



Проект по статистике (логистическая компания)

Задание 1

Менеджер логистической компании решил провести АБ-тест, выбрав в качестве метрики время доставки курьерами посылок. Для контрольной группы курьеров оставили предыдущий интерфейс навигации в приложении, а для тестовой группы внедрили новую опцию, позволяющую комбинировать несколько видов транспорта.

Полученные результаты времени доставки в минутах для тестовой и контрольной групп представлены в csv-файлах *times_couriers_control.csv* и *times_couriers_test.csv*.

Проверьте гипотезу менеджера о том, что внедрённая опция приложения уменьшила время доставки.

Подсказка: это независимые выборки.

Задание 2

В этой же компании с 1 мая часто заказывающим клиентам предоставили скидку, действующую на все заказы до конца календарного месяца, если в этом месяце клиент уже сделал четыре заказа (то есть начиная с пятого заказа).

В файле *clients_orders.csv* представлена информация о количестве заказов, которые одни и те же клиенты сделали в апреле и мае.

Проверьте гипотезу о том, что клиенты стали делать больше заказов после введения скидки.

Подсказка: это зависимая (парная) выборка.

Задание 3

В файле *couriers_education.csv* представлена анонимизированная информация о курьерах компании: образование (*education*) и количество лет работы в компании (*experience*).

Переменная *education* принимает значения:

- NS* — no school — нет аттестата об окончании школы
- MS* — middle school — закончены 9 классов школы
- S* — school — среднее образование: закончена средняя школа
- SSE* — secondary special education — среднее специальное образование (ПТУ / колледж)
- HE* — higher education — высшее образование

Переменная *experience* принимает значения:

- 0-1* — до одного года работы в компании
- 1-3* — от одного до трёх лет работы
- 3-5* — от трёх до пяти лет работы
- 5-* — от пяти лет работы

Проверьте гипотезу о том, что стаж работы в компании зависит от уровня образования.

Задание 4 (дополнительное)

Постройте 95%-е доверительные интервалы по выборкам из задания 1 для среднего времени доставки курьерами, использующими и не использующими новую опцию комбинирования нескольких видов транспорта в приложении.