Forecasting in HCSC

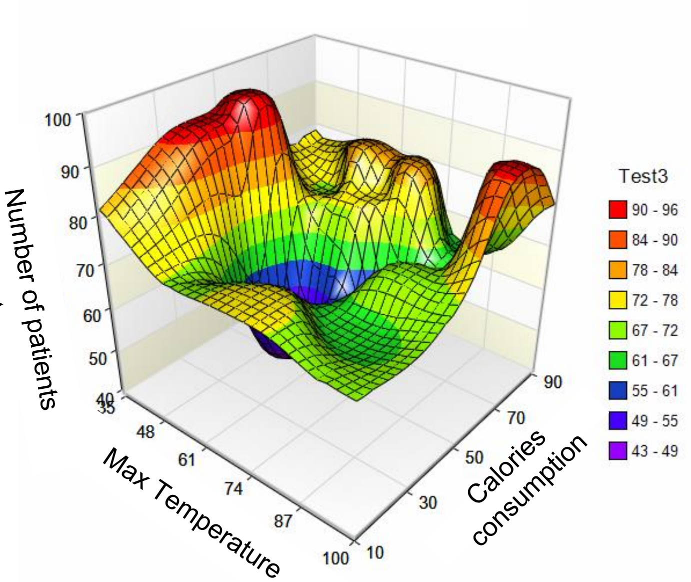
FINNISH NATIONAL AGENCY FOR EDUCATION

Co-funded by the Erasmus+ Programme of the European Union

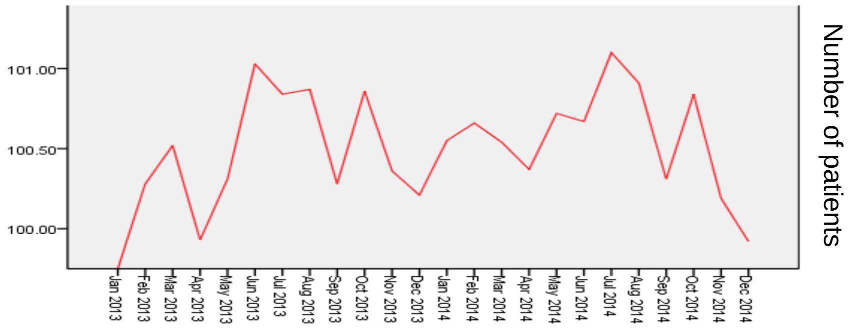
Type of preditions

Predictions

…relationship respect to other variables



…the historical data of the variable

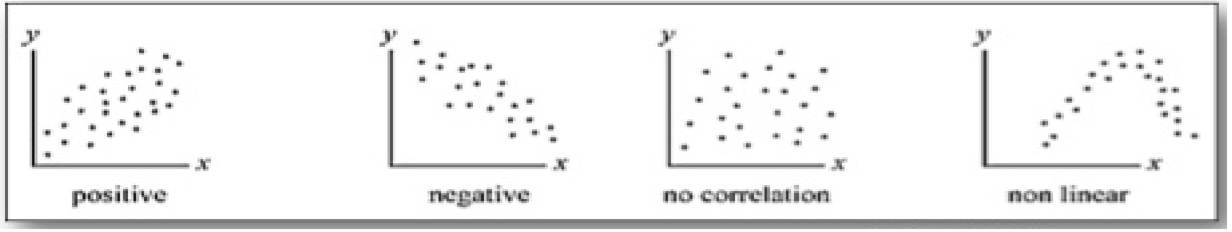


Formulating relationship between decision variables

Bivariate Analysis

Correlation Coefficient (r)

- Pearson r Coefficient: measure the strength of the linear association between continuous variables.
- Spearman rho Coefficient: is used to estimate the degree of association between two ordinal variables measuring the concordance or discordance.
- Kendall Coefficient: measures the strength of dependence between two nominal variables with same scale.



Correlation coefficient varies from -1 (Negative Correlation) to +1 (Positive Correlation)
If Correlation coefficient is 0, there is NO Correlation or association

Bivariate Analysis

Correlation Coefficient (r)

- Pearson r Coefficient: measure the strength of the linear association between continuous variables.

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})}}$$

Example in Excel…(Pills consumption_dataset)

Correlation coefficient varies from -1 (Negative Correlation) to +1 (Positive Correlation)
If Correlation coefficient is 0, there is NO Correlation or association

Independent variable

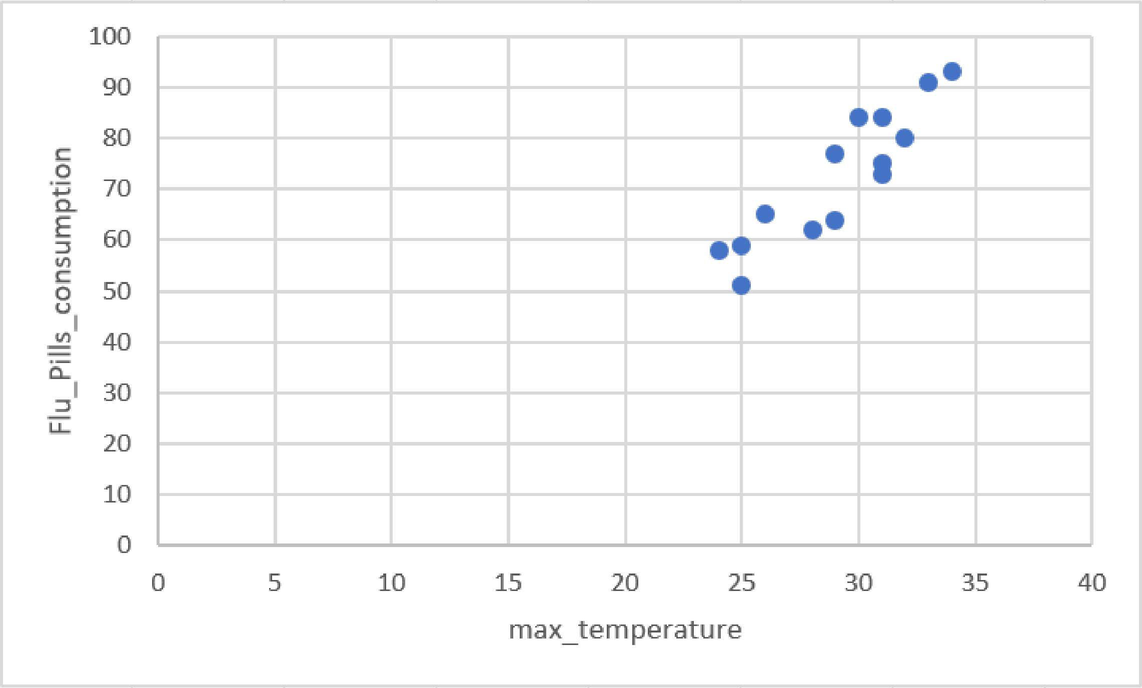
X

Y

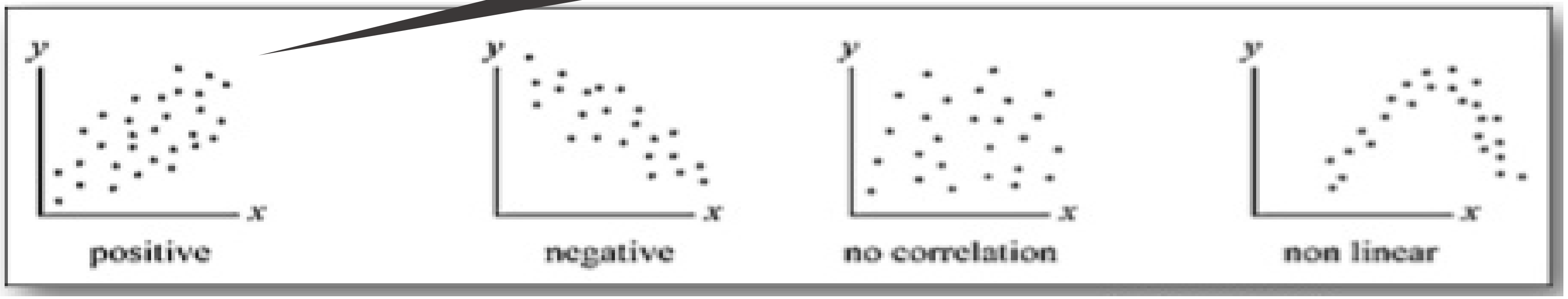
Response variable or dependent variable

Date	Max-Temperature	Pill-consumption
22/06/2017	29	77
23/06/2017	28	62
24/06/2017	34	93
25/06/2017	31	84
26/06/2017	25	59
27/06/2017	29	64
28/06/2017	32	80
29/06/2017	31	75
30/06/2017	24	58
01/07/2017	33	91
02/07/2017	25	51
03/07/2017	31	73
04/07/2017	26	65
05/07/2017	30	84

People will take more flu pills when the temperature growths?



Strong a positive relationship



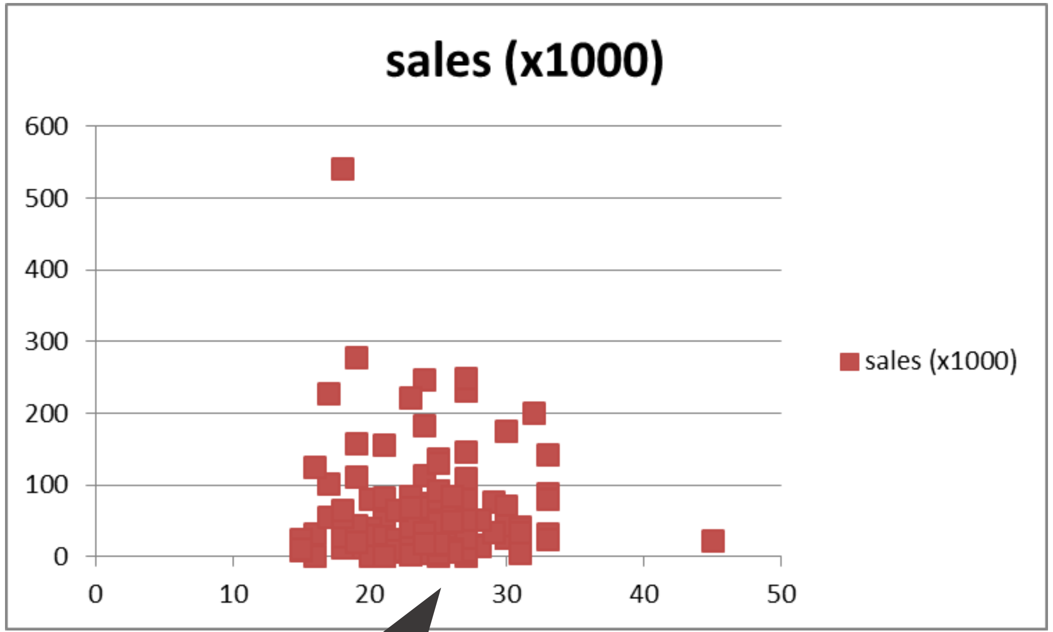
REAL STATISTICS
Add ins…

Correlation Coefficients	
Pearson	0.906923
Spearman	0.892594
Kendall	0.76156
Pearson's coeff (t test)	
Alpha	0.05
Tails	2
corr	0.906923
std err	0.121618
t	7.457156
p-value	7.66E-06
lower	0.641941
upper	1.171905

	sales (x1000)	Fuel efficiency
sales (x1000)	1	
Fuel efficiency	-0.011575691	1

Y X

sales (x1000)	Fuel efficiency
16.919	28
39.384	25
14.114	26
8.588	22
20.397	27
18.78	22
1.38	21
19.747	26
9.231	24
17.527	25
91.561	25
39.35	23
27.851	24
83.257	25
...	...



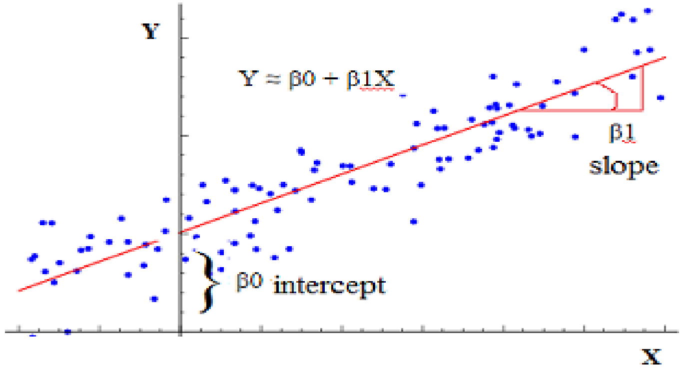
Is there any clear relationship...?

Regression analysis…

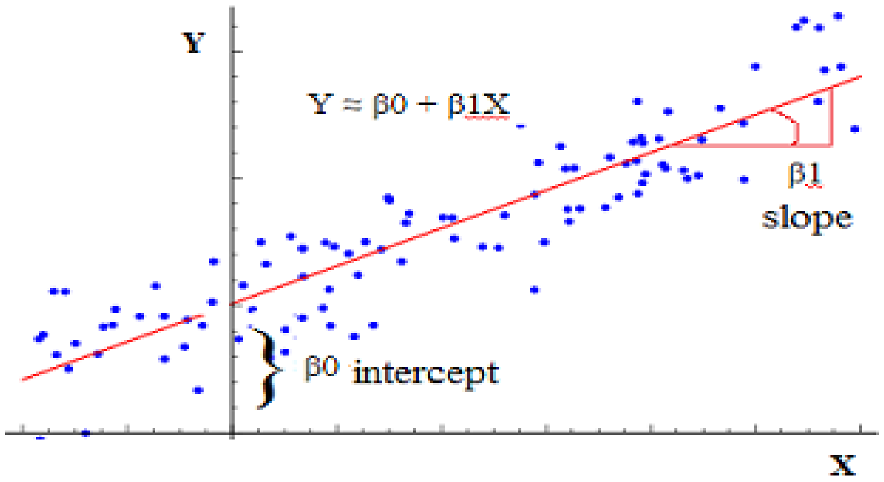
Multivariate Analysis

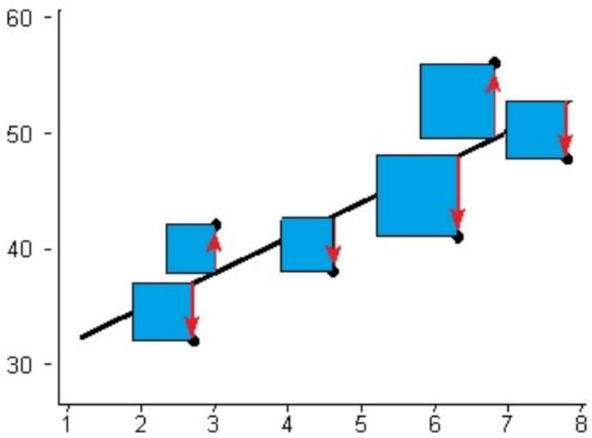
Regression Analysis

- **Linear Regression**: establishes a relationship between dependent variable and one or more independent variables using a best fit straight line (continuous).
- **Logistics Regression**: explain the relationship between dependent binary variable and one or more (continuous) independent variables.
- **Nonlinear Regression**: the relationship between the dependent and independent parameters are not linear.
- **Ordinal Regression**: explains the relationship between a dependent variable and independent variables (ordinal).



Y= dependent variable
β_0 = Y intercept
β_1 = slope coefficient
X= independent variable





Least-squares method

$$Y = \beta_0 + \beta_1 X$$

Regression equation (useful for predictions)

Independent variable

Response variable or dependent variable

Date	Max-Temperature	Pill-consumption
	X	Y
22/06/2017	29	77
23/06/2017	28	62
24/06/2017	34	93
25/06/2017	31	84
26/06/2017	25	59
27/06/2017	29	64
28/06/2017	32	80
29/06/2017	31	75
30/06/2017	24	58
01/07/2017	33	91
02/07/2017	25	51
03/07/2017	31	73
04/07/2017	26	65
05/07/2017	30	84

People will take more flu pills when the temperature growths?

Let's find the regression equation…predictions…

Strong and positive relationship…

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.906922978
R Square	0.822509288
Adjusted R Square	0.807718395
Standard Error	5.708824352
Observations	14

ANOVA					
	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1812.340466	1812.340466	55.60917	7.66141E-06
Residual	12	391.0881057	32.59067548		
Total	13	2203.428571			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-36.36123348	14.68726709	-2.475697709	0.029187	-68.36203945	-4.360427513	-68.36203945	-4.360427513
Max-Temperature	3.737885463	0.501248143	7.457155732	7.66E-06	2.645759578	4.830011347	2.645759578	4.830011347

SUMMARY OUTPUT	
Regression Statistics	
Multiple R	0.906922978
R Square	0.822509288
Adjusted R Square	0.807718395
Standard Error	5.708824352
Observations	14

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1812.340466	1812.340466	55.60917	7.66141E-06
Residual	12	391.0881057	32.59067548		
Total	13	2203.428571			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-36.36123348	14.68726709	-2.475697709	0.029187	-68.36203945	-4.360427513	-68.36203945	-4.360427513
Max-Temperature	3.737885463	0.501248143	7.457155732	7.66E-06	2.645759578	4.830011347	2.645759578	4.830011347

H0: $Y = \beta_0 + \beta_1 X \ldots \beta_1 \approx 0$

H1: $Y = \beta_0 + \beta_1 X \ldots \beta_1 \neq 0$

Result: Reject H0…regression is SIG

Rule:

If p-value <= Alpha (α) Then
  Reject H0
Else
  Accept H0

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.906922978
R Square	0.822509288
Adjusted R Square	0.807718395
Standard Error	5.708824352
Observations	14

ANOVA

	df	SS	MS	F	Significance F
Regression	1	1812.340466	1812.340466	55.60917	7.66141E-06
Residual	12	391.0881057	32.59067548		
Total	13	2203.428571			

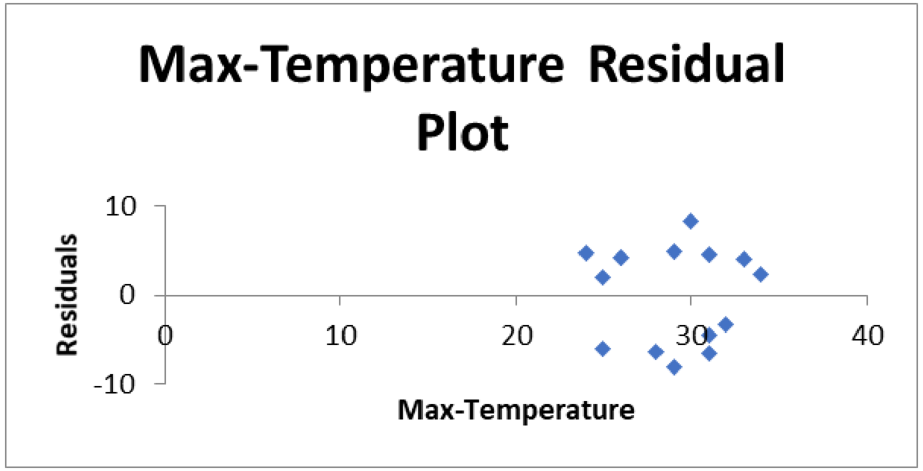
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-36.36123348	14.68726709	-2.475697709	0.029187	-68.36203945	-4.360427513	-68.36203945	-4.360427513
Max-Temperature	3.737885463	0.501248143	7.457155732	7.66E-06	2.645759578	4.830011347	2.645759578	4.830011347

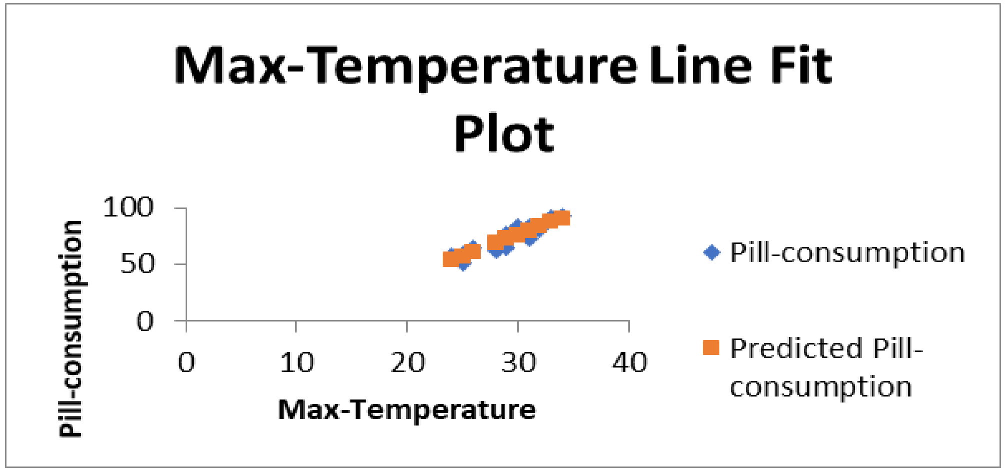
Equation: $Y = -36.3612 + 3.7379X$

Both coefficients contribute significantly to explain the Y´s variability

Pill_consumption = -36.3612 + 3.7379 x MAX-Temperature

Residual graphs





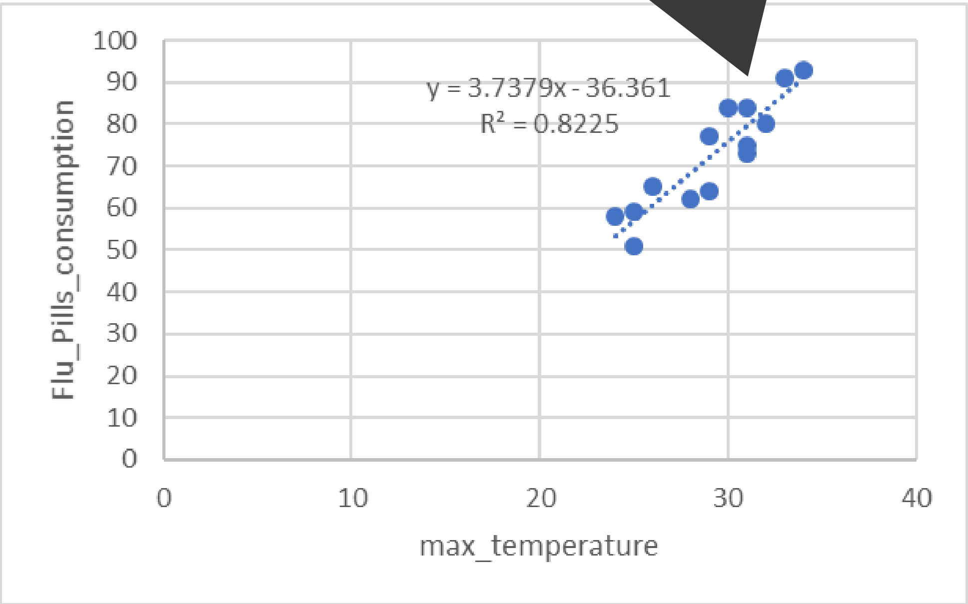
Assumptions (simple-reg.):

- Linear relationship
- Y ~ Normal Distribution
- No-autocorrelation (residuals)

Assumptions (simple-reg.):

- ❑ Linear relationship (OK)
- ❑ Y ~ Normal Distribution
- ❑ No-autocorrelation (residuals)

Clear linear relationship…



Assumptions (simple-reg.):

- ❑ Linear relationship (OK)
- ❑ Y ~ Normal Distribution (OK)
- ❑ No-autocorrelation (residuals)

Anderson-Darling Test			
Alpha	0.05	mean	72.57143
Distrib	Normal	std dev	12.54543
Method	MLE		
AD stat	0.284238		
p-value	0.630726		
crit value	0.705319		

Anderson-Darling Test:

H0: the observed variable fits to a Normal Distribution
H1: the observed variable does not fit to a Normal Distribution

P-value > 0.05 (user-defined parameter) …Then
Accept H0 (the dependent variable follows a Normal Distribution)

Assumptions (simple-reg.):

- ❑ Linear relationship (OK)
- ❑ Y ~ Normal Distribution (OK)
- ❑ No-autocorrelation (residuals) (OK)

Durbin-Watson Test	
Alpha	0.05
D-stat	1.668617028
D-lower	1.04495
D-upper	1.35027
sig	no

1.5 < D-stat < 2.5

The assumption is met

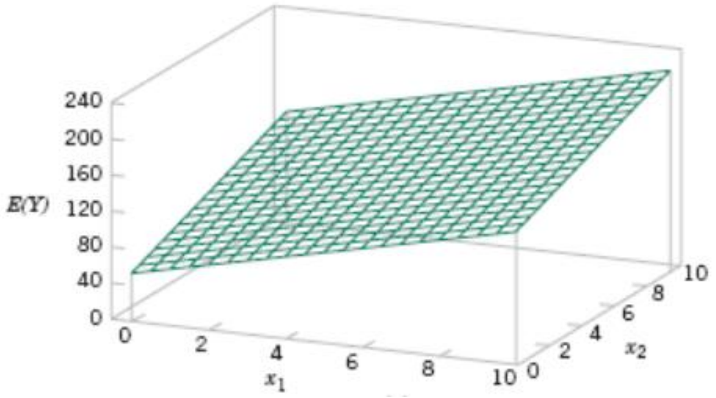
No auto-correlation (residuals)

$$d = \frac{\sum_{i=2}^{n}(e_i - e_{i-1})^2}{\sum_{i=1}^{n} {e_i}^2}$$

Multiple regression…

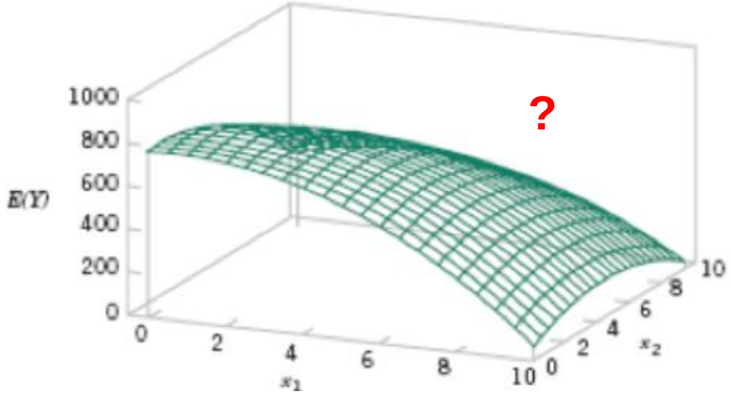
http://www.stat.ucla.edu/~hqxu/stat105/pdf/ch12.pdf

Multiple regression:



Model: $Y = \beta X + \varepsilon$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} \qquad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1b} \\ 1 & X_{21} & X_{22} & \dots & X_{2b} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \qquad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_b \end{bmatrix} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}$$



Multiple regression:

Model: $Y = \beta X + \varepsilon$

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & X_{11} & X_{12} & \dots & X_{1b} \\ 1 & X_{21} & X_{22} & \dots & X_{2b} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_b \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y$$

Seeking the coefficients…

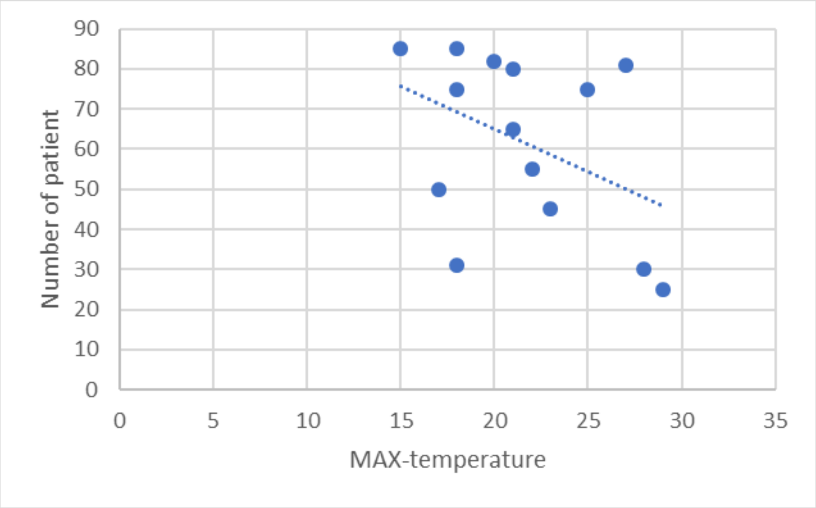
$$\hat{Y} = X\hat{\beta}$$

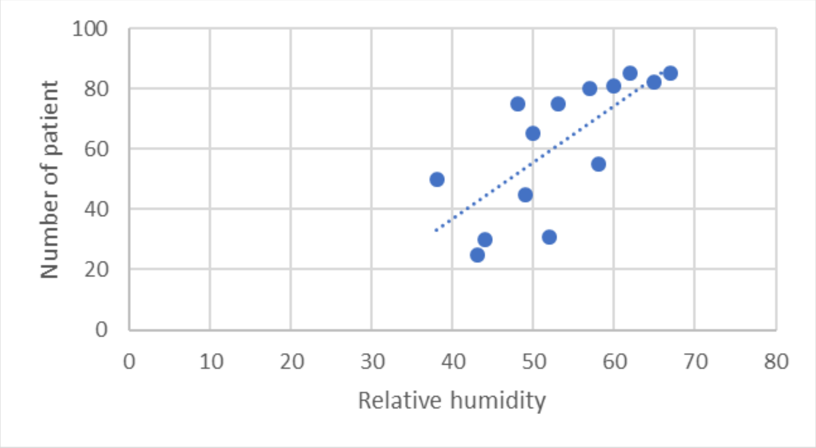
Assumptions (multiple-reg.):

- ❑ Linear relationship
- ❑ Y ~ Normal Distribution
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

X1	X2	Y
Relative humidity (%)	MAX-temperature	Number of patient
50	21	65
49	23	45
57	21	80
67	18	85
44	28	30
38	17	50
53	25	75
48	18	75
65	20	82
62	15	85
60	27	81
43	29	25
52	18	31
58	22	55

Linear relationship…





X1	X2	Y
Relative humidity (%)	MAX-temperature	Number of patient
50	21	65
49	23	45
57	21	80
67	18	85
44	28	30
38	17	50
53	25	75
48	18	75
65	20	82
62	15	85
60	27	81
43	29	25
52	18	31
58	22	55

Linear relationship…

	Relative humidity (%)	*MAX-temperature*	*Number of patient*
Relative humidity (%)	1		
MAX-temperature	-0.279773528	1	
Number of patient	0.729161317	-0.415529852	1

Assumptions (multiple-reg.):

- ❑ Linear relationship
- ❑ Y ~ Normal Distribution (YES--- α = 0.01)
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

Anderson-Darling Test	
Alpha	*0.05*
Distrib	Normal
Method	MLE
AD stat	0.796617
p-value	0.039071
crit value	0.705319

Assumptions (multiple-reg.):

- ❑ Linear relationship
- ❑ Y ~ Normal Distribution
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

Correlation Coefficients	
Pearson	-0.279773528
Spearman	-0.24751759
Kendall	-0.191059483

Pearson's coeff (t test)	
Alpha	0.05
Tails	2
corr	-0.279773528
std err	0.277147189
t	-1.009476334
p-value	0.332668518
lower	-0.88362538
upper	0.324078323

Pearson's coeff (Fisher)	
Rho	0
Alpha	0.05
Tails	2
corr	-0.27977
std err	0.27735
z	-0.95332
p-value	0.340429
lower	-0.70561
upper	0.294526

No significant correlation between independent variables…OK

Assumptions (multiple-reg.):

- ❑ Linear relationship
- ❑ Y ~ Normal Distribution
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

F-Test Two-Sample for Variances		
	Relative humidity (%)	*MAX-temperature*
Mean	53.28571429	21.57142857
Variance	74.37362637	18.87912088
Observations	14	14
df	13	13
F	3.939464494	
P(F<=f) one-tail	0.00963441	
F Critical one-tail	3.905204358	

The variance are significantly different…

Assumptions (multiple-reg.):

- ❑ Linear relationship
- ❑ Y ~ Normal Distribution
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

F-Test Two-Sample for Variances

	Relative humidity (%)	*MAX-temperature*
Mean	53.28571429	21.57142857
Variance	74.37362637	18.87912088
Observations	14	14
df	13	13
F	3.939464494	
P(F<=f) one-tail	0.00963441	
F Critical one-tail	3.905204358	

The variance are significantly different…

Z-score transformation = (x – μ) / σ

Assumptions (multiple-reg.):

- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

transformed…z-score	
Relative humidity (%)	MAX-temperature
-0.380995913	-0.131513722
-0.496951191	0.328784304
0.430691032	-0.131513722
1.590243811	-0.82196076
-1.076727581	1.479529368
-1.772459248	-1.052109773
-0.033130079	0.78908233
-0.612906469	-0.82196076
1.358333256	-0.361662734
1.010467422	-1.512407799
0.778556866	1.249380355
-1.192682859	1.709678381
-0.149085357	-0.82196076
0.54664631	0.098635291

Assumptions (multiple-reg.):

- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals – normality as well)

Variance are now equal…OK

F-Test Two-Sample for Variances		
	Relative humidity (%)	*MAX-temperature*
Mean	1.03092E-16	-3.39015E-16
Variance	1	1
Observations	14	14
df	13	13
F	1	
P(F<=f) one-tail	0.5	
F Critical one-tail	0.256068546	

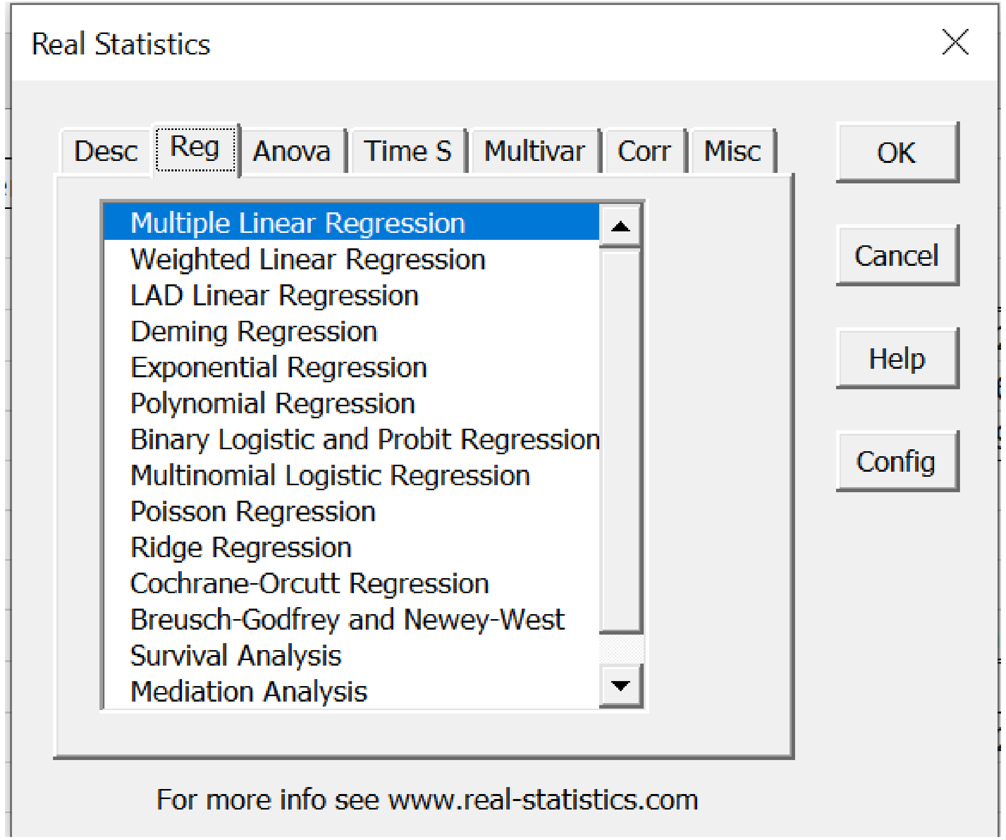
Assumptions (multiple-reg.):

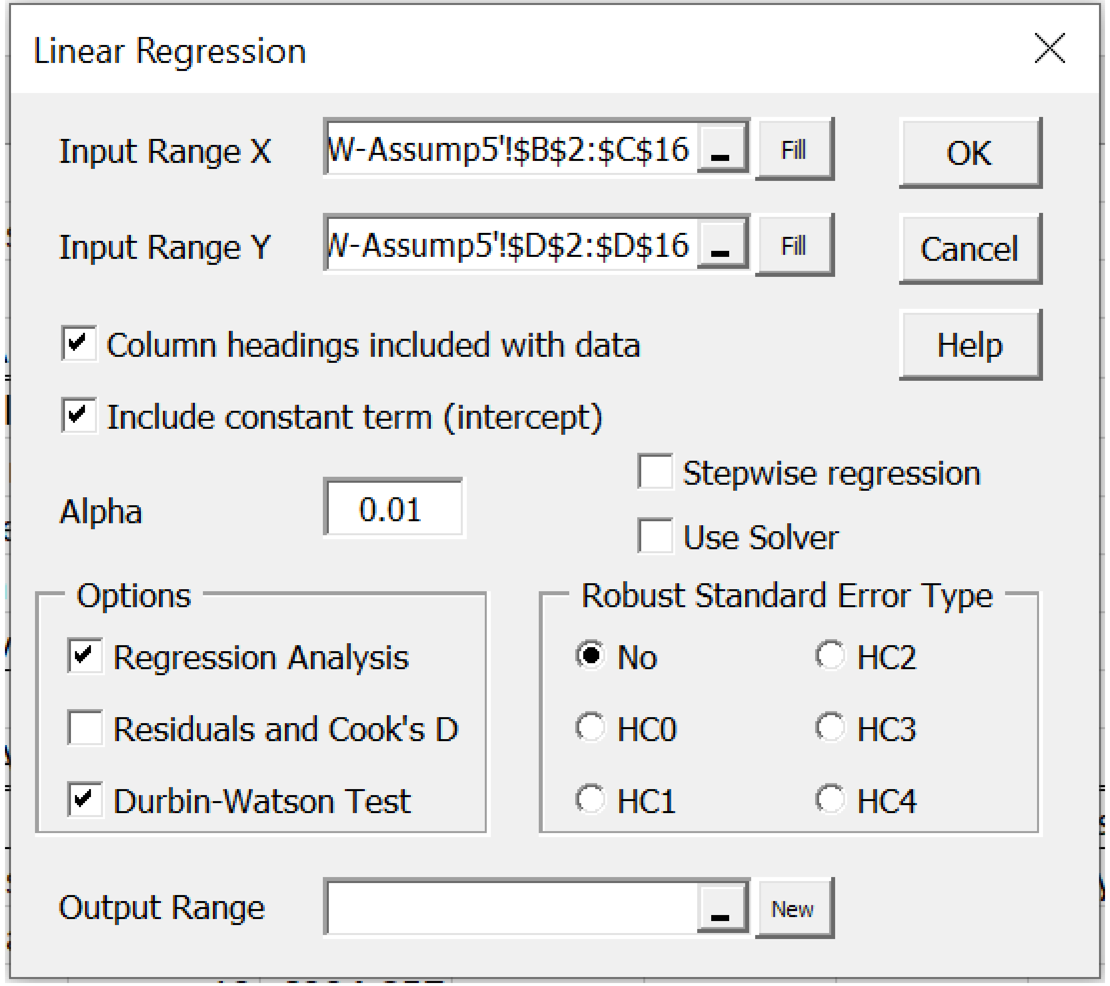
- ❑ Linear relationship
- ❑ Y ~ Normal Distribution
- ❑ No multicollinearity in the data (X-matrix)
- ❑ Homoscedasticity (X-matrix)
- ❑ No-autocorrelation (residuals)

Durbin-Watson Test	
Alpha	0.05
D-stat	1.230658
D-lower	0.90544
D-upper	1.55066
sig	unclear

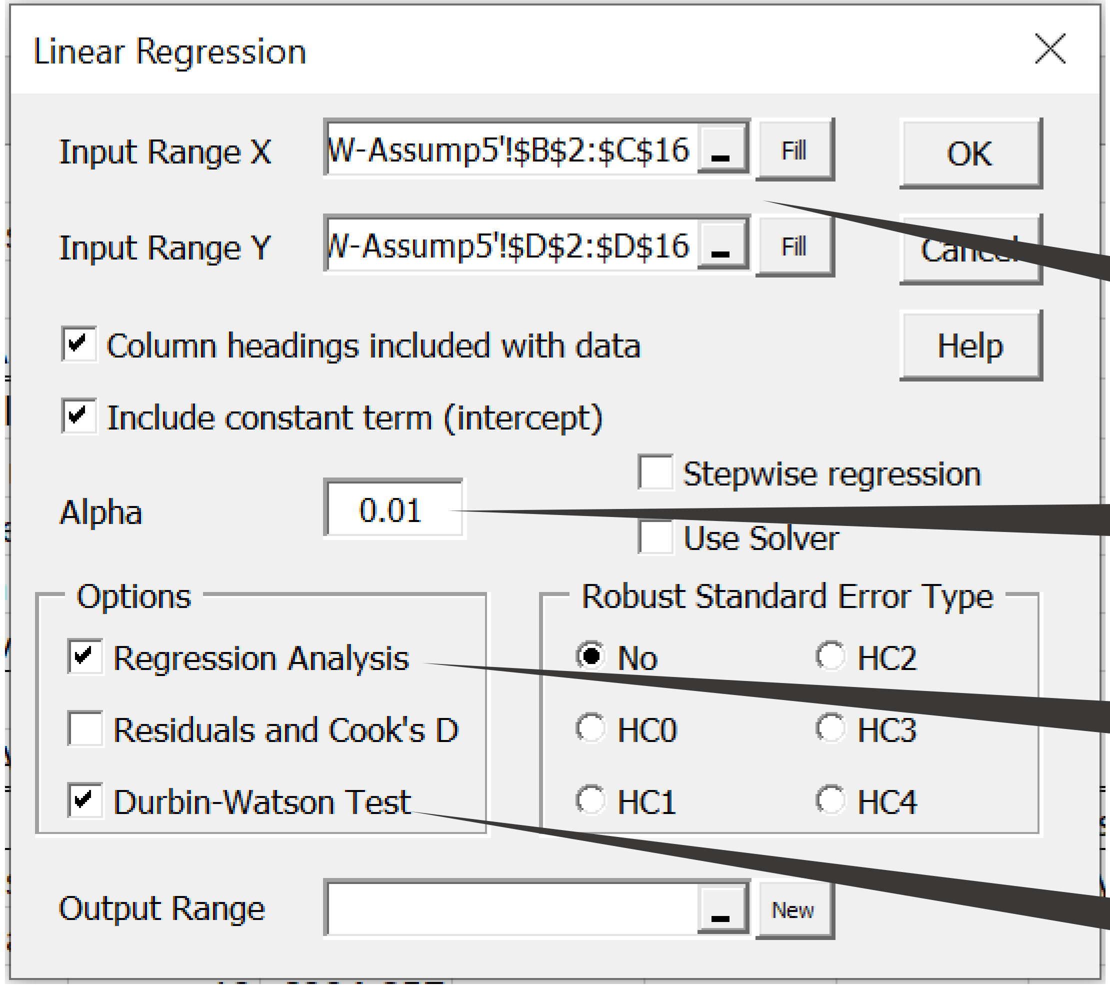
You might assume that is OK…

multiple-reg (REAL STATS-Add ins):





multiple-reg:



Setting the Xs and Y

Setting the confidence level of the regression analysis

Run the whole Regression Analysis

Auto-correlation analysis…

multiple-reg:

The two independent variables explain 76% of the variance in the dependent variable

Regression Analysis							
OVERALL FIT							
Multiple R	0.761722		AIC	79.43222			
R Square	0.580221		AICc	83.87667			
Adjusted R	0.503897		SBC	81.3494			
Standard E	15.53601						
Observatio	14						
ANOVA				Alpha	0.05		
	df	*SS*	*MS*	*F*	*p-value*	*sig*	
Regression	2	3669.814	1834.907	7.602127	0.008445	yes	
Residual	11	2655.043	241.3676				
Total	13	6324.857					
	coeff	*std err*	*t stat*	*p-value*	*lower*	*upper*	*vif*
Intercept	61.71429	4.152173	14.86313	1.26E-08	52.57541	70.85316	
Relative hu	14.66716	4.488143	3.26798	0.007492	4.788827	24.5455	1.08492
MAX-temp	-5.06201	4.488143	-1.12786	0.283378	-14.9403	4.816325	1.08492

The regression equation presents a standard error of 16 patients…

The regression is significant…meaning: you can use the regression equation for predictions…

multiple-reg:

Regression Analysis

OVERALL FIT

Multiple R	0.761722	AIC	79.43222
R Square	0.580221	AICc	83.87667
Adjusted F	0.503897	SBC	81.3494
Standard I	15.53601		
Observatic	14		

ANOVA — Alpha 0.05

	df	SS	MS	F	p-value	sig
Regressior	2	3669.814	1834.907	7.602127	0.008445	yes
Residual	11	2655.043	241.3676			
Total	13	6324.857				

	coeff	std err	t stat	p-value	lower	upper	vif
Intercept	61.71429	4.152173	14.86313	1.26E-08	52.57341	70.85316	
Relative hı	14.66716	4.488143	3.26798	0.007492	4.788827	24.5455	1.08492
MAX-temp	-5.06201	4.488143	-1.12786	0.283378	-14.9403	4.816325	1.08492

Only the intercept and the relative humidity have a significant influence on the number of patients…

multiple-reg:

Equation (daily):
Number of patients = 61.7 + 12.67*RelHumi – 5.1*MAX-temp

Regression Analysis

OVERALL FIT

Multiple R	0.761722	AIC	[illegible]
R Square	0.580221	AICc	[illegible]
Adjusted R	0.503897	SBC	[illegible]
Standard E	15.53601		
Observatio	14		

ANOVA — Alpha 0.05

	df	SS	MS	F	p-value	sig
Regression	2	[illegible]	1834.907	7.602127	0.008445	yes
Residual	11	[illegible]	241.3676			
Total	13	[illegible]				

	coeff	std err	t stat	p-value	lower	upper	vif
Intercept	61.71429	4.152173	14.86313	1.26E-08	52.57541	70.85316	
Relative hu	14.66716	4.488143	3.26798	0.007492	4.788827	24.5455	1.08492
MAX-temp	-5.06201	4.488143	-1.12786	0.283378	-14.9403	4.816325	1.08492

multiple-reg:

Equation (daily):
Number of patients = 61.7 + 12.67*RelHumi – 5.1*MAX-temp

IMPORTANT: This equation works only with z-scores of independent variables…

multiple-reg:

	Relative humidity (%)	MAX-temperature
Mean	53.28571429	21.57142857
Variance	74.37362637	18.87912088

Equation (daily):
Number of patients = 61.7 + 12.67*RelHumi – 5.1*MAX-temp

IMPORTANT: This equation works only with z-scores of independent variables…

$$Z_{revHum} = \frac{\mathbf{50} - 53.29}{\sqrt{74.37}} = -0.4$$

$$Z_{maxTem} = \frac{\mathbf{25} - 21.57}{\sqrt{18.88}} = 0.79$$

Prediction example…predict the number of patients in a day where the maximum temperature was 25 Celsius and we had 50% relative humidity…

Z-score transformation = (x – μ) / σ

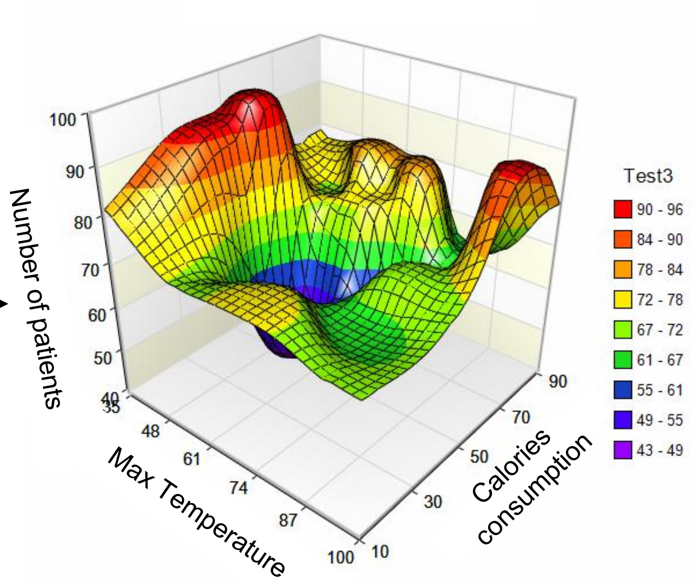
Number of patients = 61.7 + 12.67*(-0.4) – 5.1*(0.79) = 52.6 ≈ 53 patients

Time series analysis…

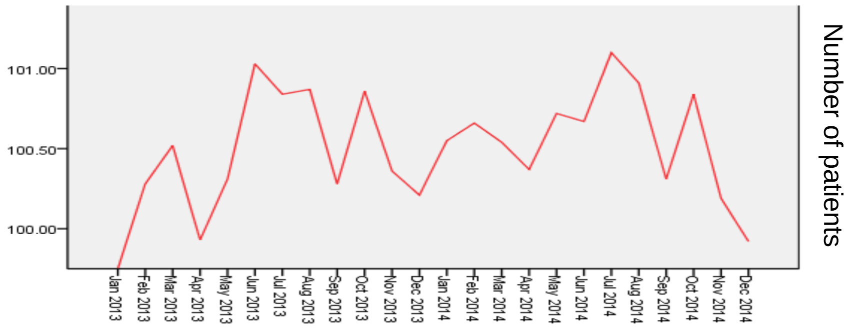
Type of preditions

Predictions

…relationship respect to other variables

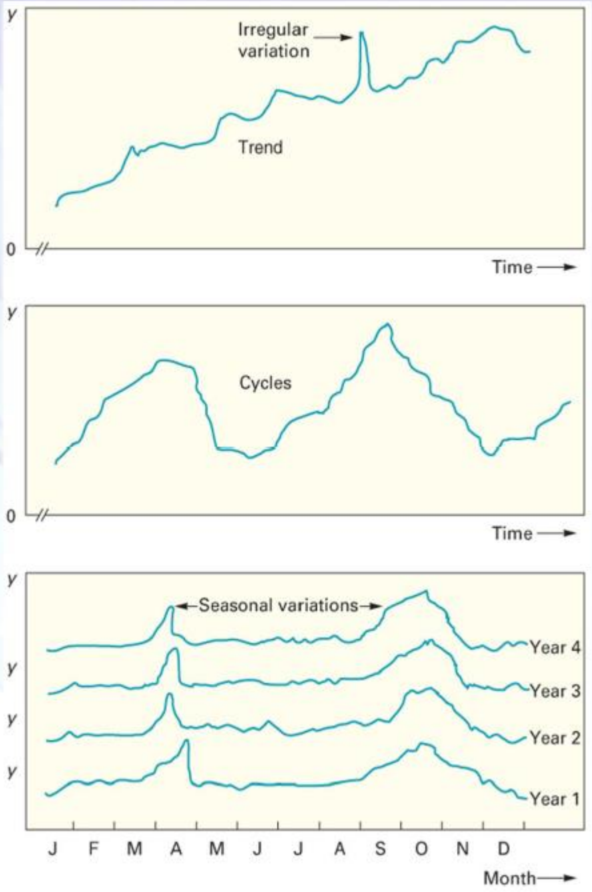


…the historical data of the variable



Time series:

- Forecasts that project patterns identified in recent time-series observations
  - **Time-series** - a time-ordered sequence of observations taken at regular time intervals
- Assume that future values of the time-series can be estimated from past values of the time-series



Time series:

Time Series Patterns

Time series can be decomposed onto several components:

(1) **Trend:** a long term increase or decrease occurs.

(2) **Seasonal:** series influenced by seasonal factors. Thus the series exhibits a behaviour that more or less repeats over a *fixed period of time*, such as a year. Such behaviour is easily demonstrated in a *seasonal plot*, where the data is plotted according to where in the seasonal cycle it was observed).

(3) **Cyclical (not addressed in this course):** series rises and falls regularly but these are *not of fixed period*. Example: economic data rises and falls according to the business cycle, but this cycle varies in length considerably.

(4) **Error:** this corresponds to random fluctuations that cannot be explained by a deterministic pattern.

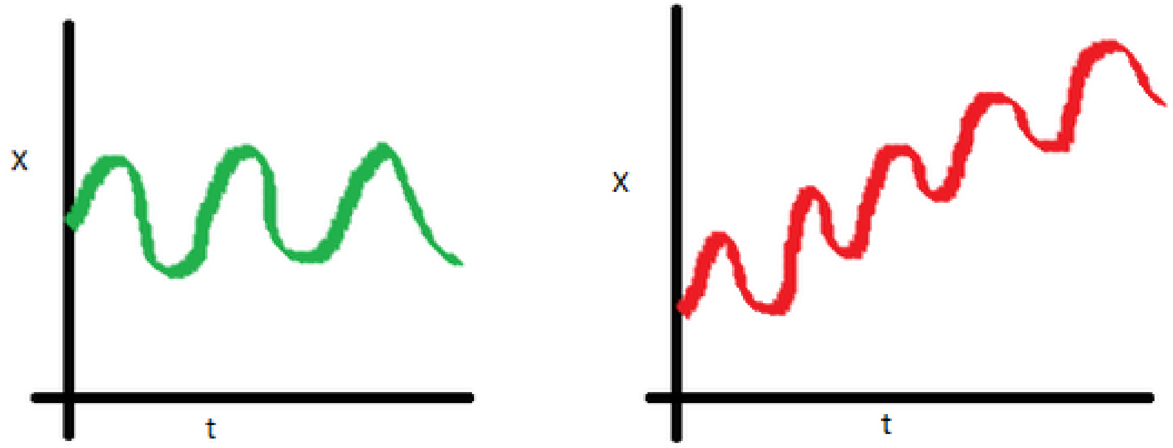
Time series:

Type of Series

- **Stationary**
- **Non-stationary**

Time series:

Constant trend

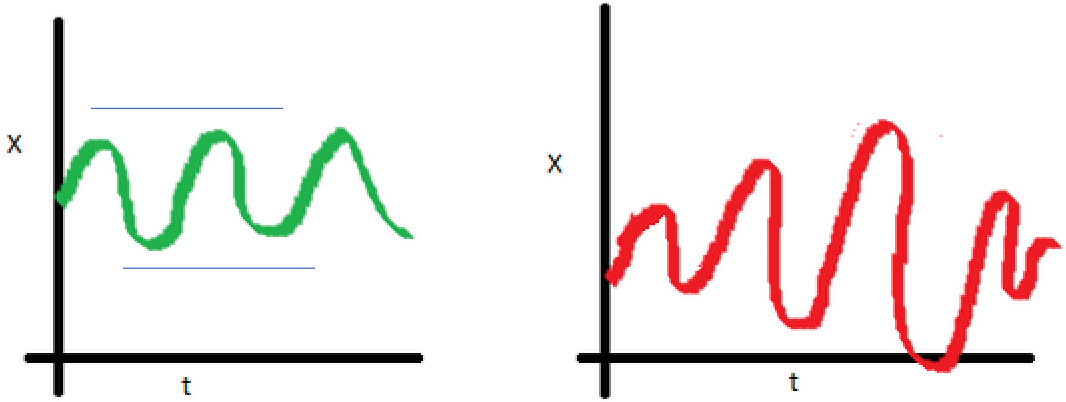


Stationary series

Non-Stationary series

Time series:

homogeneous variance

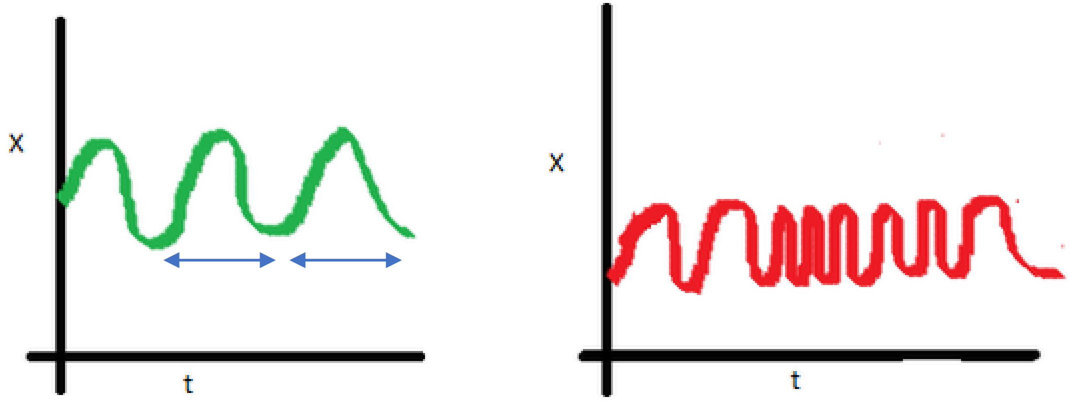


Stationary series

Non-Stationary series

Time series:

Autocovariance



Stationary series

Non-Stationary series

Time series:

Methods for non-stationary time series:

The more useful for realistic datasets

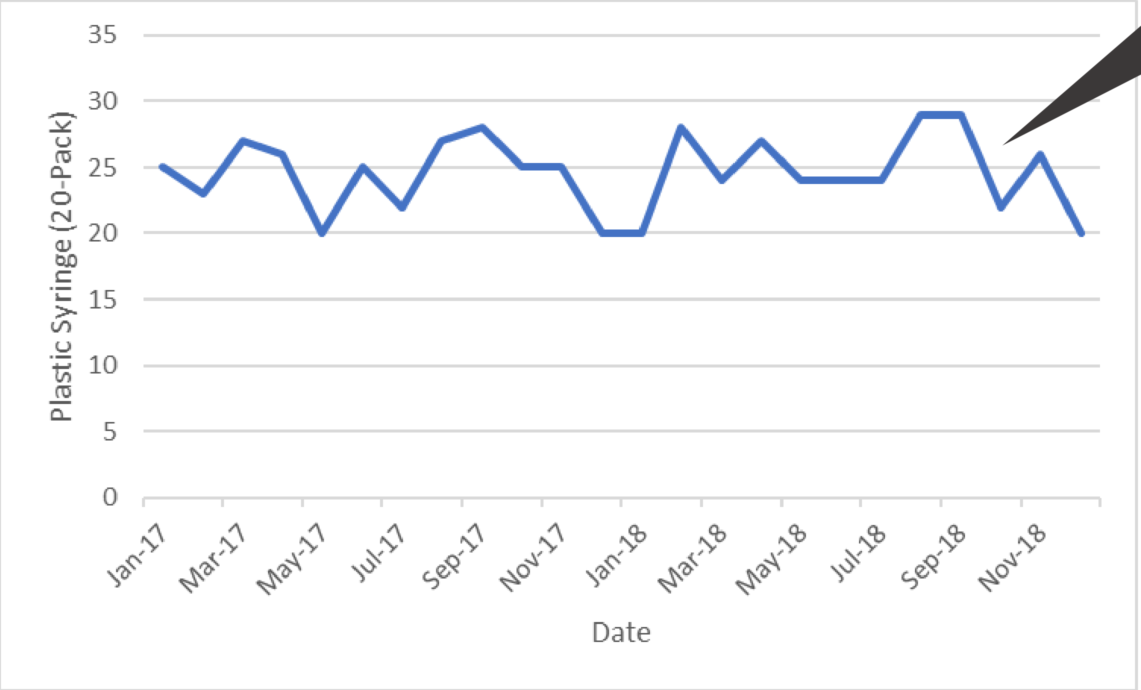
- ❑ **Moving Average (MA)**
- ❑ **Weighted Moving Average (WMA)**
- ❑ **Simple Exponential Smoothing (SES)**
- ❑ **Double Exponential Smoothing (DES)**
- ❑ **Holt Winter´s Additive (HWA)**
- ❑ **Holt Winter´s Multiplicative (HWM)**

Time series:

Date (month)	Plastic Syringe (20-Pack)
Jan-17	25
Feb-17	23
Mar-17	27
Apr-17	26
May-17	20
Jun-17	25
Jul-17	22
Aug-17	27
Sep-17	28
Oct-17	25
Nov-17	25
Dec-17	20
Jan-18	20
Feb-18	28
Mar-18	24
Apr-18	27
May-18	24
Jun-18	24
Jul-18	24
Aug-18	29
Sep-18	29
Oct-18	22
Nov-18	26
Dec-18	20

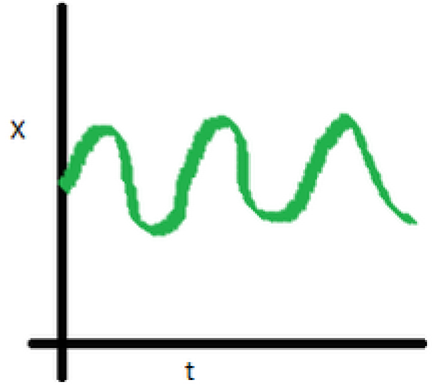
dataset

Time series:

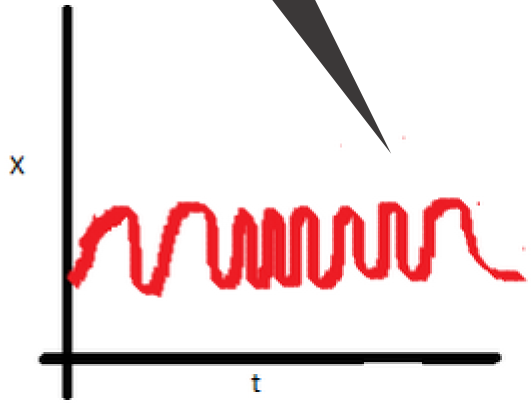


Clearly a non-stationary series...

Autocovariance

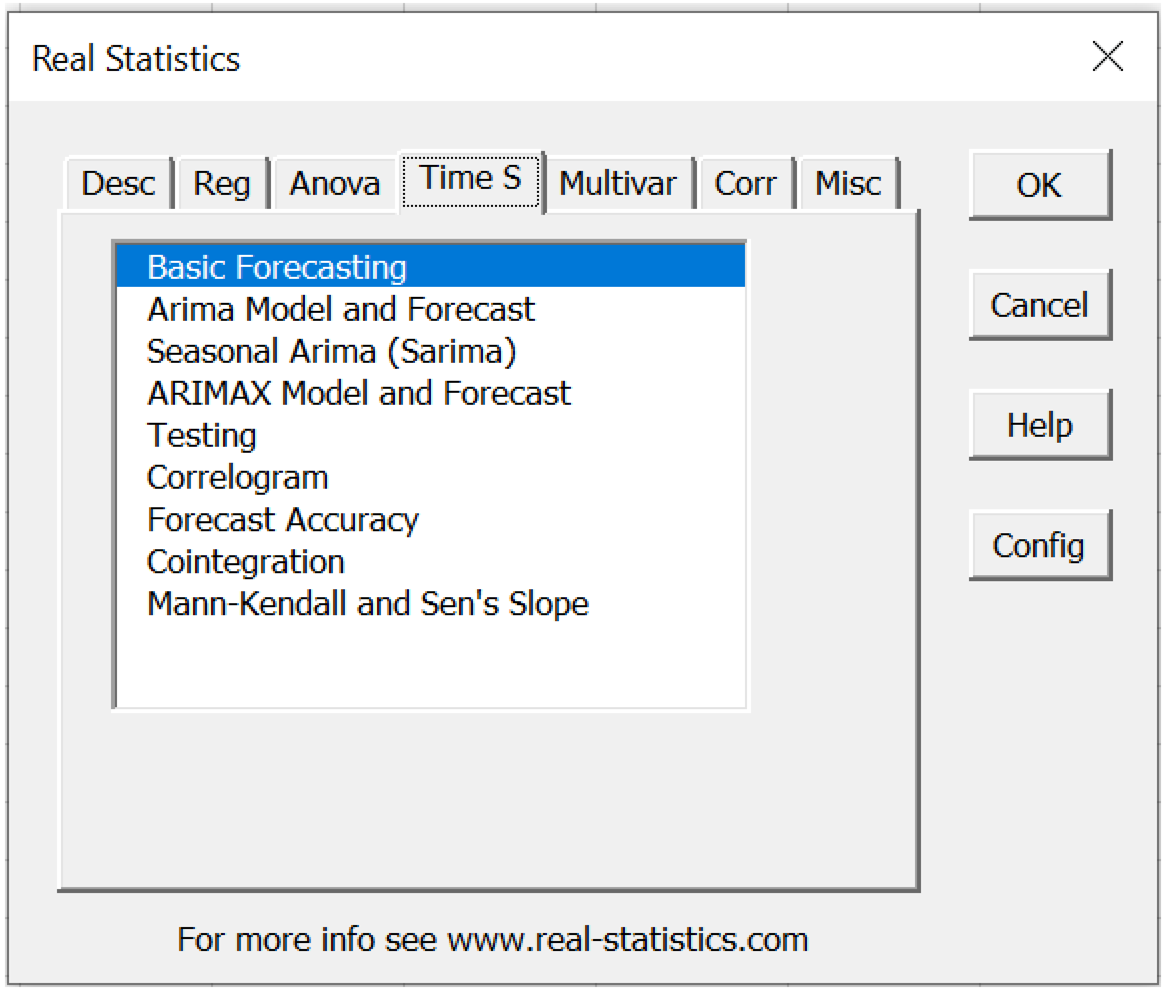


Stationary series



Non-Stationary series

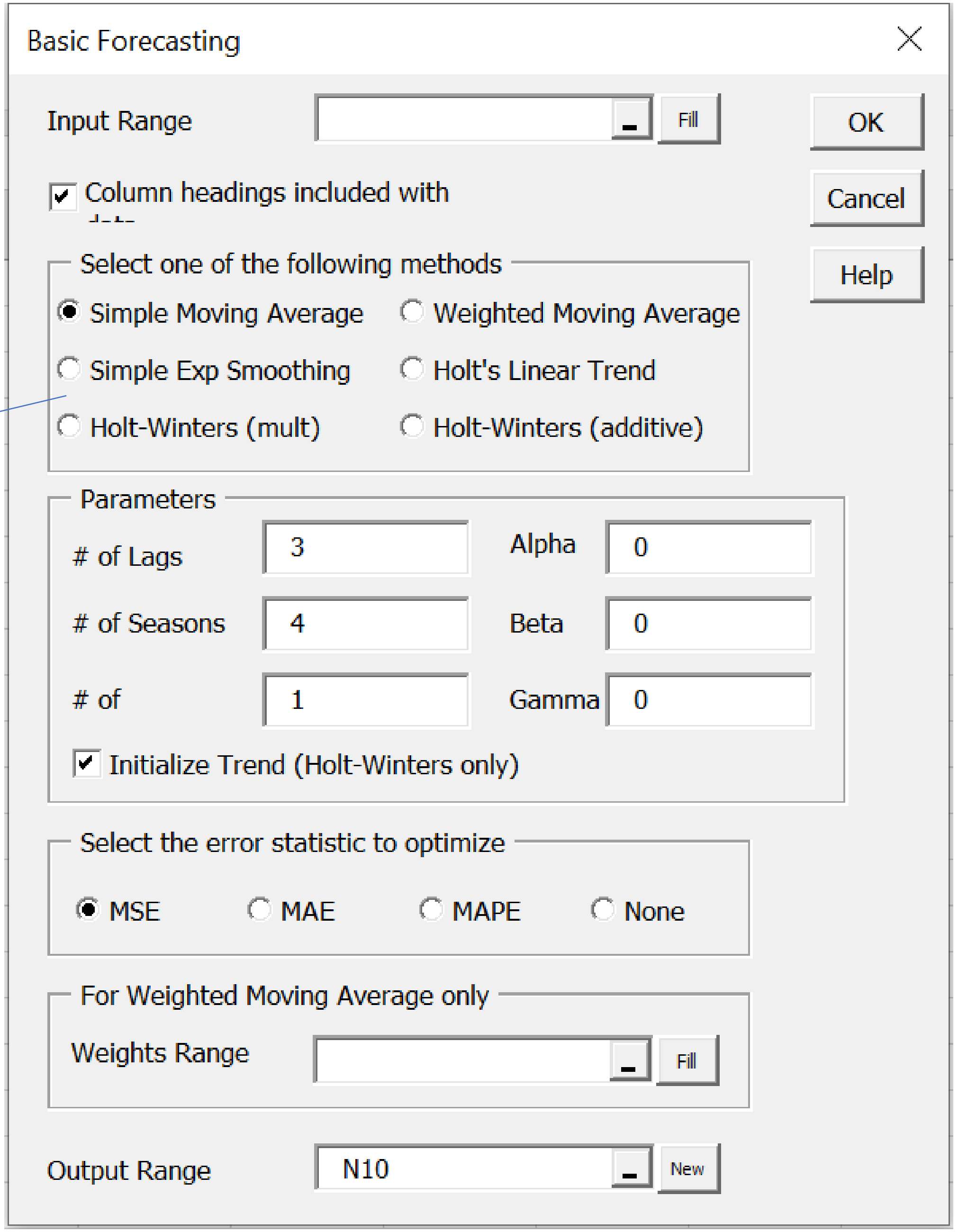
Time series:


Real Statistics
Desc
Reg
Anova
Time S
Multivar
Corr
Misc
Basic Forecasting
Arima Model and Forecast
Seasonal Arima (Sarima)
ARIMAX Model and Forecast
Testing
Correlogram
Forecast Accuracy
Cointegration
Mann-Kendall and Sen's Slope
OK
Cancel
Help
Config
For more info see www.real-statistics.com

Time series:

Methods for non-stationary time series:

- **Moving Average (MA)**
- **Weighted Moving Average (WMA)**
- **Simple Exponential Smoothing (SES)**
- **Double Exponential Smoothing (DES) (Holt´s Method)**
- **Holt Winter´s Additive (HWA)**
- **Holt Winter´s Multiplicative (HWM)**



Time series:

- **Moving Average (MA)**

- Technique that averages a number of the most recent actual values in generating a forecast

$$F_t = MA_n = \frac{\sum_{i=1}^{n} A_{t-i}}{n}$$

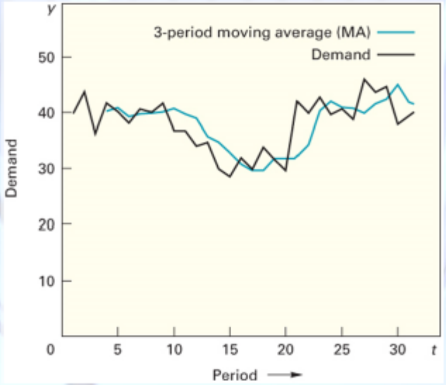
where

F_t = Forecast for time period t

MA_n = n period moving average

A_{t-1} = Actual value in period $t-1$

n = Number of periods in the moving average



As new data become available, the forecast is updated by adding the newest value and dropping the oldest and then re-computing the average

The number of data points included in the average determines the model's sensitivity

Fewer data points used-- more responsive

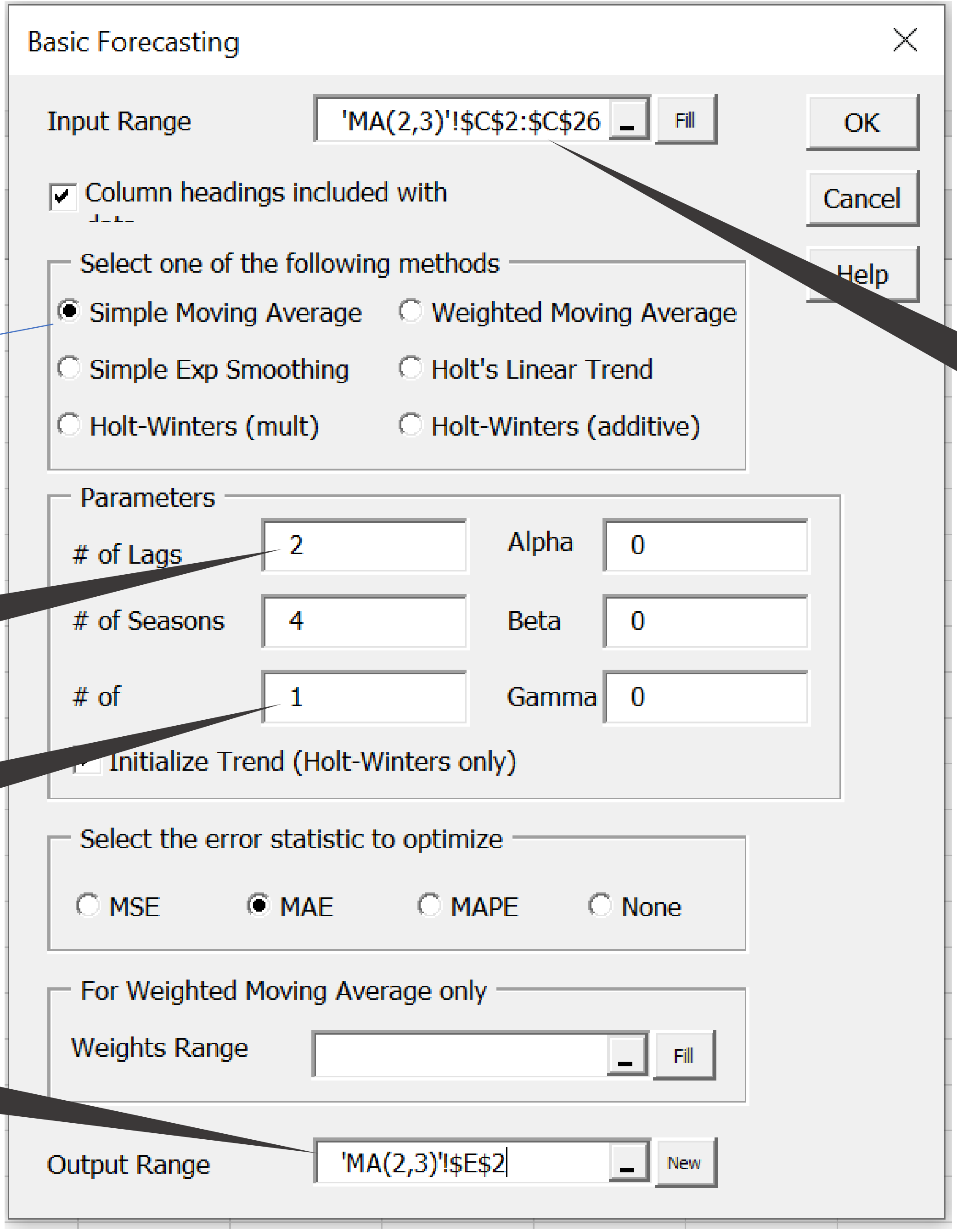
More data points used-- less responsive

User-defined parameter(s): n

Time series:

Methods for non-stationary time series:

- **Moving Average (MA)**



Range of observed variables, syringe consumption, over time…

Two(2) moving periods…

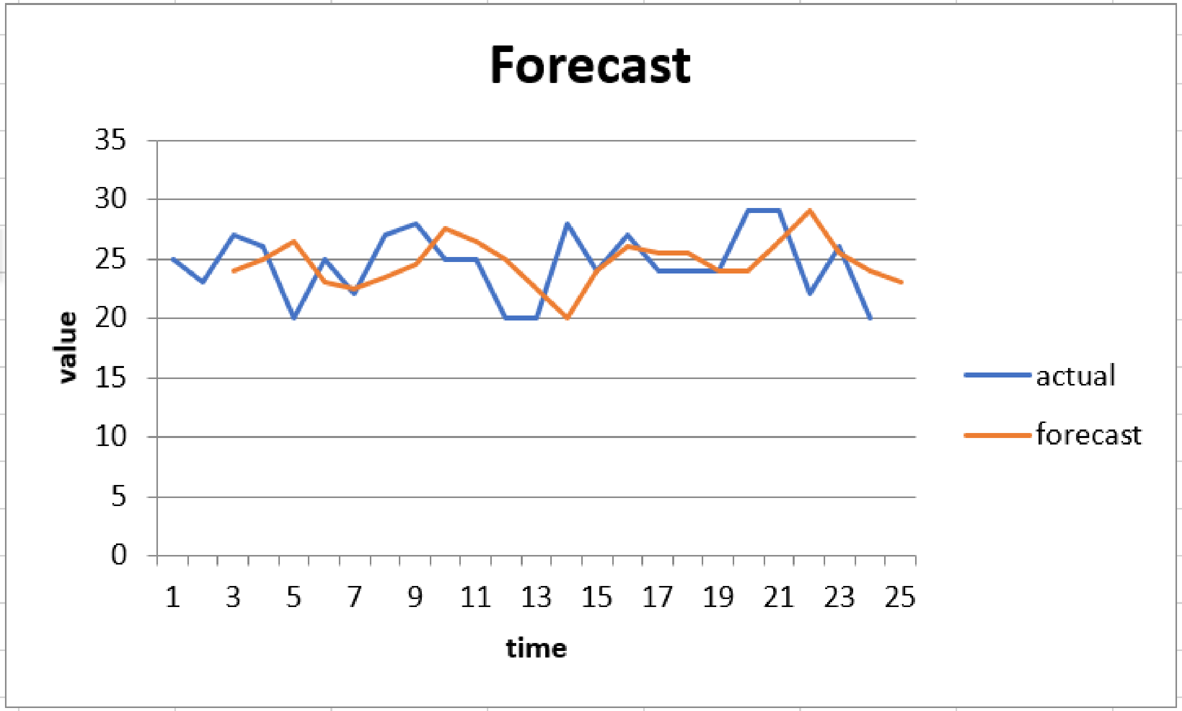
Forecasting one next period only

Print the results starting from the indicated cell…

Time series:

Simple Moving Average

Plastic Syringe (20-Pack)	*forecast*
25	
23	
27	=AVERAGE(E5:E6)
26	AVERAGE(**number1**, [number2], ...)
20	26.5
25	23
22	22.5
27	23.5
28	24.5
25	27.5
25	26.5
20	25
20	22.5



Time series:

The algorithm performance indicators…

MAE	*MSE*	*MAPE*
2.840909091	13.03409091	0.118669

Methods for non-stationary time series:

- ❑ **Moving Average (MA)**

❑ **Errors…**

$$e_t = A_t - F_t$$

Optimum

Mean Absolute Error(MAE) $$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t|$$ ≈ 0

Mean Squared Error(MSE) $$MSE = \frac{1}{n}\sum_{t=1}^{n}(A_t - F_t)^2$$ ≈ 0

Mean Absolute Percentage Error(MAPE) $$MSE = \frac{100}{n}\sum_{i=1}^{n}\frac{|A_t - F_t|}{A_t}$$ ≈ 0

Simple Moving Average

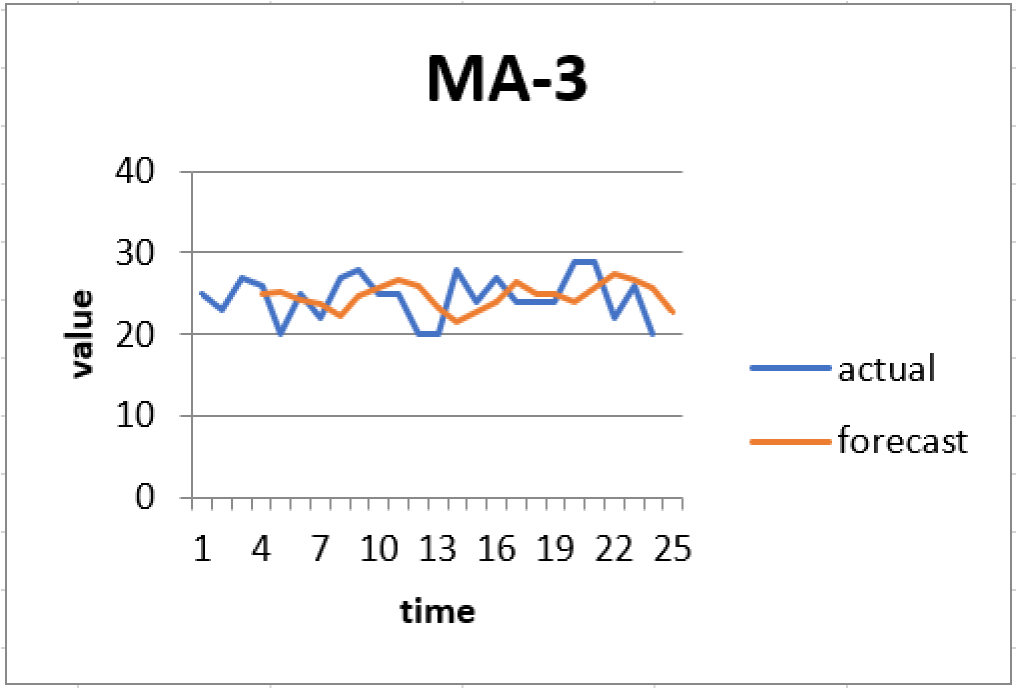
Plastic Syringe (20-Pack)	*forecast*
25	
23	
27	24
26	25
20	26.5
25	23
22	22.5
27	23.5
28	24.5
25	27.5
25	26.5
20	25
20	22.5
28	20
24	24
27	26
24	25.5

Error = 27 – 24 = 3

Time series:

Methods for non-stationary time series:

- **Moving Average (MA)…3 moving periods…**



MAE	MSE	MAPE
3.015873016	12.92063492	0.127219

MA-3 performance is worst compared to MA-2…

Learning: One can make more accurate predictions using MA-2

Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ **Weighted Moving Average (WMA)**
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ Holt Winter´s Additive (HWA)
- ❑ Holt Winter´s Multiplicative (HWM)

- The most recent values in a time series are given more weight in computing a forecast
- The choice of weights, *w*, is somewhat arbitrary and involves some trial and error

$$F_t = w_t(A_t) + w_{t-1}(A_{t-1}) + \ldots + w_{t-n}(A_{t-n})$$

where

w_t = weight for period t, w_{t-1} = weight for period $t-1$, etc.

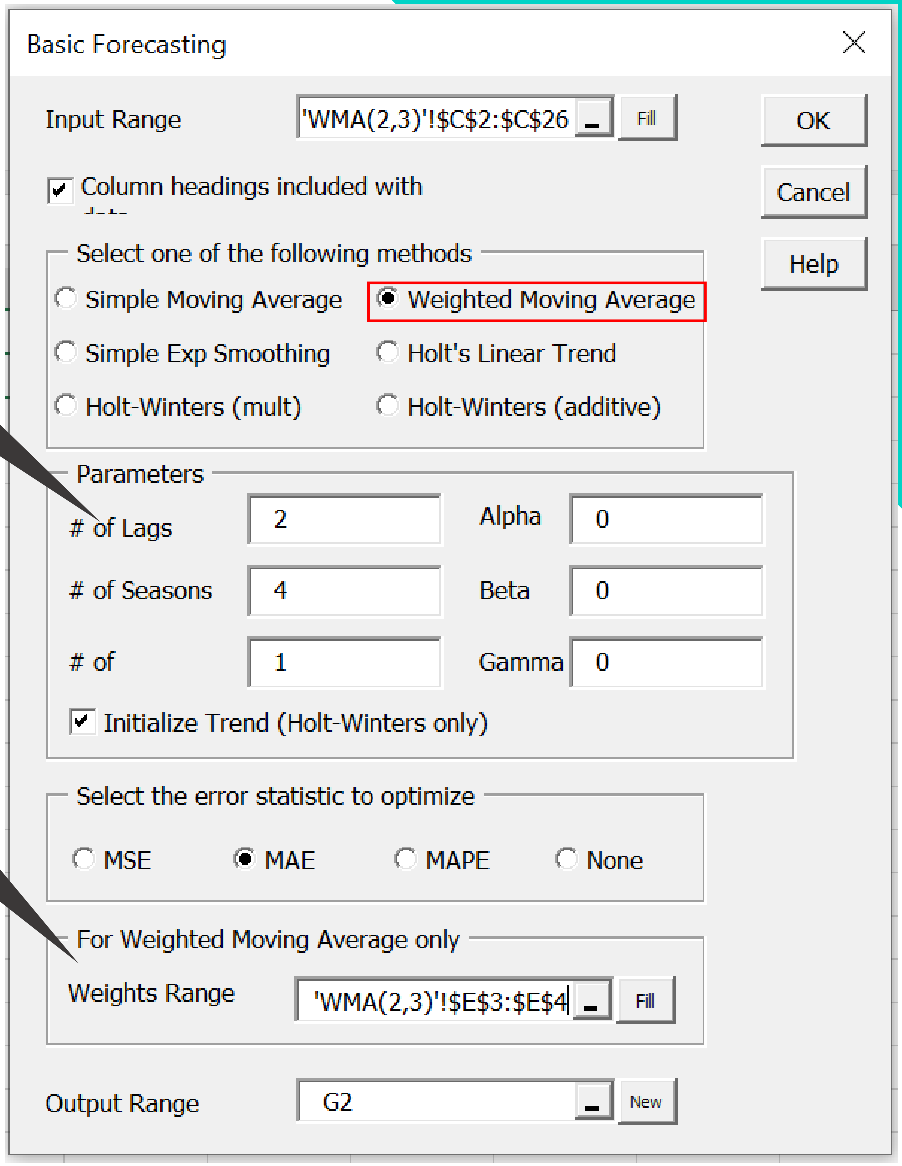
A_t = the actual value for period t, A_{t-1} = the actual value for period $t-1$, etc.

Time series:

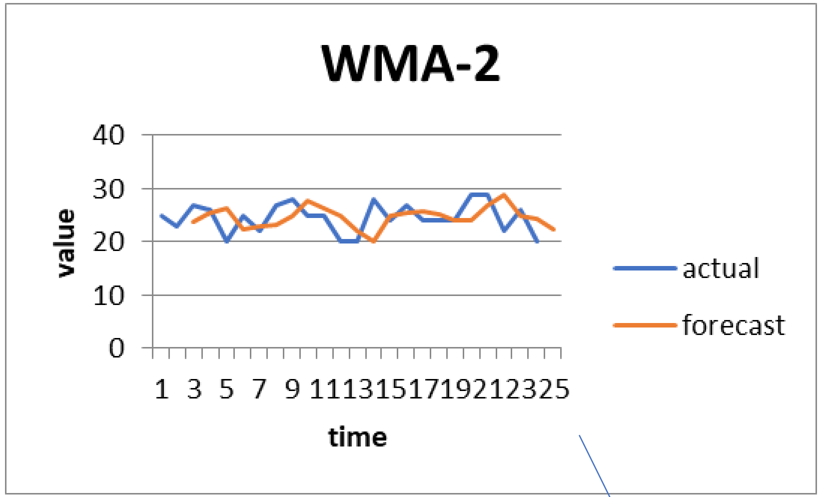
Date (month)	Plastic Syringe (20-Pack)	WMA(2)
Jan-17	25	0.4
Feb-17	23	0.6
Mar-17	27	
Apr-17	26	
May-17	20	
Jun-17	25	WMA(3)
Jul-17	22	0.2
Aug-17	27	0.3
Sep-17	28	0.5
Oct-17	25	
Nov-17	25	
Dec-17	20	
Jan-18	20	
Feb-18	28	

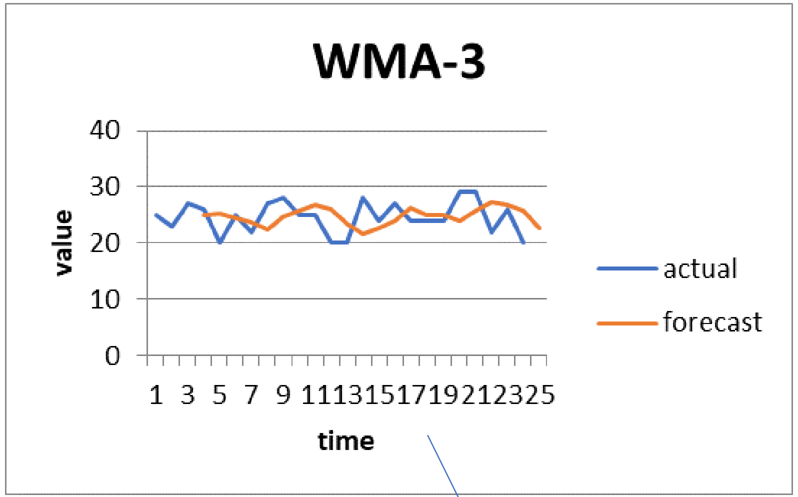
Two(2) moving periods

The weights should be defined in a column…



Time series:





MAE	*MSE*	*MAPE*
2.918181818	13.18363636	0.121934

MAE	*MSE*	*MAPE*
3.015873	12.92063	0.127219

Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ Holt Winter´s Additive (HWA)
- ❑ Holt Winter´s Multiplicative (HWM)

- A weighted averaging method that is based on the previous forecast plus a percentage of the forecast error

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

where

F_t = Forecast for period t

F_{t-1} = Forecast for the previous period

α = Smoothing constant

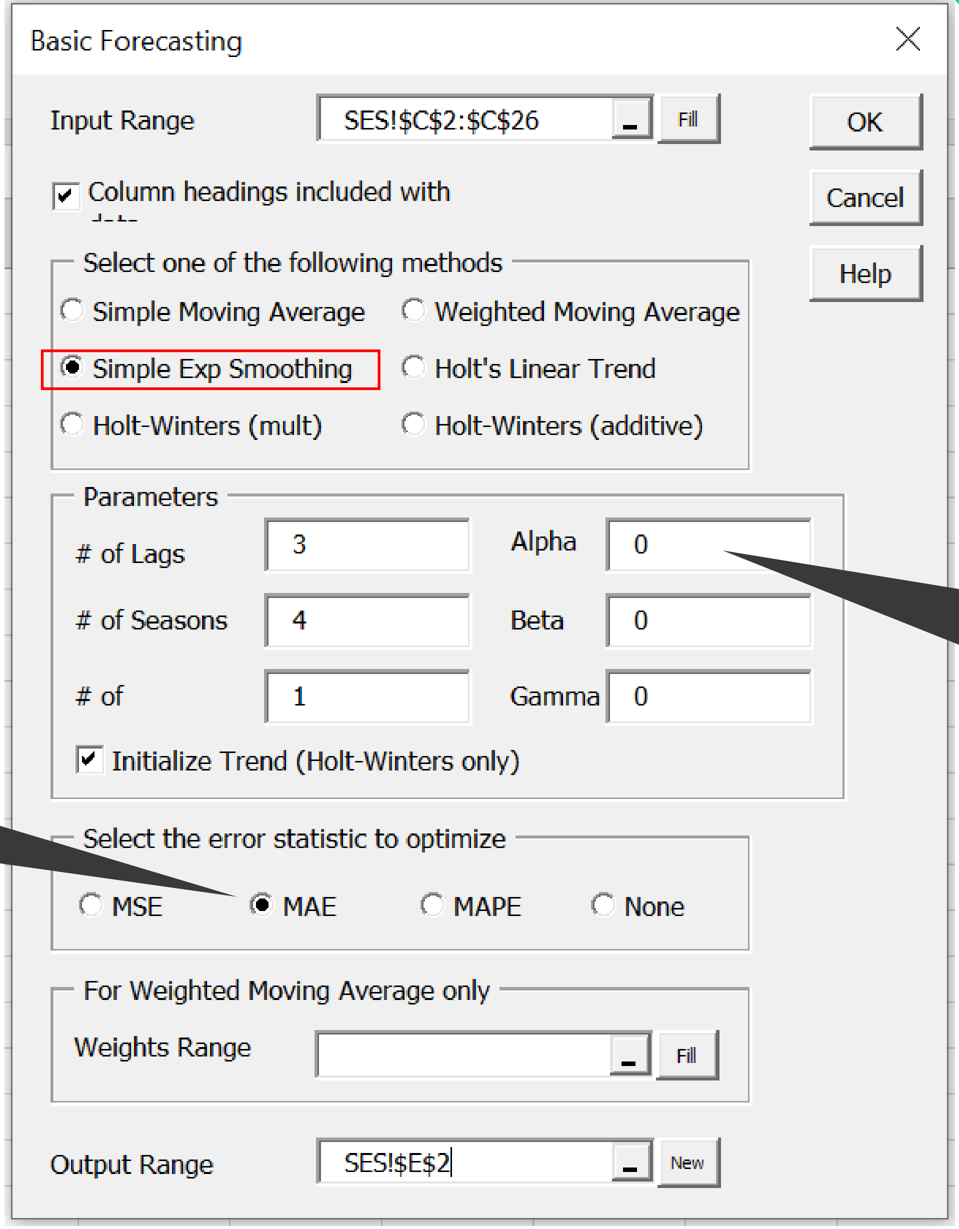
A_{t-1} = Actual demand or sales from the previous period

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

Time series:

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

How do we optimize this user-defined parameter in order to reach the best possible MAE?



Let alpha as “zero”…it will be optimize

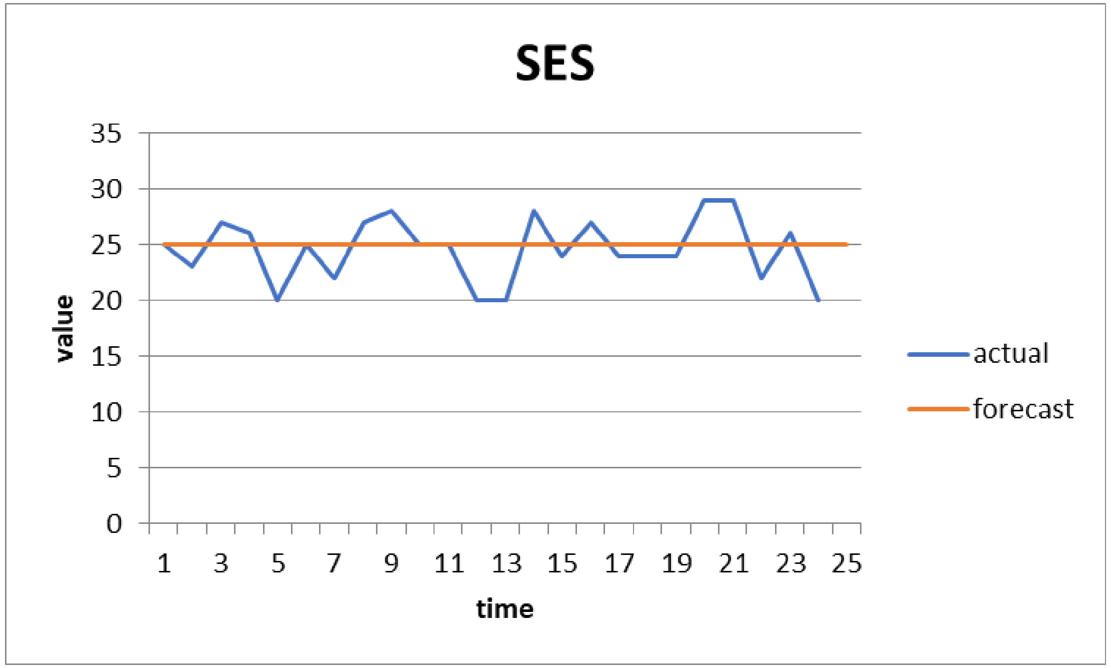
Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ **Simple Exponential Smoothing (SES)**
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ Holt Winter´s Additive (HWA)
- ❑ Holt Winter´s Multiplicative (HWM)

MAE	*MSE*	*MAPE*
2.347826087	8.260869565	0.10068

Best MAE so far...!



Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ **Double Exponential Smoothing (DES) (Holt´s Method)**
- ❑ Holt Winter´s Additive (HWA)
- ❑ Holt Winter´s Multiplicative (HWM)

Double exponential Smoothing (DES) Algorithm (also known as Holt's Linear Method) is an extension to the SES algorithm originally designed for time series with no trend nor seasonal patterns. It includes a term to model linear trends. Holt's method allows the estimates of level (L_t) and slope (b_t) to be adjusted with each new observation.

Real Stat assumes:

b1 = 0

F1 = empty

Better results...

Init: $L_1 = y_1 \quad b_1 = y_2 - y_1 \quad F_1 = y_1$ and choose $0 \le \alpha \le 1$ and $0 \le \beta \le 1$

Compute and Forecast:

$$L_t = \alpha\ y_t + (1-\alpha)\ (L_{t-1} + b_{t-1})$$
$$b_t = \beta\ (L_t - L_{t-1}) + (1-\beta)\ b_{t-1}$$
$$F_{t+1} = L_t + b_t$$

Until no more observation are available then

$$F_{n+k} = L_n + k\ b_n,\ \forall k \ge 1$$

Double Exponential Smoothing (Holt's Linear Model) Algorithm.

Note that no forecasts or fitted values can be computed until y_1 and y_2 have been observed. Also by convention, we let $F_1 = y_1$.

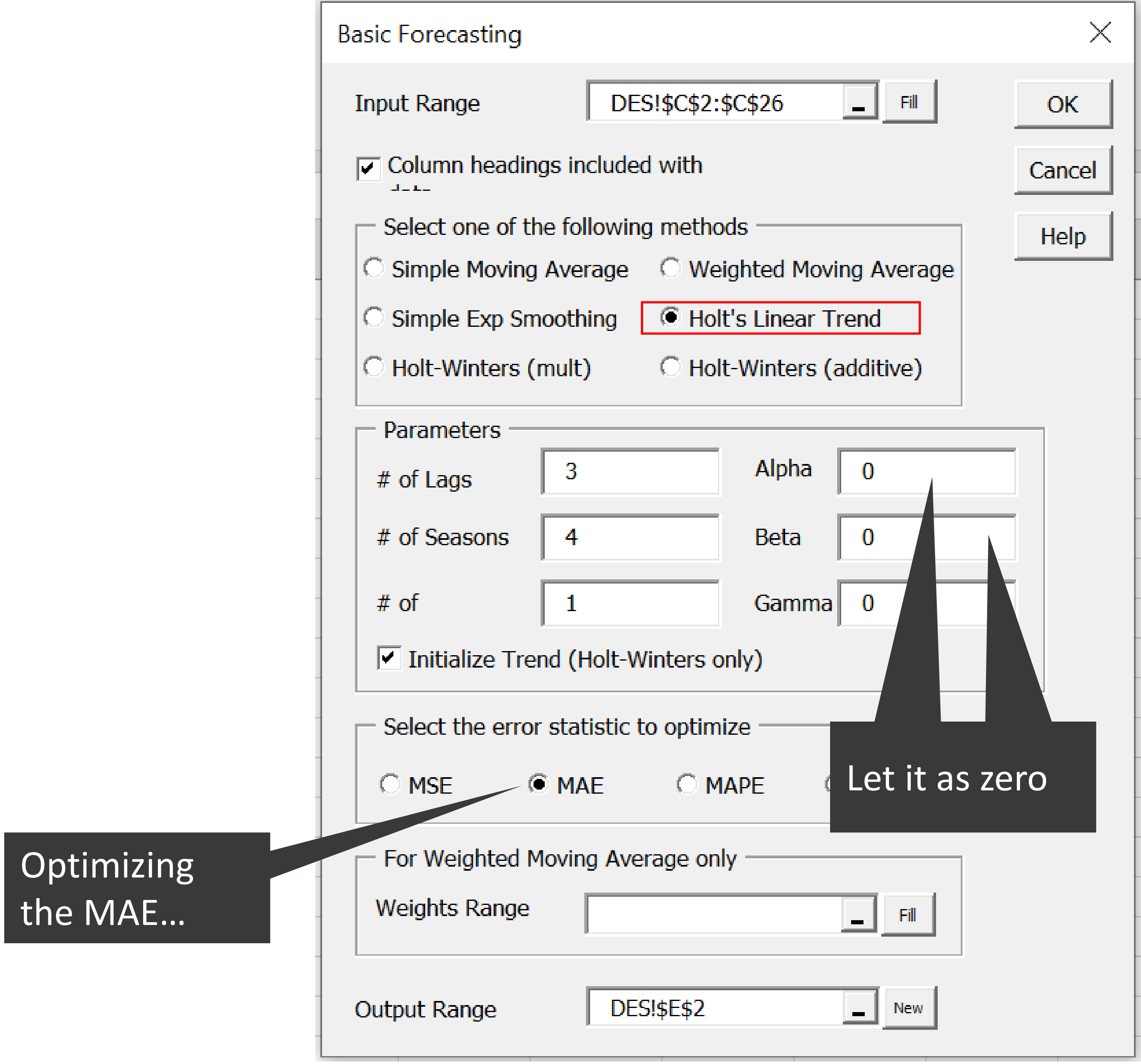
Two(2) user-defined parameters:

Alpha and Beta

Time series:

Methods for non-stationary time series:

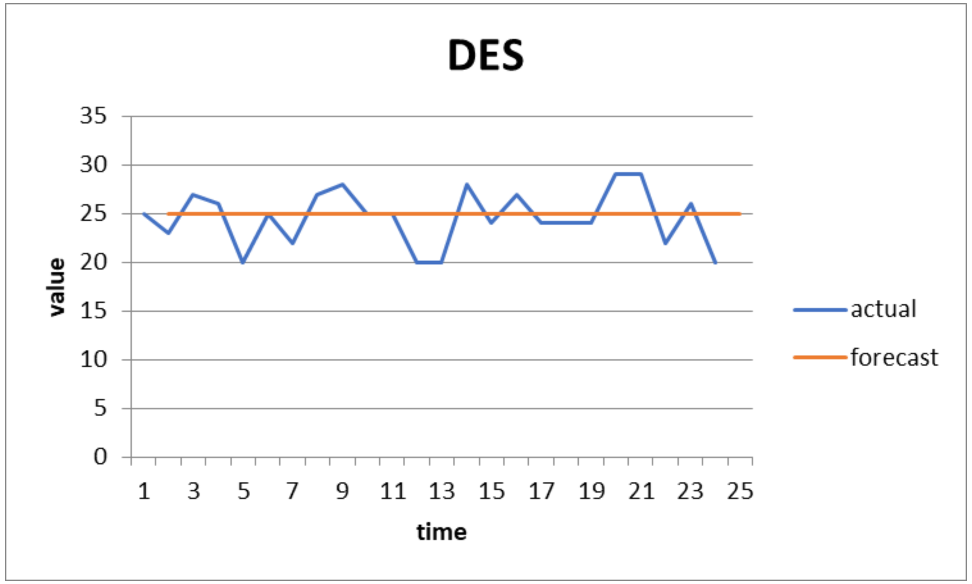
- Moving Average (MA)
- Weighted Moving Average (WMA)
- Simple Exponential Smoothing (SES)
- **Double Exponential Smoothing (DES) (Holt´s Method)**
- Holt Winter´s Additive (HWA)
- Holt Winter´s Multiplicative (HWM)



Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ **Double Exponential Smoothing (DES) (Holt´s Method)**
- ❑ Holt Winter´s Additive (HWA)
- ❑ Holt Winter´s Multiplicative (HWM)



alpha	*beta*	*MAE*	*MSE*	*MAPE*
0	0	2.347826	8.26087	0.10068

Very similar to the SES…due to there is no trend (+ or -) within the dataset…DES would performs better when data exhibits certain trend…

Time series:

Three parameters…alpha, beta and gamma

Seasonality is new

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ **Holt Winter´s Additive (HWA)**
- ❑ Holt Winter´s Multiplicative (HWM)

Init:

$$L_s = \frac{1}{s}\sum_{i=1}^{s} y_i$$

$$b_s = \frac{1}{s}\left[\frac{y_{s+1}-y_1}{s} + \frac{y_{s+2}-y_2}{s} + \cdots + \frac{y_{2s}-y_s}{s}\right]$$

$$S_i = y_i - L_s,\ i = 1,\ldots,s$$

and choose $0 \le \alpha \le 1$ and $0 \le \beta \le 1$ and $0 \le \gamma \le 1$

Compute for $t > s$:

level $L_t = \alpha\,(y_t - S_{t-s}) + (1-\alpha)\,(L_{t-1} + b_{t-1})$

trend $b_t = \beta\,(L_t - L_{t-1}) + (1-\beta)\,b_{t-1},$

seasonal $S_t = \gamma\,(y_t - L_t) + (1-\gamma)\,S_{t-s}$

forecast $F_{t+1} = L_t + b_t + S_{t+1-s}$

Until no more observationare available

and subsequent forecasts:

$F_{n+k} = L_n + k\,b_n + S_{n+k-s}$

Seasonal Holt Winter's Additive Model Algorithm (noted SHW+).

s is the length of the seasonal cycle. We have to pick the values of α, β and γ. As with the other methods (i.e. SES and DES), we can use the SSE/RMSE or MAPE to choose the best values.

Seasonality: it is traditionally defined by the number of observations within a year…

Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ **Holt Winter´s Additive (HWA)**
- ❑ Holt Winter´s Multiplicative (HWM)

Date (month)	Plastic Syringe (20-Pack)
Jan-17	25
Feb-17	23
Mar-17	27
Apr-17	26
May-17	20
Jun-17	25
Jul-17	22
Aug-17	27
Sep-17	28
Oct-17	25
Nov-17	25
Dec-17	20
Jan-18	20
Feb-18	28
Mar-18	24
Apr-18	27
May-18	24
Jun-18	24
Jul-18	24
Aug-18	29
Sep-18	29
Oct-18	22
Nov-18	26
Dec-18	20

dataset

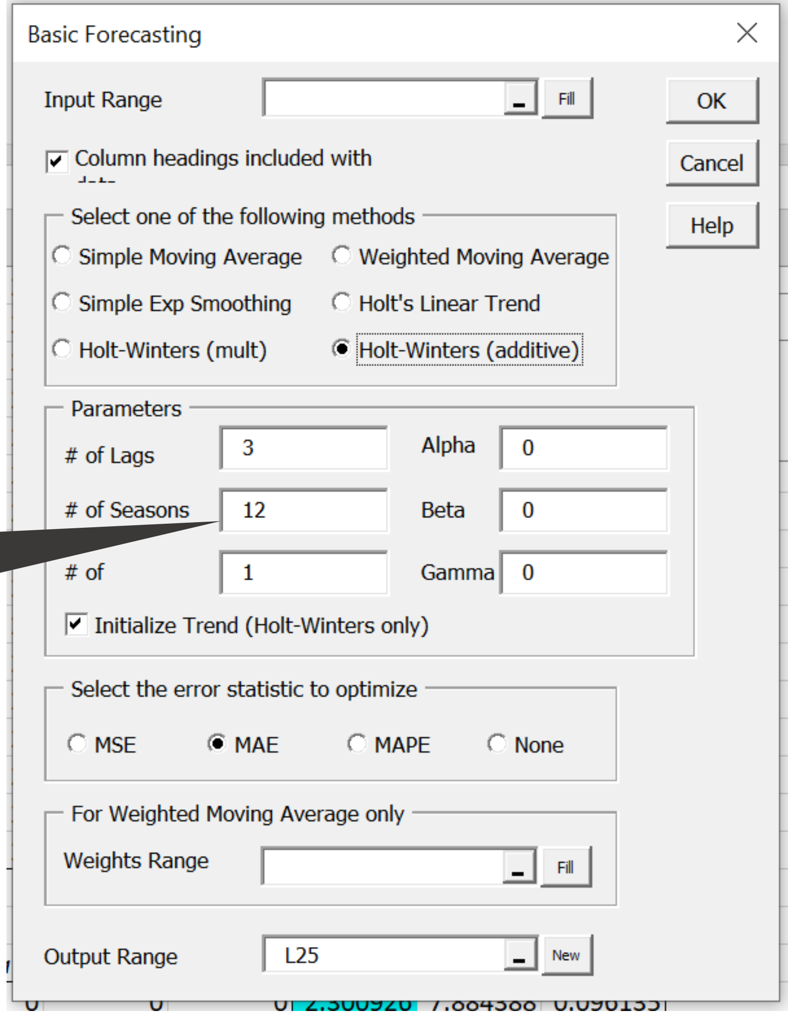
There are 12 observations every year

Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ **Holt Winter´s Additive (HWA)**
- ❑ Holt Winter´s Multiplicative (HWM)

12 observations during each revised year…

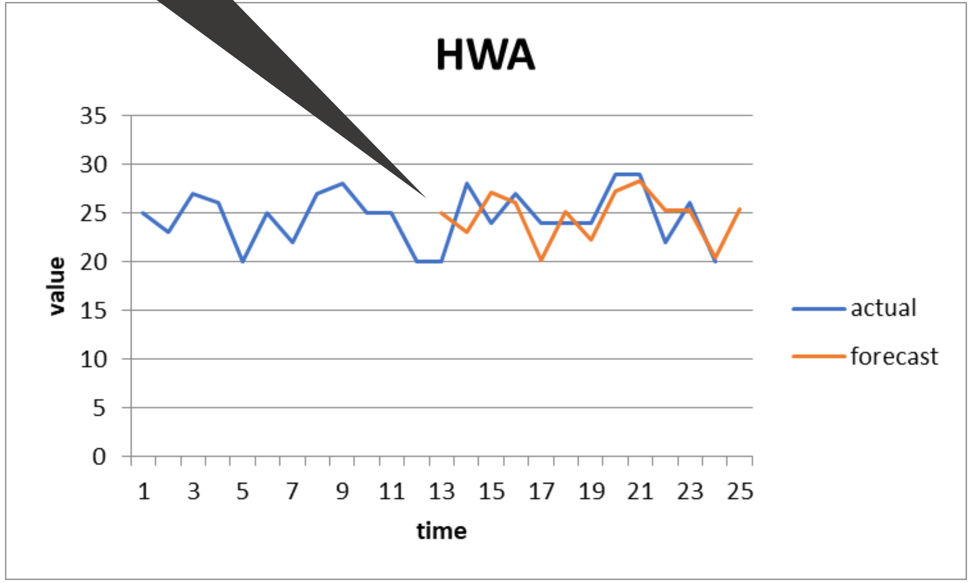


Time series:

Forecasting start right after the first "s" cycle (13th observation)

Methods for non-stationary time series:

- Moving Average (MA)
- Weighted Moving Average (WMA)
- Simple Exponential Smoothing (SES)
- Double Exponential Smoothing (DES) (Holt´s Method)
- **Holt Winter´s Additive (HWA)**
- Holt Winter´s Multiplicative (HWM)



alpha	*beta*	*gamma*	*MAE*	*MSE*	*MAPE*
0	0	0	2.300926	7.884388	0.096135
	α+γ	0			

Better that SES and DES

Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ Holt Winter´s Additive (HWA)
- ❑ **Holt Winter´s Multiplicative (HWM)**

The forecasting equation is formed by the multiplication of the series components…

Init:

$$L_s = \frac{1}{s}\sum_{i=1}^{s} y_i$$

$$b_s = \frac{1}{s}\left[\frac{y_{s+1}-y_1}{s} + \frac{y_{s+2}-y_2}{s} + \cdots + \frac{y_{2s}-y_s}{s}\right]$$

$$S_i = \frac{y_i}{L_s},\ i = 1,\ldots,s$$

and choose $0 \le \alpha \le 1$ and $0 \le \beta \le 1$ and $0 \le \gamma \le 1$

Compute for $t > s$:

level	$L_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1})$
trend	$b_t = \beta (L_t - L_{t-1}) + (1-\beta) b_{t-1},$
seasonal	$S_t = \gamma \frac{y_t}{L_t} + (1-\gamma) S_{t-s}$
forecast	$F_{t+1} = (L_t + b_t) S_{t+1-s}$

Until no more observation are available and subsequent forecasts:

$$F_{n+k} = (L_n + k \cdot b_n) S_{n+k-s}$$

Seasonal Holt Winter's Multiplicative Model Algorithm (noted SHW×).

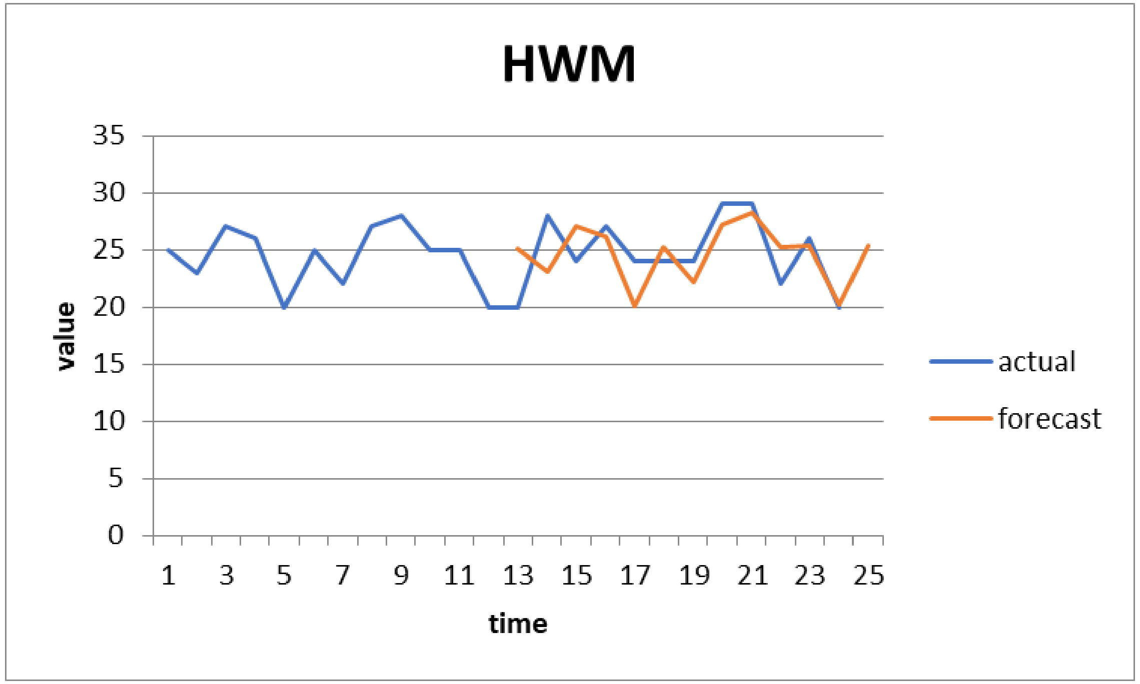
Time series:

Methods for non-stationary time series:

- ❑ Moving Average (MA)
- ❑ Weighted Moving Average (WMA)
- ❑ Simple Exponential Smoothing (SES)
- ❑ Double Exponential Smoothing (DES) (Holt´s Method)
- ❑ Holt Winter´s Additive (HWA)
- ❑ **Holt Winter´s Multiplicative (HWM)**

alpha	beta	gamma	MAE	MSE	MAPE
0	0	0	2.295317	7.902241	0.095901
	α+γ	0			

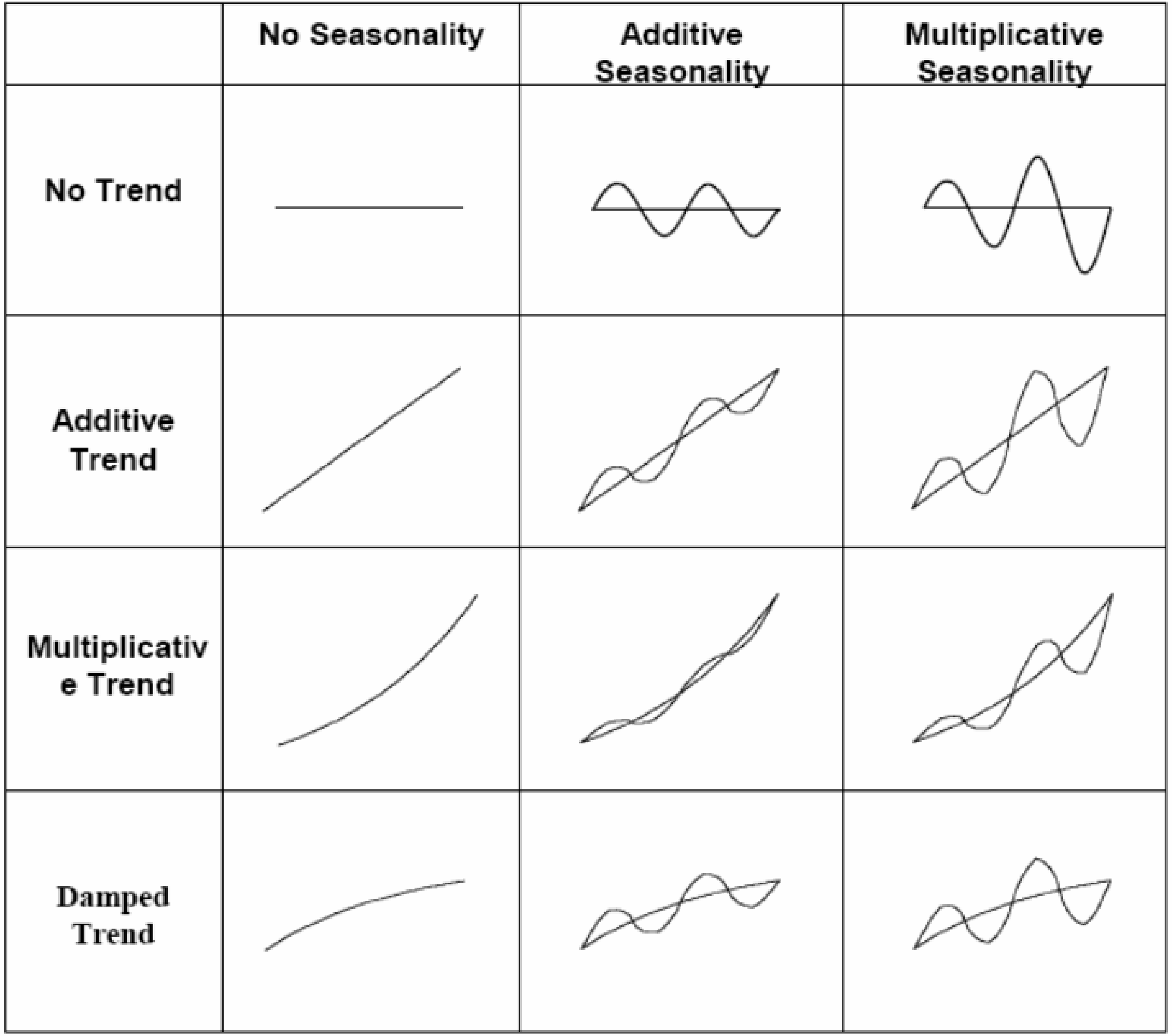
Best performance reached



Time series:

Methods for non-stationary time series:

- ❑ **Moving Average (MA)**
- ❑ **Weighted Moving Average (WMA)**
- ❑ **Simple Exponential Smoothing (SES)**
- ❑ **Double Exponential Smoothing (DES) (Holt´s Method)**
- ❑ **Holt Winter´s Additive (HWA)**
- ❑ **Holt Winter´s Multiplicative (HWM)**



HELP

Thanks...

Erasmus+ KA2 Strategic Partnership
2017-1-1FI01-KA203-034721
Healthcare Logistics Education and Learning Pahtway