

Отчет

Так как предполагается, что это первый отчет в компании, то целью было сделать общий верхнеуровневый дашборд для отслеживания ключевых показателей. После исследования того, кто и как будет использовать этот отчет, его можно будет в дальнейшем модифицировать.

Перед построением отчета было сделано:

1. Исправлена 1 ошибка в данных таблицы order_lines путем фильтрации этой строки в PowerBI.
2. Изменен тип данных столбца price в таблице order_lines с текста на десятичное число в PowerBI.
3. Создан календарь для более удобной работы с отчетом в будущем.

Найденные аномалии и ошибки:

1. Отсутствуют данные за 31.08.2017г. Можно предположить, что на данный момент фактических данных за этот день нет.
2. У 409 заказов отсутствует информация о городе клиента, сделавшего заказ. Такие заказы содержат привязку к 6 различным складам (34, 57, 61, 268, 310, 868).
3. В таблице order_lines отсутствуют записи о 28.677 заказах, так же в этих заказах нет данных о количестве заказанных товаров.

Отчет состоит из следующих визуализаций:

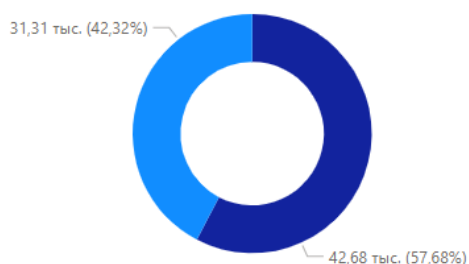
1. Название.
2. Срезы по дате, городу и конкретному складу.
3. Карточки с основными показателями и метриками.
4. Горизонтальная гистограмма. Отображает сумму продаж товаров каждой категории. Форматирование зависит от объема продаж в категории относительно всех продаж (красный – <25%, голубой – от 25% до 50%, синий - > 50%).
5. Линейный график. Отображает динамику продаж и количества заказов по дням. Пунктиром отмечено медианное значение суммы продаж.
6. Таблица. Отображает топ-20 наиболее продаваемых товаров, их название и количество проданных единиц.

В результате проведенного анализа можно сформулировать следующие выводы и гипотезы:

1. Наибольшее количество клиентов проживают в Санкт-Петербурге, сумма продаж так же выше по Санкт-Петербургу:

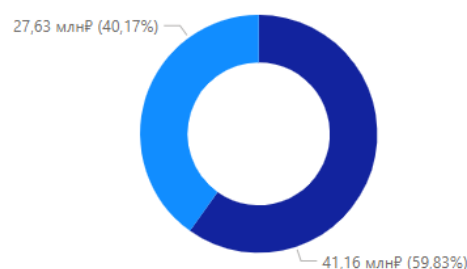
Распределение клиентов по городам

Город ● Санкт-Петербург ● Москва



Распределение суммы продаж по городам

Город ● Санкт-Петербург ● Москва



2. На графике динамики продаж можно видеть цикличность, по Пн и Вт наблюдается рост продаж и количества заказов, а со Ср до Вс – спад.



3. Начиная с 26.08.2017 наблюдается резкий спад в сумме продаж. При этом количество заказов и количество уникальных пользователей, которые совершили заказ, остаются на прежнем уровне, тогда как средний чек так же резко падает. При детальном анализе было выявлено, что количество заказов осталось на прежнем уровне, а вот количество позиций в заказе стало намного меньше (или вообще не отображалось), что и привело к падению общей суммы продаж.

Кол-во уник пользователей, сделавших заказ:



Средний чек:



4. Можно предположить, что мы имеем данные онлайн магазина и доставки продуктов и товаров ежедневного потребления. Пики в продажах по Пн и Вт могут объясняться тем, что клиенты заказывают продукты на неделю. В категориях товаров лидируют бакалея, товары ежедневного использования, то, что долго хранится, в то время как товары из категорий мясо, рыба, птица, овощи, кулинария покупают намного меньше, что может свидетельствовать о том, что при покупке таких товаров клиенты отдают предпочтение офф-лайн магазинам или рынкам, где они могут лично выбрать, какой продукт купить.
5. Для того, чтобы проанализировать резкое падение продаж и среднего чека необходимо больше данных, на основе имеющихся можно предположить, что имеются перебои с доставкой товаров на склады и клиенты не могут заказать товары в том объеме, в котором они это делали раньше (т. к. их просто нет в наличии).
6. Так же можно предполагать сбой в системе сбора и записи данных, так как при анализе заказов, созданных с 26.08.17 наблюдаются пропуски в данных о количестве заказанных товаров, то есть заказы оформляются, но в них нет данных о количестве товаров в заказе и сумме заказа. Тогда как во всех данных до 26.08.17 пропусков в данных не наблюдается (только 1 заказ за 13.08.17 так же имеет пустое количество). Соответственно, это привело к падению показателей суммы продаж и среднего чека. Всего записей с отсутствующим значением количества заказанных товаров 28.677 шт., так же записи о таких заказах отсутствуют в таблице order_lines. Необходимо проанализировать причины такого поведения системы и устранить ошибки в записи данных, если они таковыми являются.

order_date	order_id	warehouse_id	user_id	products_quantity	order_price
28 августа 2017 г	1148139	432	229483	2	209,25
28 августа 2017 г	1128359	432	228434	1	89,25
28 августа 2017 г	1171648	432	166030		
28 августа 2017 г	1140446	432	162092	4	327
28 августа 2017 г	1143096	432	279191	3	985,5
28 августа 2017 г	1156199	432	151063		
28 августа 2017 г	1163430	432	152764		
28 августа 2017 г	1153325	432	291639		
28 августа 2017 г	1161124	432	190593		
28 августа 2017 г	1162084	432	149417		
28 августа 2017 г	1169356	432	278227		
28 августа 2017 г	1159380	432	272557		
28 августа 2017 г	1170818	432	145342		
28 августа 2017 г	1146023	432	251615	8	375
28 августа 2017 г	1143883	432	266571	4	345
28 августа 2017 г	1163643	432	258567		
28 августа 2017 г	1145736	432	272557	2	164,25
28 августа 2017 г	1153127	432	243656		
28 августа 2017 г	1157812	432	249530		
28 августа 2017 г	1166867	432	291614		
28 августа 2017 г	1159814	432	299302		
28 августа 2017 г	1151150	432	274136		
28 августа 2017 г	1158617	432	249530		
28 августа 2017 г	1156493	432	149749		
28 августа 2017 г	1160534	432	236398		
28 августа 2017 г	1159499	432	245672		
28 августа 2017 г	1144483	432	155648	3	240
28 августа 2017 г	1112416	432	195346	1	66,75