF only describes the direct relationship between an observation and its lag.

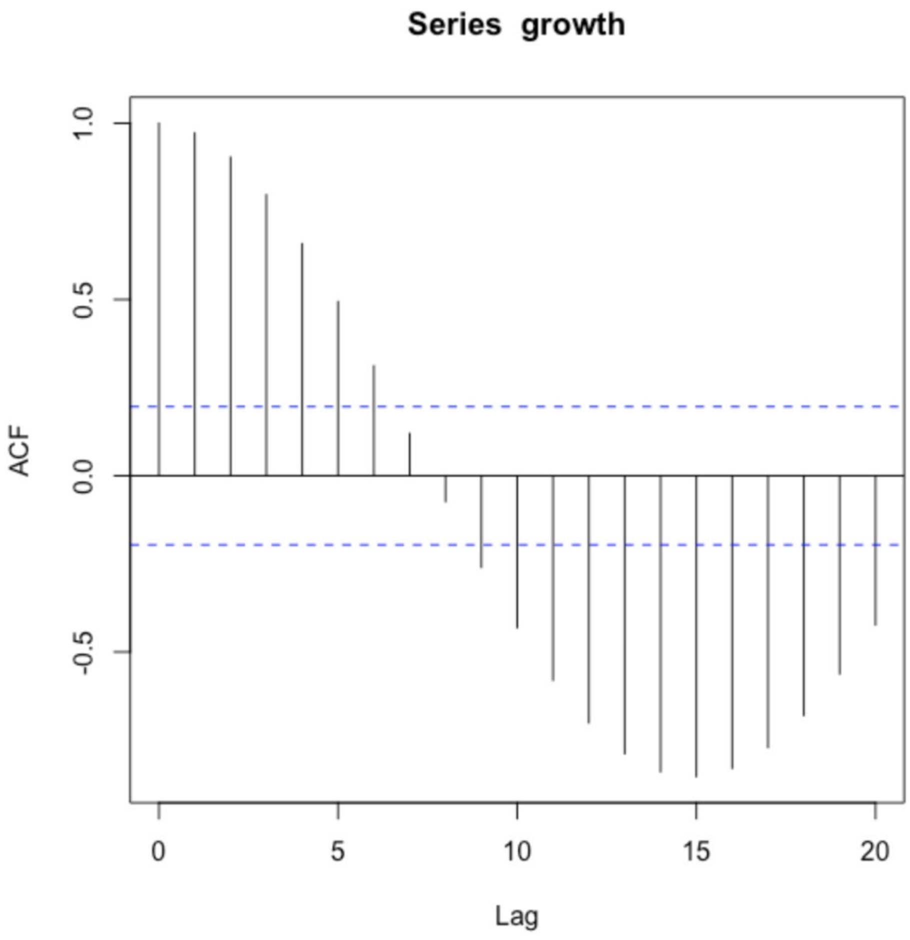
ACF and PACF – Idealized Trend

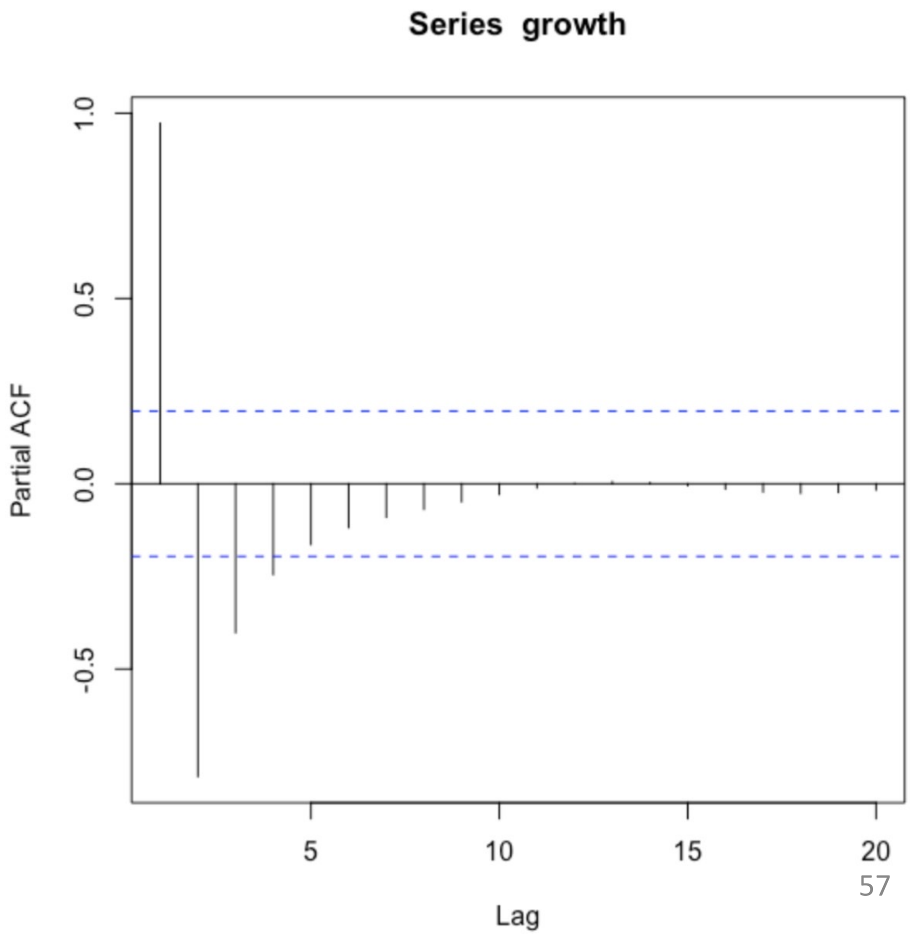


95% CI: $0 \pm \frac{1.96}{\sqrt{n}}$

- ACF is a bar chart of correlation coefficients of the time series and its lags.
- PACF is a plot of the partial correlation coefficients of the time series and its lags.

ACF and PACF – Idealized Seasonality





THANK YOU