

ЛЕКЦИЯ 1. Регрессионный анализ как путь проверки гипотез

Введение: зачем нужен регрессионный анализ

Регрессионный анализ — это один из ключевых инструментов в аналитике, позволяющий **понять, как одна переменная влияет на другую**.

Он используется в:

маркетинге — чтобы понять, влияет ли реклама на продажи;

образовании — как стресс или сон влияют на академические результаты;

здоровье — как питание или гаджеты влияют на самочувствие;

управлении — влияет ли вовлечённость сотрудников на текучесть кадров.

Во всех этих задачах регрессия отвечает на вопрос:

Есть ли зависимость? Насколько она сильна? Как её интерпретировать?

 Почему мы начинаем не с формул

Для освоения аналитики применён подход «развития младенца»:

Сначала — практика, наблюдение, ошибки и размышления.

Потом — язык, структура, теория и формулы.

Первые 7 недель вы учитесь формулировать мысль, понимать суть данных, строить гипотезу — и только затем переходите к расчётам.

Так формируется настоящий аналитик — не только технический, но и смысловой.

Шаблон: как формулировать гипотезу

Перед тем как переходить к цифрам, нужно чётко понимать **что именно вы хотите проверить**. Ваша гипотеза — это не просто утверждение, а логика

мышления.

Структура:

1. Введение:

Что вы наблюдали? Почему это вас заинтересовало?

2. Проблема:

Какая несостыковка, сложность или конфликт в этом наблюдении?

3. Предположение:

Что вы предполагаете? Какая связь может существовать?

4. Фокус-группа:

Кто объект вашего наблюдения? (возраст, поведение и т.д.)

5. Источники данных:

Как вы получите информацию? (анкеты, интервью, логи)

6. Обработка данных:

Как переведёте текстовые ответы в числа?

Пример (для ориентира):

Введение: Студенты часто жалуются на усталость, но при этом перед сном активно используют телефоны.

Проблема: Они не осознают влияние экрана на качество сна.

Предположение: Чем чаще использование экрана — тем хуже качество сна.

Фокус-группа: Студенты 1–3 курсов.

Данные: Анкетирование через Google Forms.

Обработка: Оценки по шкале от 0 до 10.

Переход от текста к числам

Чтобы провести регрессионный анализ, нужно перевести текстовые ответы из анкет в **числовой формат**. Это позволит сравнивать группы и делать выводы.

Шаг 1: Определите переменные

Зависимая переменная (Y):

Качество сна (опрос: "Как вы оцениваете общее качество сна?")

Независимая переменная (X):

Использование экрана перед сном (опрос: "Как часто вы используете экран перед сном?")

Шаг 2: Присвойте числовые оценки

Ответ	Оценка
Отличное	10
Хорошее	7
Удовлетворительное	5
Плохое	2–3
Очень плохое	0

Использование экрана:

Ответ	Оценка
Никогда	10
Иногда	6
Часто	4
Всегда	0–2

Шаг 3: Постройте таблицу

Респондент	Экран (X)	Сон (Y)
1	2	5
2	6	7
3	0	2
4	8	8
5	4	4

Эта таблица — ваш **исходный массив данных**. Вы можете в любой момент дополнить или изменить его.

Построение простой линейной регрессии

Формула линейной регрессии выглядит так:

$$y = b_0 + b_1 * x$$

Где:

y — зависимая переменная (качество сна),

x — независимая переменная (время у экрана),

b_0 — свободный коэффициент (пересечение с осью y),

b_1 — коэффициент наклона (показывает, как x влияет на y).

Формула в Excel (русский язык):

```
=ЛИНЕЙН(B2:B6; A2:A6; ИСТИНА; ИСТИНА)
```

Формула в Google Sheets (английский язык):

```
=LINEST(B2:B6, A2:A6, TRUE, TRUE)
```

Где:

A2:A6 — значения X (время у экрана),

B2:B6 — значения Y (оценка сна).

Как интерпретировать результаты

Если $b_1 < 0$ — больше экранного времени → хуже сон

Если R^2 близко к 1 — модель хорошо объясняет данные

Если $R^2 < 0.3$ — слабая связь, модель объясняет мало

Лайфхаки по работе с ИИ-инструментами

Инструмент	Для чего
Excel Copilot	Автоматический расчёт модели, графики
Google Sheets AI	Построение сводных таблиц, предложения формул
Notion AI	Резюмирование таблиц, выделение ключевых выводов
Data Analyst GPT	Объяснение коэффициентов и выводов

Что категорически запрещено:

- ✗ Подставлять данные, не собранные самостоятельно
- ✗ Копировать интерпретации без понимания
- ✗ Просить ИИ “придумать” регрессию без данных
- ✗ Вставлять результаты без проверки логики

Заключение

Регрессионный анализ — это не про уравнения, а про мышление.

Он помогает **проверять свои идеи и строить аргументированную аналитику**, основанную на наблюдении, структуре и данных.

Готовы перейти ко второй лекции — **множественная регрессия и работа с несколькими переменными?**

Регрессия — это не “магия” и не “формула ради формулы”. Это способ **проверить свою мысль, основанную на наблюдении и гипотезе.**

Ваша задача — **не просто сделать расчёт, а понять, что он говорит.**

Регрессия — это не “магия” и не “формула ради формулы”. Это способ **проверить свою мысль, основанную на наблюдении и гипотезе.**

Ваша задача — **не просто сделать расчёт, а понять, что он говорит.**