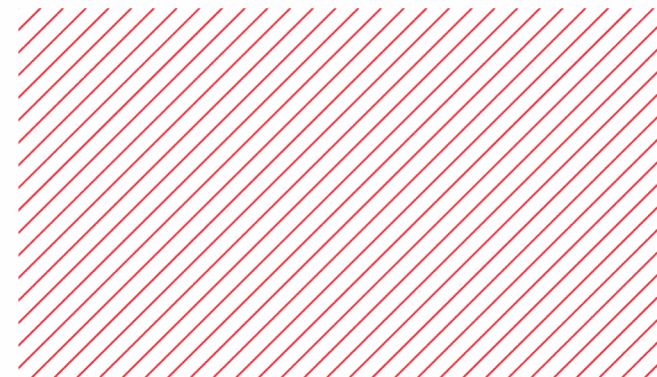


академия
больших
данных

mail.ru
group

made

HW03: Hive QL





Описание работы и критериев оценивания

Для данной задачи мы попрактикуемся в интерактивной среде Hue, где будем работать с Hive QL, создавая различные запросы.

Веб интерфейс Hue доступен по ссылке <https://demo.gethue.com/hue/accounts/login?next=/> (username:demo, password:demo)

Залогинившись, выберите Hive, как основную среду исполнения, после этого используйте базу default.

В работе, мы будем использовать две таблицы: sample_07 и sample_08, которые представляют информацию о внутреннем устройстве отделов двух компаний с указанием количества сотрудников и их зарплат.

В качестве решения, предлагается выслать код для запросов в формате .txt либо .sql

Бонусы и штрафы:

- **100%** за плагиат в решениях (всем участникам процесса)
- **30%** за посылку решения в течение недели после deadline
- **10%** за несоответствие правилам оформления задачи
- **5%** за каждую дополнительную посылку в тестирующую систему



Этапы задания

1. Вывести название компании с максимальным количеством сотрудников и зарплатой по отделам
2. Вывести список ID отделов обеих компаний, количество сотрудников в которых превышает 50000 человек
3. Для прошлого задания посчитать долю сотрудников отдела от всех сотрудников конкретной компании
4. Проверить, существуют ли отделы, работающие в одной компании, но отсутствующие в другой
5. Отсортировать отделы компаний по суммарной зарплате по убыванию. Использовать информацию по отделам обеих компаний суммарно