**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 1

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 3

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 2

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



Ваш ответ:

*На мой взгляд, первый продукт (зеленая линия) метрика Retention осталось активной вплоть до 7 дня, чуть меньше 50%. К 4 дню осталось пользователей 50%.*

*А второй продукт (красная линия) потеряла своих пользователей полностью на 5 день.*

*А начинали они одинаково и ко второму дню у обоих был спад 52% и 40% соответственно. А дальше интерес к продукту 2 у пользователь стал стремиться к 0, пока его совсем не забросили.*

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 3

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

4,96,25,32,9

Ответ 4

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 4

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:

Получила следующие результаты

Эксперимент 1:

Средний ARPU в контрольной группе: 722.46

Средний ARPU в тестовой группе: 665.74

P-value: 0.6890

Различие НЕ статистически значимо → нововведение не дало эффекта.

Эксперимент 2:

Средний ARPU в контрольной группе: 704.65

Средний ARPU в тестовой группе: 332.93

P-value: 0.0011

Различие статистически значимо → тестовая гипотеза подтверждена.

Эксперимент 3:

Средний ARPU в контрольной группе: 663.21

Средний ARPU в тестовой группе: 998.67

P-value: 0.0603

Различие НЕ статистически значимо → нововведение не дало эффекта.

**Выводы и рекомендации по результатам A/B-тестов**

После проведения анализа трех A/B-тестов для **ARPU (Average Revenue Per User)** были получены следующие результаты:

**Результаты по каждому эксперименту:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Эксперимент** | **Средний ARPU (Контроль)** | **Средний ARPU (Тест)** | **P-value** | **Вывод** |
| **№1** | 722.46 | 665.74 | 0.6890 | Различие НЕ статистически значимо → нововведение не дало эффекта. |
| **№2** | 704.65 | 332.93 | 0.0011 | Различие статистически значимо → тестовая гипотеза подтверждена. |
| **№3** | 663.21 | 998.67 | 0.0603 | Различие НЕ статистически значимо → нововведение не дало эффекта. |

**Интерпретация p-value**

* **Если p < 0.05** → Различие между группами статистически значимо → Влияние нововведения подтверждено
* **Если p ≥ 0.05** → Различия случайны → Нововведение не оказало значимого влияния

### ****Рекомендации:****

1. **Эксперимент №1 - нововведение не дало эффекта**
2. **Эксперимент №2 - нововведение дало эффект на увеличение**
3. **Эксперимент №3 – нововведение не дало эффекта**

Можно либо оставить текущую версию, либо попробовать альтернативные гипотезы.

**Дополнительные рекомендации:**

* Провести **дополнительный анализ сегментов пользователей**: возможно, нововведение работает только для определенной группы (например, новых пользователей).
* Проверить, **хватило ли мощности теста** (достаточно ли наблюдений, чтобы p-value было достоверным).
* Если **разница в ARPU мала**, можно протестировать **другие ключевые метрики** (например, Retention, конверсию в покупку).

9. По датасету с листерами посчитайте средний доход на пользователя

121.2156.470.930.7средняя здесь не применима

Ответ 2

10. По датасету с листерами посчитайте медиану возраста пользователя

27,422827,9327медиана здесь не применима

Ответ 2

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 1 и 3

*Чем линейный график я думаю plt.scatterplot , будет лучше. По отдельны точкам можно будет увидит разброс данных от сконцентрированных точек.*

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 корреляцию?  
\*возможно несколько вариантов ответа









Ответ 1

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 2

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 1

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 2

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:

По результатам получила такие значения:

χ²-статистика: 4.3378, p-value: 0.037275

Разница значима, тестовая группа действительно лучше

*Рекомендации:*

*Можно запускать вариант B, так как он немного увеличил конверсию.*

*Однако разница очень маленькая, следовательно, стоит проверить дополнительные метрики (ARPU, LTV, возврат пользователей).*