11 колонок

6 (первые) – исходные данные

5 (последние) – выходные данные

16598 значений

# Информация по столбцам

1. Рейтинг (ID)
2. Название игры
3. Платформа выпуска игры (DS(13%)/PS2(13%)/Other(74%))
4. Год выпуска игры (2009(9%)/2008(9%)/Other(83%))
5. Жанр игры (Action(20%)/Sports(14%)/Others(66%))
6. Издатель игры (EA(8%)/Activision(6%)/Other(86%))
7. Продажи в северной Америке (в миллионах) (0-41,5)
8. Продажи в Европе (в миллионах) (0-29)
9. Продажи в Японии (в миллионах) (0-10,2)
10. Продажи в остальном мире (в миллионах) (0-10,6)
11. Общий объём продаж по всему миру (0,01-82,7)

# Общая информация по датасету

Этот набор данных содержит список видеоигр с продажами более 100 000 копий. Он был создан при сканировании vgchartz.com.

Поля включают

* Рейтинг - Рейтинг общих продаж
* Имя - Название игры.
* Платформа - Платформа выпуска игр (например, ПК, PS4 и т. Д.).
* Год - Год выпуска игры.
* Жанр - Жанр игры
* Издатель - 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 остальном мире (в миллионах)
* Global\_Sales - Общий объем продаж по всему миру.

Скрипт для очистки данных доступен по адресу https://github.com/GregorUT/vgchartzScrape.

Он основан на BeautifulSoup с использованием Python.

Есть 16 598 записей. 2 записи были сброшены из-за неполной информации.