# Задание

Построить дашборд для отдела продаж. Вывести основные показатели: например, выручку по странам, дням, топ товаров, и другие графики, которые вы считаете необходимыми для отражения **текущего** положения дел.

# Легенда для выполнения

Перенесемся в прошлое. Сегодня 9 декабря 2011 года. Это последний день, за который в таблице **default.retail** есть данные.

Я попыталась построить дашборд для отдела продаж на каждый рабочий день, взглянув на который сотрудник или руководитель отдела мог бы понять, что не так, или все нормально, в чем возможная причина аномалий и что нужно проверить.

# Пояснения

## Отнесение стран по категориям по количеству заказов в год.

Я разбила страны на три категории:

* с **низкой** частотой заказов (<100);
* с **средней** частотой заказов ([100; 1000));
* c **высокой** частотой заказов (>=1000).

Это позволило уменьшить количество кривых на основных графиках выручки и количества заказов, сделав их более наглядными (Рис 1 и 2). А дополнительно сделало ситуацию с странами, покупатели которых редко совершают заказы, также более явной (Рис 3 и 4), не теряющейся на фоне «гигантов»

## Ограничение временного интервала рисунков 1 и 2 (графики ежедневной выручки и количества заказов в день для стран с высокой и средней частотой заказов в день).

Для текущего положения дел, вероятно, не стоит показывать всю картину за год или более длинный период. Я ограничила его 3-мя месяцами, как мне кажется, это достаточно наглядно отражает и динамику последних недель, но и создает контекст данных колебаний.

## Ограничение временного интервала рисунков 3 и 4 (графики ежедневной выручки и количества заказов в день для стран с низкой частотой заказов в день).

Для стран, покупатели которых редко совершают покупки, я решила ограничить временной интервал еще более жестко – до одной недели, времени достаточного для обработки заказов из этих стран (как мне кажется)

## Особенности запросов по ежедневной выручке (Рис 1 и 3).

В ежедневной выручке нет условия отбора по Quantity > 0, это сделано, чтобы учитывать отмены заказов. Некоторые отмены, по-видимому, происходят не в дату покупки. Именно так может получаться отрицательная выручка на Рис 3. Так как номера заказов и их отмен не совпадают (id отмен начинаются с «С», но их номер не совпадает с номером заказа этих товаров), мне пока не понятно как такие отмены сопоставлять с заказами, чтобы отрицательная выручка не появлялась, кроме того, это вносит некоторую ошибку в подсчет ежедневной выручки. Отмены совершенные день в день с заказом при отсутствии фильтра по Quantity > 0 анигилируются между собой и таким образом не завышают реальную выручку.

## Особенности запросов по количеству заказов в день (Рис 2 и 4).

В данных запросах также стоит проблема отмен. В данных запросах фильтр отбора по Quantity > 0 наоборот есть. Это позволяет учитывать заказ и его отмену как один заказ.

Alert (Табл. 1)

Мне кажется, что необходимой частью дашборда для оценки текущего положения дел является какая-то система предупреждений, показывающая необычное поведение каких-то параметров и побуждающее искать проблему.

Я попыталась реализовать систему предупреждения на основе сравнения текущих дневной выручки и количества заказов в день с средними показателями за последний месяц, с учетом типа дня (будни/выходные)

Предупреждение появляется при выходе текущего параметра за границы интервала:

Хср ± 2 \* stdev ,

где Хср – среднее значение параметра в данный тип дня (т.е. усреднение проводилось по ~8 выходным или ~22 рабочим дням), stdev – стандартное отклонение, вычисленное при помощи stddevPop.

Подобный Alert, к сожалению, удалось создать только для стран с высокой частотой заказов в год (UK), для стран с средним количеством заказов стандартное отклонение слишком большое для создания алерта.

## Некоторые метрики текущего дня (Табл. 2)

В этой таблице я собрала метрики, которые могут быть полезны для быстрой оценки текущего положения дел:

Так приведены средние и медианы по таким показателям как выручка на один заказ, количество товаров на один заказ. Это позволяет быстро оценить связано ли какое-то отклонение кривых выручки и количества заказов в каком-то масштабном процессе или дело в единичном выбросе.

Также приведены покупатель с макисмальной выручкой и количеством заказа, чтобы были сразу видны какие-то аномальные заказы или заказчики, которые заказывают, редко, но сразу много.

## Отмены (Табл. 3)

Как мне при мониторинге текущего положения дел важно следить за отменами. Например, выявлять причины крупных отмен, чтобы тем самым получать отклик от покупателей по тому, что их не устраивает

# Проблемы, которые обнаружились при компановке дашборда

Указанная выше проблема с отменами. В идеале необходимо, чтобы Invoice\_id для заказа и отмены с него совпадали, но отличались только буквой. Эта проблема создает сразу несколько сложностей:

- отрицательная выручка по дням (когда дата заказа с его отменой не совпадают)

- трудно посчитать среднюю выручку с заказа за день:

* если выставляешь фильтр по Quantity > 0, то туда попадают отмененные заказы, и получается некорректное значение, а если этот отмененный заказ большой – то и очень большое завышение, т.к. среднее чувствительно к выбросам
* если используешь расчет без фильтра по Quantity > 0, то отмененные частично или полностью заказы вроде компенсируют друг друга, но это приводит к росту числа в знаменателе при расчете среднего, и таким образом получается заниженный результат среднего. Я предпочла сделать так, чтобы отмененные заказы оказывали меньшее влияние на среднее, по следующим соображениям: количество заказов в день 50 – 100, количество отмен ~5%, т.е. 2-5 отмен в день. Т.е. при подсчете средней выручки знаменатель будет больше примерно на 10% (заказ+его отмена), что приведет к падению средне выручки за день (относительно реальной величины с учетом отмен) также ~ на 10%. Однако, мы не будем получать ложные сигналы о каких-то сильных выбросах из-за заказов, которые сделали, а затем сразу отменили.

# Выводы

По текущему дашборду мы видим, что график общей выручки и количества заказов за последний день сильно отклонился в меньшую сторону. Аналогично себя ведет график выручки и количества заказов для UK, следовательно, причины данного отклонения стоит искать в заказах из этой страны.

Также мы видим, что среднее значение выручки и количества предметов в заказе превосходит медиану, а также что может свидетельствовать о влиянии большого заказа покупателя из Норвегии на средние показатели.

В отмененных заказах всего 5 записей, но сумма одной из них крайне велика, по-видимому стоит уточнить у покупателя в чем причина отмены.

Также стоит предложить отделу продаж подумать над изменением нумерации отмененных заказов для уменьшения погрешности расчетов.

Если текущее значение параметра выходило за границы интервала :