

Analyse du jeu de données

« Video_games »
{

- ⌘ Quels indicateurs impactent le plus les Global_Sales
- ⌘ Quels sont les jeux qui fonctionnent le mieux (top 1%)

Quels KPI chercher?

Executive summary

	Name	Platform	Year_of_Release	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count
0	Wii Sports	Wii	2006.0	Sports	Nintendo	41.36	28.96	3.77	8.45	82.53	76.0	51.0	9.3	9100
1	Super Mario Bros.	NES	1985.0	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	NaN	NaN	9.3	9100
2	Mario Kart Wii	Wii	2008.0	Racing	Nintendo	15.68	12.76	3.79	3.29	35.52	82.0	73.0	9.3	9100
3	Wii Sports Resort	Wii	2009.0	Sports	Nintendo	15.61	10.93	3.28	2.95	32.77	80.0	73.0	9.3	9100
4	Pokemon Red/Pokemon Blue	GB	1996.0	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37	NaN	NaN	9.3	9100

↳ 16 colonnes correspondant à autant de variables:

- ↳ Name
- ↳ Platform
- ↳ Year_of_Release
- ↳ Genre
- ↳ Publisher
- ↳ NA_Sales
- ↳ EU_Sales
- ↳ JP_Sales
- ↳ Other_Sales
- ↳ Global_Sales
- ↳ Critic_Score
- ↳ Critic_Count
- ↳ User_Score
- ↳ User_Count
- ↳ Developer
- ↳ Rating

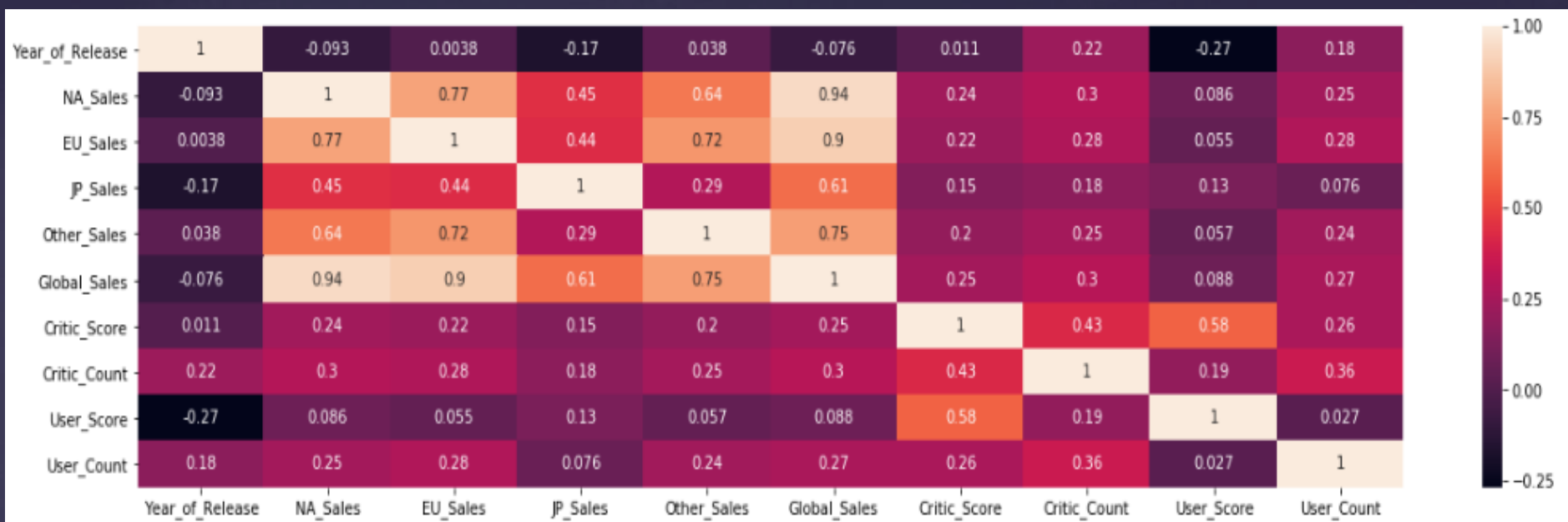
Statistiques descriptives

	Year_of_Release	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count
count	16450.000000	16719.000000	16719.000000	16719.000000	16719.000000	16719.000000	8137.000000	8137.000000	7590.000000	7590.000000
mean	2006.487356	0.263330	0.145025	0.077602	0.047332	0.533543	68.967679	26.360821	7.125046	162.229908
std	5.878995	0.813514	0.503283	0.308818	0.186710	1.547935	13.938165	18.980495	1.500006	561.282326
min	1980.000000	0.000000	0.000000	0.000000	0.000000	0.010000	13.000000	3.000000	0.000000	4.000000
25%	2003.000000	0.000000	0.000000	0.000000	0.000000	0.060000	60.000000	12.000000	6.400000	10.000000
50%	2007.000000	0.080000	0.020000	0.000000	0.010000	0.170000	71.000000	21.000000	7.500000	24.000000
75%	2010.000000	0.240000	0.110000	0.040000	0.030000	0.470000	79.000000	36.000000	8.200000	81.000000
max	2020.000000	41.360000	28.960000	10.220000	10.570000	82.530000	98.000000	113.000000	9.700000	10665.000000

⌘ Quelles conclusions en tirer?

- ⌘ 16719 jeux vidéos analysés, dont 16450 pour lesquels nous avons un Year_of_release, 8137 pour lesquels nous avons un score "critique" et 7590 pour lesquels nous avons un score "User"
- ⌘ Le dataset s'étale de 1980 à 2020 (précommandes?): 75% des jeux sont sortis après 2003, l'année moyenne de sortie est 2006,5 et la médiane est en 2007
- ⌘ On trouve des valeurs étranges pour les Global_Sales: la médiane est à 0,02 alors que le maximum est à 28,9, il y a trop de jeu pour lesquels nous n'avons pas de vente. En réalité, le minimum de Global Sales est à 0,01 donc non nul, effectivement, certains jeux ne se vendent pas DU TOUT.
- ⌘ Au niveau des scores Critiques on trouve des scores entre 0 et 100 (13 étant le minimum) avec une moyenne à 69/100 et une médiane à 71/100
- ⌘ Au niveau des scores User on trouve des scores entre 1 et 10 avec une moyenne à 7,1/10

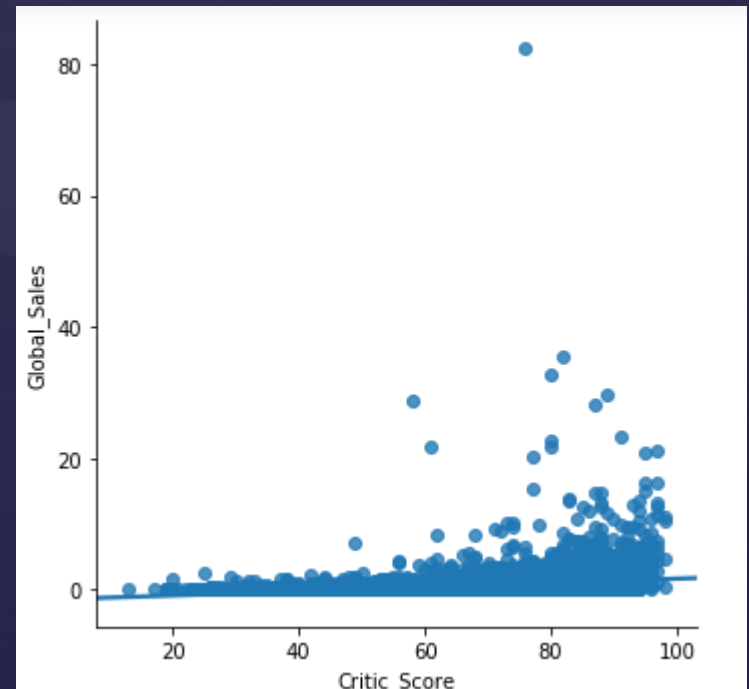
Statistiques descriptives

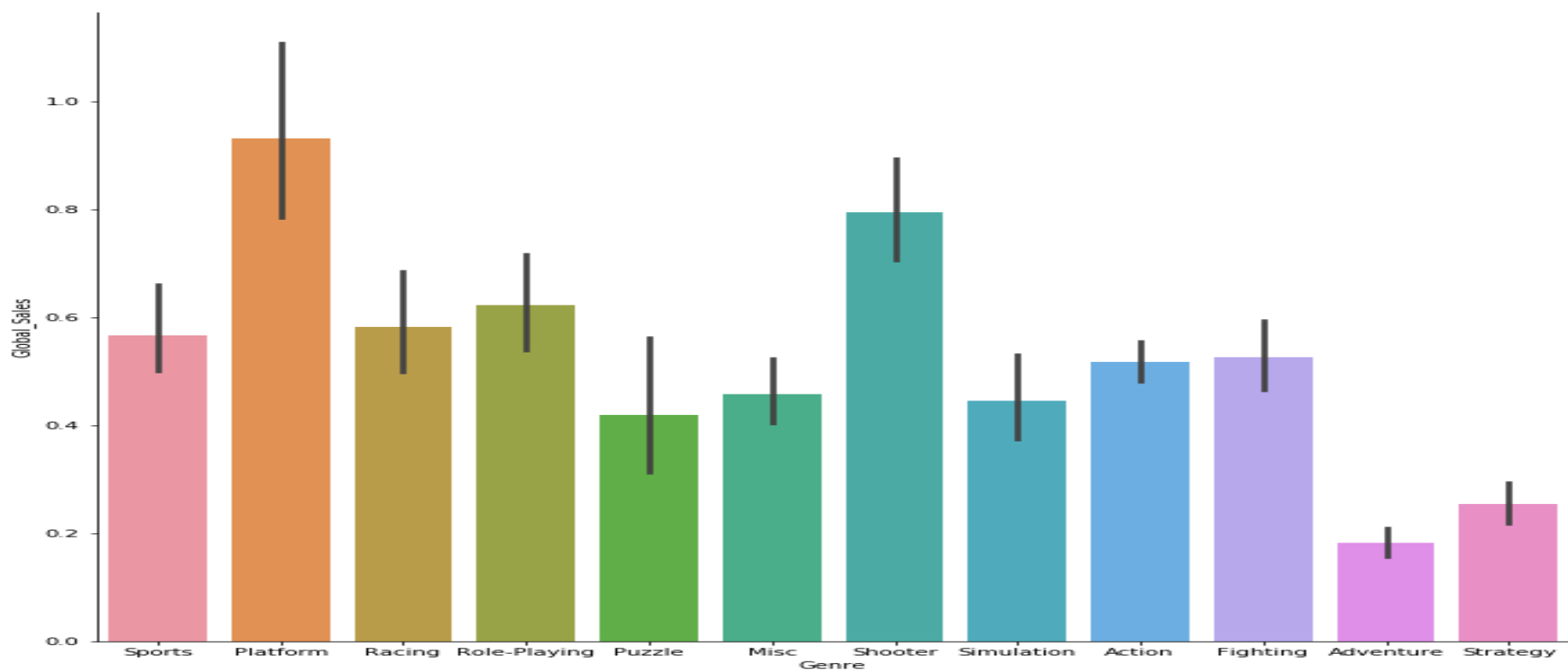


Cherchons les valeurs les plus corrélées avec Global_Sales:

- ↳ Evidemment, les Sales NA, EU, JP et Other
- ↳ Le critic_Score est assez peu corrélé à 0,25, on peut l'observer sur le graphique ci-après: un mauvais jeu peut bien se vendre.
- ↳ Le user_score est encore moins corrélé 0,08
- ↳ Le critic_count et user_count sont d'avantage corrélés (0,30) : plus un jeu a de critiques, plus il est populaire.
- ↳ **Au final, le buzz est plus important que la qualité du jeu.**

Global_Sales





Sur l'ensemble du jeu de données, les jeux les plus susceptibles de se vendre sont:

- ↳ Les jeux de plateformes (900k en moyenne)
- ↳ Les shooters (800k en moyenne)
- ↳ Les role_playing, Racing et Sports à 0,6 environ

Genre de jeu et Global_Sales

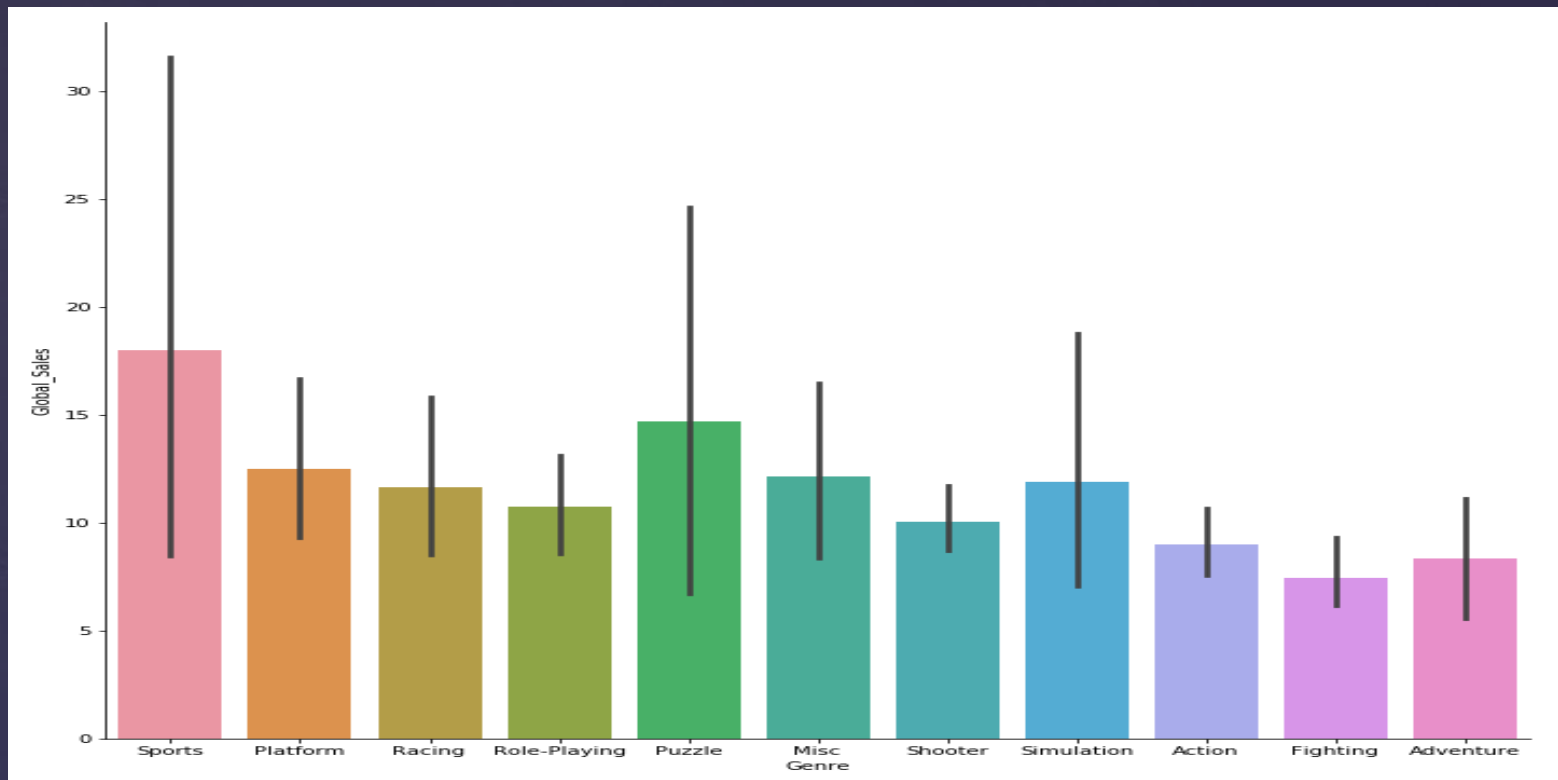
	Year_of_Release	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	Critic_Count	User_Score	User_Count
count	168.000000	168.000000	168.000000	168.000000	168.000000	168.000000	120.000000	120.000000	120.000000	120.000000
mean	2005.333333	5.369345	3.333036	1.520417	1.024167	11.246905	86.658333	55.858333	7.700000	1358.741667
std	7.393691	4.990402	2.966532	1.806232	1.269164	8.644843	9.183009	25.908467	1.389456	1588.382119
min	1982.000000	0.080000	0.000000	0.000000	0.000000	5.460000	49.000000	8.000000	2.600000	16.000000
25%	2001.000000	2.907500	1.850000	0.125000	0.450000	6.485000	82.750000	37.000000	7.300000	217.250000
50%	2007.000000	3.985000	2.575000	0.850000	0.740000	8.060000	88.000000	57.500000	8.050000	891.000000
75%	2011.000000	6.037500	3.862500	2.500000	1.130000	12.615000	94.000000	77.000000	8.625000	2052.000000
max	2016.000000	41.360000	28.960000	10.220000	10.570000	82.530000	98.000000	105.000000	9.400000	8713.000000

168 jeux représentent le 1% qui s'est le mieux vendu avec un minimum de Global Sales à 5.45M.

Ces jeux se répartissent de façon similaire aux autres jeux sur l'année de sortie (Year_of_Release)

Ces jeux ont une moyenne critique de 86,6 et un user_score de 7,7.

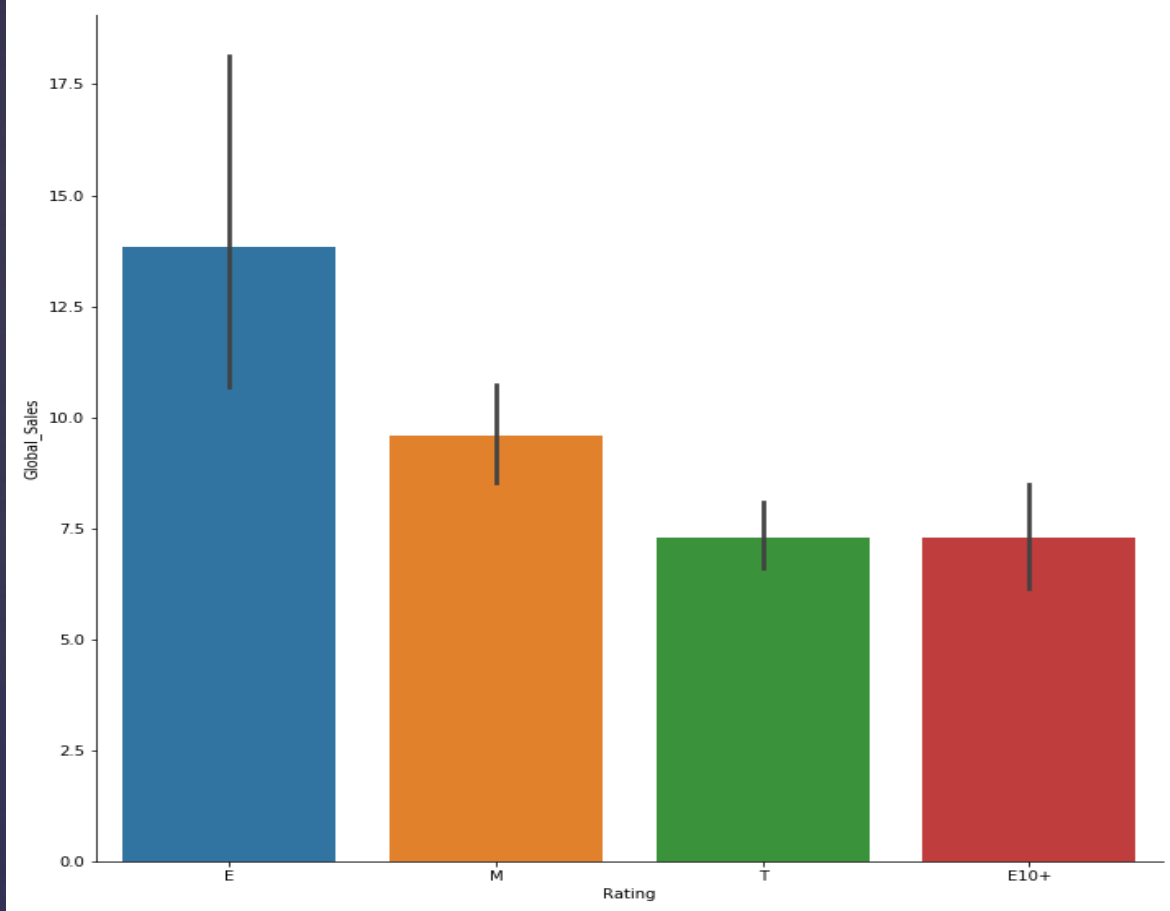
Top 1% des jeux: plus de 5M\$



On constate un changement par rapport à l'ensemble du jeu de donnée:

- ↳ Les jeux de sports dominent le top 1% des jeux vendus à environ 19M en moyenne
- ↳ Les jeux de puzzle sont 2^e à environ 15M en moyenne
- ↳ A l'inverse, les jeux d eplate-formes et les shooters ne ressortent pas particulièrement.

Genre de jeu et top 1%



Les jeux qui fonctionnent le mieux dans le Top 1% sont les genres:

- **E (Everyone)** à 13M en moyenne
- **M (Mature)** à 9M en moyenne

Rating et top 1%

168 jeux représentent le 1% qui s'est le mieux vendu avec un minimum de Global Sales à 5.45M.

Ces jeux se répartissent de façon similaire aux autres jeux sur l'année de sortie (Year_of_Release)

Ces jeux ont une moyenne critique de 86,6 et un user_score de 7,7.

Top 1% des jeux: plus de 5M\$