Linear regression - tasks

Task. 1. Work experience and productivity of the company's employees are shown in the table. Find and interpret linear regression line (Productivity- dependent variable), R square and errors. Predict the productivity when experience is 12 years.

Experience	Productivity
1	120
5	115
10	132
8	123
9	128
1	102
2	106
4	109
5	112

Hint:

	Experience, x	Productivity, y	x^2	y^2	xy
	1	120	1	14400	120
	5	115	25	13225	575
	10	132	100	17424	1320
	8	123	64	15129	984
	9	128	81	16384	1152
	1	102	1	10404	102
	2	106	4	11236	212
	4	109	16	11881	436
	5	112	25	12544	560
Sum	45	1047	317	122627	5461

Task 2. As a first step in a program to conserve energy, an economist performs an experiment to analyze the relationship between outside temperature and heating cost. During one week, she measures the average daily tempearture and the amount of money required to purchase the fuel needed to heat a four bedroom house. The results are shown in the accompanying table

Celsius Temperature X	Cost of Fuel to Heat the House ($) Y
15	4
10	5
8	4
12	3
5	7
4	10
1	14

Create a scatterplot. Find and interpret linear regression line, R square and errors.

Hint:

	X	y	x^2	y^2	xy
	15	4	225	16	60
	10	5	100	25	50
	8	4	64	16	32
	12	3	144	9	36
	5	7	25	49	35
	4	10	16	100	40
	1	14	1	196	14
Sum	55	47	575	411	267

Task. 3. Edith Benham, a stock inwestor, wants to determine how closely a firm's annual net income is related to the dolar value of its assets. Assums that she finds the following data:

Company	Assets ($) x	Net income ($) y
A	60	20
B	60	25
C	75	25
D	80	20
E	100	30
F	145	35
G	175	40
H	190	45
I	190	55
J	200	50

Create a scatterplot. Find and interpret linear regression line (Net income- dependent variable), R square and errors.

Hint:

	Assets ($) x	Net income ($) y	x^2	y^2	xy
	60	20	3600	400	1200
	60	25	3600	625	1500
	75	25	5625	625	1875
	80	20	6400	400	1600
	100	30	10000	900	3000
	145	35	21025	1225	5075
	175	40	30625	1600	7000
	190	45	36100	2025	8550
	190	55	36100	3025	10450
	200	50	40000	2500	10000
Sum	1275	345	193075	13325	50250

Task 4. A homebuilders' association lobbying for various home subsidy programs argued, that during periods of high interest rates, the numer of building permits issued decreased drastically, in turn reducing the availability of new housing . The following data were presented as part of their argument:

Interest rates [%]	Building Permits
18,00	427
10,75	1189
15,25	825
12,45	904
15,56	800
14,25	880
13,95	950
16,85	628
11,45	1027
16,95	610
17,50	582
17,00	600

Create a scatterplot. Find and interpret linear regression line, R square and errors.

Task 5.Suppose, that a meteorologist measures the daily high temperatures on two scales during one week in the winter. The measurements are shown in the table. Find and interpret simple linear regression model (F-dependent variable).

F	C
32	0
23	-5
37,4	3
50	10
40,8	6
33,8	1
41	5

Task 6. Students in a small class were polled by a survey on attempting to establish a relationship between the hours of study in the week immediately preceding a major midterm exam and the marks received on the exam. The gathered data are shown in the table. Find and interpret the linear regression model (Dependent variable- Exam Score). Predict the exam score of the student, who has studied 27 hours.

Hours of Study	Exam Score
25	93
12	57
18	55
26	90
19	82
20	95
23	95
15	80
22	85
8	61

Task 7. Advertising is often touted as the key of the success. In seeking to determine just how influential advertising is, the management of a rectently set-up retail chain has collected data over the previous 15 weeks on sales revenue and advertising expenditures. The data are shown in the table. Find and interpret the linear regression model (Dependent variable- Sales). Predict the sales for the advertising expenditures equal to 9.

Advertising expenditures	Sales
3	50
5	250
7	700
6	450
6,5	600
8	1000
3,5	75
4	150
4,5	200
6,5	550
7	750
7,5	800
7,5	900
8,5	1100
7	600

Task 8. The data are available in the file "Activities.sta". Find and interpret the linear regression model (Dependent variable- Children, Independent variable- Work).

Task 9. The data are available in the file "Activities.sta". Find and interpret the linear regression model (Dependent variable- Leisure, Independent variable- Work).

Task 10. The data are shown in the table. Find and interpret the linear regression model (Inependent variable-Hours of study, Dependent variable- Exam Score). Predict the exam score of the student, who has studied 20 hours.

Exam score	Hours of study
7	120
8	122
9	125
10	131
11	135
11,5	140
12	142
13	145
14	150
14	100
15	154
16	159
17	162
18	164
18,5	168
19	170

Task 11. A real estate agent would like to predict the selling price of single-family homes. After careful consideration, he concludes that the variables likely to be most closely related to selling price are: the size of the house, age (in 100s ft^2). of the house and lot size(in 1000s ft^2). As an experiment, he takes a random sample of fifteen recently sold houses and records the selling price (in $ 1,000s). These are shown in the accompanying table. Find and interpret the linear regression model (Dependent variable- Selling Price). Predict the selling price when house size is 100, age- 10, lot size- 5.

House size	Selling Price	Age (years)	Lot size
20	89,5	5	4,1
14,8	79,9	10	6,8
20,5	83,1	8	6,3
12,5	56,9	7	5,1
18	66,6	8	4,2
14,3	82,5	12	8,6
27,5	126,3	1	4,9
16,5	79,3	10	6,2
24,3	119,9	2	7,5
20,2	87,6	8	5,1
22	112,6	7	6,3
19	120,8	11	12,9
12,3	78,5	16	9,6
14	74,3	12	5,7
16,7	74,8	13	4,8