Excel for Statistics

Chapter 12

Linear Regression

Lecturer: Dimaz Ramananda, S.E., M.Ak.

- Linear regression is a type of data analysis that considers the linear relationship between a dependent variable and one or more independent variables.
- Simple linear regression has only one x and one y variable.
- Multiple linear regression has one y and two or more x variables.
- Example, Simple Regression (1)

employee	machine	number_sold
3	1	65
4	2	78
2	1	56
4	1	70
2	0	51
4	2	80
5	1	87
3	1	68
2	2	70
3	2	86
4	0	74
5	1	86
3	2	81
2	2	68
5	1	89
3	2	87
2	1	64
2	1	56
3	2	83
4	1	84

- Answer, Simple Regression (1) – employee & number_sold

Regression Statistics								
Multiple R	0,73352029							
R Square	0,53805202							
Adjusted R Square	0,51238825							
Standard Error	8,16002154							
Observations	20							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	1396,002874	1396,00287	20,96542653	0,000232856			
Residual	18	1198,547126	66,5859515					
Total	19	2594,55						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	48,1126437	5,972066527	8,0562806	2,21754E-07	35,56579749	60,6594899	35,5657975	60,6594899
employee	8,01149425	1,749692275	4,57880187	0,000232856	4,335527189	11,6874613	4,33552719	11,6874613

- Answer, Simple Regression (1) – machine & number_sold

Regression Statistics								
Multiple R	0,11231938							
R Square	0,01261564							
Adjusted R Square	-0,042239							
Standard Error	1,09228633							
Observations	20							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	0,274390244	0,27439024	0,229982964	0,637311612			
Residual	18	21,47560976	1,19308943					
Total	19	21,75						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	3,48780488	0,552763325	6,30976174	6,01352E-06	2,326492226	4,64911753	2,32649223	4,64911753
machine	-0,1829268	0,381442933	-0,4795654	0,637311612	-0,984308694	0,61845504	-0,9843087	0,61845504

- Example, Simple Regression (2)

Quantity	Price	Cost
6500	$ 3	$ 1.200
5200	$ 4	$ 800
7600	$ 2	$ 1.500
7200	$ 4	$ 1.400
6400	$ 5	$ 1.000
5400	$ 3	$ 900
5900	$ 2	$ 1.000
7100	$ 4	$ 1.400
6900	$ 2	$ 1.300
6200	$ 3	$ 1.200

- Answer, Simple Regression (2) – Cost & Quantity

Regression Statistics								
Multiple R	0,96124045							
R Square	0,923983202							
Adjusted R Square	0,914481103							
Standard Error	229,5204582							
Observations	10							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	5122562,874	5122562,87	97,2398977	9,42212E-06			
Residual	8	421437,1257	52679,6407					
Total	9	5544000						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	2698,802395	386,2724525	6,98678453	0,000114135	1808,056522	3589,54827	1808,05652	3589,54827
Cost	3,19760479	0,32426684	9,86102924	9,42212E-06	2,449844117	3,94536546	2,44984412	3,94536546

- Answer, Simple Regression (2) – Price & Quantity

Regression Statistics								
Multiple R	0,134331804							
R Square	0,018045034							
Adjusted R Square	-0,10469934							
Standard Error	824,9210821							
Observations	10							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	1	100041,6667	100041,667	0,147013126	0,711394562			
Residual	8	5443958,333	680494,792					
Total	9	5544000						
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	6766,666667	891,0166049	7,59432162	6,33861E-05	4711,978691	8821,35464	4711,97869	8821,35464
Price	-102,083333	266,2421344	-0,3834229	0,711394562	-716,0387962	511,87213	-716,0388	511,87213

- Example, Multiple Regression (1)

employee	machine	number_sold
3	1	65
4	2	78
2	1	56
4	1	70
2	0	51
4	2	80
5	1	87
3	1	68
2	2	70
3	2	86
4	0	74
5	1	86
3	2	81
2	2	68
5	1	89
3	2	87
2	1	64
2	1	56
3	2	83
4	1	84

- Answer, Multiple Regression (1)

Regression Statistics	
Multiple R	0,898141913
R Square	0,806658895
Adjusted R Square	0,783912883
Standard Error	5,432111295
Observations	20

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	2092,91684	1046,45842	35,46375006	8,5853E-07
Residual	17	501,633163	29,5078331		
Total	19	2594,55			

	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	33,9721749	4,92661053	6,89564857	2,59149E-06	23,57793525	44,3664146	23,5779353	44,3664146
employee	8,651334469	1,17218432	7,38052397	1,07418E-06	6,178241731	11,1244272	6,17824173	11,1244272
machine	9,277683135	1,90905575	4,85982828	0,000147075	5,249927568	13,3054387	5,24992757	13,3054387

- Example, Multiple Regression (2)

Quantity	Price	Cost
6500	$ 3	$ 1.200
5200	$ 4	$ 800
7600	$ 2	$ 1.500
7200	$ 4	$ 1.400
6400	$ 5	$ 1.000
5400	$ 3	$ 900
5900	$ 2	$ 1.000
7100	$ 4	$ 1.400
6900	$ 2	$ 1.300
6200	$ 3	$ 1.200

- Answer, Multiple Regression (2)

Regression Statistics								
Multiple R	0,96702409							
R Square	0,9351356							
Adjusted R Square	0,91660291							
Standard Error	226,655255							
Observations	10							
ANOVA								
	df	*SS*	*MS*	*F*	*Significance F*			
Regression	2	5184391,77	2592195,88	50,45872	6,9506E-05			
Residual	7	359608,234	51372,6048					
Total	9	5544000						
	Coefficients	*tandard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95,0%*	*Upper 95,0%*
Intercept	2329,41567	508,798221	4,57827008	0,00254898	1126,299059	3532,53228	1126,29906	3532,53228
Price	82,8021248	75,4764636	1,09705888	0,30891492	-95,67135131	261,275601	-95,671351	261,275601
Cost	3,28685259	0,33039107	9,94837003	2,2137E-05	2,50560186	4,06810332	2,50560186	4,06810332