**Задание 1. Решить кейс.**

Датасет **test\_analytics.csv** содержит данные по проведенным с клиентами банка коммуникациями по рекламным маркетинговым кампаниям, цель которых предложить клиенту оформить кредит. Датасет также содержит данные по оформленным договорам после коммуникаций (поле fact =1 договор оформлен, fact = 0/Nan не оформлен).

Необходимо проанализировать эффективность рекламных кампаний с клиентами на предложенных данных. Определить и рассчитать метрики эффективности.

Эффективность оценить в разрезе кампаний (названия кампаний groups) . Описание сути кампаний представлено в таблице:

|  |  |
| --- | --- |
| **groups** | **Специфика маркетинговой кампании** |
| model\_1 | кампания проводится для клиентов с высоким спросом на кредит, по рекомендации математической модели |
| process\_1 | кампания проводится для клиентов, у которых уже есть один кредит в банке и они активно пользуются кредитной картой |
| process\_2 | кампания проводится для клиентов, у которых есть кредитная карта |
| process\_3 | кампания проводится для клиентов, у которых уже есть один кредит в банке |
| process\_4 | кампания проводится для клиентов, у которых есть вклад в банке |
| process\_5 | кампания проводится для клиентов склонных к оттоку из банка |
| product\_1\_1 | клиенту предлагается оформить кредит на стандартных условиях и получить повышенный cashback следующий месяц |
| product\_1\_2 | клиенту предлагается оформить кредит в мобильном банке без дополнительных документов |
| product\_2\_1 | клиенту предлагается оформить кредит на индивидуальных предрасчитанных условиях с определенной ставкой и суммой кредита |
| product\_2\_2 | клиенту предлагается оформить кредит и кредитную карту на льготных индивидуальных условиях |
| product\_3 | клиенту предлагается оформить кредит со сниженной ставкой при условии оформления страховки |

Сравнить эффективность каналов коммуникаций (comm\_chanell), цепочек коммуникаций с клиентом в разрезе капании.

Посчитать стоимость выдачи (разные каналы сколько стоит 1 просмотр сколько стоит взятие кредита) и стоимость цепочек коммуникаций в разных каналах из расчета стоимости коммуникации с клиентом:

Успешный звонок 15 руб.

SMS 4 руб.

PUSH 0,2 руб.

(comm\_status = 1 означает успешность коммуникации и оплачивается, comm\_status = 0 не дозвон, не оплачивается).

По итогам анализа дать рекомендации повышения эффективности рекламных кампаний.

возиожно база для разных кампаний собирается по разному надо быть более эффективными на этапе сбора баз знать кто наш клиент и делать рассылку по тем кого мы считаем целевыми

Для выполнения задания использовать инструменты python. Результаты прислать в блокноте jupyter notebook с текстовыми комментариями к вычислениям и итоговыми рекомендациями.

**Задание 2. Написать запросы SQL. БД интернет-магазина фильмов и музыки. Прислать с текстовом файле.**

1.Изучите заказы, которые оформили в сентябре 2011 года. Сравните общую сумму выручки (поле total) за каждый день этого месяца: выведите день в формате '2011-09-01' и сумму. Информацию о дате заказа хранит поле invoice\_date. Оставьте в таблице только те значения суммы, которые больше 1 и меньше 10. Данные из таблицы **invoice.**

2. Посчитайте пропуски в поле с почтовым индексом billing\_postal\_code для каждой страны (поле billing\_country). Получите срез: в таблицу должны войти только те записи, в которых поле billing\_address не содержит слов Street, Way, Road или Drive. Отобразите в таблице страну и число пропусков, если их больше 10. Данные из таблицы **invoice.**

3. Нужно объединить данные двух таблиц: **track** со всеми треками в магазине и **invoice\_line** с купленными треками. Выгрузите таблицу с названием трека и числом, которое соответствует тому, сколько раз трек покупали. Учитывайте, что в одном заказе один и тот же трек может встречаться несколько раз. Если какой-либо из треков не покупали или о купленном треке нет информации — такие записи не должны войти в таблицу. Оставьте только уникальные значения и отберите первые 20 записей.

**Описание данных в таблицах:**

**invoice**

Содержит данные о заказах и выставленных счетах. Таблица включает поля:

первичный ключ invoice\_id — идентификатор счёта;

внешний ключ customer\_id (отсылает к таблице client ) — идентификатор покупателя;

invoice\_date — дата выставления счёта;

billing\_address — адрес выставления счёта;

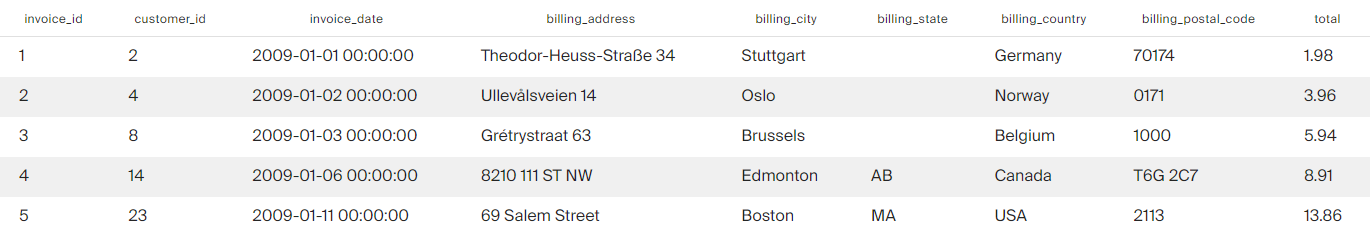
billing\_city — город выставления счёта;

billing\_state — штат или регион выставления счёта;

billing\_country — страна выставления счёта;

billing\_postal\_code — почтовый индекс адреса выставления счёта;

total — стоимость заказа в долларах.



**invoice\_line**

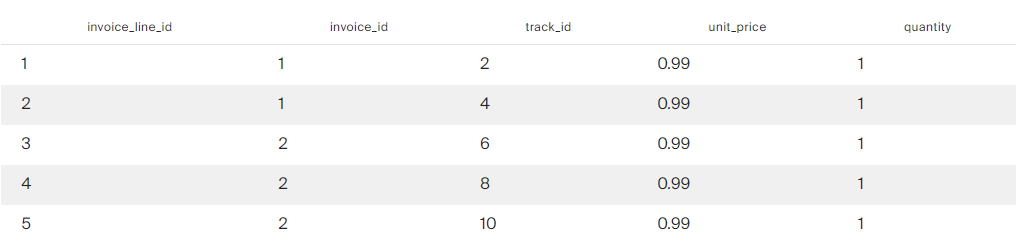
Содержит информацию о том, какие музыкальные треки входят в заказ. Таблица состоит из полей: первичный ключ invoice\_line\_id — идентификатор строки;

внешний ключ invoice\_id (отсылает к таблице invoice ) — идентификатор заказа;

внешний ключ track\_id (отсылает к таблице track ) — идентификатор трека;

unit\_price — цена трека в долларах;

quantity — количество таких треков в заказе.



**track**

Содержит данные о музыкальных треках. Таблица включает поля:

первичный ключ track\_id — идентификатор трека;

name — название трека;

внешний ключ album\_id (отсылает к таблице album ) — идентификатор альбома;

внешний ключ media\_type\_id (отсылает к таблице media\_type ) — идентификатор аудиоформата; Описание базы и ER-диаграмма для теоретической части Базового SQL 8

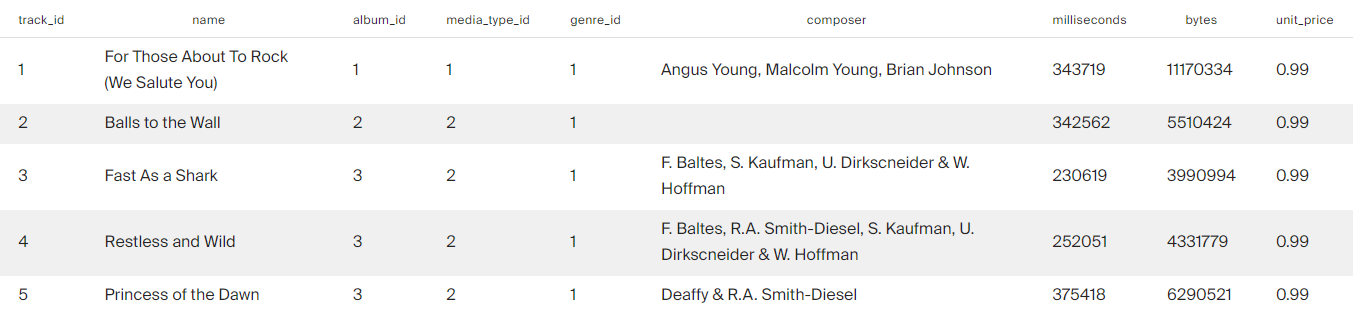
внешний ключ genre\_id (отсылает к таблице genre ) — идентификатор жанра трека;

composer — автор трека;

milliseconds — длительность трека в миллисекундах;

bytes — размер трека в байтах;

unit\_price — стоимость трека в долларах



**Схема базы данных:**

