

Университет ИТМО

Практическая работа №4  
по дисциплине «Визуализация и моделирование»

**Автор:** Логвинов Лев Анатольевич

**Поток:** 11.03.02

**Группа:** К3220

**Факультет:** ИКТ

**Преподаватель:** Чернышева А.В.

Санкт-Петербург, 2021 г.

Ссылка на датасет: <https://www.kaggle.com/gregorut/videogamesales>

Данный датасет содержит список видеоигр, у которых было продано более 100 тыс. копий. В датасете представлена информация о самих играх (название, платформа, год выпуска и т.п.), а также объемы их продаж на различных территориях (Северная Америка, Европа, Япония, остальной мир) и по всему миру в целом.

Название столбца	Данные, хранящиеся в столбце	Тип данных	Шкала
Rank	Рейтинг общих продаж	Целое число	Порядковая
Name	Название игры	Строка	Номинальная
Platform	Платформа выпуска игры	Строка	Номинальная
Year	Год выпуска игры	Целое число	Относительная
Genre	Жанр игры	Строка	Номинальная
Publisher	Издатель игры	Строка	Номинальная
NA_Sales	Продажи в Северной Америке (млн.)	Число с плавающей точкой	Относительная
EU_Sales	Продажи в Европе (млн.)	Число с плавающей точкой	Относительная
JP_Sales	Продажи в Японии (млн.)	Число с плавающей точкой	Относительная
Other_Sales	Продажи в остальном мире (млн.)	Число с плавающей точкой	Относительная
Global_Sales	Общий объем продаж по всему миру	Число с плавающей точкой	Относительная

Были получены визуализации данных как во 2 практической работе. Без углубления в цифры данные выглядят идентично на всех графиках. Это связано с тем, что датасет был лишь немного изменен (были удалены строки, содержащие пустые значения).

Гипотезы:

1. Около 5% игр имеют значение общего количества продаж сильно большее, чем у остальных 95% (в 3, 5, 10 раз больше).

Гипотеза не подтвердилась. Гораздо меньшее количество игр имеют значение общего количества продаж сильно большее, чем у остальных.

```

> print(df_n["Global_Sales"][df_n.Global_Sales > 2.05].count())
811

```

2. Около 90% игр имеют значение общего количества продаж до 1 миллиона включительно.

Видно, что почти все значения (около 90%) имеют значение от 0 до 1. В данном датасете около 16500 строк, 14233 из которых в столбце Global\_Sales имеют значение меньше 1 (миллиона). Следовательно, данная теория подтвердилась.

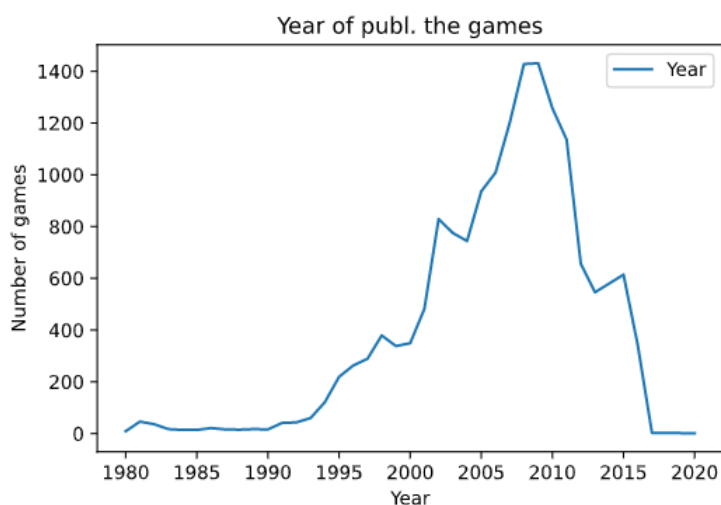
```

> df_n["Global_Sales"][df_n.Global_Sales < 1].count()
14233

```

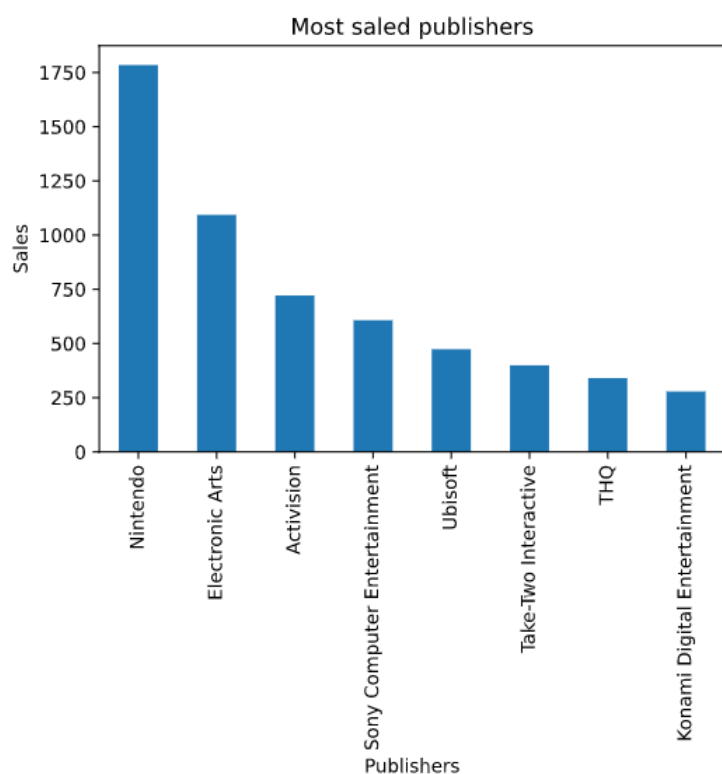
3. Пиком продажи копий игр по всему миру является 2009 год.

При визуализации данная гипотеза подтвердилась. Из графика видно, что пиковое значение он принимает при значении Year = 2009.

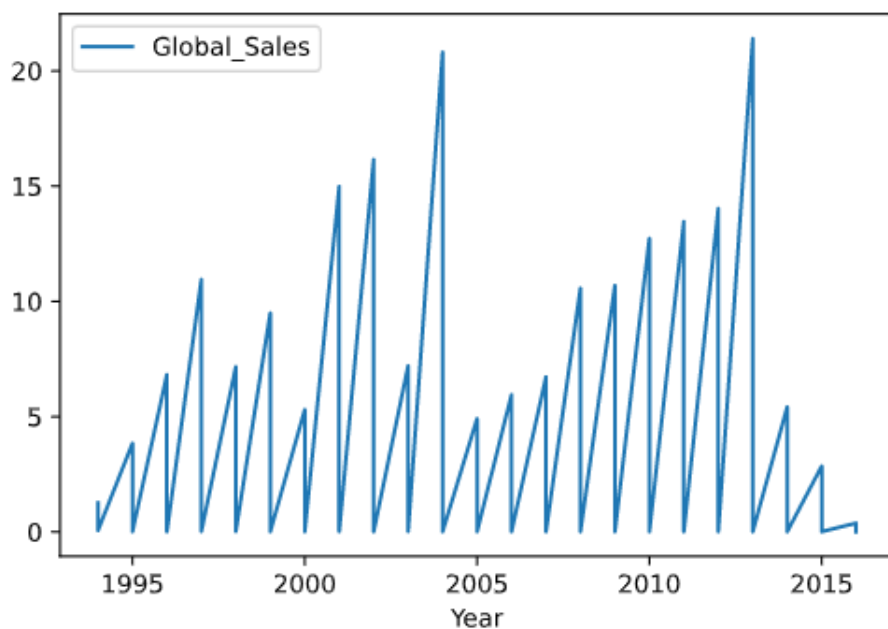


4. Наиболее популярным издателем игр, имеющим наибольшее число проданных копий, является Nintendo.

При визуализации данная гипотеза подтвердилась. Из графика видно, что наиболее популярным издателем игр является Nintendo, обгоняя 2 место почти в два раза.



5. С развитием PlayStation (PS, PS2, PS3) рос спрос на игры, выпускаемые на данную консоль.



6. Наиболее популярная игра была выпущена после 2000 года.  
Из таблицы, отсортированной по убыванию количества продаж по всему миру, видно, что самая продаваемая игра была выпущена в 2006 году (Wii Sports).

	Rank	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales
0	1	Wii Sports	Wii	2006.0	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74
1	2	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24
2	3	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82
3	4	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00
4	5	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37

## 7. Наиболее популярные в Японии игры имеют жанр Role-playing

На графике видно, что наиболее популярным игровым жанром на территории Японии является Role-Playing (зеленые столбцы)

