



дисциплина:

**Современные методы
статистического анализа
кадастровых данных**

*к.э.н., профессор кафедры землеустройства и
земельного кадастра*

Яроцкая Елена Вадимовна

Теория вероятностей — это математическая наука, изучающая закономерности случайных явлений (событий, величин, функций, процессов и др.).

Она определяет и анализирует числовые характеристики случайных событий (объектов), наиболее важными из которых являются вероятность события и математическое ожидание случайной величины.

Событие (или Случайное событие) — это любой факт, который может либо произойти, либо не произойти при выполнении некоторого комплекса условий. При этом выполнение некоторого комплекса условий отождествляется с проведением испытания (опыта).





Достоверное событие

- происходит обязательно

Невозможное событие

- не произойдет при данных условиях

Равносильные события

- если одно событие является частным случаем другого события, или одно событие влечет за собой другое событие

Несовместные (несовместимые) события

- если появление одного из них исключает появление другого (не могут произойти одновременно)

- **Вероятность события** - численная мера степени объективной возможности этого события.
- **Вероятность события A** обозначается $P(A)$.
- **Вероятность достоверного события** равна единице.
- **Вероятность невозможного события** равна нулю.
- **Вероятность случайного события** — положительное число, заключенное между нулем и единицей: $0 < P(A) < 1$

Случайная величина — это переменная, которая в результате испытания принимает одно из своих возможных значений, причем заранее не известно, какое именно, так как это зависит от случая.

$X = x_i$ есть случайное событие, характеризующееся вероятностью $P(X = x_i) = p_i$.

Закон распределения —
соотношение, устанавливающее связь
между возможными значениями
случайной величины и
соответствующими вероятностями, с
которыми случайная величина
принимает эти значения.

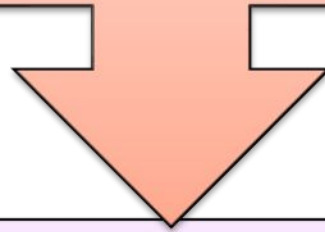
Математическое ожидание $M(X)$ — это число, характеризующее среднее значение случайной величины X

Дисперсия характеризует разброс или рассеяние значений случайной величины около ее математического ожидания

Математическая статистика —
раздел математики, изучающий методы
обработки результатов массовых
случайных явлений с целью выявления
статистических закономерностей.

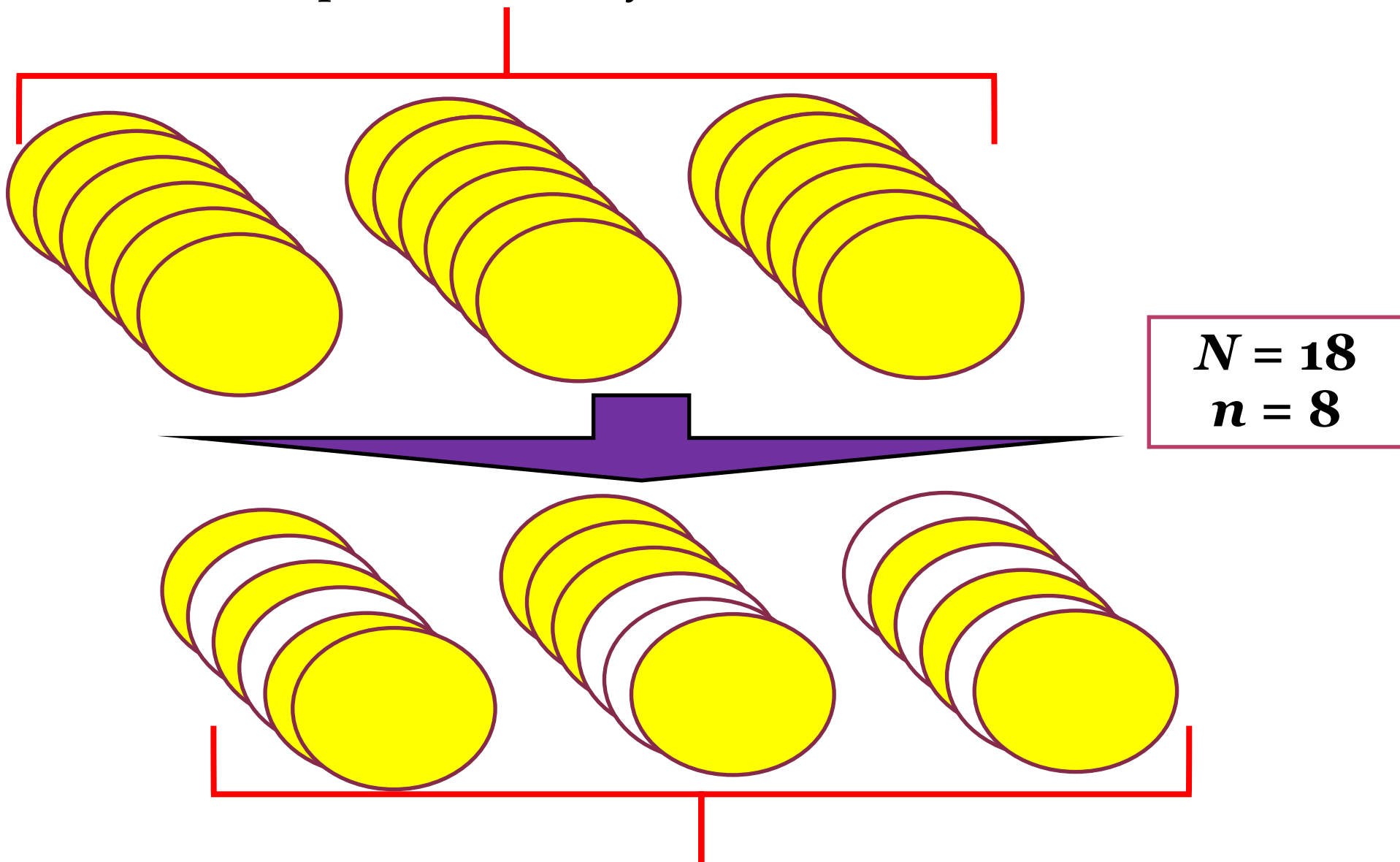
Генеральная совокупность объектов

– ЭТО СОВОКУПНОСТЬ ВСЕХ ОБЪЕКТОВ ИЗ КОТОРЫХ ПРОИЗВОДИТСЯ ВЫБОРКА



Выборочная совокупность – ЭТО СОВОКУПНОСТЬ СЛУЧАЙНО ОТОБРАННЫХ ИЗ ГЕНЕРАЛЬНОЙ СОВОКУПНОСТИ ОБЪЕКТОВ

Генеральная совокупность (N)



Выборочная совокупность (n)

Пример:

При проведении ГКО был отобран один из ценообразующих факторов для земельных участков «Расстояние между оцениваемым участком и центром населенного пункта» (X). Из генеральной совокупности земельных участков извлечена выборка n. Выборка задана в виде распределения частот:

Найти распределение относительных частот и основные характеристики вариационного ряда.

| расстояние до центра населенного пункта, км (x_i) | количество наблюдений (n) |
|---|----------------------------------|
| 4 | 2 |
| 7 | 4 |
| 8 | 5 |
| 12 | 6 |
| 17 | 3 |

Решение:

Значения x_i называются вариантами, а их последовательность, записанная в возрастающем порядке, — вариационным рядом.

Числа n_i называются частотами, а их отношения к объему выборки относительными частотами

$$W_i = \frac{n_i}{n}$$

Сумма относительных частот равна единице: $\sum W_i = 1$

При этом $\sum n_i = n$

Найдем объем выборки

$$n = 2 + 4 + 5 + 6 + 3 = 20$$

Относительные частоты равны:

$$W_1 = 2/20 = 0.1, W_2 = 4/20 = 0.2, W_3 = 5/20 = 0.25,$$

$$W_4 = 6/20 = 0.3, W_5 = 3/20 = 0.15$$

Решение:

Таблица распределения относительных частот

| расстояние до центра населенного пункта, км (x_i) | относительные частоты (W_i) |
|--|---------------------------------|
| 4 | 0,1 |
| 7 | 0,2 |
| 8 | 0,25 |
| 12 | 0,3 |
| 17 | 0,15 |

Решение:

Мода (M_0) - варианта, имеющая наибольшую частоту.

$$M_0 = 12$$

Медиана (m_e) - варианта, которая делит пополам вариационный ряд на две части с одинаковым числом вариантов в каждой.

Если число вариантов нечетно, т. е. $k = 2l + 1$, то $m_e = x_{l+1}$; если же число вариантов четно ($k = 2l$), то $m_e = (x_l + x_{l+1})/2$

В примере число вариантов нечетно $k = 2 \cdot 2 + 1$,
 $m_e = x_3 = 8$.

Размах варьирования - разность между максимальной и минимальной вариантами или длина интервала, которому принадлежат все варианты выборки:

$$R = x_{max} - x_{min}$$

$$R = 17 - 4 = 13$$

Дисперсия - это мера разброса значений случайной величины относительно среднего значения случайной величины.

Дисперсия показывает, насколько в среднем значения сосредоточены, сгруппированы около средней величины: *если дисперсия маленькая - значения сравнительно близки друг к другу, если большая - далеки друг от друга.*

Пример:

При проведении ГКО был отобран один из ценообразующих факторов для земельных участков «Расстояние между оцениваемым участком и центром населенного пункта» (X). Из генеральной совокупности земельных участков извлечена выборка n. Выборка задана в виде распределения частот:

Найти среднюю, дисперсию, среднее квадратическое отклонение.

| расстояние до центра населенного пункта, км (x_i) | количество наблюдений (n) |
|---|----------------------------------|
| 4 | 2 |
| 7 | 4 |
| 8 | 5 |
| 12 | 6 |
| 17 | 3 |

Решение:

Значения x_i называются вариантами, а их последовательность, записанная в возрастающем порядке, — вариационным рядом.

Числа n_i называются частотами, а их отношения к объему выборки относительными частотами

$$W_i = \frac{n_i}{n}$$

Сумма относительных частот равна единице: $\sum W_i = 1$

При этом $\sum n_i = n$

Найдем объем выборки

$$n = 2 + 4 + 5 + 6 + 3 = 20$$

Относительные частоты равны:

$$W_1 = 2/20 = 0.1, W_2 = 4/20 = 0.2, W_3 = 5/20 = 0.25, \\ W_4 = 6/20 = 0.3, W_5 = 3/20 = 0.15$$

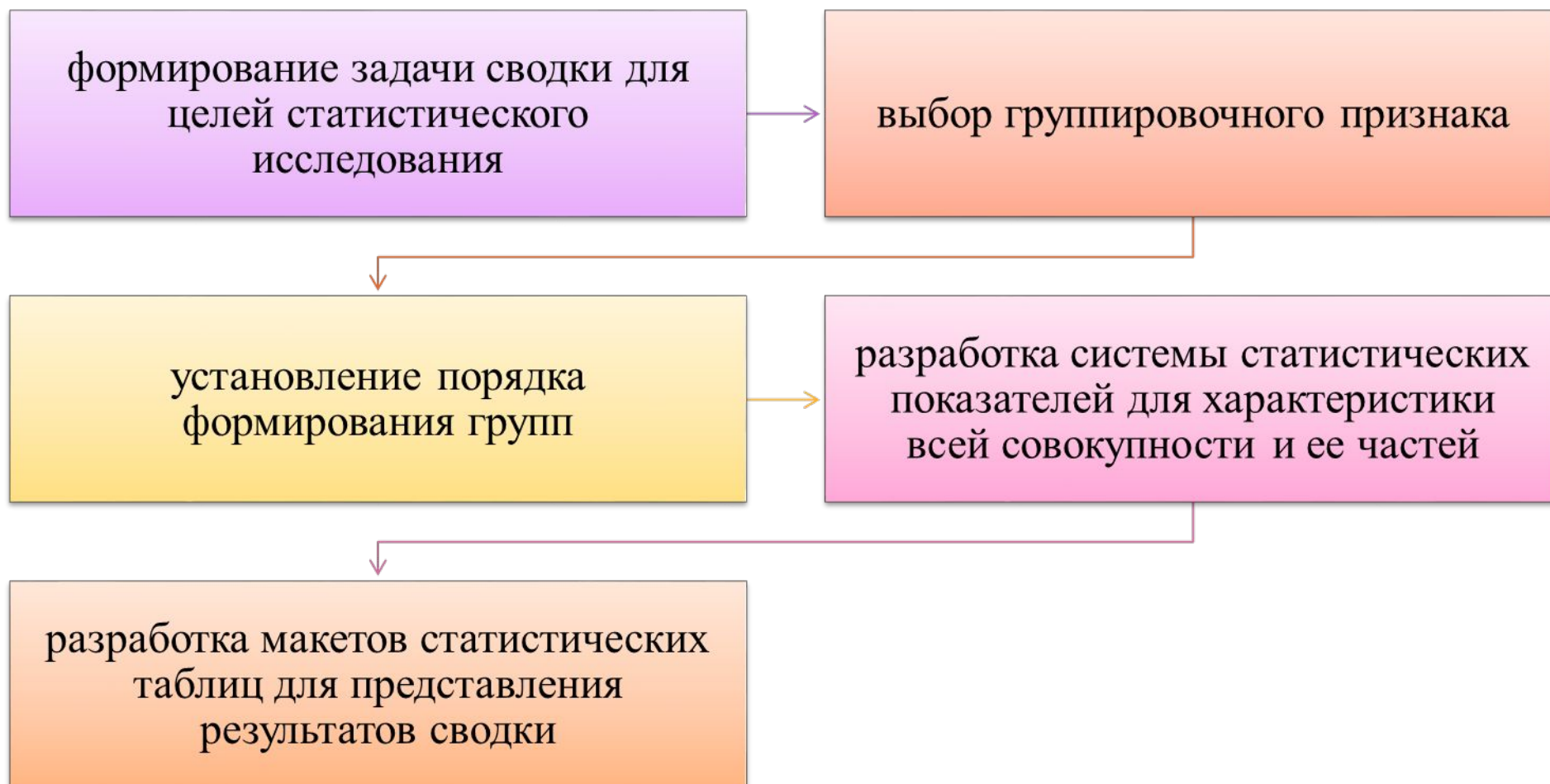
Сводка
(группировка)

- – это комплекс последовательных действий по обобщению конкретных единичных фактов, образующих совокупность, для выявления типичных черт и закономерностей, свойственных изучаемому явлению в целом.

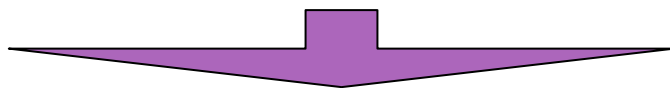
Статистическая сводка
(группировка)

- представляет собой стадию статистического исследования, когда первичные материалы сводятся вместе, образуя статистические совокупности с итоговыми обобщающими показателями в виде абсолютных, относительных и средних величин

Этапы проведения статистической сводки

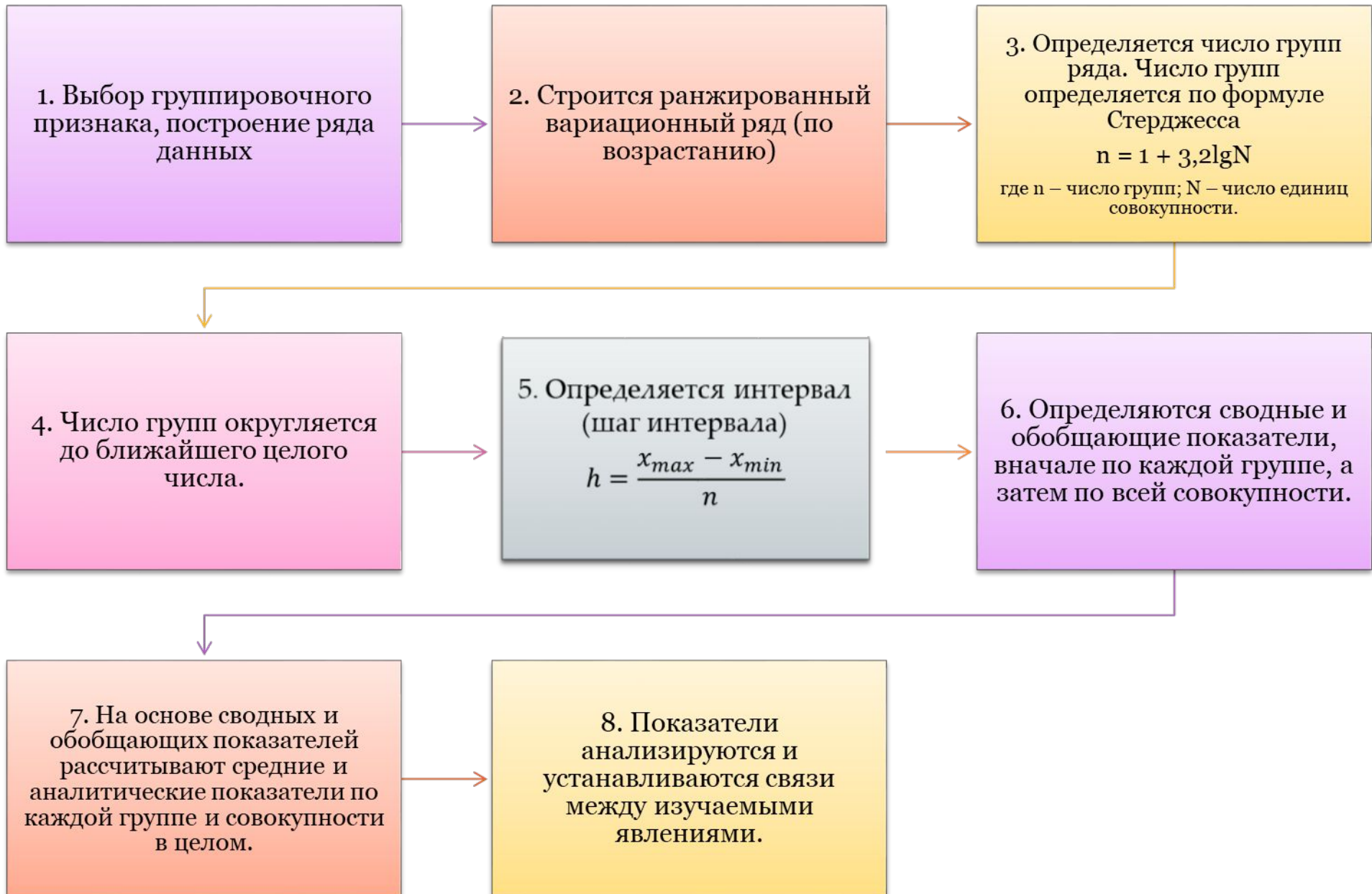


Статистической группировкой называется разделение единиц совокупности на группы, по существенным варьирующим признакам



Группировочным признаком называется признак, по которому проводится объединение единиц совокупности в отдельные группы. В качестве основания группировки используются существенные, теоретически обоснованные признаки.

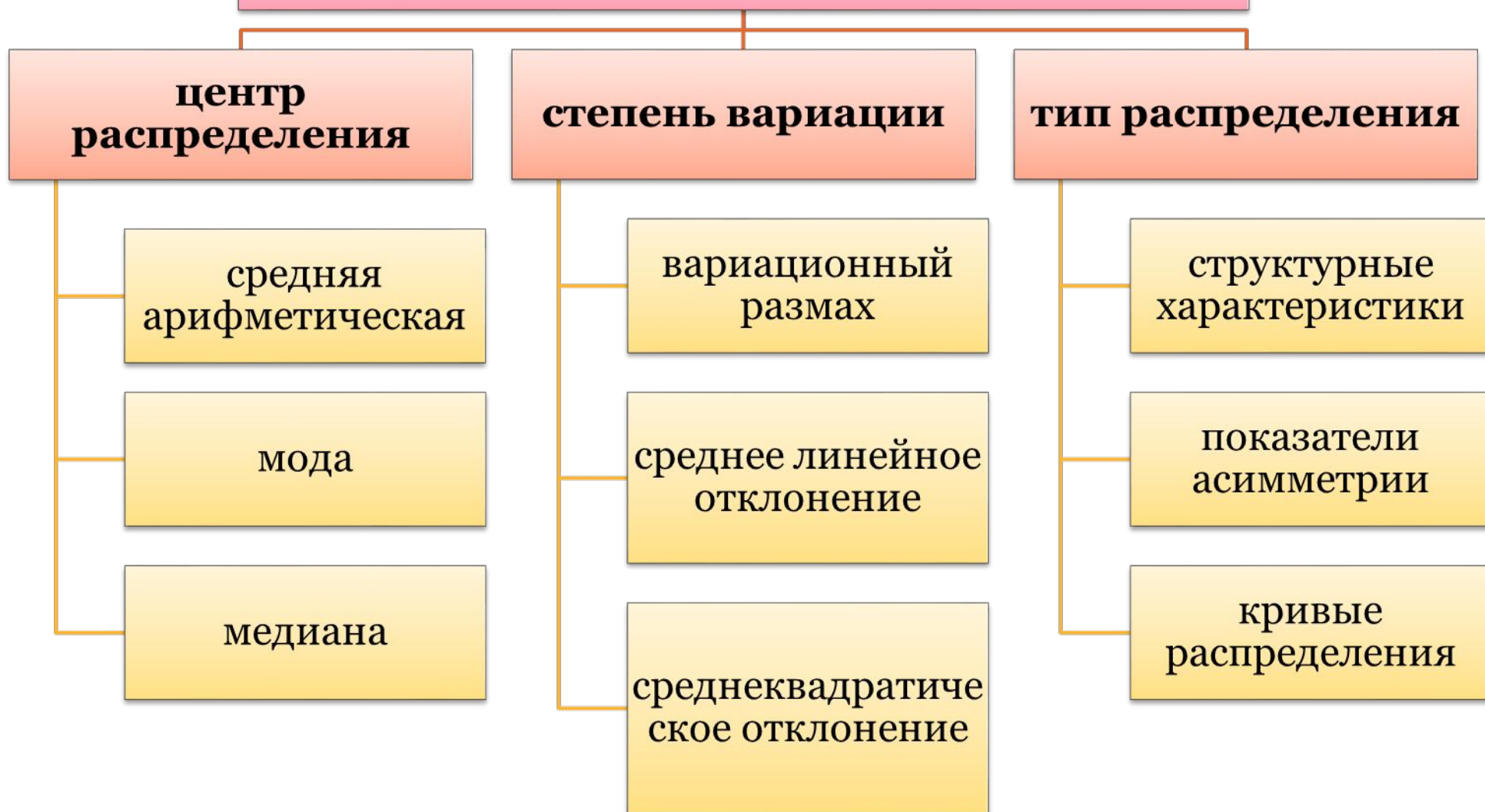
Этапы проведения статистической группировки



Вариация – это необходимое условие существования и развития массовых явлений.

Для каждого массового процесса, явления, необходимо так же соблюдение определенных условий - «допусков» - т.е. специфическая мера вариации их элементов, при которой данный процесс или явление будет существовать оптимально.

ПОКАЗАТЕЛИ ВАРИАЦИИ



ЦЕНТР РАСПРЕДЕЛЕНИЯ

| Показатель | Характеристика | Формула |
|-----------------------------------|--|--------------------------------|
| средняя арифметическая | используется в том случае, если распределение признака достаточно однородное | $\bar{x} = \frac{\sum x_i}{n}$ |

ЦЕНТР РАСПРЕДЕЛЕНИЯ

| Показатель | Характеристика | Формула |
|----------------|---|---|
| медиана | <p>величина варьирующего признака, делящая совокупность на 2 равные части - со значением признака меньше и больше медианы, находящейся в середине упорядоченного ряда.</p> <p>Для определения численного значения медианы в интервальном ряду следует сначала указать интервал нахождения медианы в интервальном ряду распределения</p> | <p>с нечетным числом признаков</p> $Me = \frac{n + 1}{2}$ <p>с четным числом признаков</p> $Me = \frac{n}{2}$ <p>значение медианы в интервальном вариационном ряду</p> $Me = x_0 + \frac{\frac{\sum f_i - S_{Me-1}}{2} \cdot i}{f_{Me}}$ <p>где:</p> <p>x_0 – низшая граница интервала, в котором находится медиана;</p> <p>S_{Me-1} – накопительная частота в интервале, предшествующем медианному;</p> <p>f_{Me} – частота медианного интервала;</p> <p>i – величина интервала;</p> <p>f_i – число наблюдений.</p> |

ЦЕНТР РАСПРЕДЕЛЕНИЯ

| Показатель | Характеристика | Формула |
|-------------|---|--|
| мода | <p>величина признака, который в изучаемом ряду встречается чаще всего. В дискретном вариационном ряду мода определяется как значение признака с наибольшей частотой.</p> <p>В интервальном ряду для определения численного значения моды необходимо сначала определить модальный интервал – интервал, к которому относится наибольшее количество участков изучаемой совокупности.</p> | $M_o = x_0 + \frac{(f_{M_o} - f_{M_{o-1}})}{(f_{M_o} - f_{M_{o-1}}) + (f_{M_{o+1}})} \cdot i$ <p>где:</p> <p>x_0 – низшая граница интервала, в котором находится мода;</p> <p>f_{M_o} – частота в модальном интервале;</p> <p>$f_{M_{o-1}}$ – частота в интервале, предшествующем модальному;</p> <p>$f_{M_{o+1}}$ – частота в последующем интервале;</p> <p>i – величина интервала</p> |

СТЕПЕНЬ ВАРИАЦИИ

| Показатель | Характеристика | Формула |
|---------------------|---|-------------------------|
| вариационный размах | показывает насколько велико различие между единицами совокупности, имеющими наименьшее и наибольшее значение, то есть вариационный размах – это разница между наибольшим и наименьшим значением варьирующего признака | $R = x_{max} - x_{min}$ |

Недостаток вариационного размаха - в общей совокупности могут встречаться значения признаков, которые или очень велики или очень малы по сравнению с основной массой и эти значения обусловлены случайными обстоятельствами. Эти значения являются аномальными, поэтому размах дает искаженную амплитуду колебания признака. Для исключения аномальных значений необходимо совокупность проверить на засоренность, то есть наличие в совокупности случайных показателей.

СТЕПЕНЬ ВАРИАЦИИ

| Показатель | Характеристика | Формула |
|---|---|---|
| <p>среднее линейное отклонение</p> | <p>показатели выступают в роли показателей силы вариации и дают обобщенную характеристику степени колеблемости признаков совокупности, т.е. показывают, на сколько, в среднем, отклоняются конкретные значения признака от его среднего значения.</p> | $\sigma = \frac{\sigma(\bar{a} - \bar{a})}{\sigma}$ |
| <p>среднеквадратическое отклонение</p> | | $G = \sqrt{\frac{\sum((x_i - \bar{x})^2 \cdot f_i)}{\sum f_i}}$ <p>где: x_i – значение признака в вариационном ряду; \bar{x} – средневзвешенное значение; f_i – количество единиц совокупности с данным значением признака; $\sum f_i$ – общее количество единиц совокупности.</p> |

СТЕПЕНЬ ВАРИАЦИИ

| Показатель | Характеристика | Формула |
|--|---|--|
| проверка на «засоренность» совокупности | Отношение средне квадратического отклонения к среднему линейному зависит от наличия в совокупности редких выделяющихся отклонений, и это отношение может служить индикатором засоренности совокупности. Для нормального закона распределения соотношение должно быть равно 1,25 или $\leq 1,25$. | индикатор наличия засоренности = $\frac{G}{D}$ |

ТИП РАСПРЕДЕЛЕНИЯ

Если соотношение фиксирует засоренность, необходимо найти признаки, которые засоряют совокупность.

Рекомендуется исключить их, снова проверить совокупность, добиваясь тем самым нормального ряда распределения.

ТИП РАСПРЕДЕЛЕНИЯ

Для определения однородности совокупности вычисляются показатели вариации, выраженные в относительных единицах

| Показатель | Характеристика | Формула |
|---------------------------------------|---|--|
| Коэффициент осцилляции | отражает относительную колеблемость крайних значений признака вокруг средней | $V_r = \frac{R}{\bar{x}} \cdot 100\%$ |
| Коэффициент вариации | мера относительного разброса случайной величины; показывает, какую долю среднего значения этой величины составляет ее средний разброс | $V_c = \frac{G}{\bar{x}} \cdot 100\%$ |
| Линейные коэффициенты вариации | позволяют сравнивать вариацию разных совокупностей, потому что в отличие от среднего линейного отклонения его значение не зависит от единиц измерения | $V_d = \frac{d}{\bar{x}} \cdot 100\%$ $V_{Me} = \frac{G}{Me} \cdot 100\%$ |

**Совокупность считается не засоренной,
если выполняются следующие условия:**

- ❖ значение медианы находится между средневзвешенной и модой;
- ❖ отношение среднеквадратического отклонения к среднему линейному отклонению $\leq 1,25$;
- ❖ коэффициенты вариации $\leq 33\%$.

Пример:

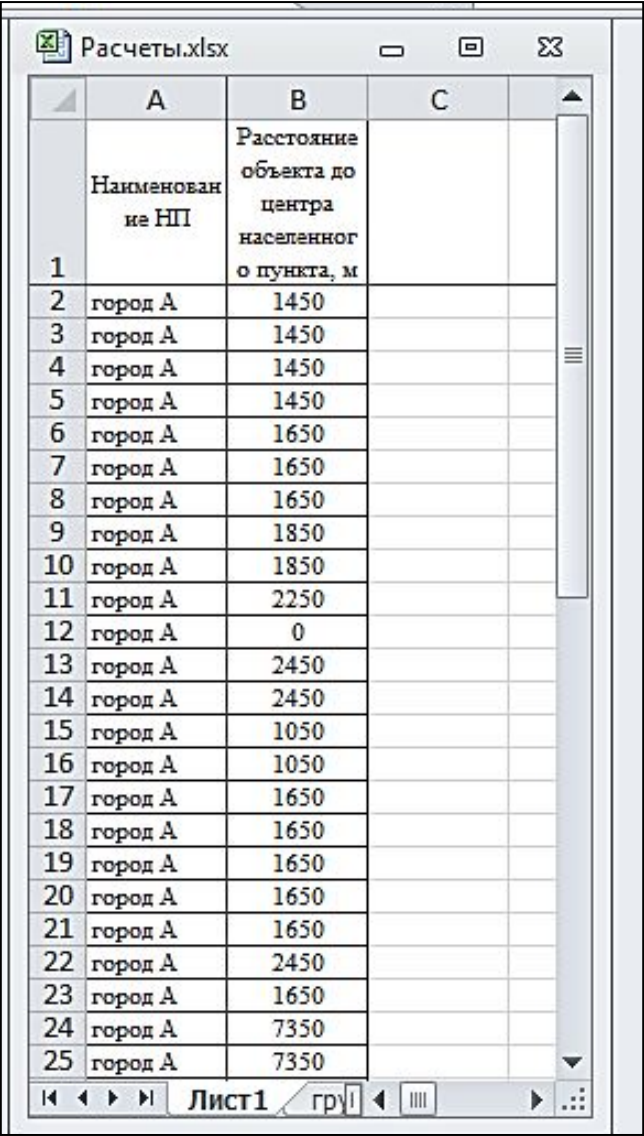
| Наименование НП | Расстояние объекта до центра населенного пункта, м | Наименование НП | Расстояние объекта до центра населенного пункта, м | Наименование НП | Расстояние объекта до центра населенного пункта, м | Наименование НП | Расстояние объекта до центра населенного пункта, м | Наименование НП | Расстояние объекта до центра населенного пункта, м |
|-----------------|--|-----------------|--|-----------------|--|-----------------|--|-----------------|--|
| город А | 1450 | город А | 0 | город А | 2450 | город А | 2650 | город А | 1050 |
| город А | 1450 | город А | 2450 | город А | 1650 | город А | 2450 | город А | 1050 |
| город А | 1450 | город А | 2450 | город А | 7350 | город А | 5750 | город А | 850 |
| город А | 1450 | город А | 1050 | город А | 7350 | город А | 1250 | город А | 850 |
| город А | 1650 | город А | 1050 | город А | 7350 | город А | 1250 | город А | 250 |
| город А | 1650 | город А | 1650 | город А | 7350 | город А | 2050 | город А | 450 |
| город А | 1650 | город А | 1650 | город А | 2650 | город А | 450 | | |
| город А | 1850 | город А | 1650 | город А | 2650 | город А | 450 | | |
| город А | 1850 | город А | 1650 | город А | 2650 | город А | 0 | | |
| город А | 2250 | город А | 1650 | город А | 2650 | город А | 1050 | | |

1. Сгруппировать выборку
2. Определить сводные и обобщающие показатели по каждой группе и по всей совокупности
3. Рассчитать средние и аналитические показатели по каждой группе и совокупности в целом
4. Проанализировать полученный результат

Решение*: * Решение будет производиться с использованием MS Excel

1. Выбор группировочного признака, построение ряда данных

Группировочный признак выбран, ряд построен



| | А | В | С |
|----|-----------------|--|---|
| 1 | Наименование НП | Расстояние объекта до центра населенного пункта, м | |
| 2 | город А | 1450 | |
| 3 | город А | 1450 | |
| 4 | город А | 1450 | |
| 5 | город А | 1450 | |
| 6 | город А | 1650 | |
| 7 | город А | 1650 | |
| 8 | город А | 1650 | |
| 9 | город А | 1850 | |
| 10 | город А | 1850 | |
| 11 | город А | 2250 | |
| 12 | город А | 0 | |
| 13 | город А | 2450 | |
| 14 | город А | 2450 | |
| 15 | город А | 1050 | |
| 16 | город А | 1050 | |
| 17 | город А | 1650 | |
| 18 | город А | 1650 | |
| 19 | город А | 1650 | |
| 20 | город А | 1650 | |
| 21 | город А | 1650 | |
| 22 | город А | 2450 | |
| 23 | город А | 1650 | |
| 24 | город А | 7350 | |
| 25 | город А | 7350 | |

2. Строится ранжированный вариационный ряд (по возрастанию)

The screenshot shows the Microsoft Excel interface with the 'Данные' (Data) tab selected. A data table is visible with columns 'Наименование НП' (Settlement Name) and 'Расстояние объекта до центра населенного пункта, м' (Distance of the object to the center of the settlement, m). A dialog box titled 'Обнаружены данные вне указанного диапазона' (Data found outside the specified range) is open, asking for action on data found outside the selected range. The 'автоматически расширить выделенный диапазон' (Automatically expand the selected range) option is selected.

| | A | B | C | D | E | F | G | H | I |
|----|---------|--|---|---|---|---|---|---|---|
| 1 | | Расстояние объекта до центра населенного пункта, м | | | | | | | |
| 3 | город А | 1450 | | | | | | | |
| 4 | город А | 1450 | | | | | | | |
| 5 | город А | 1450 | | | | | | | |
| 6 | город А | 1650 | | | | | | | |
| 7 | город А | 1650 | | | | | | | |
| 8 | город А | 1650 | | | | | | | |
| 9 | город А | 1850 | | | | | | | |
| 10 | город А | 1850 | | | | | | | |
| 11 | город А | 2250 | | | | | | | |
| 12 | город А | 0 | | | | | | | |
| 13 | город А | 2450 | | | | | | | |
| 14 | город А | 2450 | | | | | | | |
| 15 | город А | 1050 | | | | | | | |
| 16 | город А | 1050 | | | | | | | |
| 17 | город А | 1650 | | | | | | | |
| 18 | город А | 1650 | | | | | | | |
| 19 | город А | 1650 | | | | | | | |
| 20 | город А | 1650 | | | | | | | |
| 21 | город А | 1650 | | | | | | | |
| 22 | город А | 2450 | | | | | | | |
| 23 | город А | 1650 | | | | | | | |
| 24 | город А | 7350 | | | | | | | |
| 25 | город А | 7350 | | | | | | | |
| 26 | город А | 7350 | | | | | | | |
| 27 | город А | 7350 | | | | | | | |
| 28 | город А | 2650 | | | | | | | |
| 29 | город А | 2650 | | | | | | | |
| 30 | город А | 2650 | | | | | | | |
| 31 | город А | 2650 | | | | | | | |



The screenshot shows the same data table after sorting by the distance column in ascending order. The values in column B are now: 0, 0, 250, 450, 450, 450, 850, 850, 1050, 1050, 1050, 1050, 1050, 1250, 1250, 1450, 1450, 1450, 1450, 1650, 1650, 1650, 1650.

| | A | B | C |
|----|-----------------|--|---|
| 1 | Наименование НП | Расстояние объекта до центра населенного пункта, м | |
| 2 | город А | 0 | |
| 3 | город А | 0 | |
| 4 | город А | 250 | |
| 5 | город А | 450 | |
| 6 | город А | 450 | |
| 7 | город А | 450 | |
| 8 | город А | 850 | |
| 9 | город А | 850 | |
| 10 | город А | 1050 | |
| 11 | город А | 1050 | |
| 12 | город А | 1050 | |
| 13 | город А | 1050 | |
| 14 | город А | 1050 | |
| 15 | город А | 1250 | |
| 16 | город А | 1250 | |
| 17 | город А | 1450 | |
| 18 | город А | 1450 | |
| 19 | город А | 1450 | |
| 20 | город А | 1450 | |
| 21 | город А | 1650 | |
| 22 | город А | 1650 | |
| 23 | город А | 1650 | |
| 24 | город А | 1650 | |

3. Определяется число групп ряда.

$$\text{формула Стерджесса} : n = 1 + 3,2 \lg N$$

| | A | B | C | D | E |
|---|-----------------|--|---|------------------------|---------------------------------------|
| | Наименование НП | Расстояние объекта до центра населенного пункта, м | | Объем совокупности (N) | Число групп по формуле Стерджесса (n) |
| 1 | | | | | |
| 2 | город А | 0 | | =СЧЁТ(B2:B47) | =ОКРУГЛВВЕРХ((1+(3,2*LOG(D2))),0) |
| 3 | город А | 0 | | | |
| 4 | город А | 250 | | | |
| 5 | город А | 450 | | | |
| 6 | город А | 450 | | | |
| 7 | город А | 450 | | | |
| 8 | город А | 850 | | | |

4. Число групп округляется до ближайшего целого числа

Число групп $n = 6$

| | A | B | C | D | E | F |
|---|-----------------|--|---|------------------------|---------------------------------------|------------------------|
| | Наименование НП | Расстояние объекта до центра населенного пункта, м | | Объем совокупности (N) | Число групп по формуле Стерджесса (n) | Величина интервала (h) |
| 1 | | | | | | |
| 2 | город А | 0 | | 46 | 6 | 1225,00 |
| 3 | город А | 0 | | | | |
| 4 | город А | 250 | | | | |
| 5 | город А | 450 | | | | |
| 6 | город А | 450 | | | | |
| 7 | город А | 450 | | | | |
| 8 | город А | 850 | | | | |

5. Определяется интервал (шаг интервала)

Значения x_i называются вариантами, а их последовательность, записанная в возрастающем порядке, — вариационным рядом.

Числа n_i называются частотами, а их отношения к объему выборки относительными частотами

$$W_i = \frac{n_i}{n}$$

Сумма относительных частот равна единице: $\sum W_i = 1$

При этом $\sum n_i = n$

Найдем объем выборки

$$n = 2 + 4 + 5 + 6 + 3 = 20$$

Относительные частоты равны:

$$W_1 = 2/20 = 0.1, W_2 = 4/20 = 0.2, W_3 = 5/20 = 0.25,$$

$$W_4 = 6/20 = 0.3, W_5 = 3/20 = 0.15$$

Скриншот интерфейса Microsoft Excel, демонстрирующий расчет интервала (шага интервала) в ячейке F2. В строке формул отображена формула: $=((\text{МАКС}(\text{B2:B47})-\text{МИН}(\text{B2:B47}))/\text{E2})$.

| | A | B | C | D | E | F |
|---|-----------------|--|---|------------------------|---------------------------------------|----------------------------------|
| | Наименование НП | Расстояние объекта до центра населенного пункта, м | | Объем совокупности (N) | Число групп по формуле Стерджесса (n) | Величина интервала (h) |
| 1 | | | | | | |
| 2 | город А | 0 | | =СЧЁТ(B2:B47) | =ОКРУГЛВВЕРХ((1+(3,2*E2))) | =((МАКС(B2:B47)-МИН(B2:B47))/E2) |
| 3 | город А | 0 | | | | |
| 4 | город А | 250 | | | | |
| 5 | город А | 450 | | | | |
| 6 | город А | 450 | | | | |
| 7 | город А | 450 | | | | |
| 8 | город А | 850 | | | | |

5. Определяется интервал (шаг интервала)

Зная величину интервала группировки, обозначим границы групп.

The screenshot shows the Microsoft Excel interface. The ribbon is set to 'Вставка' (Insert) > 'Таблицы' (Tables). The 'Сводная таблица' (PivotTable) icon is circled in red. Below the ribbon, the formula bar shows 'M8' and 'fx'. The main grid contains a table with the following data:

| | A | B | C | D | E | F | G | H | I | J | K |
|----|-----------------|--|---|---|---|---|---|---|---|---|---|
| 1 | Наименование НП | Расстояние объекта до центра населенного пункта, м | | | | | | | | | |
| 2 | город А | 0 | | | | | | | | | |
| 3 | город А | 0 | | | | | | | | | |
| 4 | город А | 250 | | | | | | | | | |
| 5 | город А | 450 | | | | | | | | | |
| 6 | город А | 450 | | | | | | | | | |
| 7 | город А | 450 | | | | | | | | | |
| 8 | город А | 850 | | | | | | | | | |
| 9 | город А | 850 | | | | | | | | | |
| 10 | город А | 1050 | | | | | | | | | |
| 11 | город А | 1050 | | | | | | | | | |
| 12 | город А | 1050 | | | | | | | | | |
| 13 | город А | 1050 | | | | | | | | | |
| 14 | город А | 1050 | | | | | | | | | |
| 15 | город А | 1250 | | | | | | | | | |
| 16 | город А | 1250 | | | | | | | | | |
| 17 | город А | 1450 | | | | | | | | | |
| 18 | город А | 1450 | | | | | | | | | |
| 19 | город А | 1450 | | | | | | | | | |

The 'Создание сводной таблицы' (Create PivotTable) dialog box is open, showing the following options:

- Выберите данные для анализа (Select data for analysis):
 - Выбрать таблицу или диапазон (Select table or range): Лист3!\$B\$1:\$B\$47
 - Использовать внешний источник данных (Use external data source):
 - Выбрать подключение... (Select connection...)
 - Имя подключения: (Connection name:)
- Укажите, куда следует поместить отчет сводной таблицы: (Specify where to place the PivotTable report):
 - На новый лист (On new sheet)
 - На существующий лист (On existing sheet): Диапазон: Лист3!\$D\$1

Buttons: OK, Отмена (Cancel).

Строим сводную таблицу в Excel

The screenshot displays the Microsoft Excel interface with a PivotTable and the PivotTable Fields task pane. The PivotTable is located in the range D1:J22 and has the following data:

| Наименование НП | Расстояние объекта до центра населенного пункта, м | Сумма по полю Расстояние объекта до центра населенного пункта, м |
|-----------------|--|--|
| город А | 0 | 97900 |
| город А | 0 | |
| город А | 250 | |
| город А | 450 | |
| город А | 450 | |
| город А | 450 | |
| город А | 850 | |
| город А | 850 | |
| город А | 1050 | |
| город А | 1050 | |
| город А | 1050 | |
| город А | 1050 | |
| город А | 1050 | |
| город А | 1250 | |
| город А | 1250 | |
| город А | 1450 | |
| город А | 1450 | |
| город А | 1450 | |
| город А | 1450 | |
| город А | 1650 | |
| город А | 1650 | |
| город А | 1650 | |

The PivotTable Fields task pane on the right shows the following configuration:

- Выберите поля для добавления в отчет: Расстояние объекта до центра нас...
- Перетащите поля между указанными ниже областями:
 - Фильтр отчета: (empty)
 - Названия столбцов: (empty)
 - Названия строк: (empty)
 - Значения: Σ Сумма по полю Р...
- Отложить обновление макета
- Обновить

Строим сводную таблицу в Excel

Сводная таблица: Сумма по полю Расстояние объекта до центра населенного пункта, м

| Наименование НП | Расстояние объекта до центра населенного пункта, м |
|-----------------|--|
| город А | 0 |
| город А | 0 |
| город А | 250 |
| город А | 450 |
| город А | 450 |
| город А | 450 |
| город А | 850 |
| город А | 850 |
| город А | 1050 |
| город А | 1050 |
| город А | 1050 |
| город А | 1050 |
| город А | 1050 |
| город А | 1250 |
| город А | 1250 |
| город А | 1450 |
| город А | 1450 |
| город А | 1450 |
| город А | 1450 |
| город А | 1650 |
| город А | 1650 |
| город А | 1650 |

Сводная таблица: Сумма по полю Расстояние объекта до центра населенного пункта, м

| Сумма по полю Расстояние объекта до центра населенного пункта, м |
|--|
| 97900 |

Список полей сводной таблицы

Выберите поля для добавления в отчет:

- Расстояние

Добавить в значения

Перетащите поля между указанными ниже областями:

Фильтр отчета: Названия столбцов

Названия строк: Значения

Сумма по полю Р...

Отложить обновление макета Обновить

Строим группировку в Excel

The screenshot shows the Microsoft Excel interface with the 'Данные' (Data) ribbon selected. The 'Группировать' (Group) button is highlighted in the ribbon. A task pane on the right side of the window is open, displaying the 'Группировать (Shift+Alt+стрелка вправо)' dialog box. The dialog box contains a preview of the data table and instructions on how to group cells.

The data table in the background is as follows:

| Наименование НП | Расстояние объекта до центра населенного пункта, м | Сумма по полю | Расстояние объекта до центра населенного пункта, м |
|-----------------|--|-------------------|--|
| город А | 0 | 0 | 0 |
| город А | 0 | 250 | 250 |
| город А | 250 | 450 | 1350 |
| город А | 450 | 850 | 1700 |
| город А | 450 | 1050 | 5250 |
| город А | 450 | 1250 | 2500 |
| город А | 850 | 1450 | 5800 |
| город А | 850 | 1650 | 14850 |
| город А | 1050 | 1850 | 3700 |
| город А | 1050 | 2050 | 2050 |
| город А | 1050 | 2250 | 2250 |
| город А | 1050 | 2450 | 9800 |
| город А | 1050 | 2650 | 13250 |
| город А | 1250 | 5750 | 5750 |
| город А | 1250 | 7350 | 29400 |
| город А | 1450 | Общий итог | 97900 |
| город А | 1450 | | |
| город А | 1450 | | |
| город А | 1450 | | |
| город А | 1650 | | |
| город А | 1650 | | |
| город А | 1650 | | |

The task pane on the right shows the 'Группировать' dialog box with the following content:

Группировать (Shift+Alt+стрелка вправо)

Объединение диапазона ячеек таким образом, что все эти ячейки можно одновременно свернуть или расширить.

Preview of the data table:

| | | |
|---|-------|----|
| 2 | A | 2 |
| 3 | B | 5 |
| 4 | C | 7 |
| 5 | A+B+C | 14 |

Peretасщите поля между указанными ниже областями:

- Фильтр отчета
- Названия столбцов
- Названия строк
- Значения

Example configuration in the dialog:

- Названия строк: Расстояние объе...
- Значения: Сумма по полю Р...

Buttons: Отложить обновление макета, Обновить

Строим группировку в Excel

The screenshot displays the Microsoft Excel interface with a PivotTable and a dialog box for grouping data.

Панель инструментов: Включает вкладки: **Файл**, **Главная**, **Вставка**, **Разметка страницы**, **Формулы**, **Данные**, **Рецензирование**, **Вид**, **Параметры**, **Конструктор**. Основные группы команд: **Получение внешних данных**, **Подключения**, **Сортировка и фильтр**, **Работа с данными**, **Структура**, **Анализ**.

Панель формул: L12, fx, 0

Таблица данных:

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|----|---------|------|---|---|---|---|---|---|---|---|---|------|---|
| 5 | город А | 450 | | | | | | | | | | | |
| 6 | город А | 450 | | | | | | | | | | | |
| 7 | город А | 450 | | | | | | | | | | | |
| 8 | город А | 850 | | | | | | | | | | | |
| 9 | город А | 850 | | | | | | | | | | | |
| 10 | город А | 1050 | | | | | | | | | | | |
| 11 | город А | 1050 | | | | | | | | | | | |
| 12 | город А | 1050 | | | | | | | | | | | |
| 13 | город А | 1050 | | | | | | | | | | | |
| 14 | город А | 1050 | | | | | | | | | | | |
| 15 | город А | 1250 | | | | | | | | | | | |
| 16 | город А | 1250 | | | | | | | | | | | |
| 17 | город А | 1450 | | | | | | | | | | | |
| 18 | город А | 1450 | | | | | | | | | | | |
| 19 | город А | 1450 | | | | | | | | | | | |
| 20 | город А | 1450 | | | | | | | | | | | |
| 21 | город А | 1650 | | | | | | | | | | 2050 | |
| 22 | город А | 1650 | | | | | | | | | | 2250 | |
| 23 | город А | 1650 | | | | | | | | | | 2450 | |
| 24 | город А | 1650 | | | | | | | | | | 2650 | |
| 25 | город А | 1650 | | | | | | | | | | 5750 | |
| 26 | город А | 1650 | | | | | | | | | | 7350 | |
| 27 | город А | 1650 | | | | | | | | | | | |
| 28 | город А | 1650 | | | | | | | | | | | |
| 29 | город А | 1650 | | | | | | | | | | | |

Панель PivotTable: Названия строк: **Сумма по полю Расстояние объекта до центра населенного пункта**

Диалог "Группирование":

- Авто
- начинать с: 0
- по: 7350
- с шагом: 1225
- Кнопки: **ОК**, **Отмена**

Панель "Список полей сводной таблицы":

- Выберите поля для добавления в отчет:
- Расстояние объекта до центра населен...

Панель "Перетащите поля между указанными ниже областями":

- Фильтр отчета
- Названия столбцов
- Названия строк
- Значения
- Расстояние объе... (в выпадающем списке)
- Сумма по полю Р... (в выпадающем списке)
- Отложить обновление макета
- Обновить

Строим группировку в Excel, определяя виды расчетов внутри групп

The screenshot shows an Excel spreadsheet with a PivotTable. The PivotTable is structured as follows:

| Названия строк | Сумма по полю Расстояние объекта до центра населенного пункта, м | Количество по полю Расстояние объе |
|----------------|--|------------------------------------|
| город А | 0-524 | 1600 |
| город А | 525-10 | 1700 |
| город А | 1050-1 | 13550 |
| город А | 1575-2 | 20600 |
| город А | 2100-2 | 12050 |
| город А | 2625-3 | 13250 |
| город А | 5250-5 | 5750 |
| город А | 6825-7 | 29400 |
| Общи | | 97900 |

A dialog box titled "Параметры поля значений" (Value Field Settings) is open, showing the following options:

- Имя источника: Расстояние объекта до центра населенного...
- Пользовательское имя: Среднее по полю Расстояние объекта до центра населенного пунк...
- Операция: **Среднее** (Average)
- Дополнительные вычисления: (empty)

The PivotTable Fields task pane on the right shows the following configuration:

- Выберите поля для добавления в отчет:
 - Расстояние объекта до центра населен...
- Перетащите поля между указанными ниже областями:
 - Фильтр отчета: (empty)
 - Названия столбцов: Σ Значения
 - Названия строк: Расстояние объе...
 - Значения: Σ Значения
 - Сумма по полю Р...
 - Количество по п...
- Отложить обновление макета
- Обновить

6. Определяются сводные и обобщающие показатели, вначале по каждой группе, а затем по всей совокупности.

| | I | J | K | L | M | N |
|---|----------------|---|---|---|---|--|
| 1 | Названия строк | Количество по полю Расстояние объекта до центра населенного пункта, м | Сумма по полю Расстояние объекта до центра населенного пункта, м2 | Среднее по полю Расстояние объекта до центра населенного пункта, м2 | Смешенная дисперсия по полю Расстояние объекта до центра населенного пункта, м2 | Смешенное отклонение по полю Расстояние объекта до центра населенного пункта, м2 |
| 2 | 0-1049 | 8 | 3300 | 412,5 | 107678,5714 | 328,1441321 |
| 3 | 1050-2099 | 23 | 34150 | 1484,782609 | 86007,90514 | 293,2710438 |
| 4 | 2100-3149 | 10 | 25300 | 2530 | 19555,55556 | 139,8411798 |
| 5 | 5250-6299 | 1 | 5750 | 5750 | #ДЕЛ/0! | #ДЕЛ/0! |
| 6 | 6300-7350 | 4 | 29400 | 7350 | 0 | 0 |
| 7 | Общий итог | 46 | 97900 | 2128,26087 | 3548739,13 | 1883,809738 |
| 8 | | | | | | |

1. Сгруппированы показатели (5 групп вместо 6, т.к. имеется группа с интервалом 3675-4900 в которой нет значений показателя)
2. Рассчитано количество показателей в каждой группе (столбец J)
3. Рассчитана сумма показателей по каждой группе (столбец K)
4. Рассчитаны средние значения в каждой группе (столбец L)
5. Рассчитана дисперсия и отклонение по каждой группе (столбец M, N)

6. Определяются сводные и обобщающие показатели, вначале по каждой группе, а затем по всей совокупности.

Расчет показателей по генеральной совокупности

| | D | E | F | G |
|----|---------------------------------|---------------------------|------------------------|--------------------------------|
| 9 | | | | |
| 10 | средняя арифметическая | мода | медиана | среднее линейное отклонение |
| 11 | =СРЗНАЧ(B2:B47) | =МОДА.ОДН(B2:B47) | =МЕДИАНА(B2:B47) | =СРОТКЛ(B2:B47) |
| 12 | | | | |
| 13 | среднеквадратическое отклонение | размах вариационного ряда | Коэффициент осцилляции | Коэффициент вариации |
| 14 | =СТАНДОТКЛОН.Г(B2:B47) | =МАКС(B2:B47)-МИН(B2:B47) | =E14/D11 | =D14/D11 |
| 15 | | | | |
| 16 | | | | Линейные коэффициенты вариации |
| 17 | | | | =D14/F11 |
| 18 | | | | |

Расчет показателей по генеральной совокупности

| | C | D | E | F | G | H |
|----|---|---------------------------------|---------------------------|------------------------|--------------------------------|---|
| 9 | | | | | | |
| 10 | | средняя арифметическая | мода | медиана | среднее линейное отклонение | |
| 11 | | 2 128,26 | 1 650,00 | 1 650,00 | 1 240,26 | |
| 12 | | | | | | |
| 13 | | среднеквадратическое отклонение | размах вариационного ряда | Коэффициент осцилляции | Коэффициент вариации | |
| 14 | | 1 863,22 | 7 350,00 | 3,45 | 0,88 | |
| 15 | | | | | | |
| 16 | | | | | Линейные коэффициенты вариации | |
| 17 | | | | | 1,13 | |
| 18 | | | | | | |

Совокупность считается не засоренной, если выполняются условия:

Должно быть:

- ❖ значение медианы находится между средневзвешенной и модой;
- ❖ отношение среднеквадратического отклонения к среднему линейному отклонению $\leq 1,25$;
- ❖ коэффициенты вариации $\leq 33\%$.

Факт:

- ❖ значение медианы находится между средневзвешенной и модой (*истина*);
- ❖ отношение среднеквадратического отклонения к среднему линейному отклонению $\leq 1,25$ (*факт > 1,25*);
- ❖ коэффициенты вариации $\leq 33\%$ (*факт > 33%*).



Кубанский государственный
аграрный университет

Землеустроительный
факультет

Благодарю за внимание!