

Задание

Руководитель отдела сопровождения клиентов хочет получить информацию о том, насколько хорошо сотрудники его отдела (аккаунт-менеджеры) справляются с одной из своих основных задач – пролонгацией договоров с клиентами. От аналитика он хочет получить отчет о пролонгациях сотрудников за 2023 год.

В компании используется два коэффициента пролонгации:

1. Для проектов пролонгированных в первый месяц – отношении суммы отгрузки¹ проектов пролонгированных в первый месяц после завершения к сумме отгрузки последнего месяца реализации всех завершившихся в прошлом месяце проектов.
2. Для проектов, пролонгированных во второй месяц – отношение суммы отгрузки проектов, пролонгированных во второй месяц к сумме отгрузки последнего месяца проектов, не пролонгированных в первый.

То есть, если нам нужно понять, насколько хорошо менеджер пролонгировал в мае, необходимо посчитать:

1. Сумму отгрузки проектов, завершившихся в апреле (за апрель) и сумму отгрузки тех проектов завершившихся в апреле, у которых есть отгрузка в мае (за май). Коэффициент – отношение второй суммы к первой.
2. Сумму проектов, завершившихся в марте, у которых нет отгрузки в апреле (за март) и сумму отгрузки тех проектов, завершившихся в марте, у которых нет отгрузки в апреле но есть в мае (за май). Коэффициент – отношение второй суммы к первой.

Имеются два набора данных:

1. **prolongations.csv**

id – id проекта

month – последний месяц реализации проекта

AM – ФИО ответственного аккаунт-менеджера (данные первичны по отношению к financial_data)

2. **financial_data.csv:**

id – id проекта

Причина дубля – причина, почему строки с одним и тем же id встречаются несколько раз

Колонки с названием месяца – сумма отгрузки проекта в данный месяц²

Account – ФИО ответственного аккаунт-менеджера

¹ Отгрузка – сумма из таблицы financial_data.csv

² В данной колонке могут также встречаться такие значения: '**в ноль**' – отгрузка проекта в данном месяце равна 0, значит для коэффициента пролонгации нужно взять отгрузку предыдущего месяца (только если все части оплаты равны 0); '**стоп**' – проект закончился до истечения срока договора, если у проекта есть "стоп" в последний месяц реализации или ранее, то такой проект исключаем из пролонгаций; '**end**' – аналогично 'стоп'

Необходимо:

1. Рассчитать коэффициенты пролонгации для каждого менеджера и для всего отдела в целом
 - a. за каждый месяц
 - b. за год
2. Сформировать аналитический отчет в гугл-таблицах или excel, на основании которого руководитель отдела будет принимать управленческие решения (обязательно присутствие коэффициентов пролонгации, в остальном отчет можно дополнить на свое усмотрение – дополнительные метрики и визуализация приветствуются).

Отчет должен быть читабельным и понятным для руководителя отдела.

В результате ждем 2 файла:

1. Код с подробными пояснениями логики расчетов в Google Colab или Jupiter Notebook
2. Аналитический отчет для руководителя отдела (в excel/google sheets)