**Внимание!** Для выполнения тестовых заданий скачайте и откройте массив данных по ссылке:

<https://docs.google.com/spreadsheets/d/1EOEmGcBpokRfYbiNBDQs5XnWG9QGmOSwYKpKiOkhQR4/edit?usp=sharing>

1. Во вкладке "Данные об аудитории" информация о пользователях, посетивших наше приложение в ноябре. Чему равен MAU продукта?

\***MAU (Monthly Active Users)** — это метрика, используемая для измерения активности пользователей в течение одного месяца. Она показывает количество уникальных пользователей, которые взаимодействовали с продуктом, сервисом или приложением хотя бы один раз за последний месяц.

7639168141048216529

2. Используя вкладку "Данные об аудитории", посчитайте, чему будет равен DAU

\***DAU (Daily Active Users)** — это метрика, которая показывает количество уникальных пользователей, которые взаимодействовали с продуктом, приложением или сервисом хотя бы один раз в течение дня. DAU помогает понять, сколько пользователей активно пользуются продуктом каждый день.

255490560483

3. Используя вкладку "Данные об аудитории", посчитайте, чему будет равен retention первого дня у пользователей, пришедших в продукт 1 ноября

\*Retention (удержание пользователей) — это метрика, которая показывает, сколько пользователей продолжает пользоваться продуктом через определенный промежуток времени после первоначального взаимодействия. Retention можно рассчитать как процент пользователей, вернувшихся в продукт через определенное время (например, через 1 день, 1 неделю, 1 месяц) от количества всех новых пользователей.

28,3%26,6%38,5%32,7%

4. На графике изображены retention кривые 2 продуктов. Какие выводы можно сделать, глядя на них?



Ваш ответ:

5. Во вкладке "Данные об аудитории" есть информация о том, сколько объявлений посмотрел каждый пользователь (view\_adverts). Посчитайте пользовательскую конверсию в просмотр объявления за ноябрь? (в пользователях)

\* Пользовательская конверсия — это метрика, которая показывает, какой процент пользователей выполнил целевое действие по отношению к общему количеству пользователей. В контексте веб-сайтов это может быть действие, такое как просмотр объявления или клик по рекламному баннеру.

41,8%54,7%46,3%39%

6. Используя информацию из вкладки "Данные об аудитории", посчитайте среднее количество просмотренных объявлений на пользователя в ноябре

4,96,25,32,9

7. Мы провели опрос среди 2000 пользователей. Из них 500 «критики», 1200 «сторонники» и 300 «нейтралы». Посчитайте, чему будет равен NPS

\*NPS (Net Promoter Score) — это метрика, которая измеряет лояльность пользователей к компании или продукту и делит их на три группы: Сторонники (Promoters) , Нейтралы (Passives), Критики (Detractors). NPS высчитывается как (% сторонников - % критиков).

30%43%40%35%

8. Во вкладке "Данные АБ-тестов" результаты трех несвязанных АБ тестов для ARPU (общая выручка/общее количество пользователей).  
Посмотрите на результаты тестов и интерпретируйте их. Напишите значения p-value, которые вы получили.  
Подготовьте выводы и рекомендации.   
  
experiment\_num - номер эксперимента  
experiment\_group - группа, в которую попал пользователь  
user\_id - id пользователя  
revenue - выручка, которую сгенерировал пользователь, купив платную услугу продвижения

Ваш ответ:

9. По датасету с листерами посчитайте средний доход на пользователя

121.2156.470.930.7средняя здесь не применима

10. По датасету с листерами посчитайте медиану возраста пользователя

27,422827,9327медиана здесь не применима

11. Какой график лучше всего подходит для отображения разброса цен на товары в разных магазинах?  
\*возможно несколько вариантов ответа

Линейный графикКруговая диаграммаЯщик с усами (box plot)

Гистограмма

12. На каком графике бимодальное распределение?

№1

№2

№3

№4

13. Какая случайная величина имеет наибольшую дисперсию данных по следующим графикам плотности распределения?

№1

№2

№3

№4

14. На каком графике можно посчитать коррелцияю?  
\*возможно несколько вариантов ответа









15. Что значит, если при проверке гипотез мы получили p-value = 0.05?

Это означает, что нет никакой статистически значимой разницы между группами

Есть 5% вероятность случайно получить такой или еще более экстремальный результат, если нулевая гипотеза верна

Это означает, что результаты эксперимента на 95% точны

Это говорит о том, что альтернативная гипотеза верна с вероятностью 95%

16. Какой метод наиболее подходит для проверки гипотезы о равенстве средних двух выборок из нормального распределения?

t-тестХи-квадрат тестАнализ дисперсии (ANOVA)Корреляция Пирсона

17. Как интерпретировать квартили в распределении доходов пользователей?

Показывают максимальный и минимальный доход

Делят данные на четыре равные части

Указывают на наиболее часто встречающийся доход

График плотности распределения вещества во вселенной

18. Были получены следующие результаты. Коллеги просят вас подтвердить их и сделать окончательный вывод по эксперименту.

* + - * Вариант A (контрольная группа) — 100 047 501 посетитель, 1003 платежа.
      * Вариант B (тестовая группа) — 100 001 055 посетителей, 1099 платежей.

Какие рекомендации вы бы дали, основываясь на этих данных?

Ваш ответ:

1. MAU (Monthly Active Users): 7639

2. DAU (Daily Active Users) (в среднем): 560.5

3. Retention 1 дня (для 1 ноября): 26.6%

4. Выводы по графику:

* Голубая кривая (продукт 1) показывает более высокое удержание пользователей на протяжении всех 7 дней. Пользователи возвращаются чаще и остаются дольше.
* Красная кривая (продукт 2) быстро теряет пользователей — к 5 дню почти все ушли. Это говорит о низком качестве пользовательского опыта или интереса.
* Разрыв между кривыми увеличивается с каждым днём, что указывает на стабильное преимущество продукта 1 в удержании аудитории.

5. Конверсия в просмотр объявления: 46.3%

6. Среднее количество просмотренных объявлений на пользователя: 2.9

7. NPS (Net Promoter Score): 35.0%

8. Эксперимент 1

• Средняя выручка (Test): 665.74

• Средняя выручка (Control): 722.46

• p-value: 0.689

Вывод: Разница статистически незначима. Изменения не повлияли

на выручку. Рекомендация: оставить текущую версию.

Эксперимент 2

• Средняя выручка (Test): 332.93

• Средняя выручка (Control): 704.65

• p-value: 0.0011

Вывод: Статистически значимое ухудшение в тестовой группе.

Рекомендация: отклонить тестовое изменение, оно снижает доход.

Эксперимент 3

• Средняя выручка (Test): 998.67

• Средняя выручка (Control): 663.21

• p-value: 0.0603

Вывод: Тестовая версия показывает лучше результат, но p-value

немного выше 0.05. Рекомендация: провести дополнительный сбор

данных, возможно, при увеличении выборки результат станет

значимым.

9. Средний доход на пользователя: 156.5

Правильный вариант: 156.4 (округлено в тесте)

10. Медиана возраста пользователя: 28.0 Правильный вариант: 28

11. Ящик с усами (box plot) и Гистограмма

12. На каком графике бимодальное распределение?

Бимодальность — это две выраженные “горки” на графике плотности.

Ответ зависит от изображения графиков (ожидается визуальный выбор).

Если есть 4 графика (например, №1–4), то правильный будет тот, где

видно 2 пика.

13. Какая случайная величина имеет наибольшую дисперсию данных по

следующим графикам плотности распределения? Наибольшая

дисперсия — это график с наиболее широким и “плоским”

распределением.

14. На каком графике можно посчитать корреляцию?

Правильный вариант: Диаграмма рассеяния (scatter plot) — только на ней

можно оценить степень линейной связи между двумя переменными.

15. Есть 5% вероятность случайно получить такой или еще более

экстремальный результат, если нулевая гипотеза верна

16. Ответ: t-тест

17. Ответ: Делят данные на четыре равные части

18. По данным:

• Группа A: 100 047 501 посетитель, 1003 платежа

• Группа B: 100 001 055 посетителей, 1099 платежей

Рассчитаем конверсии:

• A: 1003 / 100047501 ≈ 0.001002

• B: 1099 / 100001055 ≈ 0.001099

Разница минимальна, при таком большом размере выборки нужна

статистическая проверка, но абсолютное изменение невелико.

Вывод: несмотря на чуть более высокую конверсию у варианта B,

улучшение незначительное. Рекомендуется либо собрать

дополнительные данные, либо не внедрять изменение без

подтверждения статистической значимости.