***Часть 1 практическая:***

**Задание 1.1**

В приложении вы найдете 2 таблицы, это кусок базы для «прогнозирования оттока».

В «Таблице 1» содержатся некоторые данные о поведении Клиентов за три месяца

В «Таблице 2» данные об оттоке этих Клиентов в 4-м месяце.

Необходимо из исходных таблиц (см. вкладки «Таблица 1»  и «Таблица 2»  во вложении) получить  таблицу со следующими характеристиками (описание показателей на вкладке «Словарь показателей»):

         В таблицу должны попасть  клиенты со сроком жизни (CL\_AGE) больше 90 дней на март 2016г.

         В таблицу должны попасть только клиенты с ненулевым количеством интернет сессий (ACC\_COUNT\_INTER)  в марте 2016г.

         Таблица должна иметь следующие показатели по каждому идентификатору клиента:

o   Возраст клиента в марте 2016 г.

o   Показатели № 4-20 за месяца января-марта 2016 г.  (данные за каждый месяц по каждому показателю отдельно)

o   Суммарное значение за три месяца  для показателей № 4, 13, 18 (сумма  рассчитывается по каждому показателю в отдельности)

o    Сумма максимального платежа за три месяца

o   Среднее значение за три месяца для показателей № 9, 10, 12

o   Показатель, характеризующий наличие хотя бы одной аварии за три месяца (1 – аварии были, 0 – не было)

o   Показатель, характеризующий наличие хотя бы одно перехода из добросовестного плательщика в недобросовестного за 3 мес. (1 – был переход, 0 – не было)

o   Показатель – таргет с листа «Таблица 2»

Итоговый результат: написанный скрипт SQL , который  исходную таблицу преобразует  в вышеописанный вид + таблица в excel с результатами запуска скрипта

Для написания и тестирования скрипта можно воспользоваться  СУБД MYSQL, которая доступна по ссылке [https://free4.beget.com/phpMyAdmin/index.php](https://vk.com/away.php?to=https%3A%2F%2Ffree4.beget.com%2FphpMyAdmin%2Findex.php)

Логин к базе данных: e099284b\_alena

Пароль: gyCK91\*y

**Задание 1.2**

Ниже представлена таблица о количестве абонентов и статусе их активности через 12 месяцев

Абоненты разделены на сегменты, в зависимости от подключенных ими дополнительных услуг на январь 2018 года

Проанализируйте предоставленные данные и сделайте вывод о влиянии доп. услуг на отток абонентов.

Приложите excel файл с расчетами

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **По состоянию на январь 2018** | | | **Количество абонентов** | | |
| **Есть антивирус** | **Есть мультирум** | **Есть ТВ-пакет** | **Всего абонентов было активно в январе 2018** | **Активно абонентов  через 12 месяцев** | **В оттоке абонентов через 12 месяцев** |
| **1** | **1** | **1** | 4000 | 3600 | 400 |
| **1** | **1** | **0** | 2000 | 1700 | 300 |
| **1** | **0** | **1** | 12000 | 9000 | 3000 |
| **1** | **0** | **0** | 10000 | 7400 | 2600 |
| **0** | **1** | **1** | 40000 | 34000 | 6000 |
| **0** | **1** | **0** | 30000 | 25000 | 5000 |
| **0** | **0** | **1** | 200000 | 140000 | 60000 |
| **0** | **0** | **0** | 320000 | 200000 | 120000 |
| **Общий итог** |  |  | **618000** | **420700** | **197300** |

***Часть 2 теоретическая:***

Давайте порассуждаем на тему «прогнозирования оттока клиентов»

Кратко опишите:

1.       Каким образом стоит подготовить/обработать данные перед началом моделирования?

2.       Варианты типов моделей, которые могут быть использованы для решения задачи предсказания оттока. А так же результат, который можем получить после применения этих моделей.

3.       Какие данные вы бы еще добавили для моделирования? В каких разрезах?

4.       На какие оценки вы бы опирались для проверки качества модели?

5.       Какие действия с клиентами можно провести на основании результатов моделирования для проактивного предотвращения оттока?

Результат:  около 1-2 страниц текста А4 для теоретической части