

EXERCISE - STATISTICS FOR AI

Summer Semester 2025 (Mag. Thomas Forstner)

366.591

366.592

366.593

366.594

366.595

43. A real estate agent wants to predict apartment prices based on their size (in square meters). He has data from 10 apartments, including their size and sale price. He wants to perform a linear regression to model the relationship between the size of the apartment and the sale price.

apartment i	1	2	3	4	5	6
size [square meters]	50	120	90	80	140	90
sale price [100 Euro]	150	360	270	300	420	240

- a) Calculate an appropriate summary statistic to quantify the linear relationship between these two parameters. Pearson correlation: 0.951
- b) What is the expected sale price of a 60 square meter apartment? 19030 €
- c) Calculate an appropriate quality of fit statistic for the linear regression model. R²: 0.905

44. A teacher wants to predict students' exam points based on the number of hours they have studied. She has data from 8 students, including the number of hours they have studied and their achieved points on the exam.

student i	study hours	achieved points
1	2	40
2	11	100
3	7	90
4	5	55
5	4	46
6	3	45
7	8	70
8	8	92

- a) Calculate an appropriate summary statistic to quantify the linear relationship between these two parameters. Pearson correlation: 0.927
- b) Based on a linear regression model, how many points can be expected on the exam given 10 hours of study time? 96.75 points

45. The entry-level salaries and the salaries paid after 6 years for a population of 100 employees were investigated. The following summary statistics are given.

	mean	variance	covariance
entry-level salary in Euro	1239	927	826
salary after 6 years in Euro	1534	1519	

It is also known that there is a linear relationship between the entry-level salaries and the salaries after 6 years.

- a) State the intercept and the slope of a linear regression model for predicting the salary after 6 years based on the entry salary. **intercept** = 430.094 ; **slope** = 0.8910
- b) Based on this linear regression model, what is the expected salary after 6 years for a person with an entry salary of 1306 Euro? 1593.74

46. A car owner wants to estimate the market price of his old car. He examines the relationship between the age of the cars offered and the price paid for these cars. He gets the table below:

	mean	standard deviation
age [years]	5.2	2.2
price [1000 Euro]	12.0	3.0

He also knows that the Bravais-Pearson correlation coefficient between age and price paid is -0.85.

The car, which the owner wants to sell, is 3 years old. What price can he expect assuming a simple linear relationship between age and price? 14550 €

47. In a psychological study about children and adolescents, the development of reading speed, which is measured as the time that is used for reading a specific, simple text, is investigated.

The reading speed in seconds of 8 children and adolescents of different ages is given in the table below.

age (X)	7	6	8	9	14	12	17	15
reading speed (Y)	360	250	195	178	90	85	78	72

- a) State the intercept and the slope of a linear regression model for predicting reading speed based on age. **intercept** = 402.465 ; **slope** = -21.724
- b) Based on this linear regression model, what is the expected reading time of a 56 year old person? -813.85 s
48. In a small study, the systolic blood pressure and the cholesterol levels of 8 patients were examined. Furthermore, a linear relationship between these two parameters can be assumed.
- | | | | | | | | | |
|----------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| blood pressure [mmHg] | 149 | 142 | 149 | 141 | 132 | 168 | 161 | 158 |
| cholesterol levels [mg/dl] | 179 | 182 | 183 | 189 | 165 | 195 | 194 | 189 |
- a) Calculate an appropriate summary statistic to quantify the linear relationship between these two parameters. Pearson correlation: 0.821
- b) Based on a linear regression model, what is the expected cholesterol level for a patient with a blood pressure of 145 mmHg? 181.17 mg/dl
- c) Calculate an appropriate quality of fit statistic for the linear regression model. R²: 0.674
49. John has five shirts, eight pairs of pants, and three pairs of shoes. In how many ways can he dress if he wants to wear one shirt, one pair of pants, and one pair of shoes? 120
50. How many permutations can be created with all the letters in the word STATISTICS?
50400

51. What is the probability of rolling a die and getting a number that is at least “3”? 66.667%

We are always considering a fair six-sided die with the following possible outcomes: 1, 2, 3, 4, 5, and 6.

52. A parade float can be composed of 7 sections in such a way that there are two flower sections, four music sections, and one banner section in this parade float. How many different possibilities are there theoretically (each section must be included)? 105

53. Three candidates X, Y, and Z are running for a position. The probability that X will win is 3 times greater than that of Y, and the probability that Y will win is 2 times greater than that of Z. Calculate the winning probability for each candidate.

P(X) = 66.667% ; **P(Y)** = 22.223% ; **P(Z)** = 11.112% ;

54. A delegation of 4 people out of 10 candidates is sent to the capital for a public meeting.

- a) How many ways are there to select this delegation? 210
- b) How many possibilities are there, if two of the candidates do not want to be sent to the meeting together and the delegation still needs 4 people? 182
- c) How many possibilities are there, if two of the candidates are friends and only want to be sent together to the meeting and the delegation still needs 4 people? 98

55. In how many ways can 12 boys and 8 girls sit in a row ...

- a) if the boys and girls want to sit together. $2 \cdot 12! \cdot 8!$ * I asked on Moodle if this format is allowed
- b) only the girls want to sit together. $13! \cdot 8!$

(Of course, not every boy and girl is the same, so you can distinguish between the boys and girls.)

Please keep the formal guidelines for submitting the homework assignments in mind to avoid losing points unnecessarily.