

Лабораторная работа №3

Визуализация в python

Пояснение к заданию:

1. Задание требуется выполнять jupyter notebook.
2. Для подключения к базе данных clickhouse можно использовать библиотеку clickhouse_driver.
3. Для подключения к базе можно использовать следующий код

```
from clickhouse_driver import Client
```

```
client = Client('oleg.orbita.work', port = 9000, user=student_ml,  
password=CGRV8zemLGgsfdz7Uc6ZJeisGTcPQa, settings={'use_numpy':  
True})
```

Задание

4. Загрузите данные из таблицы game_sales, которая находится в базе students_datas.

Расшифровка:

- Name – название игры,
- Platform – игровая платформа (PC, PSP, X360 и др.),
- Year – год выпуска игры,
- Genre – жанр игры,
- Publisher – издатель игры,
- NA_Sales – продажи в Северной Америке (в миллионах),
- EU_Sales – продажи в Европе (в миллионах),
- JP_Sales – продажи в Японии (в миллионах),
- Other_Sales – продажи в остальных странах мира (в миллионах),
- Global_Sales – объем продаж по всему миру.

рассчитайте необходимые показатели и визуализируйте информацию, используя различные инструменты.

2. Проанализируйте полученные графики и сделайте выводы.

Игры каких жанров были наиболее популярны до 2000 года, а какие после? Оцените популярность жанров по количеству выпущенных игр и по объему продаж по всему миру. Для визуализации полученных результатов используйте столбиковые диаграммы. Замечание. Одна и та же игра может встречаться в выборке несколько раз, т.к. она может быть выпущена на нескольких платформах.

2. Отобразите на графике общее число видеоигр, выпущенных в каждом году.
3. Определите трех издателей, выпустивших наибольшее количество видеоигр. Изобразите количество выпущенных издателями видеоигр для каждой платформы на столбиковой диаграмме (можно использовать диаграмму с накоплением).

4. Отобразите на круговых диаграммах доли суммарного объема продаж с 1980г. до 2000г. и с 2000г. до 2020г. в Северной Америке, Европе, Японии от объема продаж по всему миру.