**Отчет по финальной работе курса “Метрики, гипотезы, точки роста”**

**1.** **Описание данных: что отражают, их качество и полнота, некорректности и аномалии.**

Для анализа представлены данные о транзакциях в магазине. Общее количество транзакций(строк) в датасете - 54214. Датасет состоит из колонок с наименованиями: номер чека, код товара (далее штрих-код или штрих), наименование товара, количество,дата и время транзакции, цена за единицу, ID покупателя, страна.

Данные содержат большое количество пропусков как то:

* 13571 строк не указан CustomerID, что составляет 25,03% всех строк датасета.
* 130 строк не указано наименование товара.
* 225 строк в датасете с ценой 0,0
* 41 строка вместо наименования прописано: “damaged, throw - can’t sell, found e.t.c.”
* в столбце “InvoiceDate” дата и время в разных форматах

**2. Описание применённого подхода к анализу данных: ограничения, накладываемые данными, и соответствующие интерпретации допущения.**

Разные форматы даты и времени а также отсутствие ID покупателя позволяют сделать предположение, что данные взяты из нескольких источников: физический магазин и интернет-магазин. Если это только интернет-магазин, можно предположить, что покупатели без ID - те, кто в итоге не совершил покупку.

Допущения в интерпретации данных:

* В колонке с количеством есть отрицательные значения. Я приняла решение считать это как убыток. (возврат, порча товара, недостача)
* Вести расчет пользовательских метрик только по тем строкам, где есть ID покупателя.
* Для заполнения строк с пропусками сформирован “ассортимент” на основании имеющегося датасета. Пропуски заполнены отсутствующими данными путем сопоставления их штрих-кодов со штрихами в “ассортименте”.

Дополнения, внесенные в датасет:

* столбец “Costs” - рассчитана на основании количества товара и цены за единицу.
* столбец min max в InvoiceDate и вычисленный на их основе столбец Lifetime

**3. Описание бизнес-заказчика, его гипотетических болей и целей развития бизнеса**.

Бизнес заказчика – магазин, занимающийся продажей украшений для дома или праздников. Магазин, вероятно, небольшой.(54201 транзакций за 374 дня - 144 в день)

Так как основная цель любого бизнеса – получение максимальной прибыли, то и наш заказчик не исключение. Учитывая количество транзакций одна из проблем – привлечение клиентов и их удержание. В датасете указана страна, возможно заказчик отслеживает популярность товаров, чтобы закупать соответствующее.

**4. Описание стейкхолдеров отчёта с обоснованием выбранных метрик (системы метрик).**

На основе предоставленных данных, выведены следующие метрики:

* Выручка за период;<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=n6nlwDM9j-oF&line=1&uniqifier=1>
* Убыток за период;
* Сумма среднего чека;
* Наиболее прибыльная группа товаров по выручке:<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=uOUCkpTUoWGQ&line=1&uniqifier=1>
* Наименее прибыльная группа товаров по выручке:

<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=xLWT8l6fKze-&line=1&uniqifier=1>

* Наиболее и наименее продаваемый товар по количеству;

<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=2qNb6rzLISmR&line=1&uniqifier=1>

* Среднее время жизни пользователей;<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=DrsWOkSHzCNF&line=2&uniqifier=1>
* Прибыль, принесённая каждым пользователем

<https://colab.research.google.com/drive/1tsHpGwxzTfphEa7khenle4zRJ0dnGUap#scrollTo=uOUCkpTUoWGQ&line=1&uniqifier=1>

На основании этого основные стейкхолдеры отчета:

* директор и(или) собственник;
* специалисты отдела продаж (ответственный за продажи специалист);
* специалист отвечающий за маркетинг;
* технический специалист (ответственный за загрузку/выгрузку информации для разъяснения как формируется датасет и откуда в нем пропуски).

**5. Рекомендации по модели данных бизнес-заказчика: что вы считаете важным отслеживать в дополнение к тому, что уже есть в датасете.**

Рекомендации что отслеживать в дополнение к имеющемуся датасету:

* Количество посещений сайта день/неделя/месяц
* Количество уникальных пользователей день/неделя/месяц
* Средняя выручка одного пользователя
* LTV
* У некоторых клиентов время жизни 0. Они раз купили и не вернулись. Можно предложить оценивать количество вернувшихся клиентов.
* Продажи по группам товаров месяцам/неделям, так сможем оценивать сезонные товары. Чем повысим прибыль в сезон.
* Объем продаж, и прибыль месяц/квартал/год
* Количество товаров на складе (сроки годности если есть такие товары, у которых он ограничен)
* Число посещающих сайт. Конверсию в платящих и тех кто не окончил покупку.
* Если имеется бюджет, выделяемый на рекламу, можно порекомендовать отслеживать маркетинговые метрики: CAC(Стоимость привлечения одного клиента), ROMI (возврат на маркетинг), Количество показов рекламы, Количество кликов по рекламе, CTR. (Последняя рекомендация скорее как задел на будущее.)

**6. Сформулированные выводы и рекомендации для бизнес-заказчика.**

Исходя из цели максимизации прибыли я бы предложила бизнес-заказчику сосредоточиться в первую очередь на увеличении объемов продаж и увеличении среднего чека.

Что касается среднего чека - решение стоит ли поднимать цены можно принять после изучения цен конкурентов: если у заказчика цены значительно ниже, чем у конкурентов, то стоит поднимать цены. Либо поработать с увеличением количества позиций в одном чеке. (через обновление и расширение ассортимента или рекламных акций набери на сумму - получи скидку 15% или купи n товаров, получи i-товар в подарок)

Чтобы товар больше и чаще покупали стоит изучить клиента, кто хочет этот товар покупать и где этих людей можно найти (блоги, сообщества, форумы).

Пропуски и неясные данные в датасете это направление тоже стоит отработать. Как формируются данные о транзакциях, где хранятся, откуда в них берутся пропуски.

Покупатели без ID/ Если мое предположение верно, что это часть из физического магазина, часть из интернет: для того чтобы учесть покупателей физического магазина, можно внедрить какие-то клиентские карты с программой лояльности.(хотя бы кэшбэк с тратой в этом же магазине)

Однако. хочу отметить, что рекламные акции и карты лояльности - это не первоочередная задача.