Kompetensi dasar 2

Regression & Correlation Analysis with Minitab

Analisis regresi sederhana

- Contoh data

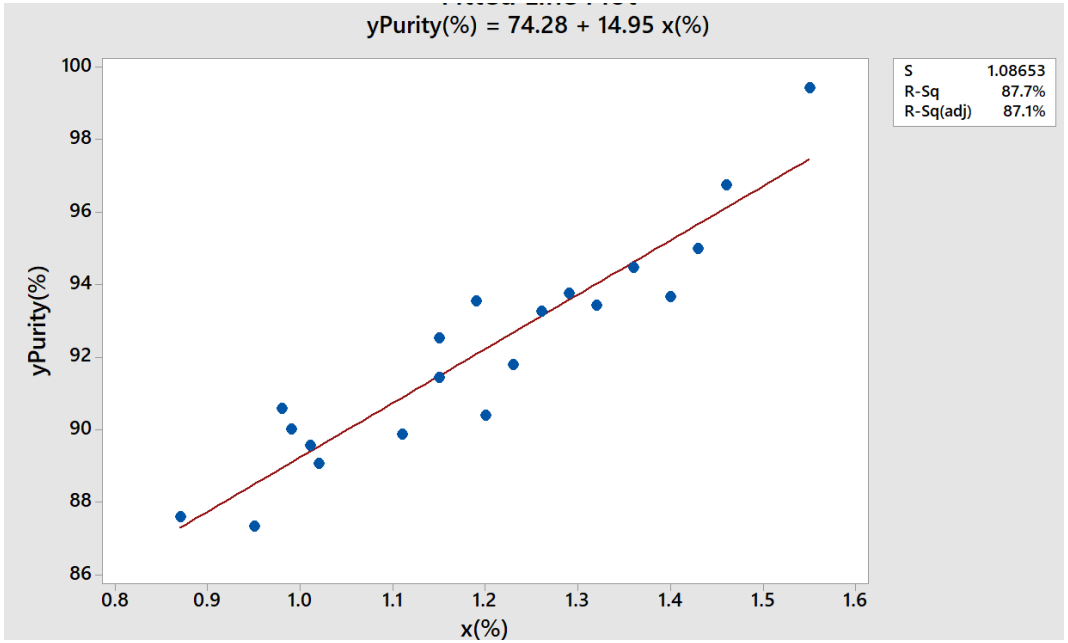
X: tingkat hidrokarbon (%)

Y: kejernihan air (%)

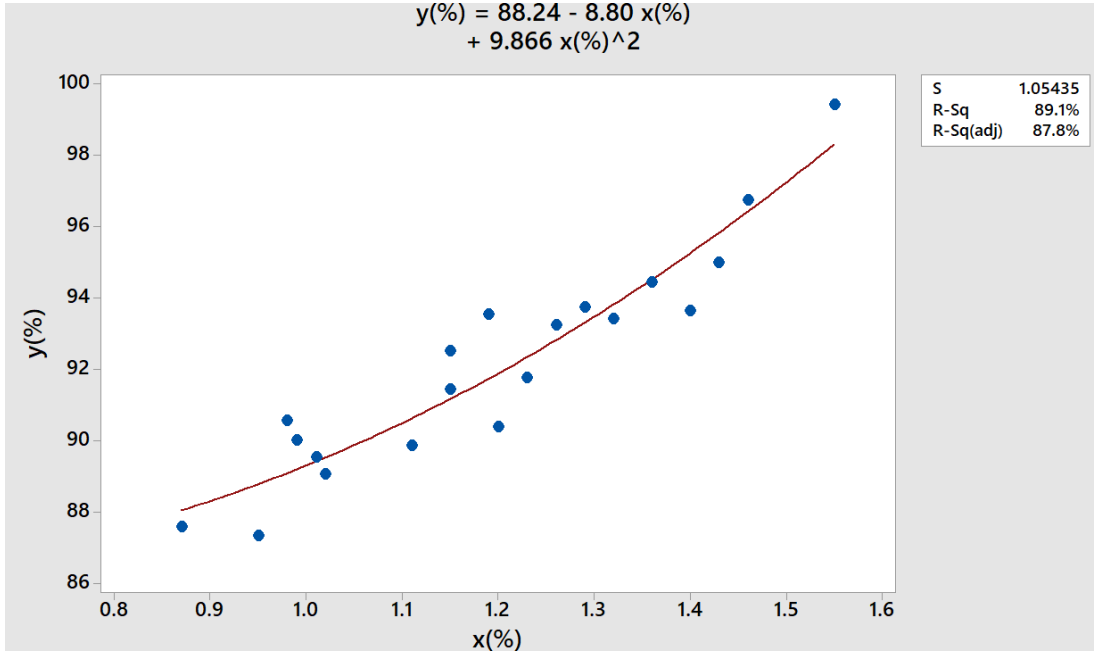
	x(%)	y(%)
1	0.99	90.01
2	1.02	89.05
3	1.15	91.43
4	1.29	93.74
5	1.46	96.73
6	1.36	94.45
7	0.87	87.59
8	1.23	91.77
9	1.55	99.42
10	1.40	93.65
11	1.19	93.54
12	1.15	92.52
13	0.98	90.56
14	1.01	89.54
15	1.11	89.85
16	1.20	90.39
17	1.26	93.25
18	1.32	93.41
19	1.43	94.98
20	0.95	87.33

Cek data

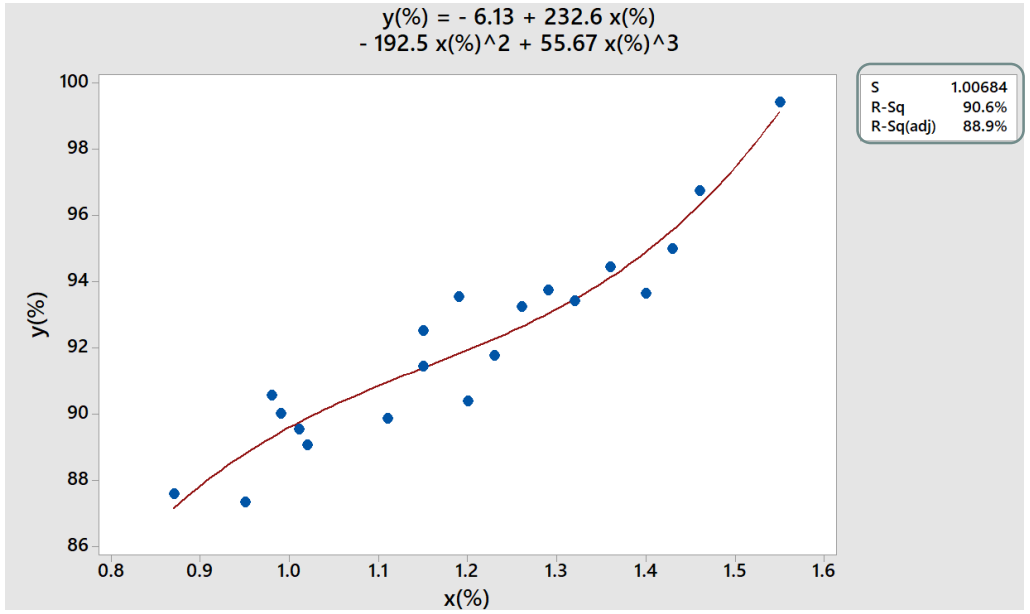
- Linier



→ Kuadratik



→ Kubik



R-sq terbesar

Regression Analysis: y(%) versus x(%) : Linier

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	152.127	152.127	128.86	0.000
x(%)	1	152.127	152.127	128.86	0.000
Error	18	21.250	1.181		
Lack-of-Fit	17	20.656	1.215	2.05	0.506
Pure Error	1	0.594	0.594		
Total	19	173.377			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
1.08653	87.74%	87.06%	84.13%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	74.28	1.59	46.62	0.000	
x(%)	14.95	1.32	11.35	0.000	1.00

Regression Analysis: y(%) versus x(%): Linier

```
Regression Equation
y(%) = 74.28 + 14.95 x(%)


Fits and Diagnostics for Unusual Observations
                                Std
Obs    y(%)     Fit  Resid  Resid
  9  99.420  97.452  1.968   2.07  R
R  Large residual
```

Polynomial Regression Analysis: y(%) versus x(%) : Kuadratik

The regression equation is
y(%) = 88.24 - 8.80 x(%) + 9.866 x(%)^2

Model Summary

S	R-sq	R-sq(adj)
1.05435	89.10%	87.82%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	154.479	77.2394	69.48	0.000
Error	17	18.898	1.1116		
Total	19	173.377			

Sequential Analysis of Variance

Source	DF	SS	F	P

Polynomial Regression Analysis: y(%) versus x(%) : Kuadratik

Linear	1	152.127	128.86	0.000
Quadratic	1	2.352	2.12	0.164

Polynomial Regression Analysis: y(%) versus x(%) : Kubik

The regression equation is
y(%) = - 6.13 + 232.6 x(%) - 192.5 x(%)^2 + 55.67 x(%)^3

Model Summary

S	R-sq	R-sq(adj)
1.00684	90.64%	88.89%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	157.157	52.3858	51.68	0.000
Error	16	16.220	1.0137		
Total	19	173.377			

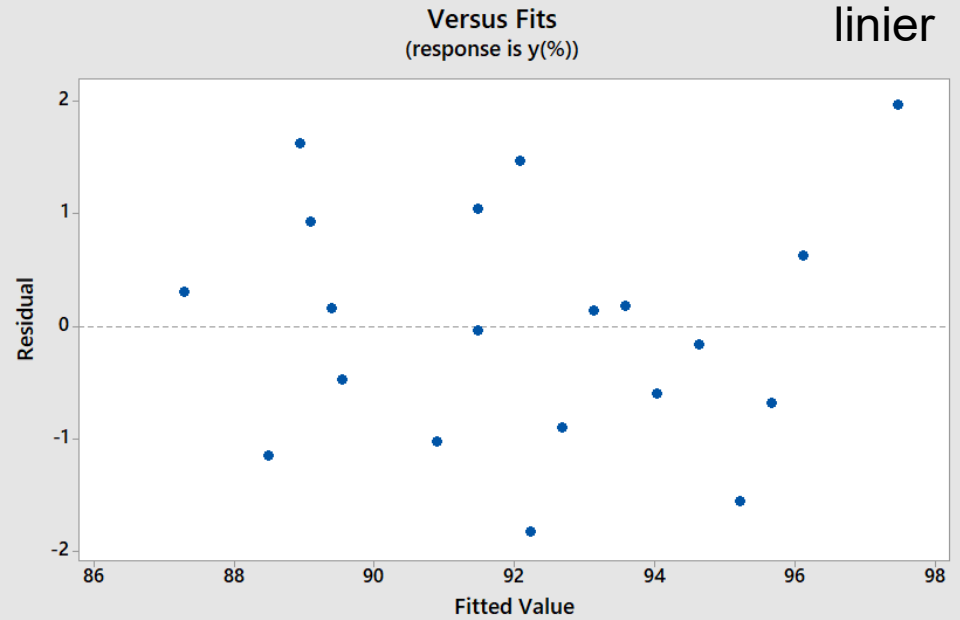
Polynomial Regression Analysis: y(%) versus x(%): Kubik

Sequential Analysis of Variance

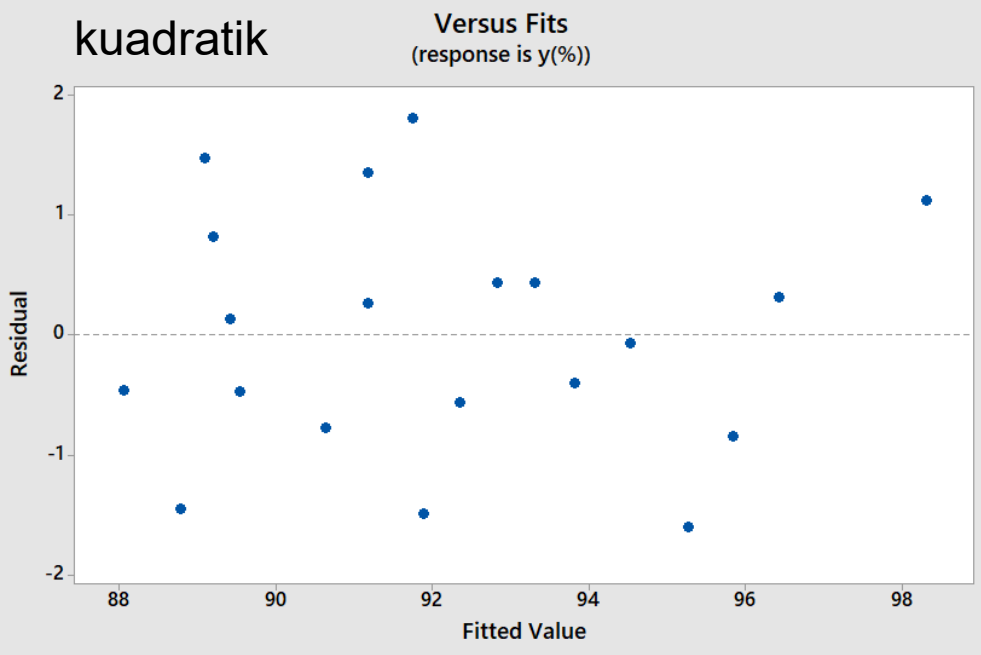
Source	DF	SS	F	P
Linear	1	152.127	128.86	0.000
Quadratic	1	2.352	2.12	0.164
Cubic	1	2.678	2.64	0.124

Plot residual vs Fitted Value

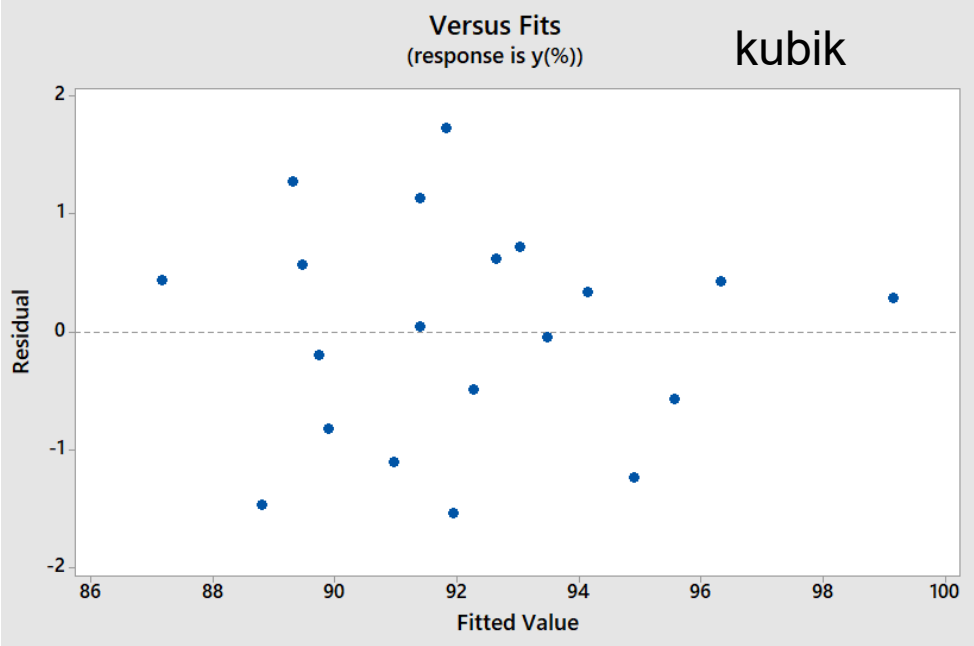
linier



kuadratik



kubik

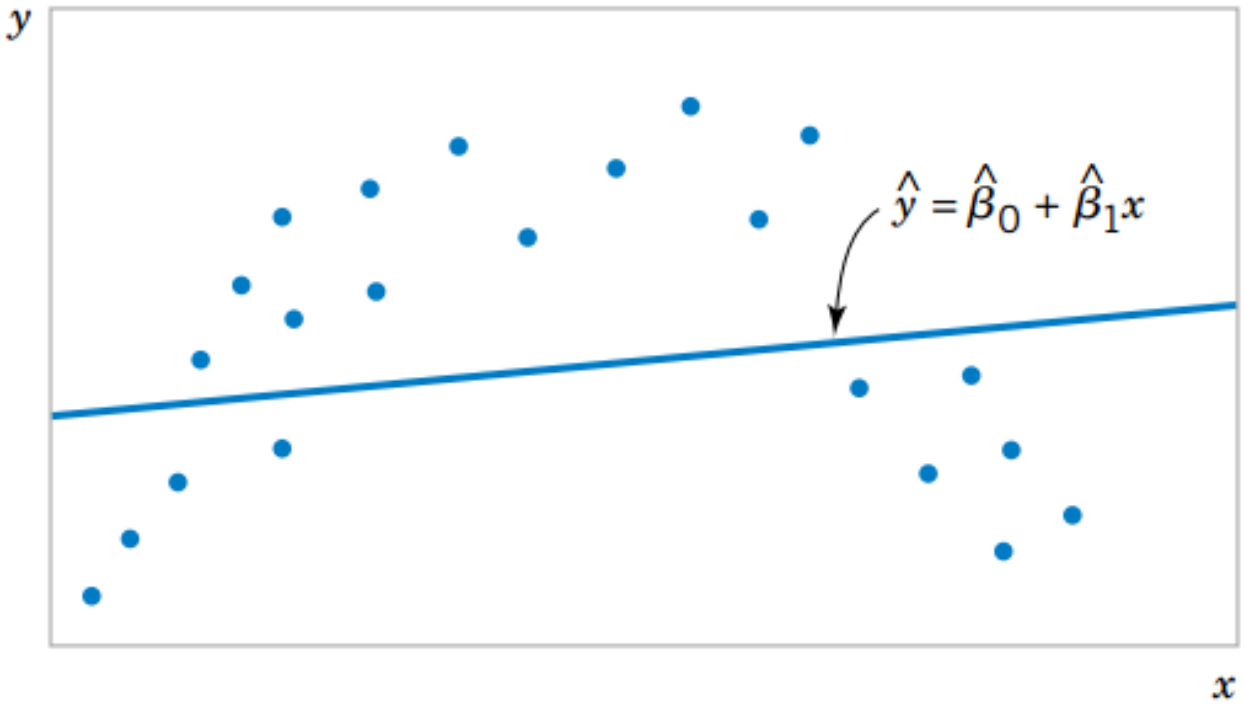


Correlation: x(%), y(%)

Correlations	
Pearson correlation	0.937
P-value	0.000

Lack of Fits

- The danger of using a regression model that is a poor approximation of the true functional relationship



H_0: The simple linear regression model is correct.

H_1: The simple linear regression model is not correct.

Contoh lain

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	3.493	3.4928	6.66	0.022
x	1	3.493	3.4928	6.66	0.022
Error	14	7.337	0.5241		
Lack-of-Fit	7	4.301	0.6144	1.42	0.329
Pure Error	7	3.037	0.4338		
Total	15	10.830			

- Since the lack-of-fit which has a *P*-value of $P = 0.329$, we cannot reject the hypothesis that the tentative model adequately describes the data.
- In addition, since the *P*-value for the statistic associated with significance of regression is $P = 0.022$, we conclude that $\beta 1 \neq 0$

	x	y
1	1.0	2.3
2	1.0	1.8
3	2.0	2.8
4	3.3	1.8
5	3.3	3.7
6	4.0	2.6
7	4.0	2.6
8	4.0	2.2
9	5.0	2.0
10	5.6	3.5
11	5.6	2.8
12	5.6	2.1
13	6.0	3.4
14	6.0	3.2
15	6.5	3.4
16	6.9	5.0

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.723938	32.25%	27.41%	10.90%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1.697	0.473	3.59	0.003	
x	0.259	0.100	2.58	0.022	1.00

Regression Equation

y = 1.697 + 0.259 x