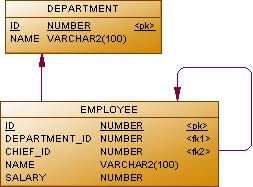
1. **SQL**

Есть БД следующего вида (для примера есть sqlite-версия во вложении):



|  |  |
| --- | --- |
| **Кодификатор:** |  |
| DEPARTMENT.NAME | Название отдела |
| EMPLOYEE.ID | ID сотрудника |
| EMPLOYEE.CHIEF\_ID | ID руководителя |
| EMPLOYEE.NAME | Имя сотрудника |
| EMPLOYEE.SALARY | Зарплата |

Напишите следующие SQL-запросы:

* 1. Вывести список сотрудников, получающих заработную плату выше 50000. Формат вывода - список ID
  2. Вывести названия всех отделов, где количество сотрудников, зарабатывающих более 50000, менее пяти. Вывести поля: Название отдела, количество сотрудников.
  3. Вывести список сотрудников с минимальной заработной платой в своих отделах в формате ID, ФИО, Название отдела

1. **API**

С помощью запроса к методу /characters API <https://anapioficeandfire.com/>, нужно составить список актёров (см. в ответе API ключ playedBy), чьи персонажи из Вестероса и живы (culture=Westeros, isAlive=true).

1. **Статистика**

Был проведен A/B-тест на конверсию. Объемы групп одинаковые, по 500 клиентов. В группе А случилось 99 целевых событий; в группе В - 101. Какие выводы мы можем сделать на основе этих данных? Какая дополнительная информация может быть полезна, на что еще стоит обратить внимание?

1. **Classic ML**

Во вложении sample.xlsx содержится таблица с данными по 5000 клиентов в следующей кодификации:

* No : Порядковый номер клиента
* predictor1-7\* : Фичи (предикторы) по которым необходимо построить модель.
* Target : Целевая бинарная переменная, обозначающая факт дефолта клиента

Задачи:

1. Необходимо построить модель, позволяющую предсказывать вероятность дефолта клиента. Предлагается построить 2 конкурирующие модели (например, логистическая регрессия и дерево решений) и произвести выбор наилучшей модели на основе предложенного вами показателя качества.
2. Продемонстрируйте силу отдельных предикторов, используя предложенный вами метод. Постройте модель используя только 3 предиктора. Допускается создание своих фичей на основе имеющихся.
3. Напишите класс транформер совместимый с пайплайном sklearn, который проведет отбор (или преобразование) предикторов. То есть количество полученных предикторов будет параметром этого класса (возможно и увеличение количества фичей). Сделайте sklearn пайплайн, где одним из шагов будет применяться этот класс.