

# «Информационные технологии в экономике» практика по описательным статистикам

старший преподаватель Суязова Светлана Андреевна

# Непрерывная количественная шкала (интервалов, отношений)

Примеры данных в непрерывной шкале: стоимостные показатели

Допустимые операции: все арифметические

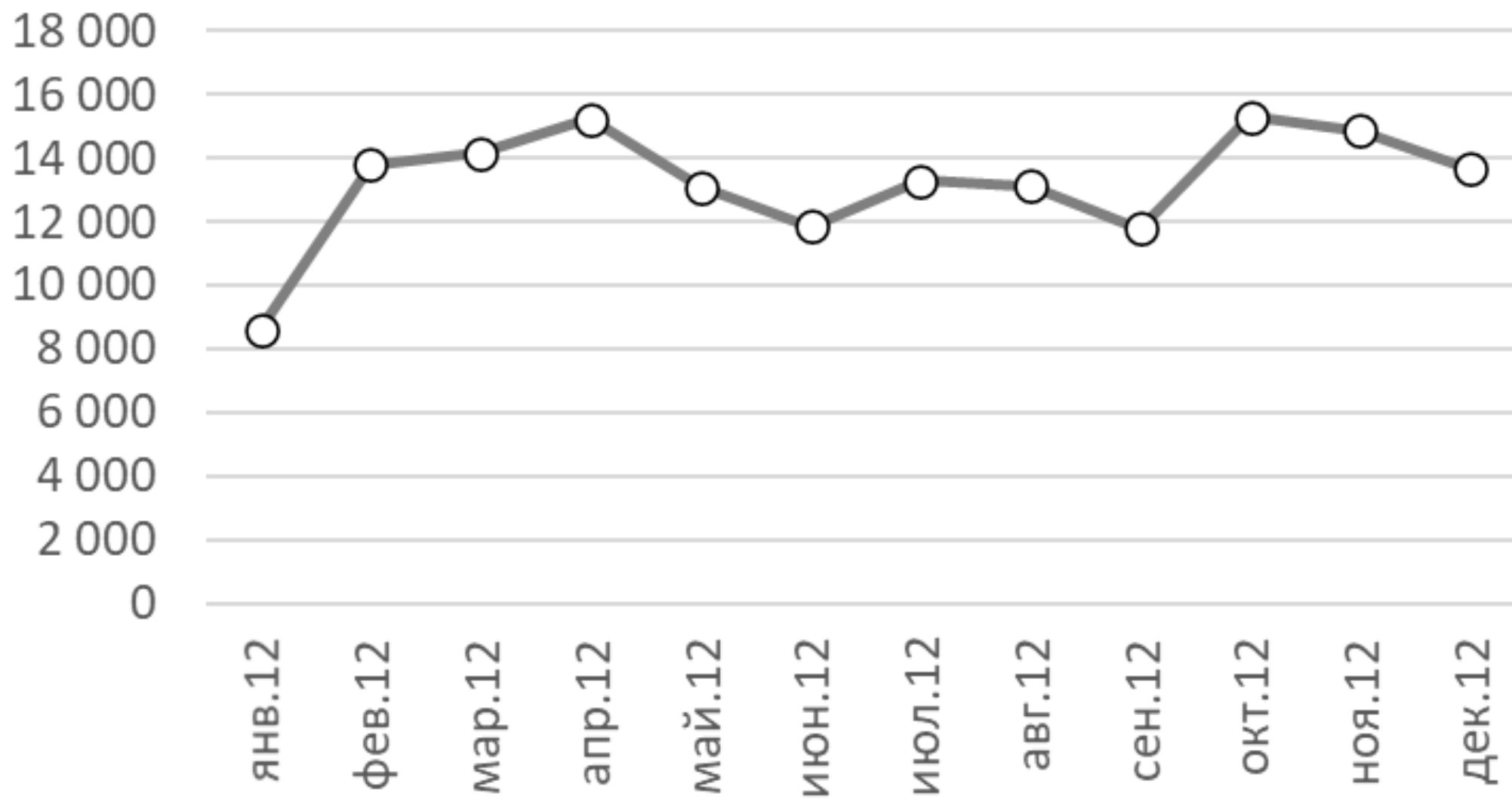
Что измеряем	Как измеряем
Характеристики центра распределения	Среднее (арифметическое), медиана, мода
Разброс относительно центра распределения	Дисперсия, стандартное отклонение, размах вариации
Форма распределения	Коэффициент асимметрии, Коэффициент эксцесса

## Исследуем непрерывный показатель

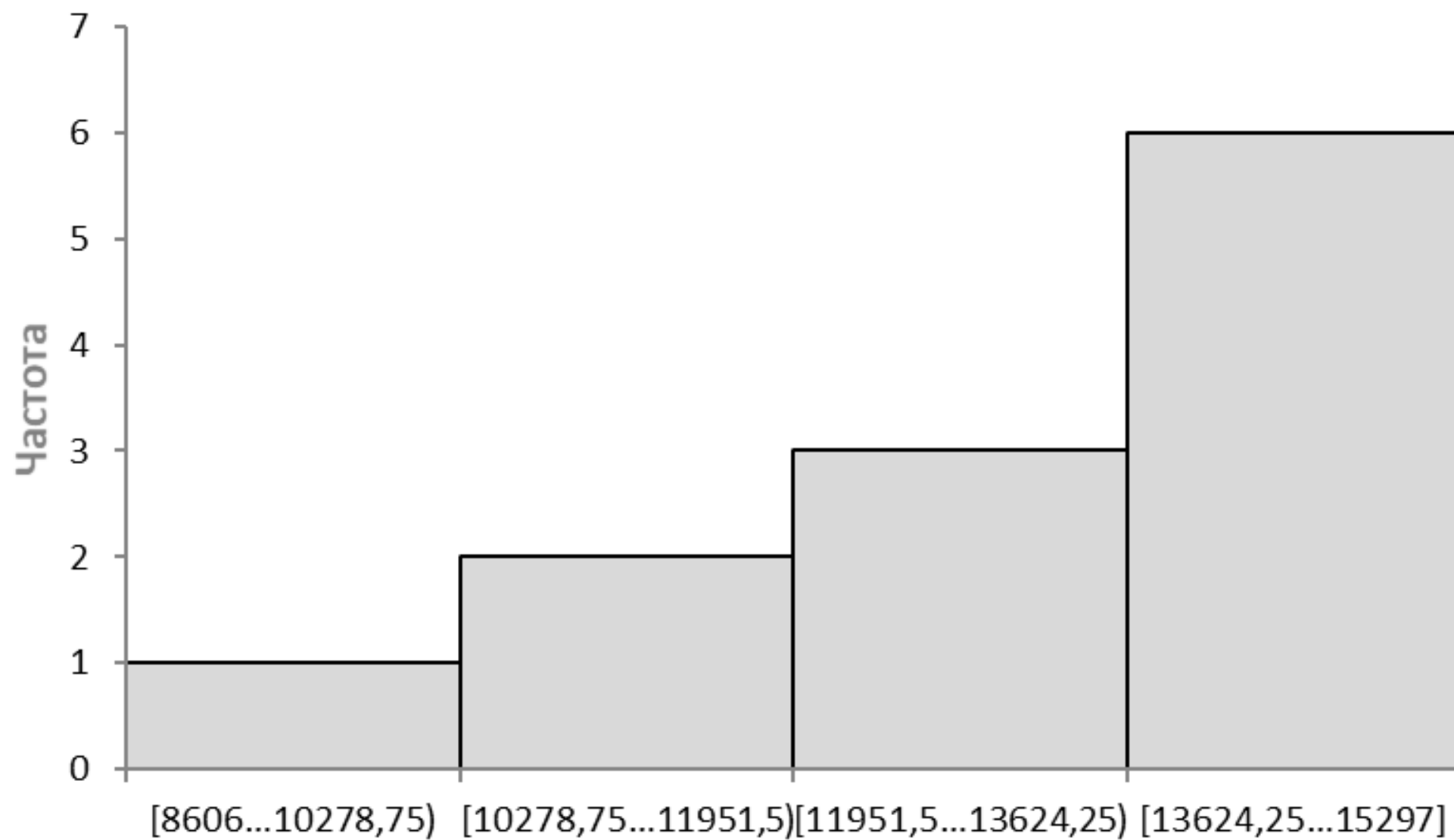
1. Рассчитать описательные статистики для показателя **X** – величины вновь созданных организаций в 2012 году.
2. Постройте гистограмму показателя **X**.
3. Постройте усиковую диаграмму показателя **X** («ящик с усами»).

месяц	X
январ.12	8 606
февр.12	13 802
мар.12	14 160
апр.12	15 235
май.12	13 097
июн.12	11 854
июл.12	13 293
авг.12	13 125
сен.12	11 800
окт.12	15 297
ноя.12	14 860
дек.12	13 662

График динамики X



Гистограмма X



# Порядковая шкала

Данные в порядковой шкале: рейтинги (необязательно числовые)

Допустимая операция: упорядочивание по возрастанию и убыванию

Что измеряем	Как измеряем
Характеристики центра распределения	Медиана, мода
Разброс относительно центра распределения	размах вариации, межквартильный интервал
Форма распределения	Максимум, минимум, квартили

## Исследуем порядковый показатель

Рассчитать описательные статистики для показателя **X** – бальной оценки разнообразия контента паблика в социальной сети.

№ подписчика	X
1	8
2	7
3	6
4	8
5	7
6	10
7	9
8	10
9	5
10	7

# Работа с данными в порядковой шкале

- Рейтинги, балльные оценки
- Разрешённые сравнения позиций: равно-не равно, больше-меньше
- Приводятся к одному знаменателю с помощью *совершенной ранжировки  $n$  объектов*:
  1. Лучшая оценка имеет ранг 1, худшая – ранг  $n$ .
  2. Ранги располагаются по убыванию качества оценок.
  3. Группа из одинаковых оценок получает средний ранг группы.
  4. Проверка:  $\sum_{i=1}^n r_i = n(n + 1)/2$



# Номинальная шкала

Примеры данных в номинальной шкале: качественные показатели

Допустимые операции: сравнение равно/не равно

Что измеряем	Как измеряем
Характеристики центра распределения	Мода
Разброс относительно центра распределения	нет
Форма распределения	нет

## Исследуем показатель в номинальной шкале

Рассчитать описательные статистики для показателя **X** – теги, приписанные постам паблика «Вконтакте» за январь.

№ п/п	Тег	№ п/п	Тег
1	имнорм	9	реклама
2	реклама	10	имнорм
3	неестественное	11	имнорм
4	имнорм	12	фанатское
5	фанатское	13	неестественное
6	фанатское	14	реклама
7	неестественное	15	имнорм
8	неестественное	16	реклама

# Корреляция в количественной шкале

Коэффициент корреляции Пирсона, функция Excel «КОРРЕЛ».

Лежит в интервале  $[-1, 1]$ . Интерпретация – [шкала Чеддока](#):

Теснота связи	Значение коэффициента корреляции	
	Прямая связь	Обратная связь
Слабая	0,1...0,3	-0,3...-0,1
Умеренная	0,3...0,5	-0,5...-0,3
Заметная	0,5...0,7	-0,7...-0,5
Высокая	0,7...0,9	-0,9...-0,7
Весьма высокая	0,9...0,99	-0,99...-0,9

# Корреляция в порядковой шкале

Проверка согласованности с помощью рангового коэффициента Спирмена  $\rho_S$ , где  $d_i^2$  – квадраты разностей совершенных рангов.

$$\rho_S = 1 - \left( \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \right)$$

Лежит в интервале  $[-1, 1]$ . Интерпретация – [шкала Чеддока](#)