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

<https://docs.google.co8m/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

*ответ ( 560 )*

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%

*ответ ( 26,6 % )*

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



Ваш ответ:

* **Первая кривая:**

На начало 100%, но затем кривая быстро падает на 5-й день и вероятность устояться **на низком уровне**, это может привести к началу высокого уровня активности, но к низкому удержанию в основных трендах.

По всей вероятности, продукт не представляет достаточной ценности и его функционал ограничен.

* **Вторая кривая**:

Медленно падает со 2-го дня, но сохраняется стабильное удержание на 5–7-й день, можно сказать, что второй продукт обладает **более стабильным удержанием**. Кривая показывает постепенное снижение, что может привести к **хорошему удержанию юзеров** в более широком диапазоне, даже если юзеры начинают уходить, продукт продолжает удерживать часть своей активности, это говорит о наличии контента или функциях, которые способствуют возвращению юзеров.

* **Заключение:**

**Первая кривая**: необходимо уделить внимание стратегии удержания, улучшению контента, дополнительных возможностей за счет более сложной механики обработки новых функций.

**Вторая кривая**: выглядит более стабильно, но для того, чтобы держать уровень, не забывать повысить мотивацию время от времени дополнительными плюшками…

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%

*ответ ( 46,3 % )*

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

Ваш ответ:

#### Эксперимент 1:

* **ARPU контрольной группы** : 722,46
* **ARPU тестовой группы** : 665,74
* **p-значение** : 0,688966

**Вывод** : Значение p > 0,05, указывает на наличие статистически значимых признаков между контрольной и тестовым алгоритмом.

**Рекомендация** : Изменения в тестовой группе не влияют на стандартный ARPU, поэтому их не стоит внедрять.

#### Эксперимент 2:

* **ARPU контрольной группы :** 704,65
* **ARPU тестовой группы** : 332,93
* **p-значение** : 0,001128

**Вывод :** Значение p < 0,05, свидетельствует о статистически значимых различиях между контрольной и тестовой измерениями, однако в данном случае ARPU тестовой группы значительно ниже, чем у контрольной (332,93 против 704,65), что указывает на возможное негативное влияние изменений на выручку.

**Рекомендация** : поскольку изменения привели к снижению ARPU, не рекомендуется их внедрять, несмотря на статистическую оригинальность.

#### Эксперимент 3:

* **ARPU контрольной группы :** 663,21
* **ARPU тестовой группы :** 998,67
* **p-значение :** 0,060315

**Вывод :** Значение p близко к 0,05, но все же > 0,05, что означает наличие статистически значимых символов. ARPU тестовой группы выше, чем у контрольной группы, но данные не подтверждаются.

**Рекомендация:** Хотя ARPU в тестовой группе выше, показатель не является статистически значимым, и изменения не имеют достаточного основания для развития.

#### Подведем итог:

* На основании проведенных тестов A/B **эксперимент 2** показывает отрицательный эффект изменений на выручку, и такие изменения не следует вносить.
* **Эксперименты 1 и 3** не являются характерными особенностями инновационного эффекта, поэтому изменения в этих тестах также не рекомендуется вносить.

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 (также может быть полезна для визуализации распределения цен)

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, но №2 тоже подходит*

№1

№2

№3

№4

14. На каком графике можно посчитать коррелцияю?

*Ответ изначально Scatter Plot, и второй вариант Correlation Heatmap*

\*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:

### ****Пояснение:****

* **Если p-значение ≤ 0,05** наблюдаются ., мы отвергаем нулевую гипотезу и делаем вывод, что конверсии в двух группах наблюдаются.
* **Если значение p > 0,05** , мы не отвергаем нулевую гипотезу и делаем вывод, что разница в конверсиях статистическая незначима.

#### **Подведем итог:**

* Если значение **p меньше 0,05** , это будет означать, что группа B (тестовая группа) показала статистическое улучшение в конверсии по сравнению с последовательностью A (контрольной). В этом случае **рекомендуется внедрить изменения, которые находились в группе** **B**.
* Если значение **p** **больше 0,05** , то анализировать разницу между отклонениями и **не стоит делать вывод о значительном улучшении** от изменений в группе **B**. В таком случае можно оставить функционал без изменений.