

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.

☐ 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.

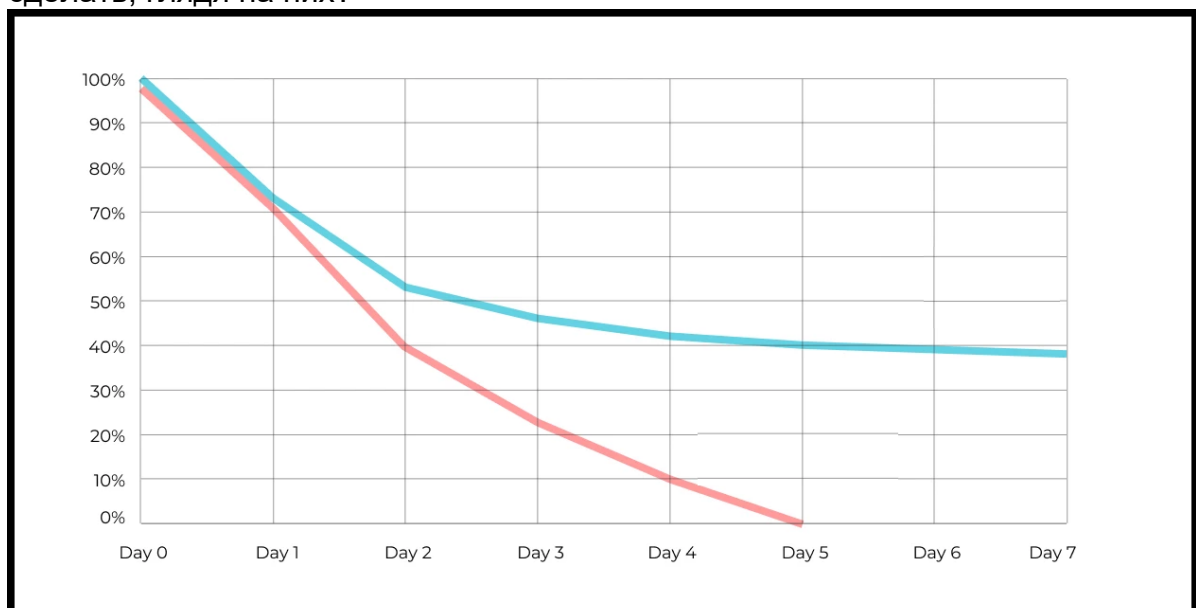
☐ 255 ☐ 490 ☒ 560 ☐ 483

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.

☐ 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: можно сделать вывод, что красный продукт полностью не оправдал ожидания пользователей и те кто попробовал его в «нулевой» день, полностью отказались к пятому дню. Синий продукт, сохранил половину «нулевых» пользователей к 6 дню, далее идет плавное уменьшение. Но в целом ситуация намного лучше чем у красных.

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

☐ 4,9 ☐ 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).

☐ 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:

При H0: у тестовой и контрольной группы ARPU равны; H1: у тестовой группы ARPU больше, чем у контрольной. То есть это тип выборка vs выборка, направление one tail.

Test1_p_value = 0.65, test2_p_value = 0.999, test3_p_value = 0.03.

Лишь в третьем тесте мы можем отклонить нулевую гипотезу и сказать, что действительно ARPU повысился у тестовой группы в результате введения фич. Рекомендуется собрать больше данных и провести тесты еще раз.

9. По датасету с листерами посчитайте средний доход на пользователя

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

10. По датасету с листерами посчитайте медиану возраста пользователя

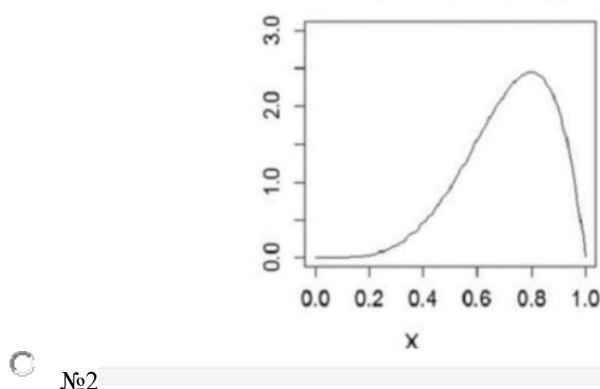
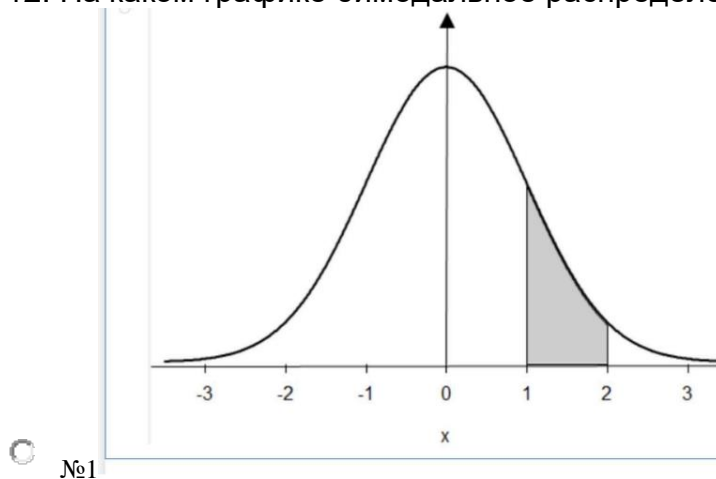
☐ 27,42
 ☒ 28
 ☐ 27,93
 ☐ 27
 медиана здесь не применима

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

*возможно несколько вариантов ответа

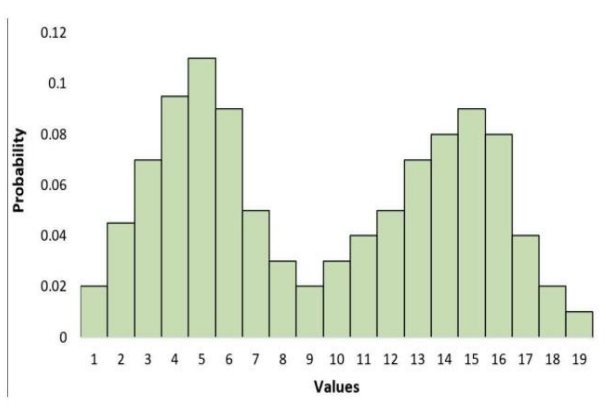
☐ Линейный график
 ☐ Круговая диаграмма
 ☒ Ящик с усами (box plot)
 ☒ Гистограмма

12. На каком графике бимодальное распределение?

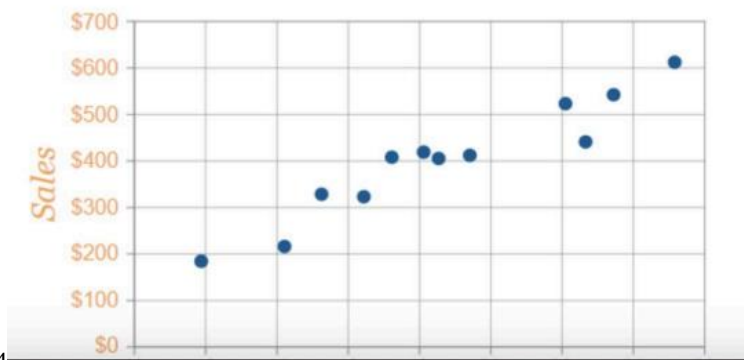




№3



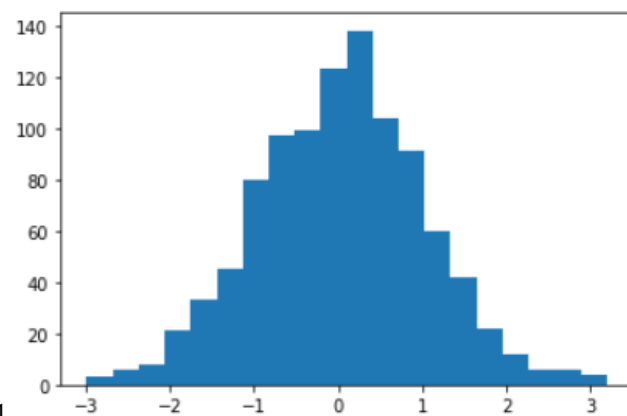
№4



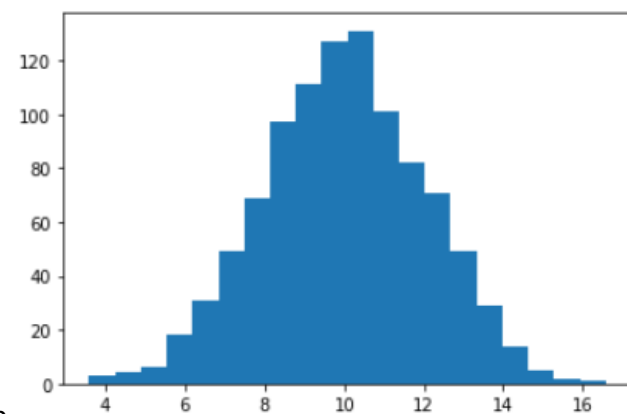
13. Какая случайная величина имеет наибольшую дисперсию данных по следующим графикам плотности распределения?

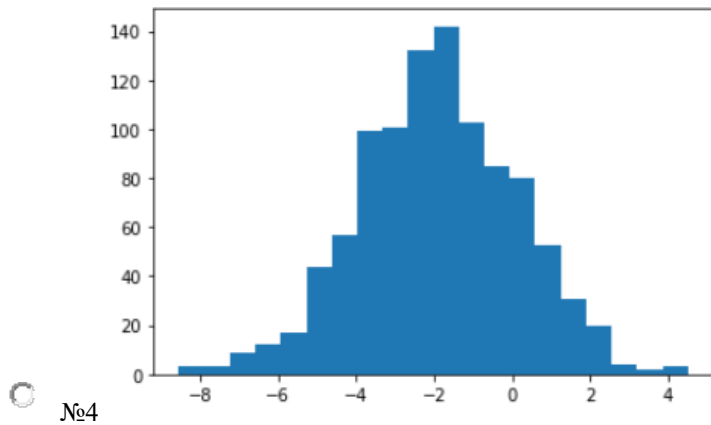
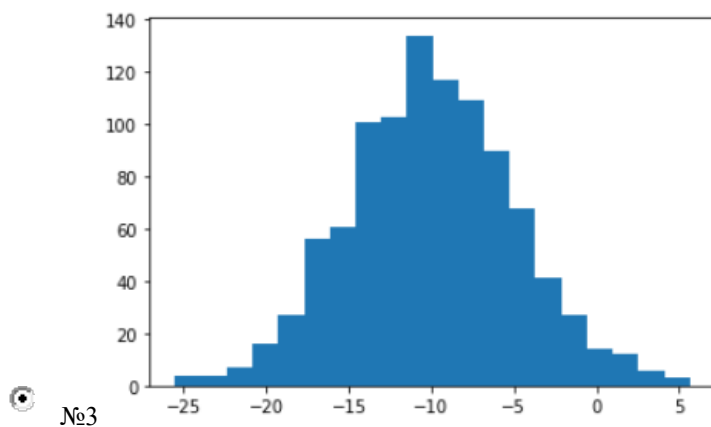


№1

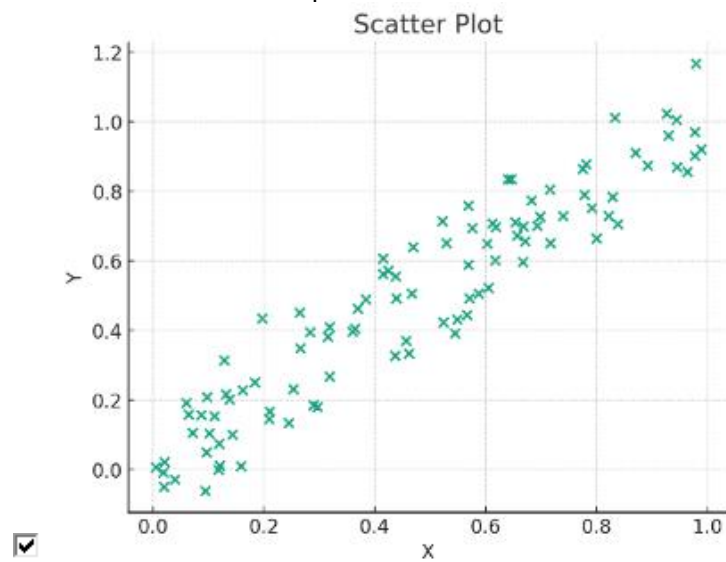


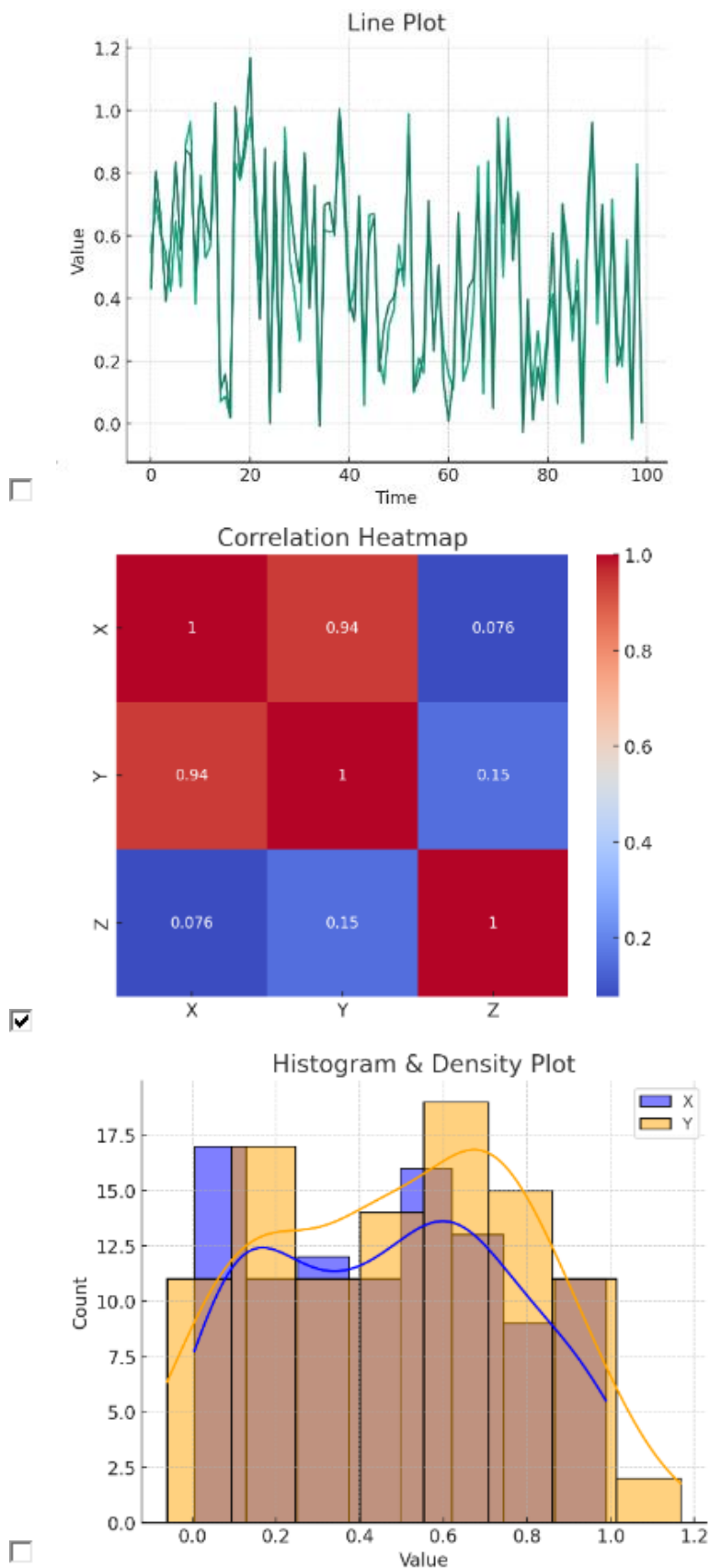
№2





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





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

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

17. Как интерпретировать квантили 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 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:

Нужно выяснить конверсию. Если провести односторонний тест one tail и взять как h_1 : у тестовой группы конверсия повысилась. $P\text{-value} = 0.017$. При уровне значимости 5% можно отклонить нулевую гипотезу и с уверенностью сказать что у тестовой группы конверсия повысилась.