

# Тема 8

## **МЕТОДЫ ОБРАБОТКИ И АНАЛИЗА ПЕРВИЧНОЙ СОЦИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ**

# ПЛАН

1. Проблемы количественного измерения в социологии
2. Этапы и процедуры первичной обработки данных
3. Методы математической и статистической обработки баз данных
4. Проверка гипотез
5. Табличное и графическое представление данных

# Понятие шкалы

- **Шкалы** – это алгоритмы, по которым производится отображение изучаемых социальных объектов в ту или иную числовую математическую систему.
- Шкалы подразделяются на метрические (интервальная и пропорциональная) и неметрические (номинальная и ранговая).

# Типы шкал

- **Номинальная:** с помощью номинальной шкалы мы измеряем такие переменные, которые в принципе не могут количественно отличаться друг от друга.
- **Ранговая:** данные рангового уровня измерений включают в себя категории наблюдения, которые размещены по порядку (от большего значения измеряемого признака к меньшему его значению или наоборот – от меньшего к большему).
- **Интервальная:** значения переменных, измеряемых с помощью интервальных шкал, изменяются непрерывно, они представляют собой численные величины.
- **Пропорциональная шкала** – та же интервальная шкала, но с зафиксированной точкой начала отсчета.

# Выявление центральных тенденций

- Уточнения того, какие из значений переменных встречаются в линейных распределениях наиболее часто, а значит, определяют общие (или центральные) закономерности
- Соответствующие такому анализу таблицы называют *линейными* или *одномерными* распределениями
- *Модой* или *модальной категорией* называется то значение переменной, которое встречается среди данных наиболее часто
- *Медиана* – категория или значение в распределении значений, лежащих выше и ниже того уровня, на который приходится половина всех частот
- Критериями центральной тенденции для интервального и пропорционального уровней измерений выступают и мода, и медиана, и *среднее арифметическое* – сумма значений переменной, поделенная на число значений

# Показатели разброса

*Дисперсия* ( $\sigma^2$ ) представляет собой сумму квадратов отклонений от среднего, поделенную на число отметок:

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{N}.$$

*Среднеквадратическое отклонение* (S) представляет собою корень квадратный из дисперсии:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{N}}.$$

Чем больше разброс точек данных вокруг среднего, тем выше значение  $\sigma^2$  и S. Это означает, что если все данные одинаковы, то  $\sigma^2$  и S равны нулю.

# Шаги для вычисления $\sigma^2$ и S

- 1) вычислить среднее;
- 2) вычислить разности между средним и каждым из значений;
- 3) возвести в квадрат разности, вычисленные на этапе 2;
- 4) умножить квадраты разностей на частоты наблюдений каждого из значений;
- 5) просуммировать квадраты разностей, вычисленные на этапе 4;
- 6) разделить сумму квадратов из этапа 5 на N; это равняется дисперсии;
- 7) извлечь квадратный корень из числа, вычисленного на этапе 6; это и будет среднеквадратическое отклонение.