

Разведочный анализ (Exploratory analysis)

Разделы: [Бизнес-задачи](#)

Разведочный анализ — это предварительный анализ данных с целью выявления наиболее общих зависимостей, закономерностей и тенденций, характера и свойств анализируемых данных, законов распределения анализируемых величин. Применяется для нахождения связей между переменными в ситуациях, когда отсутствуют (или недостаточны) априорные представления о природе этих связей.

Как правило, при разведочном анализе учитывается и сравнивается большое число признаков, а для поиска закономерностей используются самые разные методы.

Термин «разведочный анализ» был впервые введен математиком из Принстонского университета Дж. Тьюки. Он также сформулировал основные **цели** данного анализа:

- Максимальное «проникновение» в данные.
- Выявление основных структур.
- Выбор наиболее важных переменных.
- Обнаружение отклонений и аномалий.
- Проверка основных гипотез (предположений).
- Разработка начальных моделей.

Результаты разведочного анализа не используются для выработки управленческих решений. Их назначение — помощь в разработке наилучшей стратегии углубленного анализа, выдвижение гипотез, уточнение особенностей применения тех или иных математических методов и моделей. Без разведочного анализа углубленный анализ данных будет производиться практически «вслепую».

К основным методам разведочного анализа относится процедура анализа распределений переменных, корреляционный анализ с целью поиска коэффициентов, превосходящих по величине определенные пороговые значения, факторный анализ, дискриминантный анализ, многомерное шкалирование, визуальный анализ гистограмм и т.д.

Предварительное исследование данных может служить лишь первым этапом в процессе их анализа, и пока результаты не подтверждены на других выборках или на независимом множестве данных, их следует воспринимать самое большее как гипотезу. Если результаты разведочного анализа говорят в пользу некоторой модели, то ее правильность можно затем проверить, применив к новым данным.

