**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

Ответ: 7639

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:

На графике показано, что продукт A (голубая линия) удерживает пользователей лучше, чем продукт B (красная линия). У продукта A на всех днях выше процент возвращающихся пользователей, и снижение происходит более плавно. К 4–7 дню retention стабилизируется примерно на уровне 40%, что говорит о сформировавшейся лояльной аудитории. В то время как у продукта B пользователи быстро уходят, и уже к 5 дню почти никто не возвращается. Это может свидетельствовать о низкой ценности продукта B для пользователей или проблемах с качеством пользовательского опыта.

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:

Эксперимент 1:

- Средний ARPU (test): 665.74

- Средний ARPU (control): 722.46

- p-value: 0.689

Нет статистически значимого различия.

Эксперимент 2:

- Средний ARPU (test): 332.93

- Средний ARPU (control): 704.65

- p-value: 0.0011

Статистически значимое различие. Рекомендуется внедрение.

Эксперимент 3:

- Средний ARPU (test): 998.67

- Средний ARPU (control): 663.21

- p-value: 0.0603

Нет статистически значимого различия.

В первом эксперименте среднее значение ARPU в тестовой группе оказалось ниже, чем в контрольной, но разница статистически незначима (p-value = 0.689), следовательно, тестируемое изменение не повлияло на поведение пользователей. Во втором эксперименте тестовая группа показала значительно более низкое ARPU по сравнению с контрольной, и эта разница является статистически значимой (p-value = 0.0011), что говорит о негативном эффекте изменений — их не стоит внедрять. В третьем эксперименте тестовая группа демонстрирует более высокий ARPU, чем контрольная, однако p-value = 0.0603, то есть статистическая значимость не достигнута. Несмотря на положительную динамику, решение о внедрении требует дополнительной проверки или повторного теста с большей выборкой.

9. По датасету с листерами посчитайте средний доход на пользователя

121.2156.470.930.7средняя здесь не применима  
Отвтет: 156.4

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?

Ответ: №3

№1

№2

№3

№4

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

Ответ: №3

№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:

У тестовой группы немного больше платежей, чем у контрольной, но при этом количество пользователей в обеих группах очень большое — около 100 миллионов. Разница в платежах слишком маленькая, чтобы с уверенностью сказать, что тест действительно лучше. Чтобы это проверить точно, нужно использовать статистический тест, но предварительно видно, что отличия незначительные. Поэтому пока не стоит внедрять изменения, так как нет уверенности, что они действительно помогают.