

Analyse Exploratoire et Visualisation de Données

Introduction

Le jeu de données que nous avons retenu pour ce projet contient les sorties et ventes de jeux vidéo de 1980 à début 2016. Chaque ligne correspond à un jeu vidéo ainsi que le total de ses ventes depuis sa sortie. Nous avons pour informations son nom, son année de sortie, sa plateforme, son genre et son éditeur. Nous avons également le total des ventes en Europe, au Japon et en Amérique du Nord, chiffrés en millions d'exemplaires.

<https://www.kaggle.com/gregorut/videogamesales>

Nous pouvons nous poser plusieurs questions concernant les ventes de jeux vidéo, notamment :

- Quel est le top 10 des jeux vidéo les plus vendus de tous les temps ?
- Quel éditeur est le plus productif / favori des joueurs ? L'est-il encore aujourd'hui ?
- Les joueurs d'un continent achètent-ils plus de jeux venant d'éditeurs de leur propre continent ?
- Sur estimations linéaires, quel sera le genre de jeu le plus vendu en 2016 ?

Nous répondrons à chacune de ces questions en constituant une analyse du jeu de données par le biais de calcul de données en R, avec à l'appui des graphiques générés avec le package ggplot2.

Axe 1 : Top 10 des jeux les plus vendus

Le premier axe sur lequel nous avons travaillé concerne les ventes d'un jeu vidéo par rapport aux autres. En étudiant le nombre de ventes dans le monde pour chaque jeu vidéo, nous avons pu établir un top 10 des jeux les plus vendus de tous les temps. La figure ci-dessous montre un top 10 pour toutes les années confondues.

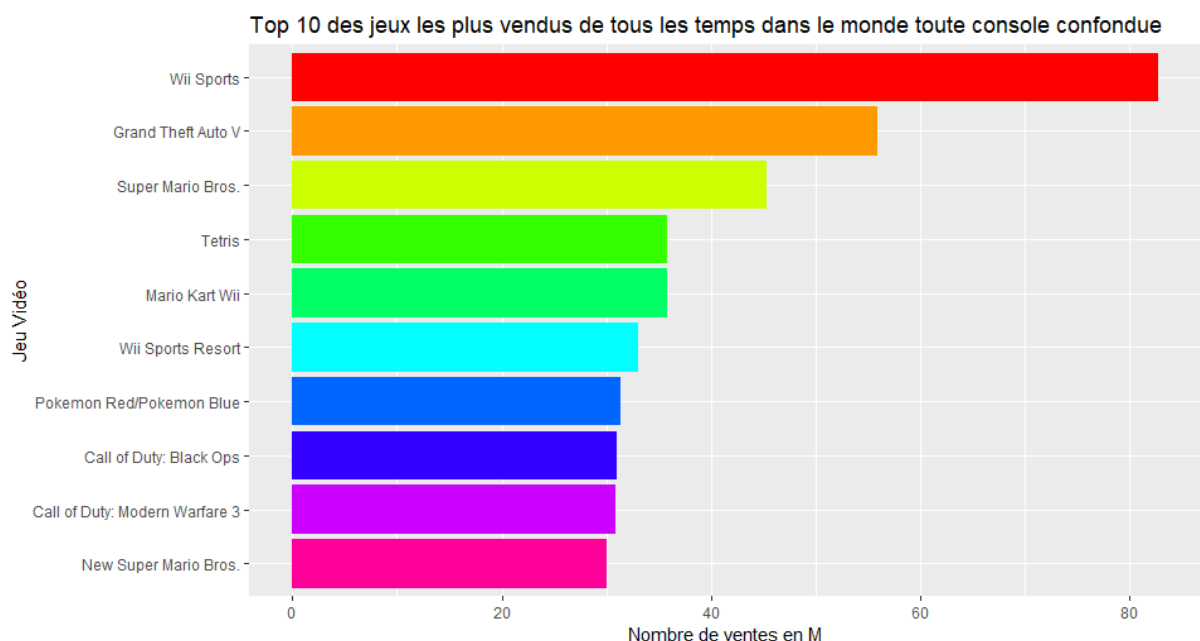


Fig. 1 : Top 10 des jeux vidéo les plus vendus de tous les temps dans le monde toute console confondue

On constate que le jeu le plus vendu est sans conteste Wii Sports. Cela peut s'expliquer par le fait qu'il était inclus pour tout achat de la Wii, la plateforme sur laquelle il fut déployé. On peut aussi constater que ce classement est majoritairement occupé par des jeux destinés aux consoles Nintendo. Ce classement permet également de mettre en évidence qu'il existe des jeux multiplateformes s'étant vendus bien plus que d'autres, comme par exemple GTA V en seconde place du classement ou Tetris qui est sorti sur beaucoup de consoles (également sur certaines consoles non référencées dans notre jeu de données car nous avons ici les consoles « principales » de jeu vidéo).

Axe 2 : Editeurs les plus productifs / éditeurs favoris

Dans cette section, nous nous sommes penchés sur la question des éditeurs : ceux produisant le plus de jeux vidéo, et ceux en vendant le plus.

Tout d'abord, on s'intéresse aux éditeurs de jeux les plus productifs (Fig. 2 ci-dessous). Ceux-ci sont classés en fonction du nombre de jeux publiés, toutes années et plateformes confondues. L'éditeur Namco Bandai Games devance ainsi tous les autres éditeurs avec un nombre de publications s'élevant à près de 800 jeux. Etonnant pour un éditeur qui ne produit pas de consoles lui-même. Nous retrouvons également Konami et EA qui sont tout aussi bien classés. Il n'est cependant pas surprenant de retrouver Nintendo et Sony dans ce top 5.

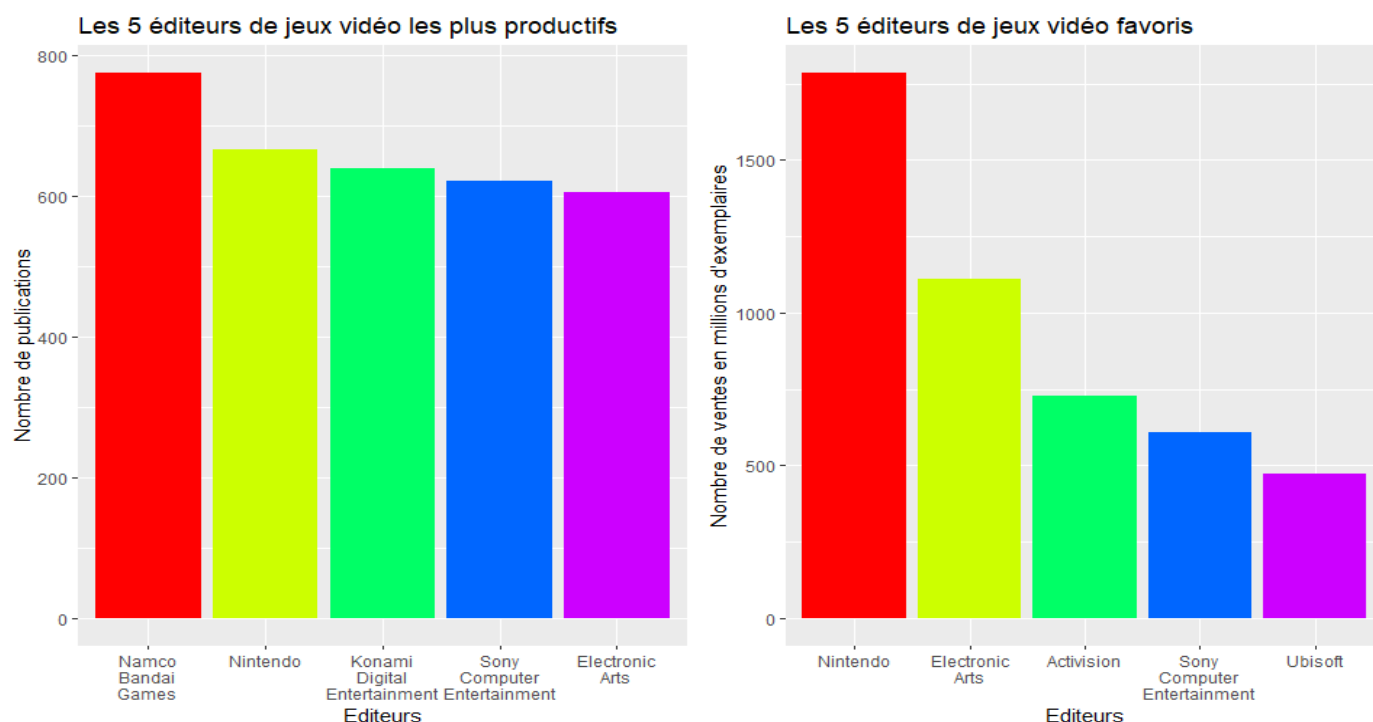


Fig. 2 : Les 5 éditeurs les plus productifs / Les 5 éditeurs favoris

On cherche par la suite à déterminer les éditeurs « favoris » des joueurs. Cette fois ci, les éditeurs sont classés en fonction du total d'unités de jeux ayant été vendues. On estime que le nombre de ventes est lié au fait que tel éditeur connaît plus de succès auprès des joueurs qu'un autre et les attire davantage. On suppose aussi que plus un éditeur est productif, plus il doit faire de ventes.

Ici, c'est le célèbre éditeur Nintendo qui se démarque le plus. Comme nous pouvons le constater sur la figure (Fig. 2), il réalise le nombre de ventes le plus haut et peut donc être considéré en tant qu'éditeur de jeux « favori » de tous les temps. Cela étant, des éditeurs comme Activision et Ubisoft remontent dans ce classement grâce à leurs ventes. Cela met en évidence le fait qu'il ne suffit pas à un éditeur de produire un maximum de jeux pour faire des ventes, la qualité de ces derniers est tout aussi importante. Petite fierté pour Ubisoft qui, étant une société française, arrive à trouver sa place dans ce classement.

Le fait qu'un éditeur soit le plus populaire de tous les temps ne signifie pas forcément qu'il l'est encore aujourd'hui. Ainsi nous allons regarder de plus près l'évolution des ventes des éditeurs « favoris » au fil des années, afin de déterminer s'ils ont su maintenir leur titre. La figure suivante montre l'évolution des ventes pour les 5 éditeurs favoris au cours des 20 dernières années.

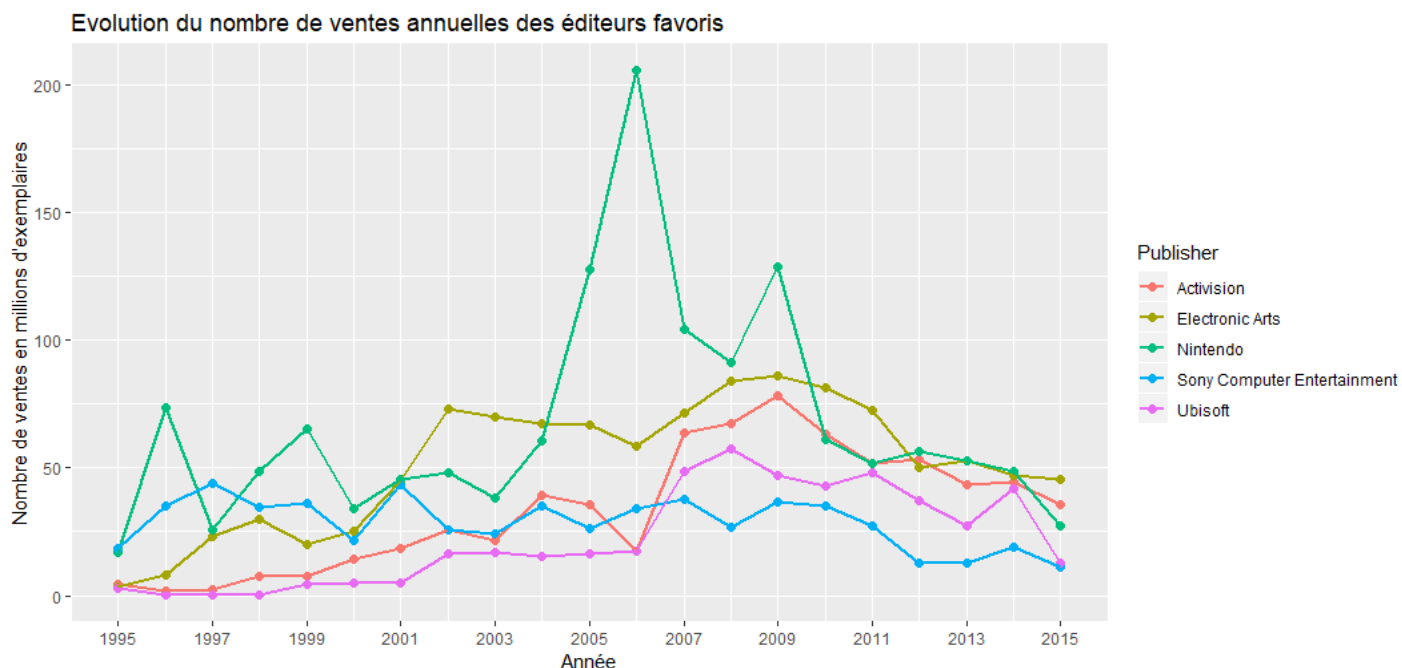


Fig. 3 : L'évolution des ventes des 5 éditeurs favoris au cours des 10 dernières années

Le graphe montre que Nintendo a connu ses plus grands succès auparavant mais qu'il semble petit à petit réduire son total de ventes par années. Peut-être ne mérite-t-il plus sa place de numéro 1 aujourd'hui... Il est à noter qu'il ne fait plus le monopole du jeu vidéo comme avant, ce qui permet à d'autres éditeurs d'être davantage sur le devant de la scène. On constate par exemple que des éditeurs comme EA ont su maintenir un bon rythme de ventes en quasi permanence, se classant toujours dans le top tiers.

On peut finalement remarquer que depuis un certain temps ces éditeurs font de moins en moins de ventes chaque année. Cela peut s'expliquer par plusieurs facteurs, comme par exemple la montée en popularité de petits studios indépendants (le jeu vidéo est plus accessible aujourd'hui qu'avant en termes de développement). On peut aussi penser au piratage de jeux vidéo qui pourrait impacter les ventes de façon générale.

Axe 3 : Origine des jeux vidéo consommés par un continent

Nous allons nous intéresser ici à 4 des éditeurs favoris vus précédemment. Nintendo et Sony sont deux éditeurs japonais, tandis que EA et Activision sont des sociétés américaines. On peut supposer que les japonais consommeraient davantage de jeux japonais et inversement pour les américains. Pour les deux continents, nous comparerons le total des ventes de jeux de provenance japonaise et américaine (en estimant que les 4 éditeurs favoris sont suffisamment représentatifs des ventes générales). Nous analyserons les ventes sur les périodes 1996-2005 et 2006-2015 afin de voir si cela a pu évoluer au cours du temps (les périodes d'avant 1995 n'étant pas représentatives car certains de ces éditeurs n'étaient pas aussi bien implantés dans le milieu du jeu vidéo qu'ils ne le sont aujourd'hui).

Nous précisons bien que dans le cas présent nous n'avons pris en compte ni les jeux en provenance d'autres éditeurs américain et japonais moins populaires, ni les éditeurs d'origine européenne (comme Ubisoft) ou autre. Beaucoup d'éditeurs différents étaient présents dans le jeu de données et leur nationalité n'était pas référencée.

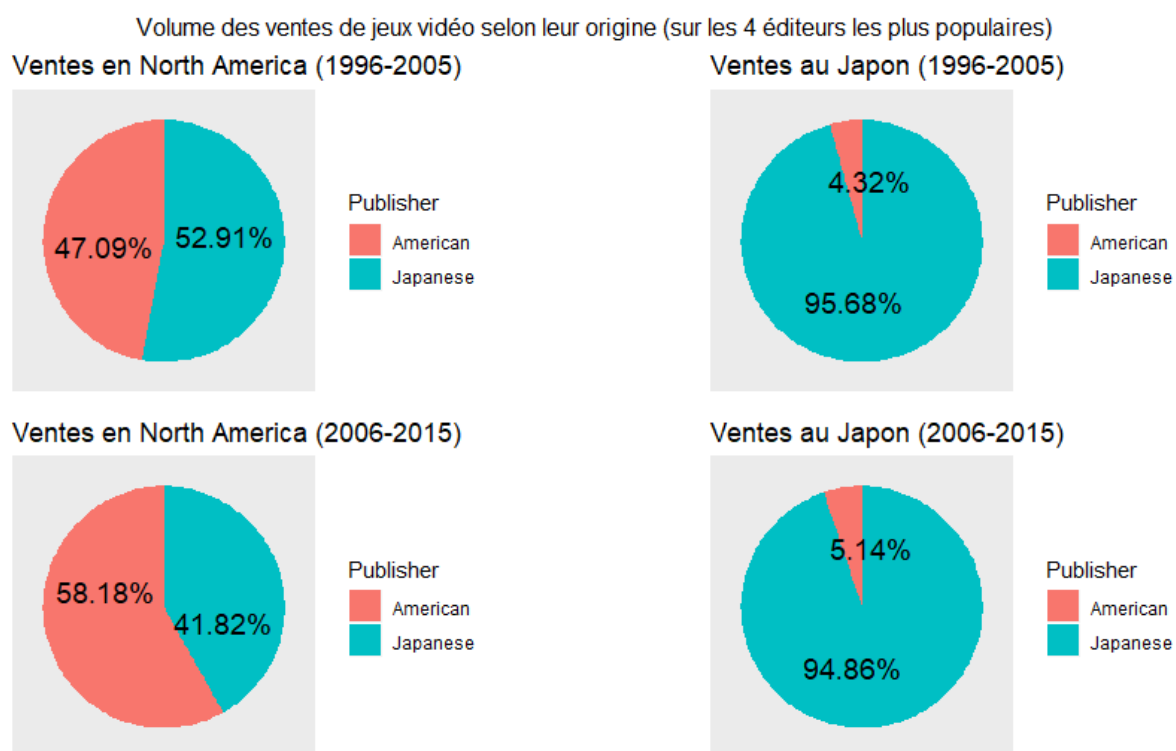


Fig. 4 : Volume des ventes de jeux vidéo selon leur origine (sur les 4 éditeurs les plus populaires)

En additionnant les ventes des 4 éditeurs sur les deux continents et pour les deux périodes choisies, nous pouvons observer et comparer le volume que représentent les jeux américains et japonais. Pour l'Amérique du Nord, bien que la tendance soit en davantage faveur des jeux américains aujourd'hui, les proportions restent tout de même assez équilibrées. Il est possible que cette tendance perdure et que les américains consomment davantage de jeux en provenance de chez eux à l'avenir. En revanche, pour les japonais, le résultat est flagrant : ils consomment majoritairement des jeux d'origine japonaise et ce depuis toujours. L'évolution des parts est à peine visible et on peut imaginer que cela sera encore le cas pour les prochaines années.

Ces résultats montrent que les japonais semblent très attachés à leur propre culture, tandis que les américains semblent plus ouverts aux arrivages extérieurs. Deux continents avec deux mentalités très différentes...

Axe 4 : Genres les plus populaires et prédictions

Le but de cette analyse est d'apprécier l'évolution de la popularité de certains genres de jeux par rapport à d'autres. Par estimation linéaire, nous souhaitons aussi tenter de prévoir la popularité d'un genre pour l'année à venir (**2016** car l'année n'était pas finie au moment de la sortie du jeu de données). Nous avons donc rassemblé le total des ventes globales pour chaque genre et pour chaque année depuis 1996 (afin de constater une évolution sur 20 ans).

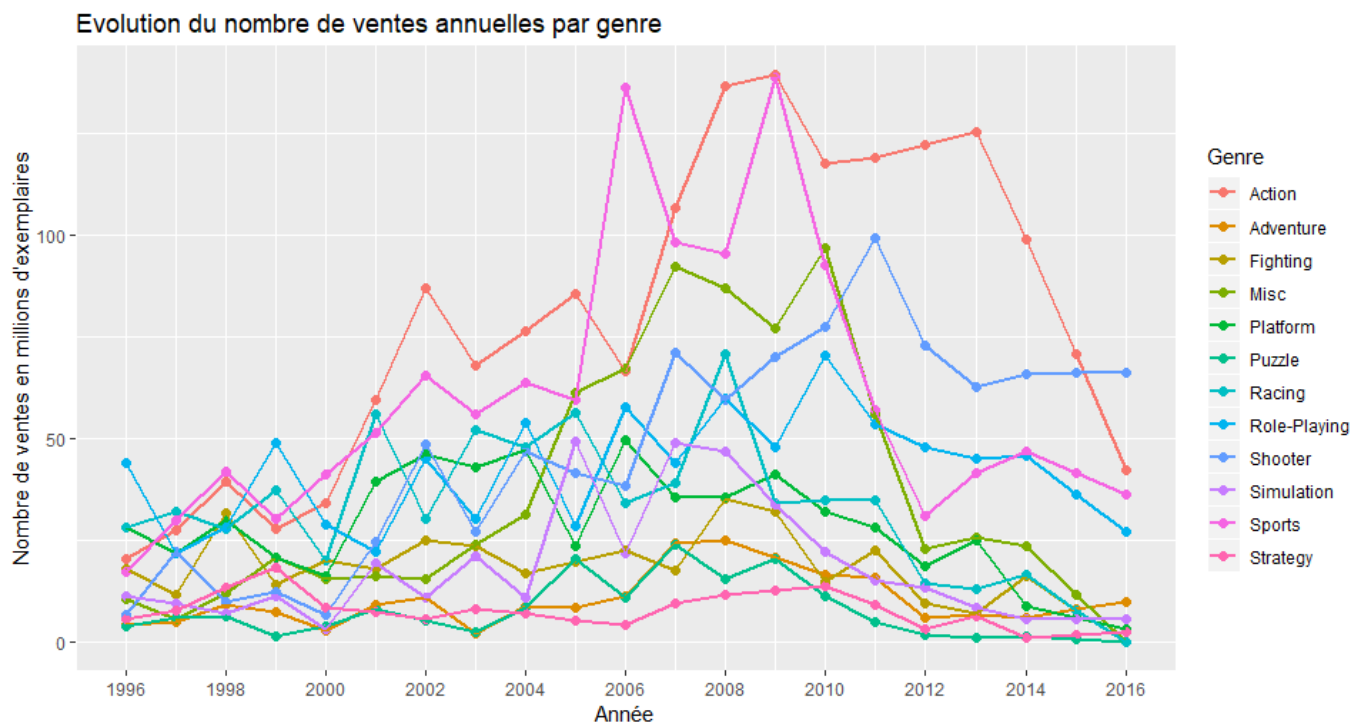


Fig. 5 : Evolution du nombre de ventes annuelles de jeux vidéo par genre

Premier constat : Wii Sport aura influencé tous nos calculs... Pour rappel, la Wii étant sortie en 2006, et le jeu Wii Sport étant inclus avec la console, cela fit grandement remonter les ventes de jeux de sport en 2006 (le pic de 2009 correspondant à la baisse du prix de la console à l'approche de Noël, augmentant ainsi ses ventes cette année-là).

Nous pouvons constater plusieurs choses sur ce graphique. Certains genres comme la stratégie ont toujours eu un public assez restreint, tandis que des genres comme la plateforme se sont essouffés et touchent moins de public qu'avant. La plus grosse chute est celle des jeux d'action qui ont fortement diminué ces dernières années bien que très populaires auparavant. Enfin, les shooters ont continué à gagner en popularité petit à petit au fil des années.

Pour nos estimations de 2016, les jeux d'actions devraient continuer à chuter et passer en dessous de la barre des 50M de ventes. Les shooters resteraient en augmentation lente et deviendraient ainsi le genre le plus populaire. Pour certains genres, la diminution des ventes serait telle qu'ils se retrouveraient à 0 ventes en 2016, ce qui ne serait certainement pas le cas. Cela met en évidence les limites de fiabilité de notre estimation linéaire.

Conclusion

Travailler sur ce jeu de données nous aura permis de nous intéresser à un sujet qui partage notre quotidien tout en apprenant davantage. Nous avons pu observer concrètement des tendances de jeux vidéo selon les générations et comment cela se reflétait en termes de ventes.

Au-delà du jeu de données, nous avons pu explorer les possibilités offertes par l'analyse de données comme par exemple la prévision de données ou le croisement de données pour vérifier des suppositions. Nous avons pris plaisir à nous approprier ce sujet et à tester diverses choses grâce à ce qui avait été appris en cours et en TP. Ce fut pour nous une très bonne expérience d'étude en cette cinquième année SDBD.