

Формулы:

Выборочное среднее:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$$

где x_i — длина слов из выборки.

Выборочная дисперсия:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2$$

Исправленная дисперсия:

$$S_o^2 = S^2 \frac{1}{n-1} = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{X})^2$$

Исправленное стандартное отклонение:

$$S = \sqrt{S^2}$$

Задание 1

Пункт 1

x_i	1	2	3	4	5	6	7	8
n_i	20	18	18	18	18	26	25	19
x_i	9	10	11	12	13	14	15	18
n_i	15	10	2	6	1	2	1	1

где n_i это количество слов длиной x_i .

Пункт 2

У нас есть данные о количестве букв в словах и их частоте. Размах:

$$R = X_{max} - X_{min} = 18 - 1 = 17$$

По правилу Стерджеса:

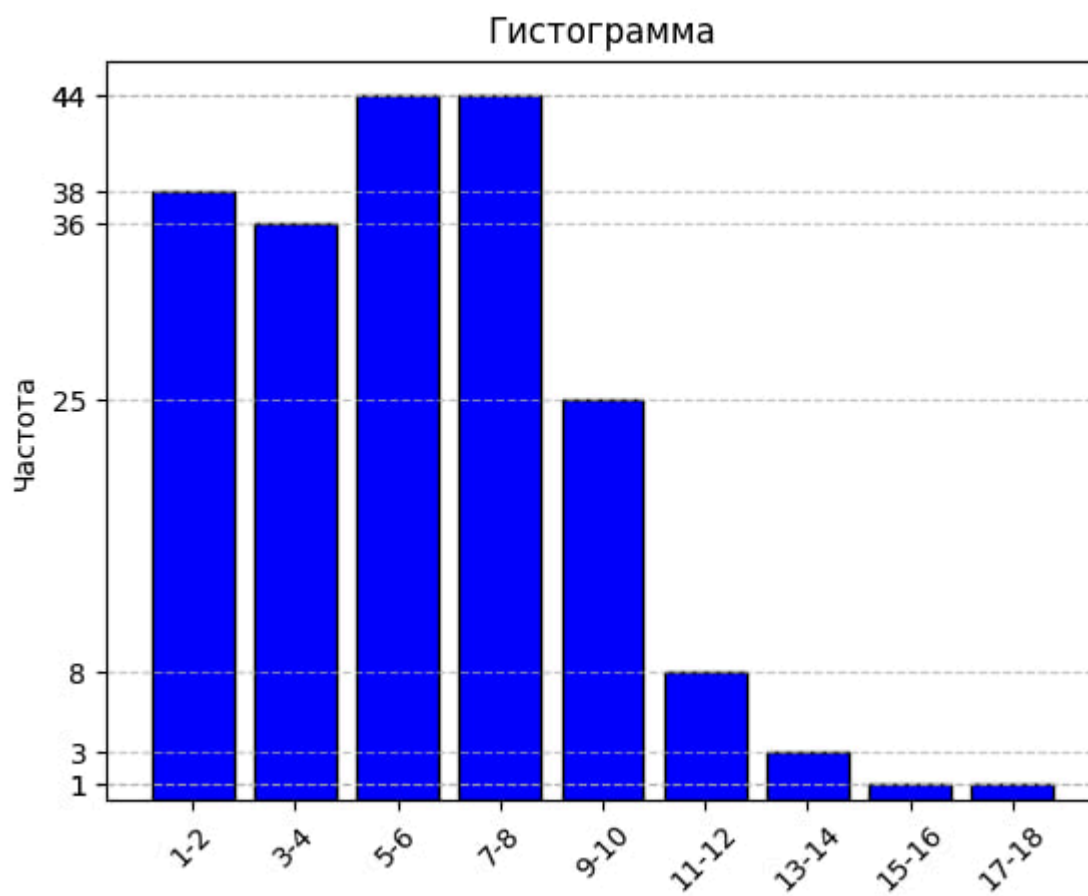
$$k = 1 + 3.322 * \log_{10} 200 \approx 8$$

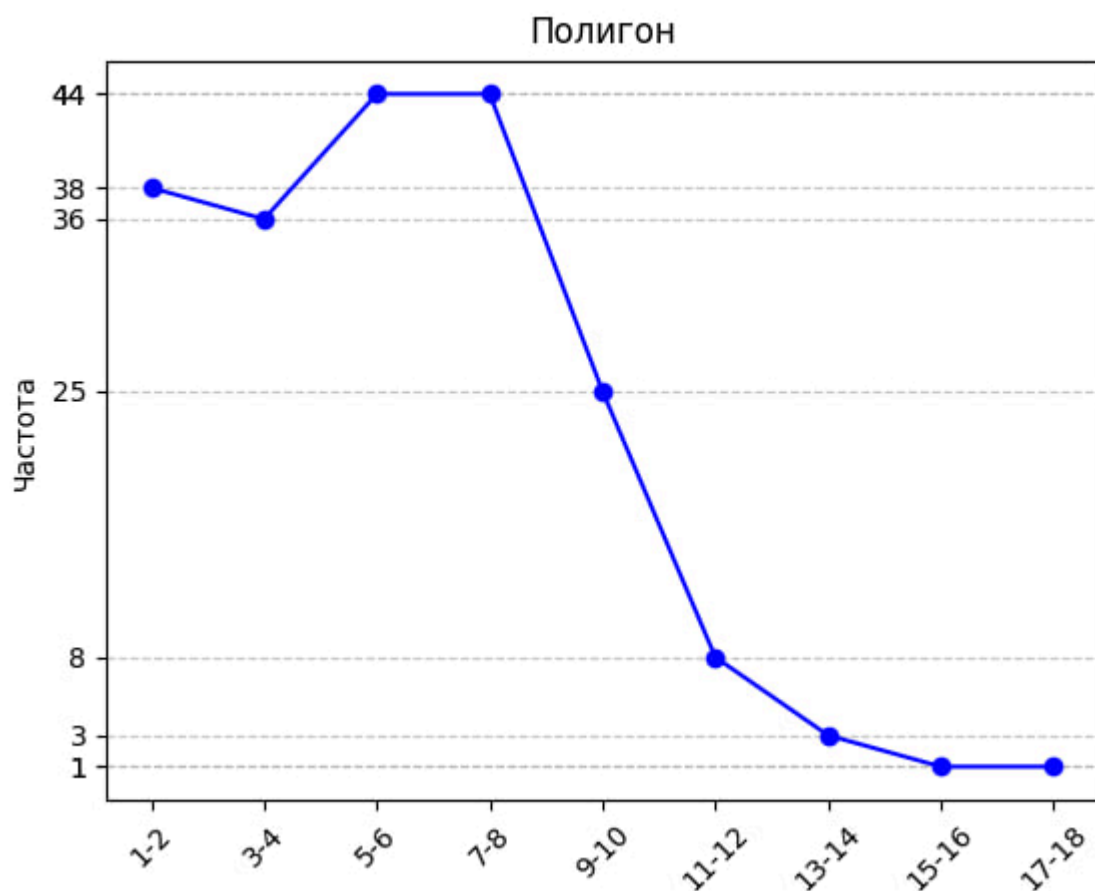
Ширина интервала:

$$h = \frac{17}{8} \approx 2$$

Интервал (длина слова)	Частота
1 – 2	38
3 – 4	36
5 – 6	44
7 – 8	44
9 – 10	25
11 – 12	8
13 – 14	3
15 – 18	2

Пункт 3





①

$$F^*(x) = \begin{cases} 0, & x < 3 \\ 0,19, & 3 \leq x < 5 \\ 0,37, & 5 \leq x < 7 \\ 0,59, & 7 \leq x < 9 \\ 0,81, & 9 \leq x < 11 \\ 0,935, & 11 \leq x < 13 \\ 0,975, & 13 \leq x < 15 \\ 0,99, & 15 \leq x < 18 \\ 1, & x > 18 \end{cases}$$



Пункт 4

a)

$$\begin{aligned}
 \sum (x_i \cdot f_i) &= (1 \cdot 20) + (2 \cdot 18) + (3 \cdot 18) + (4 \cdot 18) + (5 \cdot 18) + (6 \cdot 26) + (7 \cdot 25) + (8 \cdot 19) + (9 \cdot 15) + \\
 &+ (10 \cdot 10) + (11 \cdot 2) + (12 \cdot 6) + (13 \cdot 1) + (14 \cdot 2) + (15 \cdot 1) + (18 \cdot 1) = \\
 &= 20 + 36 + 54 + 72 + 90 + 156 + 175 + 152 + 135 + 100 + 22 + 72 + 13 + 28 + 15 + 18 = 1158 \\
 \bar{X} &= \frac{1158}{200} = 5.79
 \end{aligned}$$

b)

x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i * (x_i - \bar{x})^2$
1	20	-4.79	22.94	458.8
2	18	-3.79	14.36	258.5
3	18	-2.79	7.78	140.0
4	18	-1.79	3.20	57.6
5	18	-0.79	0.62	11.2
6	26	0.21	0.04	1.0
7	25	1.21	1.46	36.5
8	19	2.21	4.88	92.7
9	15	3.21	10.30	154.6

x_i	n_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$n_i * (x_i - \bar{x})^2$
10	10	4.21	17.74	177.4
11	2	5.21	27.12	54.2
12	6	6.21	38.55	231.3
13	1	7.21	51.98	52.0
14	2	8.21	67.40	134.8
15	1	9.21	84.92	84.9
18	1	12.21	149.06	149.1

$$S^2 = \frac{2094.5}{200} = 10.47$$

с)

$$S_o^2 = \frac{200}{199} 10.47 = 10.52$$

d)

$$S = \sqrt{10.52} \approx 3.24$$

е)

Мода — это значение, встречающееся чаще всего. В данном случае, самая высокая частота — 26 у числа 6.

Пункт 5

Значение из Excel:

$$\bar{X} = 5.79$$

$$S^2 = 10.48$$

$$S_o^2 = 10.52$$

$$S = 3.24$$

Наши вычисления совпадают с расчётом через формулы Excel

Задание 2

Пункт 1

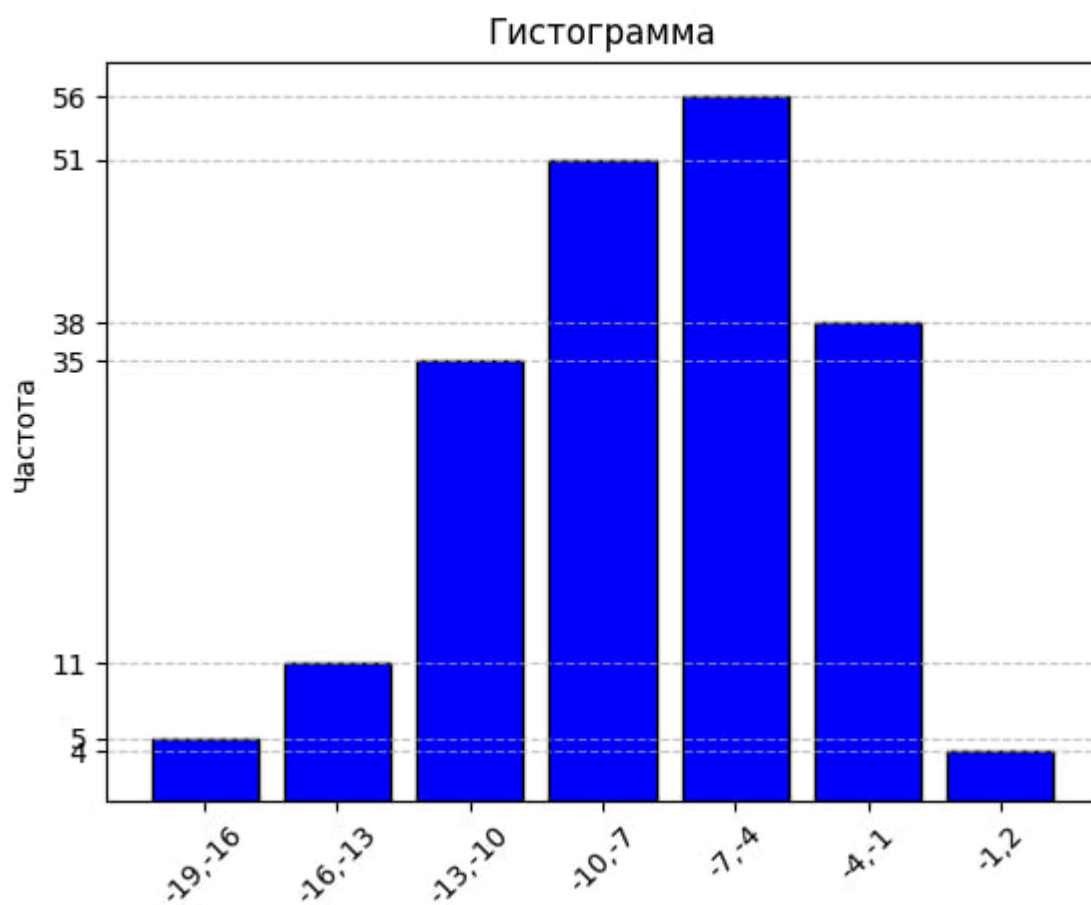
x_i	-19 - -18	-18 - -17	-17 - -16	-16 - -15	-15 - -14	-14 - -13	-13 - -12	-12 - -11	-11 - -10	-10 - -9
n_i	1	3	1	3	4	4	13	9	13	17
x_i	-9 - -8	-8 - -7	-7 - -6	-6 - -5	-5 - -4	-4 - -3	-3 - -2	-2 - -1	-1 - 0	1 - 2
n_i	18	16	21	17	18	17	14	7	3	1

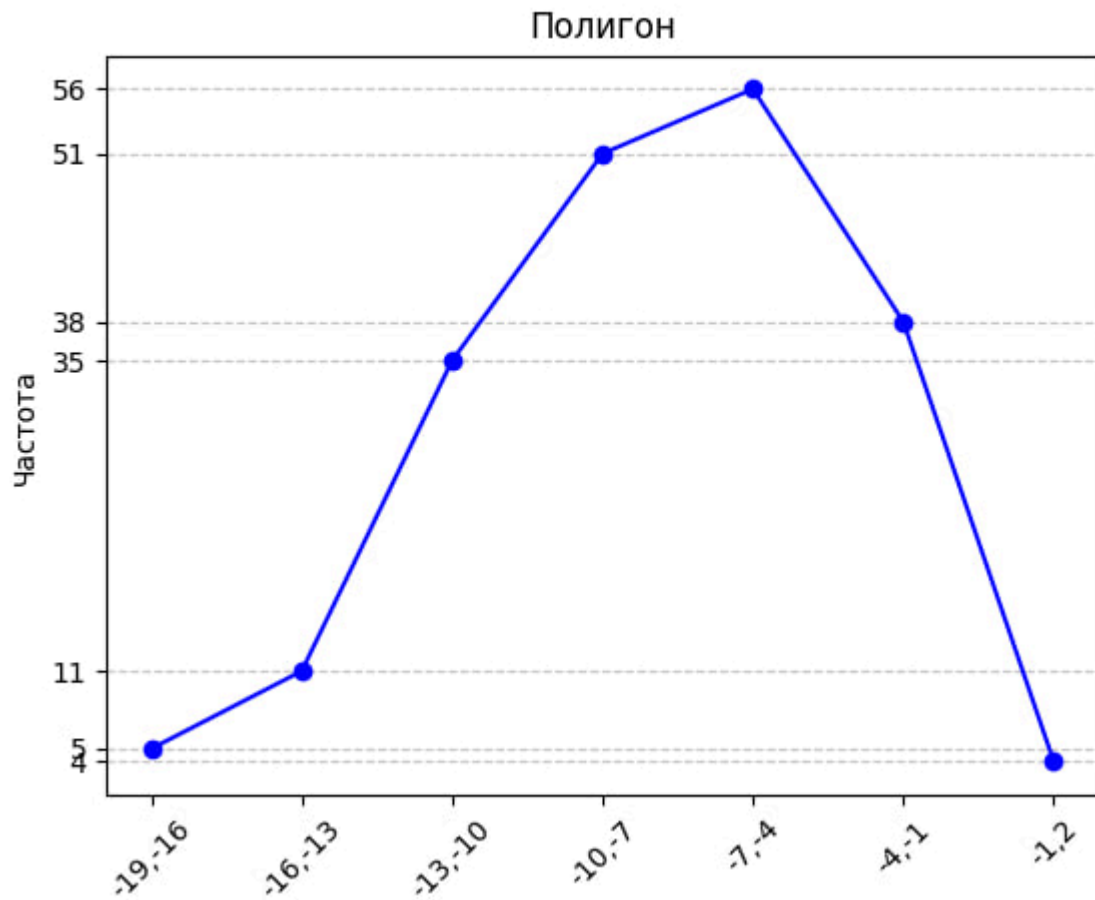
Пункт 2

Значения для расчёта таблице тебе при схожем n

Интервал	Частота n_{in_ini}
$[-19, -17)$	4
$[-17, -15)$	4
$[-15, -13)$	8
$[-13, -11)$	22
$[-11, -9)$	30
$[-9, -7)$	34
$[-7, -5)$	38
$[-5, -3)$	35
$[-3, 1)$	24

Пункт 3





②

$$F^*(x) = \begin{cases} 0, & x < -16 \\ 0,025, & -16 \leq x < -13 \\ 0,08, & -13 \leq x < -10 \\ 0,255, & -10 \leq x < -7 \\ 0,51, & -7 \leq x < -4 \\ 0,79, & -4 \leq x < -1 \\ 0,98, & -1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$



Пункт 4

a)

$$\bar{X} = \frac{-1499}{200} = -7.495$$

b)

$$S^2 = \frac{3058.995}{200} = 15.3$$

c)

$$S_o^2 = \frac{200}{199} 15.3 = 15.4$$

d)

$$S = \sqrt{15.4} \approx 3.9$$

e)

Мода — это значение, встречающееся чаще всего. В данном случае, самая высокая частота — 21 у числа $[-7 - -6)$

Пункт 5

Значение из Excel:

$$\bar{X} = -7.495$$

$$S^2 = 15.295$$

$$S_o^2 = 15.37$$

$$S = 3.92$$

Наши вычисления совпадают с расчётом через формулы Excel