## Тестовое задание на позицию “Инженер данных”

**Обязательные требования:**

Результат выполненного задания должен быть представлен в виде SQL запросов решающий предложенные задачи. Одна задача, один запрос. Запросы необходимо передать HR’у в виде текстового файла с комментариями, какой запрос какую задачу решает. Задание может быть выполнено с использованием следующих баз данных Clickhouse или PostgreSQL.

Структура таблиц и данные необходимые для решения заданий, представлены в виде sql файлов по ссылкам ниже (перед выполнением необходимо выполнить данные запросы).

* <https://drive.google.com/file/d/147lK8Ii7q8plgg7Ep4G_nF9KttNjdth_/view?usp=sharing> - для выполнения задания с использование базы данных Clickhouse
* <https://drive.google.com/file/d/173EX4y_BftD50HjnJcrYyrjElJsWWjLV/view?usp=sharing> - для выполнения задания с использование базы данных PostgreSQL

Описание структуры таблиц

manager\_departments - таблица содержащая информацию о принадлежности менеджера к отделу

* email - email адрес менеджера
* department - отдел к которому относится менеджер

payments - таблица содержащая информацию о поступлениях

* id - уникальный идентификатор поступления
* value - сумма поступления
* client\_id - идентификатор клиента осуществивший платеж
* client\_name - название клиента осуществивший платеж
* payment\_date - дата поступления
* manager\_name - ФИО менеджера ответственного за клиента
* manager\_email - email адрес менеджера ответственного за клиента

**Дополнительные требования:**

Требования перечисленные ниже не являются обязательными, но их выполнение (в том числе частичное) будет расцениваться как дополнительный плюс и повысит шансы на трудоустройство:

* Использование Clickhouse в качестве базы данных
* Использование git (в данном случае результат можно передать в виде ссылки на публичный репозиторий github, bitbucket)
* Использование docker для запуска базы данных (приложить docker-compose файл)
* Использование инструмента трансформации данных <https://www.getdbt.com/>

**Задание №1**

На основе таблицы с поступлениями, необходимо найти платежи с максимальными суммами в январе 2023 года. Результат должен быть представлен в виде таблицы со следующими полями:

* Название клиента
* Дата поступления
* Сумма поступления

Пример исходной таблицы поступлений payments

+--+------+---------+-----------+------------+------------+------------------------+

|id|value |client\_id|client\_name|payment\_date|manager\_name|manager\_email |

+--+------+---------+-----------+------------+------------+------------------------+

|1 |531320|50 |Клиент #50 |2023-01-09 |Manager #23 |nmanager23@profitbase.ru|

|2 |515660|41 |Клиент #41 |2023-01-09 |Manager #16 |nmanager16@profitbase.ru|

|3 |528350|264 |Клиент #264|2023-01-10 |Manager #23 |nmanager23@profitbase.ru|

|4 |878000|109 |Клиент #109|2023-01-10 |Manager #12 |nmanager12@profitbase.ru|

|5 |508100|21 |Клиент #21 |2023-01-12 |Manager #23 |nmanager23@profitbase.ru|

|6 |515660|124 |Клиент #124|2023-01-13 |Manager #11 |nmanager11@profitbase.ru|

|7 |543065|203 |Клиент #203|2023-01-14 |Manager #19 |nmanager19@profitbase.ru|

|8 |552674|23 |Клиент #23 |2023-01-17 |Manager #7 |nmanager7@profitbase.ru |

+--+------+---------+-----------+------------+------------+------------------------+

Пример результирующей таблицы

+-----------+------------+------+

|client\_name|payment\_date|value |

+-----------+------------+------+

|Клиент #1 |2023-01-09 |531320|

|Клиент #2 |2023-01-09 |515660|

|Клиент #3 |2023-01-10 |528350|

+-----------+------------+------+

**Задание №2**

Необходимо определить сколько принес денег каждый отдел. Таблица с поступлениями хранит информацию с имением и email адресе менеджера. Информация об отделе хранится в отдельной таблице-справочнике manager\_departments. Связь этих двух таблиц осуществляется по полю email, но формат почты отличается. В каких-то случаях email менеджера представлен в формате n.ssssssss@domain.ru, где n это первая буква имени, ssssssss фамилия менеджера и между ними “.” (знак точки). В каких-то случаях формат имеет вид nssssssss@domain.ru, без точки между именем и фамилией. Необходимо объединить две таблицы с предварительным приведением почты к единому виду и дальнейшей агрегацией. Если для менеджера не определен отдел (указано значение NULL), то таких менеджеров необходимо отнести в группу “Отдел не определен”

Пример исходной таблицы поступлений payments

+--+------+---------+-----------+------------+------------+------------------------+

|id|value |client\_id|client\_name|payment\_date|manager\_name|manager\_email |

+--+------+---------+-----------+------------+------------+------------------------+

|1 |531320|50 |Клиент #50 |2023-01-09 |Manager #23 |nmanager23@profitbase.ru|

|2 |515660|41 |Клиент #41 |2023-01-09 |Manager #16 |nmanager16@profitbase.ru|

|3 |528350|264 |Клиент #264|2023-01-10 |Manager #23 |nmanager23@profitbase.ru|

|4 |878000|109 |Клиент #109|2023-01-10 |Manager #12 |nmanager12@profitbase.ru|

|5 |508100|21 |Клиент #21 |2023-01-12 |Manager #23 |nmanager23@profitbase.ru|

|6 |515660|124 |Клиент #124|2023-01-13 |Manager #11 |nmanager11@profitbase.ru|

|7 |543065|203 |Клиент #203|2023-01-14 |Manager #19 |nmanager19@profitbase.ru|

|8 |552674|23 |Клиент #23 |2023-01-17 |Manager #7 |nmanager7@profitbase.ru |

+--+------+---------+-----------+------------+------------+------------------------+

Пример исходной таблицы с отделами manager\_departments

+-------------------------+-----------------+

|email |department |

+-------------------------+-----------------+

|n.manager10@profitbase.ru|ТОП-менеджмент |

|n.manager11@profitbase.ru|Клиентский сервис|

|n.manager12@profitbase.ru|Отдел продаж |

|n.manager13@profitbase.ru|null |

+-------------------------+-----------------+

Пример результирующей таблицы

+------------------+-------+

|department |sum |

+------------------+-------+

|ТОП-менеджмент |1095739|

|Клиентский сервис |515660 |

|Отдел продаж |878000 |

|Клиентский сервис |2083430|

+------------------+-------+

**Задание №3**

Необходимо дополнить таблицу с поступлениями полем “состояние клиента”. Если платеж для клиента является первым, то состояние клиента должно быть “Новый”. Для всех остальных платежей данного клиента, состояние должно быть “Действующий”.

Пример исходной таблицы поступлений payments

+--+------+---------+-----------+------------+------------+------------------------+

|id|value |client\_id|client\_name|payment\_date|manager\_name|manager\_email |

+--+------+---------+-----------+------------+------------+------------------------+

|1 |531320|50 |Клиент #50 |2023-01-09 |Manager #23 |nmanager23@profitbase.ru|

|2 |515660|41 |Клиент #41 |2023-01-09 |Manager #16 |nmanager16@profitbase.ru|

|3 |528350|264 |Клиент #264|2023-01-10 |Manager #23 |nmanager23@profitbase.ru|

|4 |878000|109 |Клиент #109|2023-01-10 |Manager #12 |nmanager12@profitbase.ru|

|5 |508100|21 |Клиент #21 |2023-01-12 |Manager #23 |nmanager23@profitbase.ru|

|6 |515660|124 |Клиент #124|2023-01-13 |Manager #11 |nmanager11@profitbase.ru|

|7 |543065|203 |Клиент #203|2023-01-14 |Manager #19 |nmanager19@profitbase.ru|

|8 |552674|23 |Клиент #23 |2023-01-17 |Manager #7 |nmanager7@profitbase.ru |

+--+------+---------+-----------+------------+------------+------------------------+

Пример результирующей таблицы

+--+------+---------+-----------+------------+------------+------------------------+-----------------+

|id|value |client\_id|client\_name|payment\_date|manager\_name|manager\_email |client\_state |

+--+------+---------+-----------+------------+------------+------------------------+-----------------+

|1 |531320|50 |Клиент #50 |2023-01-09 |Manager #23 |nmanager23@profitbase.ru|Новый |

|2 |515660|41 |Клиент #41 |2023-01-09 |Manager #16 |nmanager16@profitbase.ru|Новый |

|3 |528350|264 |Клиент #264|2023-01-10 |Manager #23 |nmanager23@profitbase.ru|Новый |

|4 |878000|109 |Клиент #109|2023-01-10 |Manager #12 |nmanager12@profitbase.ru|Действующий |

|5 |508100|21 |Клиент #21 |2023-01-12 |Manager #23 |nmanager23@profitbase.ru|Действующий |

|6 |515660|124 |Клиент #124|2023-01-13 |Manager #11 |nmanager11@profitbase.ru|Новый |

|7 |543065|203 |Клиент #203|2023-01-14 |Manager #19 |nmanager19@profitbase.ru|Новый |

|8 |552674|23 |Клиент #23 |2023-01-17 |Manager #7 |nmanager7@profitbase.ru |Действующий |

+--+------+---------+-----------+------------+------------+------------------------+-----------------+

**Задание №4**

Необходимо сформировать таблицу в которой будет отражено сколько поступило денег в каждом месяце и как росла общая выручка с учетом предыдущих месяцев.

Пример исходной таблицы поступлений payments

+--+------+---------+-----------+------------+------------+------------------------+

|id|value |client\_id|client\_name|payment\_date|manager\_name|manager\_email |

+--+------+---------+-----------+------------+------------+------------------------+

|1 |531320|50 |Клиент #50 |2023-01-09 |Manager #23 |nmanager23@profitbase.ru|

|2 |515660|41 |Клиент #41 |2023-01-09 |Manager #16 |nmanager16@profitbase.ru|

|3 |528350|264 |Клиент #264|2023-01-10 |Manager #23 |nmanager23@profitbase.ru|

|4 |878000|109 |Клиент #109|2023-01-10 |Manager #12 |nmanager12@profitbase.ru|

|5 |508100|21 |Клиент #21 |2023-01-12 |Manager #23 |nmanager23@profitbase.ru|

|6 |515660|124 |Клиент #124|2023-01-13 |Manager #11 |nmanager11@profitbase.ru|

|7 |543065|203 |Клиент #203|2023-01-14 |Manager #19 |nmanager19@profitbase.ru|

|8 |552674|23 |Клиент #23 |2023-01-17 |Manager #7 |nmanager7@profitbase.ru |

+--+------+---------+-----------+------------+------------+------------------------+

Пример результирующей таблицы

+-------------+-----------------+-------------------------+

|period |revenue\_by\_month |revenue\_cumulative |

+-------------+-----------------+-------------------------+

|Январь 2023 |2961430 |2961430 |

|Февраль 2023 |1058725 |4020155 |

|Март 2023 |552674 |4572829 |

+-------------+-----------------+-------------------------+