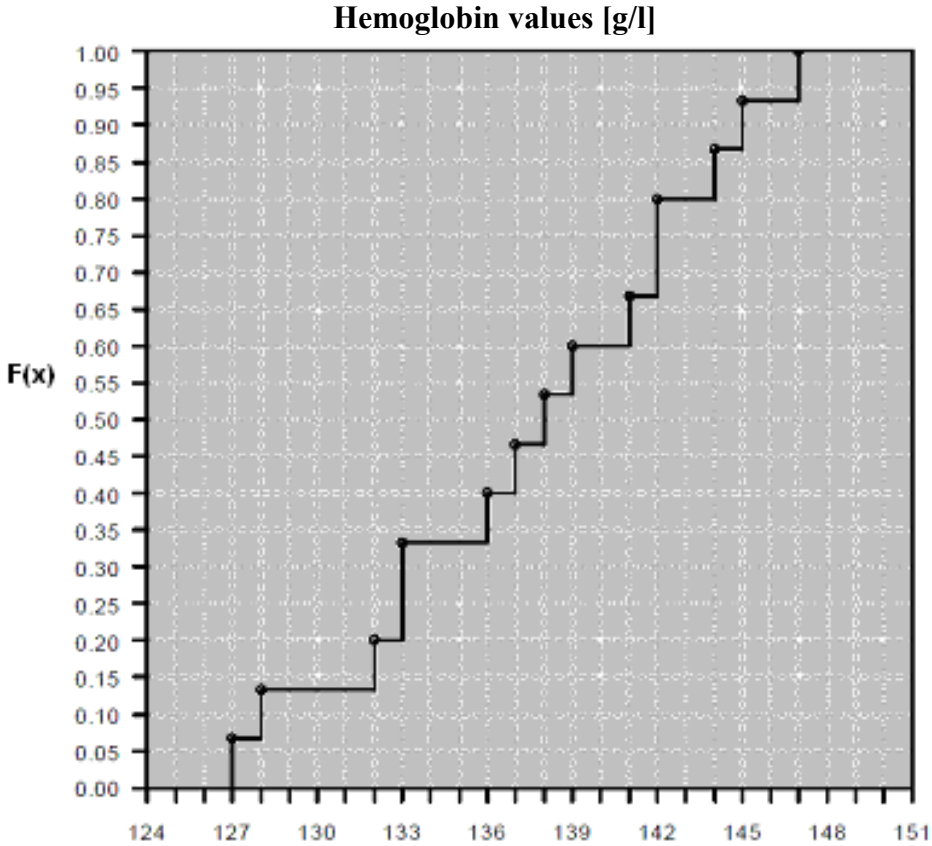
EXERCISE - STATISTICS FOR AI
Summer Semester 2025 (Mag. Thomas Forstner)

366.591 366.592 366.593 366.594 366.595

33. In a certain study, the hemoglobin values [g/l] of 15 healthy people were measured. The cumulative distribution function of these people is shown in the graphic below.



a) Use the above graph to determine the minimum hemoglobin value, the maximum hemoglobin value and the corresponding quartiles of the hemoglobin values.

minimum: 127 g/L, **maximum:** 147 g/L

1st quartile: 133 g/L, **2nd quartile:** 138 g/L, **3rd quartile:** 142 g/L

b) Below which hemoglobin value are 60% of all measured values? 139 g/L

34. In the table below, the body-mass-index (BMI) of a population of 80 people is given.

class limits	frequency
[16 ; 20)	13
[20 ; 24)	21
[24 ; 28)	19
[28 ; 32)	9
[32 ; 36)	18

a) Calculate the standard deviation for the given data. 5.549

b) Above what value lies 60% of the distribution mass? 23.619

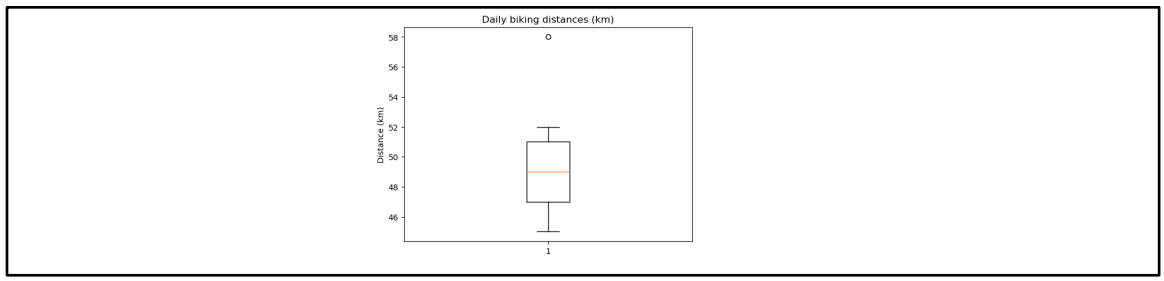
35. Somebody has recorded the daily driven distances in kilometers of a 9-day biking trip:

45, 50, 48, 52, 47, 49, 51, 46, 58

a) State the final results of all summary statistics necessary for creating a boxplot.

Min: 45, Max: 58, Median: 49, Q1: 46.5, Q3: 51.5

b) Draw a boxplot to represent the distances traveled in the box below.



36. Someone has a population of 60 rabbits. After weighing all these rabbits, she creates the following frequency table of the weight in grams of this population of rabbits.

weight class $[e_i\text{-}1;\ e_i)$	**frequency** f_i
1000 ; 1100	16
1100 ; 1200	18
1200 ; 1300	15
1300 ; 1400	11
1400 ; 1500	0

a) Calculate the arithmetic mean for this given data. 1185 g

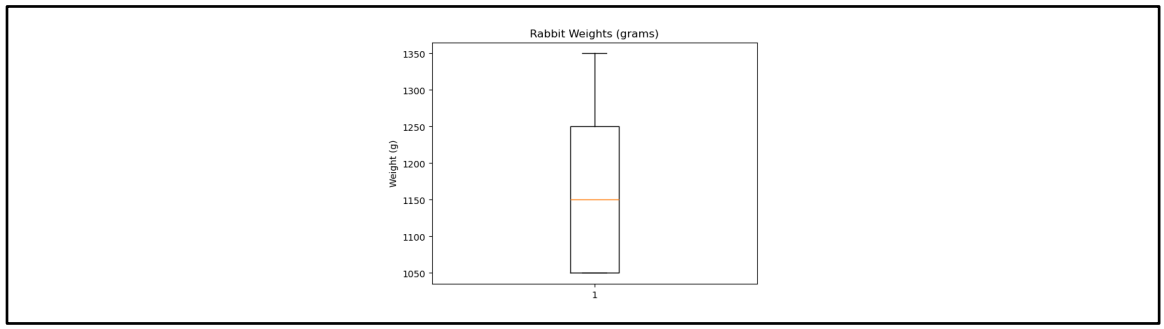
b) Calculate the standard deviation of the weight for this given data. 106.184 g

c) Calculate the interquartile range of this given data. 179.583

d) State the final results of all summary statistics necessary for creating a boxplot.

Min: 1000, Max: 1400, Median: 1177.78, Q1: 1093.75, Q3: 1273.33

e) Draw a boxplot for the given data in the box below.



37. A travel agency in the USA charges the following prices in US Dollars for one-week package holidays in a popular California holiday resort, depending on the hotel chosen:

580, 670, 720, 760, 810, 740

a) Calculate the variance of these given prices based on the US Dollar values. 5322.222

b) Calculate the variance of the corresponding Euro prices, assuming a conversion rate of 1 US dollar = 0.93 euro. 4603.190

38. A teacher examines the number of hours six students spend on learning and the points achieved on a test. The data is as follows:

    learning hours: 10 4 6 8 10 12
    points on test: 90 50 60 70 50 95

    a) Calculate the empirical covariance of the two given variables. 40.333
    b) Calculate the Bravais-Pearson correlation coefficient. 0.699

39. A fitness trainer analyses the relationship between body weight (in kg) and body fat (in %) in five people. The data is as follows:

    body weight [kg]: 75 70 80 80 85
    body fat [%]: 20 18 22 23 24

    Calculate an appropriate measure to quantify the linear relationship between these two variables. Person Correlation Coefficient: 0.983

40. In the table below, the speed limit on interstates (in kilometers per hour) and the number of deaths per 100 million kilometers driven in 5 countries are given.

Country	Top speed [km/h]	Number of Deaths
Denmark	130	4.1
Japan	100	4.7
Canada	110	4.3
Netherlands	120	5.1
Italy	130	6.1

    a) Calculate the correlation coefficient according to Bravais-Pearson. 0.329
    b) How does the value of the Bravais-Pearson correlation coefficient change, if the speed limit is not in kilometers per hour, but in miles per hour? No since the conversion is linear and the correlation is coefficient invariant

41. Two friends (A and B) are rating movies regarding the entertainment factor of these movies. The question, is to what extent do these two people agree on their rating, i.e. do they have similar tastes in movies? The rating scale is based on the school grading system, i.e. 1 means the movie is considered very good, 5 means very bad.

Movie	Rating person A	Rating person B
The Shawshank Redemption	3	2
The Hunt for Red October	3	2
E.T.	5	4
Blade Runner	2	1
Wag the Dog	3	4

    Calculate an appropriate summary statistic to answer this question. Spearman rank correlation coefficient: 0.825

42. Two food critics (A and B) are rating 5 restaurants regarding their overall dining experience. The rating scale is from 1 (miserable) to 10 (excellent). The judgments of the food critics A and B can be found in the table below. Use an appropriate correlation coefficient to assess the relationship between the judgments of these two food critics.

Restaurant	Judgment A	Judgment B
Pasta La Vista	8	9
Nacho Average Taco	9	8
Wok This Way	9	6
Brewed Awakening	10	9
Thai Tanic	7	4

    appropriate correlation coefficient: Spearman rank correlation coefficient: 0.526

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