Practicals 9: Regression, correlation, and general linear models

1. Snowflakes can have very different sizes, which might depend on the temperature of the ice crystals from which they are formed. In winter 2021, mean snowflake size and ice crystal temperature were recorded during 14 snowing events in Brno. The resulting data are summarized in the table below:

Snowing event No.	mean snowflake ice crystal temperature (°C)	mean snowflake diameter (mm)
1	-18	6.9
2	-20	7.1
3	-5	8.5
4	-12	8
5	-8	8.7
6	-9	8.5
7	-2	10.2
8	-14	7.1
9	-8	8.4
10	-17	6.2
11	-2	9.9
12	-4	9.3
13	-3	10.2
14	-12	7.6

How does the snowflake size depend on temperature?
Perform a statistical analysis, support it with a figure, and present a conclusion.
In addition to the statistical model, report the regression equation describing the dependence of snowflake size on temperature.
What can we say about the snowflake size at 5°C?

2. The Intelligence Quotient (IQ) of Czech Ministry of Interior Affairs employees was measured. These people were also asked about the average amount of beer they drink daily. The results were as follows:

IQ	beer (liters)
69	2
84	3.5
95	4.6
98	1.2
65	8
105	1.1
87	2.3
91	3.2
94	1
111	0.5
110	8
75	1

Is the amount of beer consumed correlated with the intelligence of Czech Ministry of Interior Affairs employees?

3. 530 School children took part in a swimming competition. Their height was also measured in addition to swimming speed. Is swimming speed significantly affected by children height? The data are available in Excel spradsheed on sheet 3.

A. Dependence of THC concentration in blood on the amount of cannabis smoked was analyzed in one person who smoked different amounts of dried cannabis from the same source. The intervals between measurements were long enough to decrease of THC concentration to 0 before each trial.

Smoking event	THC (mg/ l blood)	Cannabis DW (g)
1	10.1	5.3
2	3	1.2
3	8.7	3.8
4	12.3	8.5
5	20.8	9.1
6	5.9	3.1
7	10.1	4.5
8	12.3	8.5
9	5.9	6.5
10	10.1	7.8

Does THC concentration in the blood depend on the amount of cannabis smoked?
Perform a statistical analysis and illustrate it with a figure.

B. 20 books published in recent years were randomly selected in a bookshop. The numbers of pages were counted for each book and the age of the author was retrieved. The resulting data were as follows:

author age	number of pages
57	568
41	302
23	102
56	574
85	600
57	162
74	128
85	405
61	201
35	129
38	204
62	305
45	450
41	275
43	320
75	401

56	230
51	222
31	188
48	196

Does the author's age have an effect on the thickness of books?
Perform a statistical analysis and illustrate it with a figure.

C. The relationship between the mean age of children in a family and the height of the Christmas tree was studied. The resulting data were the following:

tree height (m)	age
2.2	3.5
3.1	4.2
0.8	15.8
2.5	7.6
1.4	12.8
1.7	16.4
1.2	15.3
2.8	6.5
0.9	19.5
1.6	5.6

Does the age of children in the family affect the height of the Christmas tree bought by the parents?

D. During a field survey, 10 frogs were captured, measured (body length and body mass), and released. The following data were obtained:

Frog	body mass (g)	body length (mm)
1	7	56
2	10	71
3	11	80
4	8	53
5	9	61
6	14	91
7	8	64
8	11	79
9	12	85
10	8	62

Is there any correlation between body mass and length in frogs? What is the proportion of variability shared by the two variables?

E. The relationship between car mass and fuel consumption was studied. The resulting data are summarized in the table below.

car mass (kg)	fuel consumption (liter per 100km)
1540	6.5
1100	4.7
1230	4.3
2410	5.9
1890	6.5
1220	5.3
1080	4.7
2340	7.1
3030	7.8
1930	8.2
1460	5.6

Does the fuel consumption depend on car mass? Perform a statistical analysis and illustrate it by a figure.