**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.

Правильный ответ 7 639

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.

Правильный ответ 560

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.

Правильный ответ 26,6%

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:

Сравнение продуктов:

Обе кривые (синяя и красная) стартуют с 100% на Day 0 — это означает, что все пользователи активны в день установки.

Синяя кривая: Убывает медленно, что говорит о лучшем удержании пользователей. Пользователи продолжают возвращаться в продукт на протяжении всей недели.

Красная кривая: Убывает резко, особенно в первые дни. Это указывает на низкое удержание — пользователи быстро теряют интерес.

Выводы:

Продукт с синей кривой демонстрирует более высокую вовлечённость и ценность для пользователей. Продукт с красной кривой требует улучшений в онбординге, пользовательском опыте или функциональности, чтобы повысить удержание.

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.

Ответ 46,3%

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

Ответ 2,9

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%

Ответ 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

9. По датасету с листерами посчитайте средний доход на пользователя

121.2156.470.930.7средняя здесь не применима

Ответ 156,4

10. По датасету с листерами посчитайте медиану возраста пользователя

27,422827,9327медиана здесь не применима

Ответ 28

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

Ответ Ящик с усами (box plot), Гистограмма

12. На каком графике бимодальное распределение?

№1

№2

№3

№4

Ответ № 3

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

№1

№2

№3

№4

Ответ №3

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









Ответ: Scatter Plot (первый график), Line Plot (второй график) (при анализе временных данных), Correlation Heatmap (третий график) (готовый результат корреляции).

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%

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

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

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

Ответ t-тест

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

Ответ Делят данные на четыре равные части

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:

### Результаты:

-Конверсия в группе A (контроль): **0.0000100** (или **0.001%**).

-Конверсия в группе B (тест): **0.0000110** (или **0.0011%**).

-Z-статистика: **2.10**.

-P-value: **0.035**.

### Интерпретация:

**P-value = 0.035**: Это значение меньше стандартного уровня значимости 0.05, что говорит о наличии статистически значимых различий между группами.

**Z-статистика > 0**: Указывает, что конверсия в тестовой группе B выше, чем в контрольной группе A.

### Рекомендации:

-Тестовая группа B показала лучший результат по конверсии, и разница статистически значима.

-Рекомендуется внедрить изменения, протестированные в группе B, так как они улучшают конверсию.

-Дополнительно можно проанализировать долгосрочное влияние этих изменений, например, на удержание пользователей или выручку