1. Ваш продукт - это школа английского языка <http://english22.tilda.ws/>

Проанализируйте конкурентов на рынке (SkyEng, Lingvoleo и прочие) и сформулируйте гипотезу по улучшению первого экрана лендинг.

1. Первый экран – замена обложки на уникальную и запоминающуюся позволит уменьшить bounce rate на 10%
2. Первый экран – добавление шапки с контактами, ссылками и графиком работы колл-центра позволит увеличить конверсию на 5%
3. Первый экран – добавление меню для быстрого понимания структуры представленной на лендинге информации позволит уменьшить bounce rate на 5%
4. Дополнительно – добавление блока о компании, сравнение с конкурентами (индивидуальный подход)
5. Дополнительно – иллюстрация учебного процесса - видео с учениками
6. Подведите результаты эксперимента в экселе по следующим данным: **ab\_stats.csv.**

Стат значимо ли отличается ARPPU в двух группах ? Какие рекомендации дадите менеджеру?

Рекомендации менеджеру: В тестовой группе нет статистически значимых отличий по удержанию, при этом ARRPU заметно ниже. Попробовать альтернативные варианты экспериментов - исходный вариант лучше.

**3)** Мы хотим провести А/Б-тест для трех источников трафика. Нынешняя конверсия равна 5%, мы ожидаем прирост в 0,2%. Уровень доверия 97% и уровень мощности 87%. Всего на наш продукт заходит 40 000 пользователей в месяц.

**3.1)** За сколько дней мы сможем протестировать гипотезу? И что вы можете посоветовать по результатам подсчета?

1162 дней (больше 3 лет) – с учетом выборки в 1,522,706 человек (в каждом варианте: 507,569).

Слишком большая длительность теста, поэтому необходимо либо скорректировать ожидания от прироста, либо снизить порог достоверности/мощность.

**3.2)** Допустим в задаче нет проблемы с количеством посетителей на сайт, тогда подведите результаты тестирования, если у нас следующие результаты по количеству конверсии:

1) 25 000

2) 30 000

3) 32 000

Статистически значимые отличия отсутствуют – все варианты попадают в доверительный интервал друг друга.

**4)** Вы решили сравнивать метрику CPA в двух группах.

Размер выборки - 2350 элементов в каждой группе. Для проверки нормальности распределения на выборке в 2350 наблюдений применили , критерий Шапиро-Уилка и получили p-value, равный 0.00002, alpha = 5%. Какой бы вывод мы могли сделать в данном случае? В этом случае какой статистический критерий для проверки первоначальной гипотезы тут лучше всего подойдёт и почему ?

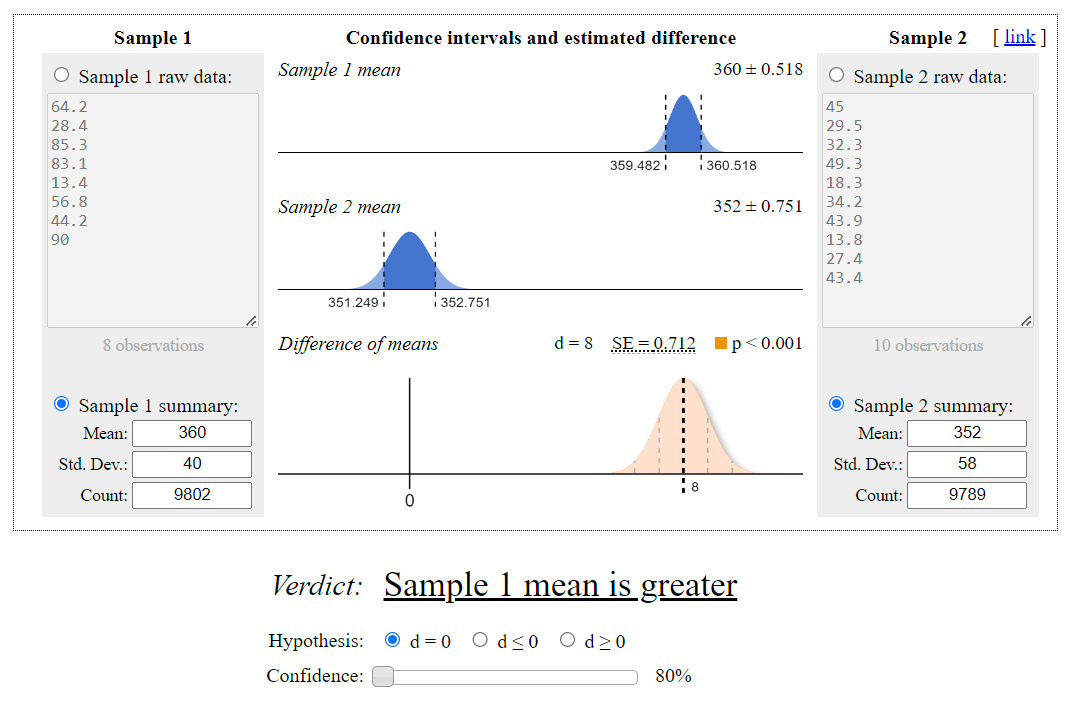
P-value заметно ниже уровня значимости – значит, распределение данных не нормальное.

Соответственно, лучше будет применить непараметрический критерий Манна-Уитни, т.к. он не требует нормального распределения данных для их проверки в двух независимых метриках. Также он менее чувствителен к выбросам и отклонениям от нормальности, т.к. использует сравнение медиан значений.

**5)** Мы провели АБ-тест на увеличение average timespent per user. По итогам тестирования мы получили следующие данные. Является ли результат статистически значимым с уровнем доверия 80%? Какую версию мы выкатим на продакшн?

A) Средняя - 360, отклонение - 40, количество - 9802

B) Средняя - 352, отклонение - 58, количество - 9789



С уровнем доверия 80% результат является статистически значимым – медиана варианта 1 по т-тесту больше, чем варианта 2 – пользователь проводит больше времени на сайте. Значит, именно вариант 1 будет отправлен на продакшн.

**6)** Создайте техническую архитектуру проекта по аб тестированию продукта он-лайн кинотеатра с учетом кросс-девайс аналитики по следующей гипотезе: Если договориться с банком о 99% кэшбэке на подписку первого месяца, то это повысит конверсию в подписку на 30%, благодаря упрощенному принятию решения со стороны пользователя.

На схеме необходимо отобразить:

1) Управленческий процесс по договоренностям с внешними партнерами

2) Архитектуру данных с указанием систем, из которых будем скачивать данные

3) Внутрикомандное взаимодействие