Метод рангов (метод Спирмена)

Y - исследуемый признак #2
X - исследуемый признак #1

		X	Y	Ранг x_1	Ранг y_1	d	d^2
Вариант - 2		100	200	2	1	1	1
Вариант - 1		250	150	1	2	1	1

порядковый номер — Ранг
Разность рангов — d

x_1 - величина признака X принятая за порядковый номер (250 > 100 ~ 250 = 1)
y_1 - величина признака Y принятая за порядковый номер (200 > 150 ~ 200 = 1)

Метод квадратов (метод Пирсона)

Y - исследуемый признак #2
X - исследуемый признак #1

dy^2 - квадрат отклонения по признаку #2
dx^2 - Квадрат отклонения по признаку #1

X	Y	dx	dy	dx x dy	dx^2	dy^2
100	200	6	7	8	9	10
259	150	6	7	8	9	10

dx - отклонение каждой варианты от среднего значения по признаку №1
dy - отклонение каждой варианты от среднего значения по признаку №2
dx * dy - произведение отклонений каждой варианты по признаку №1 и №2

Mx - средняя величина по признаку X
Mx - сумма вариант / количество = 350/2 = 175

My - средняя величина по признаку X
My = 350/2 = 175

Σx - сумма всех вариант по признаку X
Σy - сумма всех вариант по признаку Y
$\Sigma dx * dy$ - сумма произведений отклонений

Алгоритм работы с методом квадратов

Σx - сумма всех вариант по признаку X - по армосферному давлению
Σy - сумма всех вариант по признаку Y - по систолическому давлению
Mx - среднее значение армосферного давления
My - среднее значение систолического давления
$\Sigma dx * dy$ - сумма произведений отклонений по двум признакам
Σdx^2 - сумма квадратов средних отклонений по атмосферному давлению
Σdy^2 - сумма квадратов средних отклонений по систолическому давлению
$dx^2 * dy^2$ - произведение квадратов средних отклонений по двум признакам
$\sqrt{(dx^2 * dy^2)}$ - вычисляем корень из произведения квадратов средних отклонений
r_{xy} - вычисляем коэффициент корреляции
r_{xy}^2 - возводим в квадрат коэффициент корреляции
M_r - вычисляем среднюю ошибку коэффициента корреляции
t - вычисляем t-критерий Стьюдента

Метод квадратов

r_{xy} — Коэффициент корреляции

$$r_{xy} = \frac{\sum d_x \times d_y}{\sqrt{d_x^2 \times d_y^2}}$$

m_r — Средняя ошибка коэффициента корреляции

$$m_r = \sqrt{\frac{1 - r_{xy}^2}{n - 2}}$$

$$t = \frac{r_x}{m_r}$$

Метод рангов

p — Достоверность

$$p = \frac{6 \times \sum d^2}{n (n^2 - 1)}$$

m_p — Коэффициент ранговой корреляции

$$m_p = \sqrt{\frac{1 - p^2}{n - 1}}$$

$$t = \frac{p}{m_p}$$

=СУММ(A1:A2)
=СРЗНАЧ(A1:A2)
=МОДА.ОДН(A1:A2)
=МЕДИАНА(A1:A2)
=КОРРЕЛ(A1:A2;A3:A4)

Оценка корреляционной связи

Сила связи	Прямая (+)	Обратная (-)
Полная	+1	-1
Сильная	От +1 до 0,7	От -1 до -0,7
Средняя	От +0,7 до 0,3	От -0,7 до -0,3
Слабая	От +0,3 до 0	От -0,3 до 0
Отсутствует связь	0	0

Группировки

Статистическая группировка - это процесс образования **однородных групп** на основе расчленения статистической совокупности на части или объединения изучаемых единиц в частные совокупности **по существенным для них признакам**, каждая из них характеризуется системой статистических показателей.

Структурная группировка позволяет изучить структуру совокупности, структурные сдвиги.
Аналитическая группировка позволяет выявить связи и зависимости между отдельными признаками явления.
Основание группировки - признаки, по которым производится распределение единиц изучаемой совокупности на группы, называются **группированными признаками**, или **основанием группировки.**

В зависимости от вида группировочных признаков бывают:
1. по атрибутивным признакам
2. по количественным признакам

В зависимости от степени сложности бывают:
1. простые (по 1 признаку)
2. сложные (по 2 + признакам)
3. комбинационные
4. многомерные

Корреляционная связь

Корреляционная связь – соотношение, при котором значению каждой средней величины **одного признака** соответствует несколько значений **другого взаимосвязанного с ним признака**.
Коэффициент корреляции r_{xy} – число, которое дает **представление о направлении и силе связи между явлениями**.
Прямолинейная связь – относительно равномерное изменение средних значений одного признака при равных изменениях другого.

Криволинейная зависимость – при равномерном изменении одного признака могут наблюдаться возрастающие и убывающие средние значения другого признака.

частота сердечных сокращений → **Прямая** → температура тела, t

количество знаний по теме → **Обратная** → количество плохих оценок

Экстенсивный

Указывает на **отношение** части к целому.

Характеризует только **состав** совокупности

Пример: отношение числа студентов лечебного факультета к общему количеству студентов
***Одна среда - студенты**

Показатель соотношения

Указывает на отношение **двух самостоятельных** совокупностей

Пример: отношение количества терапевтических коек к численности населения (две не связанные совокупности). В результате поиска значения узнаем обеспеченность населения терапевтическими койками
***Две разные среды (койки - люди)**

Интенсивный

Указывает на **распространенность**
Используются для **характеристики явления**

Пример: отношение зарегистрированных заболеваний (**явление**) к общей численности населения (**среда**)

Показатель наглядности

Указывает на **сколько %**, **во сколько раз произошло** уменьшение / увеличение

Сравнение в динамике

Пример: считаем количество госпитализаций в 2 городах. Берем за 100% исходное значение показателя (за первый год), а остальные показатели (за последующие 5 лет) пересчитываем в процентах по отношению к нему, аналогично делаем для другого города. Сравниваем и делаем вывод, где показатель растет,а где падает.

Соотношение из формулы Стерджесса

N	15 - 24	25 - 44	45 - 89	90 - 179	180 - 359	360 - 719
n	5	6	7	8	9	10

N - число единиц совокупности, n - количество групп