**Задание 1.**

Провести анализ рынка услуг доступа в сеть Интернет базе фиксированной связи (тарифные планы и услуги для пользователей) для корпоративного сегмента Беларуси, отметить слабые и сильные стороны портфолио операторов, предложить активности по улучшению конкурентоспособности тарифных планов A1. Обосновать. Оформить в виде слайдов (.pptx) – 2-3 слайда.

**Задание 2.**

Построить диаграммы на основании данных в файле task\_2.xlsx:

a. Абонентская база:

Динамика числа абонентов – всего и в разрезе тарифных планов

b. Начисления:

Динамика ARPU (выручка за отчетный месяц, деленная на средневзвешенное число абонентов по состоянию на начало и конец отчетного месяца) всего и в разрезе каждого ТП.

Отразить структуру начислений на диаграмме динамики среднего счета клиента (всего, в разрезе каждого ТП не обязательно).

c. Потребление голосового трафика:

Динамика общего объема потребленных минут.

Динамика AMOU (объем потребленных минут за отчетный месяц, деленный на средневзвешенное число абонентов по состоянию на начало и конец отчетного месяца).

Структура исходящего голосового трафика по направлениям (в среднем за последние 3 месяца).

d. Потребление интернет-трафика:

Динамика общего объема потребленного интернет-трафика.

Динамика MbOU (объем потребленных MB за отчетный месяц, деленный на средневзвешенное число абонентов по состоянию на начало и конец отчетного месяца).

Оформить в виде слайдов (.pptx)

**Задание 3.**

Дано описание 4 таблиц данных и задание. Предложить скрипт SQL, который сформирует выборку по абонентам, которые:

1. Подключились в декабре (с 1 по 31), не реактивации.
2. Тип клиента Business
3. Подключились на один из тарифных планов «Бизнес-план», «Бизнес-план 1.0», «Бизнес-план 2.0», «Бизнес-план 3.0», «Бизнес-план 4.0» и продолжают пользоваться сегодня.
4. Подключили один из пакетов «Безлимитный интернет 512», «Безлимитный интернет 1», «Безлимитный интернет 2» в дату подключения абонента и продолжают пользоваться на сегодня.

Скрипт должен формировать таблицу со следующими полями:

1. Дата подключения абонента
2. ID абонента
3. Название тарифного плана, которым абонент пользуется на сегодня
4. Название услуги Безлимитного интернета, которым абонент пользуется на сегодня
5. Название компании, к которой принадлежит абонент\*
6. Количество абонентов с голосовым типом тарифного плана в статусе активный у компании, к которой принадлежит этот абонент.\*

\* опционально, будет плюсом

Таблица GROSS содержит информацию о событиях по всем абонентам, которые приравниваются к активации абонентов: новые подключения и реактивации.

Таблица TARIFF содержит информацию о Тарифных планах по каждому абоненту. Если у абонента было несколько тарифных планов, будет несколько записей. В одно и то же время у 1 абонента может быть активен только 1 тарифный план.

Таблица SERVICE содержит информацию о дополнительных пакетах по каждому абоненту. У абонента может быть несколько пакетов одновременно или ни одного. Но 1 и тот же пакет не может быть назначен в одно время несколько раз.

Таблица SUBSCRIBER содержит информацию об абонентах компании на конец каждого месяца. Т.е. записей по каждому абоненту столько, сколько месяцев.

Описание полей предоставленных таблиц:

|  |  |  |  |
| --- | --- | --- | --- |
| **Название таблицы** | **Поля** | **Описание поля** | **Тип данных** |
| GROSS | DATE\_ACT | Дата активации абонента | DATE |
| GROSS | SUB\_ID | ID\_абонента, уникальное значение для каждого абонента | BIGINT |
| GROSS | PP\_ID | ID тарифного плана, уникальное для каждого тарифного плана | BIGINT |
| GROSS | CUSTOMER\_TYPE | Тип клиента (Business или Private) | VARCHAR(50) |
| GROSS | REACTIVATION | Если причина активации реактивация, то значение 1, иначе - 0. | BYTEINT |
| TARIFF | PP\_ID | ID тарифного плана, уникальное для каждого тарифного плана | BIGINT |
| TARIFF | PP\_NAME | Коммерческое название Тарифного плана (например, Бизнес-план, Бизнес-план 2.0 и др.) | VARCHAR(100) |
| TARIFF | SUB\_ID | ID\_абонента, уникальное значение для каждого абонента | BIGINT |
| TARIFF | PP\_START\_DATE | Дата начала обслуживания на тарифном плане. Если у абонента было несколько тарифных планов (была смена), то по этому абоненту будет несколько записей с разной датой старта. | TIMESTAMP(0) |
| TARIFF | PP\_END\_DATE | Дата окончания обслуживания на тарифном плане. Если у абонента было несколько тарифных планов (была смена), то по этому абоненту будет несколько записей с разной датой окончания. Актуальный тариф будет иметь временную отметку окончания 2999-01-01 00:00:00 | TIMESTAMP(0) |
| SERVICE | SERVICE\_START\_DATE | Дата подключения пакета | TIMESTAMP(0) |
| SERVICE | SERVICE\_STOP\_DATE | Дата отключения пакета (если отключение было), если не было, стоит временная отметка отключения 2999-01-01 00:00:00 | TIMESTAMP(0) |
| SERVICE | SUB\_ID | ID\_абонента, уникальное значение для каждого абонента | BIGINT |
| SERVICE | SERVICE\_NAME | Коммерческое название пакета (например Безлимитный интернет 512) | VARCHAR(250) |
| SUBSCRIBER | MONTH | Месяц, на конец которого, сделана запись по данному абоненту (например, закрытый декабрь - 2020-12-01) | DATE |
| SUBSCRIBER | COMPANY\_NAME | наименование компании | VARCHAR(100) |
| SUBSCRIBER | SUB\_ID | ID\_абонента, уникальное значение для каждого абонента | BIGINT |
| SUBSCRIBER | CUSTOMER\_TYPE | Тип клиента (Business или Private) | VARCHAR(50) |
| SUBSCRIBER | STATUS | 1 - активный абонент на указанный месяц, 0 - неактивный в указанный | BYTEINT |
| SUBSCRIBER | PP\_TYPE\_ID | Тип тарифного плана абонента (Voice или Fixed) | VARCHAR(50) |

**Задание 4.**

Дана таблица task\_4.csv со следующими полями:

subscriber\_id – id абонента

birthday – дата рождения абонента

gender – пол

handset\_vendor – марка используемого телефонного аппарата

region – область регистрации абонента

price\_plan – наименование тарифного плана

price\_plan\_monthly\_fee – абонентская плата

ARPU – средний счет абонента за месяц

Используя Python:

1. Приведите описательную статистику по абонентской плате и среднему счету, представленные в таблице.
2. Добавьте поле с указанием возраста абонента (полных лет, целое число)
3. Добавьте поле с флагом для тех абонентов, у кого счет за месяц более чем на 15% превышает размер абонентской платы (1 - превышает более чем на 15%, 0 – не превышает).
4. Визуализируйте и проанализируйте распределение абонентов по:

- возрасту и счету

- областям и полу (в разрезе пола)

- по счету, тарифу и марке телефона с отражением степени разброса данных по счету

наиболее подходящим на Ваш взгляд типом диаграммы.

Какие еще диаграммы распределения имеет смысл (будет интересно) проанализировать по предоставленным данным? Приведите эти диаграммы.