

**Бизнес-кейсы и анализ данных – Самостоятельная работа****Листок. АД7_дз****По лекции:** Практические задачи по статистике**Составитель:** Дмитрий Ким**Дедлайн:** 7 августа 2024 года - 8 августа 2024 годаЗадачи в этом листке можно сдавать **в любом порядке.****Бонус за сданную задачу:** 25 HSE.

Представим, что вы стали мастером по изготовлению мармеладных животных. В вашем ассортименте представлены мармеладные кот Кейнс, хомяк, капибара, панда, лев и другие милые и вкусные фигурки. Вы уже продавали эти сладости своим друзьям и знакомым и теперь решили расширяться, запустив продажи через интернет. Для этого вы оформили статус индивидуального предпринимателя, создали сайт на платформе Tilda, подключили рекламу и начали продвижение в Интернете. Ваши основные каналы продвижения — это поисковик Яндекс, лента в VK и Telegram-каналы. Вам доступны данные от сервиса сквозной аналитики, который отслеживает все аспекты вашей рекламной кампании. Ответьте на следующие вопросы:

Задача 1. Какой рекламный канал наиболее эффективен?

Определите, через какой канал рекламы клиенты делают больше всего покупок, где средний чек выше, и сколько денег приносит каждый клиент. Уточните за какой период вы рассматриваете наблюдения. Попытайтесь выяснить, какой канал наиболее успешен для привлечения покупателей.

Задача 2. Как часто бывают корзины разной стоимости?

Изучите, распределение частности покупок по стоимости корзины. Какие интересные выводы можно сделать на основе этого распределения?

Задача 3. Отличается ли поведение клиентов в разных регионах?

Проанализируйте корзины в разрезе регионов: общая выручка с региона, выручка с человека, средний чек, количество заказов за период. Постройте графики. Какие интересные выводы можно сделать на основе этих данных?

Задача 4. Какие есть, если имеются таковые, временные тренды продаж?

Постройте графики изменений общего количества продаж и суммы покупок во времени. Анализируйте, как меняется спрос и выявите возможные сезонные тренды.

Задача 5. В каких статусах находятся доставки по регионам?

Сравните успешность доставки заказов между разными регионами. Выясните, возможно какой-то регион является проблематичным.

Задача 6. Как можно сегментировать клиентов по сумме покупок?

Разделите клиентов на группы по их тратам. Опишите характеристики каждой группы, (включая предпочтительные способы оплаты и регионы.)

Задача 7. Как связаны данные между собой?

Проверьте, есть ли корреляционная связь между переменными. Проинтерпретируйте результаты. (подсказка: используйте коэффициент корреляции Пирсона)

База данных: ссылка <https://bit.ly/4dv6yji> и файл на сайте

Примечание: База искусственная, поэтому данные чистые, но все равно хороший тон проверить их.

Описание переменных:

purchase_id: Уникальный номер каждой покупки.

client_id: Уникальный номер каждого клиента.

purchase_date: Дата, когда была совершена покупка.

amount: Общая стоимость покупки.

item_cnt: Количество товаров в покупке.

region: Регион, откуда был сделан заказ.

payment_method: Способ оплаты заказа (например, карта или наличные).

delivery_status: Текущий статус доставки заказа (например, доставлено или в пути).

ad_channel: Канал рекламы, через который клиент узнал о вашем магазине (Яндекс, VK, Telegram).