Statistics
Exercise Set 1

Exercise 1
Exercise 2
Exercise 3
Exercise 4
Exercise 5

Exercise Set 1: One-Dimensional Statistical Variable Set

Exercise Set 1

First Year, Semester 1 · 2026/2027

N.SAHLI · First Year, Semester 1 · 2026/2027

Exercise 1

A household survey of gas consumption yielded the following results:

Gas consumption	Frequency (f_i)
0 – 9	1
10 – 19	2
20 – 29	1
30 – 39	5
40 – 49	8
50 – 59	16
60 – 69	19
70 – 79	20
80 – 89	17
90 – 99	11

1. Complete the above statistical table (fic(incresing and decreaisng)ect...).
2. What is the average gas consumption of these households? Calculate the variance, standard deviation, and coefficient of variation.
3. Give Q_1, Q_3 and MAD.
4. Plot the histogram and the ogive of the above distribution, then determine the mode and median of the series graphically.
5. Give the modal class, then calculate the mode value.

Exercise 2

A researcher collected data on the daily commute times (in minutes) for 50 employees at a company. The data is as follows:

25, 30, 30, 35, 40, 45, 50, 50, 50, 55, 60, 65, 70, 70, 70, 75, 80, 85, 90, 90, 95, 100, 100, 100, 105, 110, 115, 120, 120, 125, 130, 135, 140, 140, 145, 150, 150, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210.

1. Calculate the following measures:
   (a) Mean
   (b) Median
   (c) Mode
   (d) Range
   (e) Variance and standard deviation
   (f) Interquartile range (IQR)
2. Is the data symmetric, positively skewed, or negatively skewed? Justify your answer.

Statistics
Exercise Set 1

Exercise 1
Exercise 2
Exercise 3
Exercise 4
Exercise 5

3. Create a frequency distribution table (first, give the number of classes).
4. Draw a histogram based on the frequency distribution and construct a box plot.
5. If the researcher removes the top 10% of the data (the longest commute times), how does this affect the mean and standard deviation? Recalculate these measures after removing the top 10%.

Exercise 3

We give the following statistical table:

Class	Frequency	Relative frequency	Cumulative relative frequency
$[20;30[$	100	0.125	0.125
$[30;40[$	140	0.175	0.300
$[40;x_1[$	125	0.15625	0.45625
$[x_1;70[$	200	0.250	0.70625
$[70;100[$	180	0.225	0.93125
$[100;x_2[$	55	0.06875	1

1. Given that the median of this distribution is equal to 56.8, calculate x_1.
2. Given that the arithmetic mean of the population studied is equal to 60.5, calculate the bound x_2.
3. Plot the histogram of the distribution and calculate its mode.

Exercise 4

The following table gives the grouped data (class intervals) of a certain character:

Class interval	$[11;17[$	$[17;19[$	$[19;21[$	$[21;23[$	$[23;25[$	$[25;35]$
Frequency	3	5	6	5	2	1

1. The classes do not all have the same amplitude. Reset the intervals so that they all have equal amplitude, and give the corresponding frequencies.
2. Calculate the cumulative frequencies.
3. Calculate the median, the first quartile and the third quartile. What does the coefficient $Q_3 - Q_1$ represent?
4. Confirm the results of the previous question using the analytical formula.
5. Calculate the coefficient of skewness. Interpret the result.

Exercise 5

The following statistical table is given:

Variable x_i	$0-2$	$2-4$	$4-6$	$6-8$	$8-10$	Total
Frequency	n_1	5	6	n_4	n_5	20

We know that the arithmetic mean is equal to 4.7 and that the variance is equal to 5.71. Determine the value of the median of this series.