**Учебный проект.**

В ходе данного проекта нам предстоит исследовать продажи некого бизнеса.

Анализ будет производиться с помощью:

Сбор необходимых данных посредством PostgreSQL - подключимся к БД в DBeaver

Визуализацияя в Preset Superset

Консолидация графиков и выводов в Google Slides.

База данных состоит из четырех таблиц.

**customers** — таблица покупателей.

• customer\_id — ID покупателя

• first\_name — имя

• middle\_initial — инициал отчества

• last\_name — фамилия

**employees** — таблица сотрудников отдела продаж:

• employee\_id — ID сотрудника

• first\_name — имя

• middle\_initial — инициал отчества

• last\_name — фамилия

**products**— таблица товаров:

• product\_id — ID товара

• name — название

• price — цена

**sales**— таблица с данными о проданных товарах:

• sales\_id

• sales\_person\_id — ID сотрудника, продавшего товар

• customer\_id — ID покупателя

• product\_id — ID товара

• quantity — количество

• sale\_date — дата продажи.

Исследование будет состоять из следующих шагов:

**1. Анализ отдела продаж**

Компания хочет проанализировать работу продавцов. Чтобы в этом помочь,

предоставим несколько отчетов о прошедшей работе.

Первый отчет о десятке лучших продавцов. Таблица состоит из трех колонок -

данных о продавце, суммарной выручке с проданных товаров и количестве

проведенных сделок, и отсортирована по убыванию выручки:

⁃ name — имя и фамилия продавца

⁃ operations - количество проведенных сделок

⁃ income — суммарная выручка продавца за все время

Второй отчет содержит информацию о продавцах, чья средняя выручка за сделку

меньше средней выручки за сделку по всем продавцам. Таблица отсортирована по

выручке по возрастанию.

⁃ name — имя и фамилия продавца

⁃ average\_income — средняя выручка продавца за сделку с округлением до

⁃ целого

Третий отчет содержит информацию о выручке по дням недели. Каждая запись

содержит имя и фамилию продавца, день недели и суммарную выручку.

Отсортируйте данные по порядковому номеру дня недели и name

⁃ name — имя и фамилия продавца

⁃ weekday — название дня недели на английском языке

⁃ income — суммарная выручка продавца в определенный день недели,

⁃ округленная до целого числа

**2. Анализ покупателей**

Поможем компании проанализировать данные о покупателях и покупках. Для того,

чтобы оказать помощь в этом деле, вам необходимо предоставить следующие

данные:

Первый отчет - количество покупателей в разных возрастных группах: 16-25, 26-

40 и 40+. Итоговая таблица должна быть отсортирована по возрастным группам и

содержать следующие поля:

⁃ age\_category - возрастная группа

⁃ count - количество человек в группе

Во втором отчете предоставьте данные по количеству уникальных покупателей и

выручке, которую они принесли. Сгруппируйте данные по дате, которая

представлена в числовом виде ГОД-МЕСЯЦ. Итоговая таблица должна быть

отсортирована по дате по возрастанию и содержать следующие поля:

⁃ date - дата в указанном формате

⁃ total\_customers - количество покупателей

⁃ income - принесенная выручка

Третий отчет следует составить о покупателях, первая покупка которых была в

ходе проведения акций (акционные товары отпускали со стоимостью равной 0).

Итоговая таблица должна быть отсортирована по id покупателя. Таблица состоит

из следующих полей:

⁃ customer - имя и фамилия покупателя

⁃ sale\_date - дата покупки

⁃ seller - имя и фамилия продавца

Задачи

1. Подготовьте в файл age\_groups.csv с возрастными группами покупателей

2. Подготовьте в файл customers\_by\_month.csv с количеством покупателей и

выручкой по месяцам

**3. Дашборд**

Результаты проведенного исследования продавцов и покупателей следует

представить в виде дашборда, который наилучшим образом отразит полученные

результаты.

Задачи

1. Загрузите следующие csv-файлы в Preset:

o top\_10\_total\_income.csv

o lowest\_average\_income.csv

o day\_of\_the\_week\_income.csv

o age\_groups.csv

o customers\_by\_month.csv

o special\_offer.csv

2. Составьте дашборд по имеющимся данным

**4. Презентация**

Полученные данные необходимо представить в виде презентации с анализом

полученных результатов. Презентация должна быть интересной, наглядной и

понятной для аудитории, не знакомой с дата-аналитикой.