

Příklad: Jednoduchá/Vícenásobná lineární regrese:

	Y	X1	X2
Prodej	Roční tržby tis.	Poč. kolemjdoucí	Velikost prodejny
1	7800	12	90
2	10500	20	150
3	5700	11	100
4	12000	30	180
5	8100	15	120
6	9600	17	90
7	12900	27	200
8	6600	13	100
9	19500	55	320
10	15600	45	220
11	11400	29	170
12	9000	15	145
13	10800	24	170
14	9900	22	130
15	7200	11	120
16	10560	16	140
17	11280	18	150
18	11700	20	190
19	12300	23	190
20	10320	31	170
21	8040	16	130
22	8760	19	140
23	10920	21	170
24	11940	24	160
25	12360	29	170

Úkoly:

- Vyjádřete (lineární) závislost tržeb na 4 prediktorech (poč. kol
- Na základě modelu prognózuje tržby nové prodejny s těmito p

X3

X4

Prům. plat prodavačů	Přítomnost konkurence v místě
10.0	1
17.1	0
10.5	1
20.8	0
12.4	1
15.7	1
23.2	1
12.1	1
26.3	0
24.8	0
20.5	0
13.8	1
16.2	0
15.4	1
13.1	1
14.6	1
15.9	0
20.5	0
21.3	1
14.3	1
12.6	1
14.2	1
17.4	1
21.1	1
22.1	1

emjdoucích, velikost prodejny, průměrný plat prodavačů, pi

parametry: počet kolemjdoucích = 20, velikost prodejny = 20

Excel	Statistická terminologie
Násobné R	= R - koeficient korelace
Hodnota spolehlivosti R	= R^2 - koeficient determinace
Nastavená hodnota spolehlivosti R	= R^2_{adj} - upravený koeficient
Chyba stř. hodnoty	= s^2 - směrodatná chyba (oc

řítomnost konkurence)

00, průměrný plat prodavačů = 15, přítomnost konkurence = 1.

se		
t determinace		

odhad směrodatné odchyly náhodné složky)

Jednoduchá lineární regrese

Zjistěte závislost ročních tržeb na počtu kolemjdoucích (!?)

Zjistěte závislost ročních tržeb na velikosti prodejny (!?)

Zjistěte závislost velikosti prodejny na počtu kolemjdoucích (!?)

Zjistěte závislost průměrného platu prodavačů na velikosti tržeb (

x	y
Poč. kolemjdoucích/hod.	Roční tržby tis. Kč
12	7800
20	10500
11	5700
30	12000
15	8100
17	9600
27	12900
13	6600
55	19500
45	15600
29	11400
15	9000
24	10800
22	9900
11	7200
16	10560
18	11280
20	11700
23	12300
31	10320
16	8040
19	8760
21	10920
24	11940
29	12360

x_i	y_i
12	7800
20	10500
11	5700
30	12000
15	8100
17	9600
27	12900
13	6600
55	19500
45	15600
29	11400
15	9000
24	10800
22	9900
11	7200
16	10560
18	11280
20	11700
23	12300
31	10320
16	8040
19	8760
21	10920

součet
průměr

24	11940
29	12360

!?)

Velikost prodejny m2
90
150
100
180
120
90
200
100
320
220
170
145
170
130
120
140
150
190
190
170
130
140
170
160
170

Maticové násobení: SOUČIN.MATIC

3 sloupce 2 sloupce 2 sloupce

4 řádky 3 řádky

1	2	3	3	4	
4	5	6	2	5	
7	8	9	1	6	
10	11	12			

=

$A = (4 \times 3)$ "krát" $B = (3 \times 2)$ = (4×2)

AB
BA nelze

A=

10	2	3
4	5	6
8	10	12

A-1=



4 řádky

A.A-1 =



1	12	90	10.0	1
1	20	150	17.1	0
1	11	100	10.5	1
1	30	180	20.8	0
1	15	120	12.4	1
1	17	90	15.7	1
1	27	200	23.2	1
1	13	100	12.1	1
1	55	320	26.3	0
1	45	220	24.8	0
1	29	170	20.5	0
1	15	145	13.8	1
1	24	170	16.2	0
1	22	130	15.4	1
1	11	120	13.1	1
1	16	140	14.6	1
1	18	150	15.9	0
1	20	190	20.5	0
1	23	190	21.3	1
1	31	170	14.3	1
1	16	130	12.6	1
1	19	140	14.2	1
1	21	170	17.4	1
1	24	160	21.1	1
1	29	170	22.1	1

= F

$$F^T F =$$



$$(F^T F)^{-1} =$$



$$(F^T F)^{-1} \cdot F^T y =$$

- absolut. Člen
- Poč. kolemjdoucích/hod.
- Velikost prodejny m2
- Prům. plat prodavačů/mě
- Přítomnost konkurence v

$$(F^T F)^{-1} \cdot F^T =$$

VÝSLEDEK

Regresní statistika

Násobn 0.9693
Hodnot: 0.9395
Nastave 0.9274
Chyba s 780.55
Pozorov 25

ANOVA

	Rozdíl	SS	MS	F	znamnost F
Regrese	4	2E+08	5E+07	77.655479	7E-12
Rezidua	20	1E+07	609262		

Celkem 24 2E+08

	<i>Koeficienta</i>	<i>stř. hod.</i>	<i>t stat</i>	<i>Hodnota P</i>	<i>dolní 95%</i>	<i>horní 95%</i>	<i>dolní 95,0%</i>	<i>horní 95,0%</i>
Hranice	1642.6	932.47	1.7616	0.0934172	-302.5	3587.7	-302.5	3587.7
Soubor	81.899	36.855	2.2222	0.0379675	5.0213	158.78	5.0213	158.78
Soubor	19.893	8.4958	2.3416	0.0296686	2.1716	37.615	2.1716	37.615
Soubor	241	70.482	3.4193	0.0027174	93.979	388.03	93.979	388.03
Soubor	-171.8	399.21	-0.43	0.6715393	-1005	660.94	-1005	660.94

%

VÝSLEDEK

<i>Regresní statistika</i>	
Násobné R	0.969282
Hodnota spolehlivosti R	0.939508
Nastavená hodnota spole	0.92741
Chyba stř. hodnoty	780.5522
Pozorování	25

ANOVA

	<i>Rozdíl</i>	<i>SS</i>	<i>MS</i>
Regrese	4	1.89E+08	47312508
Rezidua	20	12185234	609261.7
Celkem	24	2.01E+08	

	<i>Koeficientyba stř. hod.</i>		<i>t stat</i>
Hranice	1642.6	932.4707	1.761601
Poč. kolemjdoucích/hod.	81.9	36.85457	2.222208
Velikost prodejny m2	19.9	8.495774	2.341568
Prům. plat prodavačů/měs	241.0	70.48227	3.419335
Přítomnost konkurence v	-171.8	399.2127	-0.430355

$$Y = 1642,6 + 81,9 \cdot X_1 + 19,9 \cdot X_2 + 9231.694$$

REZIDUA

<i>Pozorování</i>	<i>Roční trž</i>	<i>Rezidua</i>
1	6654.054	1145.946
2	10385.77	114.2307
3	6891.591	-1191.59
4	12693.27	-693.267
5	8074.959	25.04107
6	8437.261	1162.739
7	13252.04	-352.043
8	7440.992	-840.992

9	18851.32	648.6756
10	15681.49	-81.4919
11	12340.13	-940.133
12	8909.698	90.30175
13	10894.33	-94.3299
14	9570.19	329.8096
15	7916.067	-716.067
16	9084.932	1475.068
17	9932.769	1347.231
18	12000.92	-300.915
19	12267.61	32.39049
20	10837.91	-517.912
21	8403.992	-363.992
22	9234.226	-474.226
23	10766.03	153.9659
24	11704.5	235.4955
25	12553.93	-193.934

<i>F</i>	<i>namnost F</i>
77.655479	7E-12

<i>Hodnota P</i>	<i>olní 95</i>	<i>brní 95</i>	<i>lní 95,</i>	<i>lní 95,0%</i>
0.0934	-302	3588	-302	3588
0.0380	5.021	158.8	5.021	158.8
0.0297	2.172	37.62	2.172	37.62
0.0027	93.98	388	93.98	388
0.6715	-1005	660.9	-1005	660.9

- 241*X3

Roční tržby tis. Kč	Poč. kolemjdoucích/hod.	Velikost prodejny m2	Prům. plat prodavačů/měs.
7800	12	90	10.0
10500	20	150	17.1
5700	11	100	10.5
12000	30	180	20.8
8100	15	120	12.4
9600	17	90	15.7
12900	27	200	23.2
6600	13	100	12.1
19500	55	320	26.3
15600	45	220	24.8
11400	29	170	20.5
9000	15	145	13.8
10800	24	170	16.2
9900	22	130	15.4
7200	11	120	13.1
10560	16	140	14.6
11280	18	150	15.9
11700	20	190	20.5
12300	23	190	21.3
10320	31	170	14.3
8040	16	130	12.6
8760	19	140	14.2
10920	21	170	17.4
11940	24	160	21.1
12360	29	170	22.1

- = R - koeficient korelace
- = R² - koeficient determinace
- = R²_{adj} - upravený koeficient determinace
- = s² - směrodatná chyba (odhad směrodatné odchylky náhodné složky)

<i>MS</i>	<i>F</i>	<i>Významnost F</i>
47312508 609261.7	77.65548	6.82042E-12

<i>t stat</i>	<i>Hodnota P</i>	<i>Dolní 95%</i>	<i>Horní 95%</i>	<i>Dolní 95,0%</i>	<i>Horní 95,0%</i>
1.761601	0.093417	-302.4573713	3587.74	-302.457	3587.74
2.222208	0.037967	5.0212676	158.7758	5.021268	158.7758
2.341568	0.029669	2.171569106	37.6153	2.171569	37.6153
3.419335	0.002717	93.97909072	388.0258	93.97909	388.0258
-0.43035	0.671539	-1004.545876	660.9396	-1004.55	660.9396