

ПИШУ ОТВЕТ ТАК КАК КОГДА НАЖИМАЮ НЕ СОХРАНЯЕТСЯ!

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. **ОТВЕТ 7639**

☐ 7639 ☐ 16814 ☒ 10482 ☐ 16529

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. **ОТВЕТ 560**

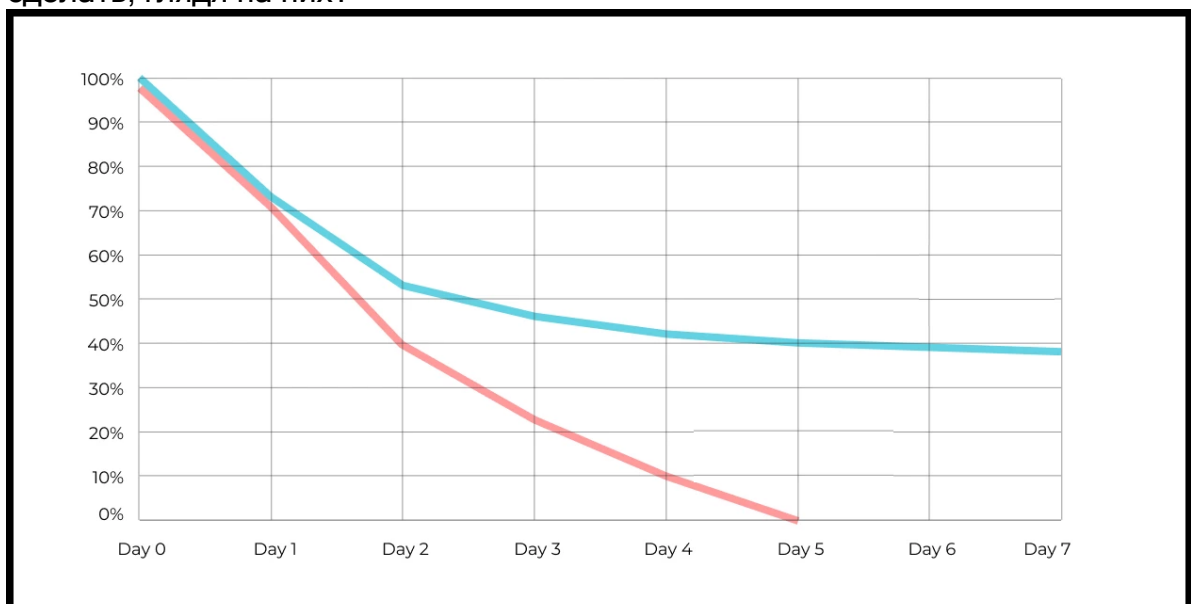
☐ 255 ☐ 490 ☐ 560 ☐ 483

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. **ОТВЕТ 28,3**

☐ 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:

- Продукт с голубой кривой показывает значительно лучшее удержание пользователей, особенно начиная с Day 2 и далее.
- Продукт с розовой кривой теряет пользователей очень быстро, и уже на 5-й день практически никто не возвращается.
- Если это A/B-тест, то голубая версия (A) явно выигрывает, так как обеспечивает более устойчивое вовлечение пользователей.
- Продукту с розовой кривой требуется анализ причин оттока — возможно, проблемы с пользовательским опытом, ошибками, контентом или ценностью.

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 54,7

☐ 41,8% ☒ 54,7% ☐ 46,3% ☐ 39%

6. Используя информацию из вкладки "Данные об аудитории", посчитайте среднее количество просмотренных объявлений на пользователя в ноябре ОТВЕТ 2,9

☐ 4,9 ☐ 6,2 ☐ 5,3 ☒ 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). ОТВЕТ 35

☐ 30% ☐ 43% ☐ 40% ☒ 35%

8. Во вкладке "Данные AB-тестов" результаты трех несвязанных AB 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:

Анализ трёх A/B тестов показал, что в первом эксперименте разницы в ARPU между контрольной и тестовой группами нет (p-value = 0.6890), следовательно, внедрение изменений нецелесообразно.

Во втором тесте выявлено значительное снижение выручки в тестовой группе (p-value = 0.0011), поэтому изменение следует отклонить. В третьем эксперименте наблюдается рост ARPU в тестовой группе, и хотя p-value = 0.0603 не достигает традиционного уровня значимости, результат близок к нему, что указывает на перспективность теста и необходимость его повторного

проведения с большей выборкой. Рекомендуется отклонить изменения из экспериментов 1 и 2, а эксперимент 3 повторить на расширенной выборке для подтверждения положительного эффекта.

9. По датасету с листерами посчитайте средний доход на пользователя **ОТВЕТ 156,4**

- ☐ 121.2 ☒ 156.4 ☐ 70.9 ☐ 30.7 ☐ средняя здесь не применима

10. По датасету с листерами посчитайте медиану возраста пользователя **ОТВЕТ 28**

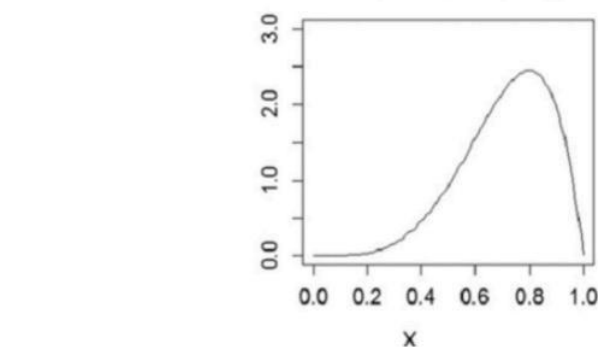
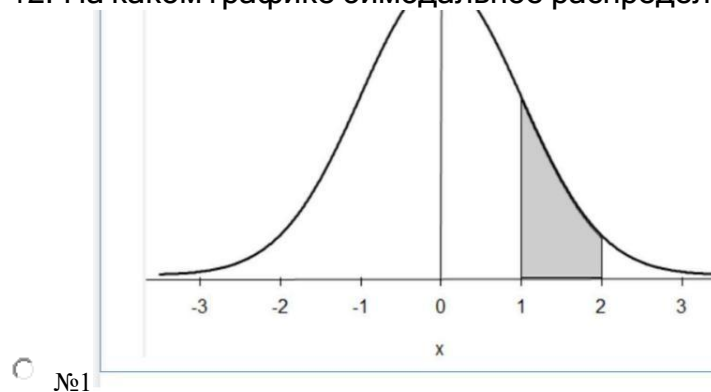
- ☐ 27,42 ☐ 28 ☐ 27,93 ☐ 27 ☐ медиана здесь не применима

11. Какой график лучше всего подходит для отображения разброса цен на товары в разных магазинах? **ОТВЕТ ГИСТОГРАММА**

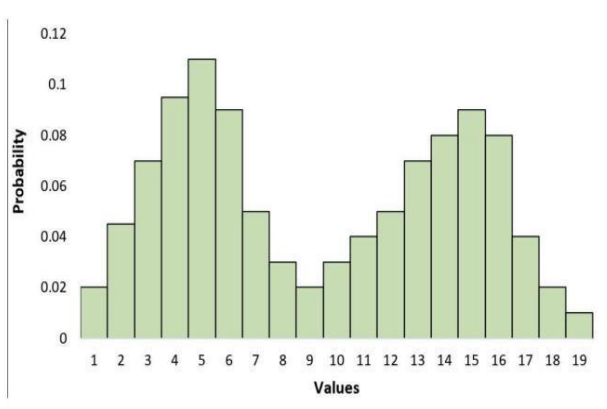
*возможно несколько вариантов ответа

- ☐ Линейный график ☐ Круговая диаграмма ☐ Ящик с усами (box plot)
☐ Гистограмма

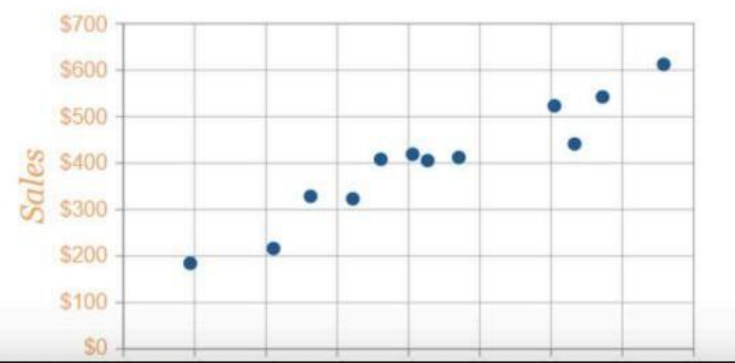
12. На каком графике бимодальное распределение? **ОТВЕТ 3 ГРАФИК**



☐ №3

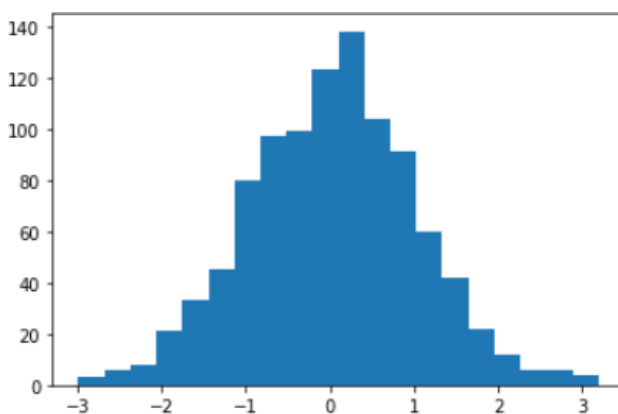


☐ №4

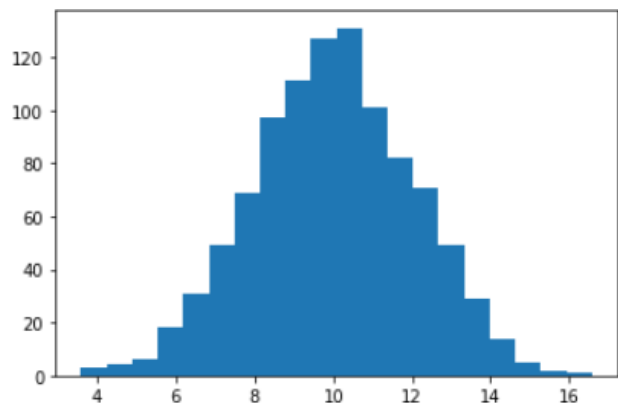


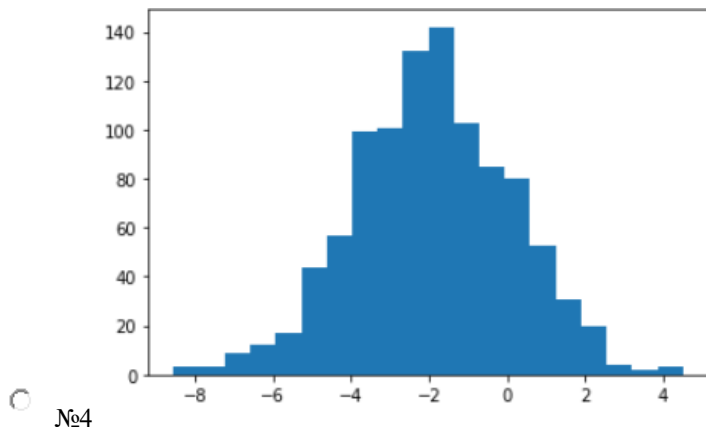
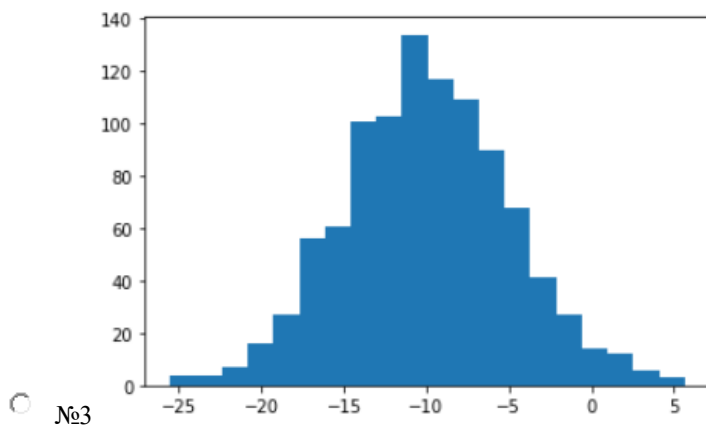
13. Какая случайная величина имеет наибольшую дисперсию данных по следующим графикам плотности распределения? **ОТВЕТ 3 ГРАФИК**

☐ №1

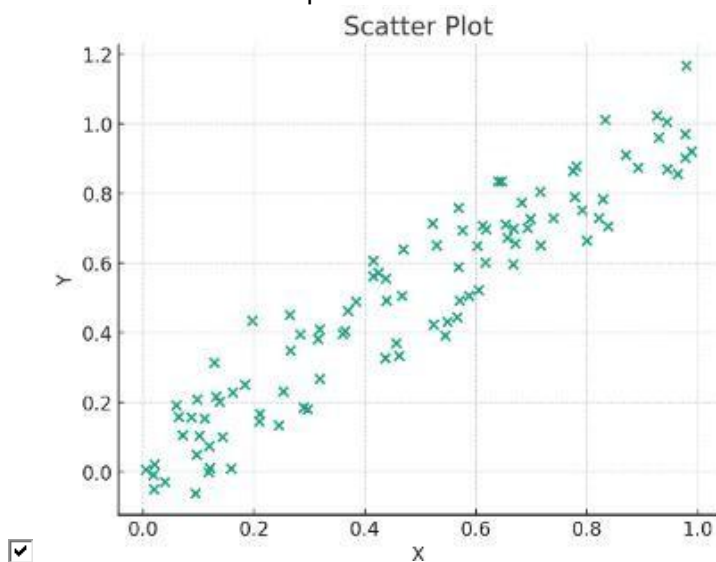


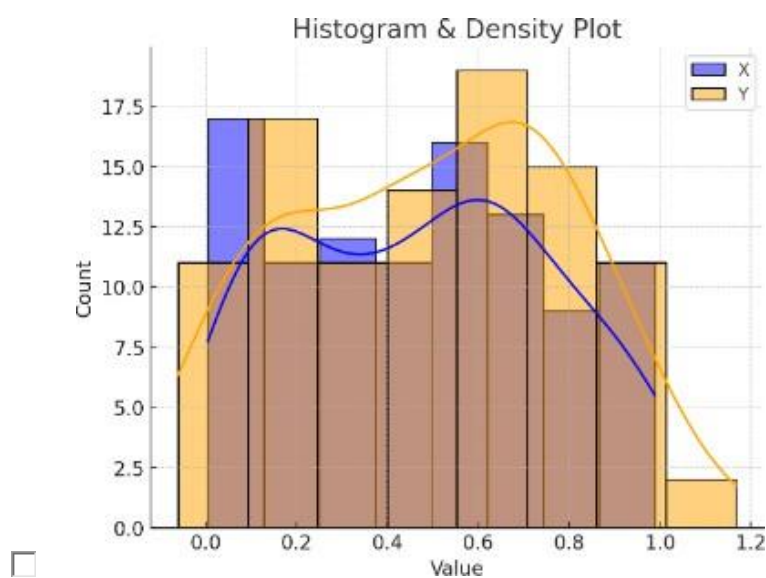
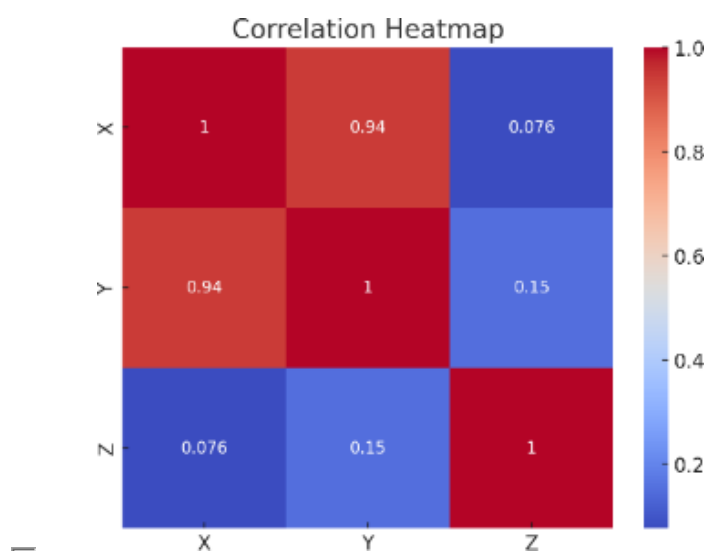
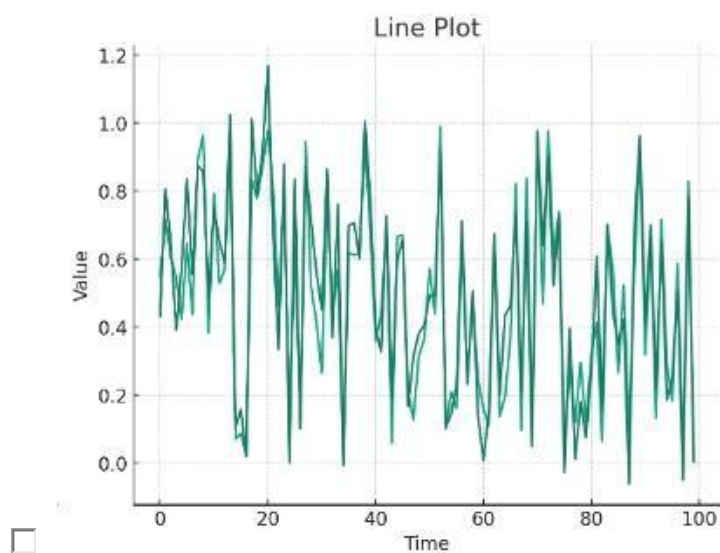
☐ №2





14. На каком графике можно посчитать корреляцию? **ОТВЕТ 1 ГРАФИК**
 *возможно несколько вариантов ответа





15. Что значит, если при проверке гипотез мы получили $p\text{-value} = 0.05$? **ОТВЕТ 2**
ВАРИАНТ

- ☐ 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? **ОТВЕТ 1 ВАРИАНТ**

- ☒ t-тест
- ☐ Хи-квадрат тест
- ☐ Анализ дисперсии (ANOVA)
- ☐ Корреляция Пирсона

17. Как интерпретировать квартили в распределении доходов пользователей? **ОТВЕТ 2 ВАРИАНТ**

- ☐ 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 А (контрольная группа) – 100 047 501 посетитель, 1003 платежа.
- Вариант В (тестовая группа) – 100 001 055 посетителей, 1099 платежей.

Какие рекомендации вы бы дали, основываясь на этих данных?

Ваш ответ:

Разница в конверсии между вариантами А и В минимальна. Необходимо провести z-тест для пропорций. Если разница статистически незначима – оставить вариант А. Если значима – оценить бизнес-эффект, так как выигрыш очень небольшой.