

Synthèse projet Hadoop

HALLER François - BROSSARD Florian

1) Introduction

Le monde du jeu vidéo représente aujourd'hui un marché financier important. Des jeux qui étaient autrefois développés au fond d'un garage par un binôme de passionnés sont désormais rentrés dans une phase plus importante, où certains développements de jeu impliquent plus d'un millier de personnes.

Le bouche à oreille a laissé place aux publicités et les avis des joueurs sont aujourd'hui légion. Les critiques spécialisées sont apparues et l'univers du jeu vidéo, tout comme son économie, a évolué vers quelque chose de plus important et de plus complexe.

La sortie d'un jeu vidéo et sa réussite dépendent aujourd'hui plus que jamais de sa popularité. Les journalistes et la presse spécialisée sont donc présents pour faciliter la prise de décision pour les futurs consommateurs. Il est donc important de se demander pourquoi certaines affaires d'articles et d'éloges sponsorisés ont eu lieu et si la presse spécialisée vise à protéger et aider le consommateur ou si cela relève plus de la publicité cachée.

2) Problématique

Dans ce contexte, notre curiosité nous a poussée à nous poser la question suivante:

La note des critiques spécialisées impactent-elles la vente des jeux vidéos ?

3) Méthodologie

Pour répondre à cette question, nous avons décidé d'étudier la vente de jeux vidéos. Pour ce faire, nous avons récupéré le jeu de données suivant:

https://www.kaggle.com/rush4ratio/video-game-sales-with-ratings/#Video_Games_Sales_as_at_22_Dec_2016.csv

Ce jeu de données, fourni par Kaggle, contient le chiffre de vente, la note critique, la note publique, la plateforme et le type d'un jeu.

Elle répertorie plus de 16000 entrées (16719) de rapports de vente des jeux entre 1980 et 2016.

4)Algorithme

Pour répondre à notre question, nous avons réalisé 5 MapReduce.

Dans un premier temps, nous parcourons les fichiers et récupérons les informations qui nous intéressent (il s'agit ici du nom du jeu, de la plateforme du jeu, de son chiffre de vente, de sa note critique).

Une fois nos clés créées, un second MapReduce filtre les résultats pour retirer les jeux ne possédant pas de note de la critique ou ayant des notes de la critique indéterminées (TBD : to be determined).

Un troisième MapReduce est alors utilisé pour faire un ratio, ce ratio permet de voir le nombre vente par point obtenu par la critique. Il est obtenu en faisant la formule suivante:

$$\text{Ratio} = 100 * (\text{nombre de vente} / \text{note critique})$$

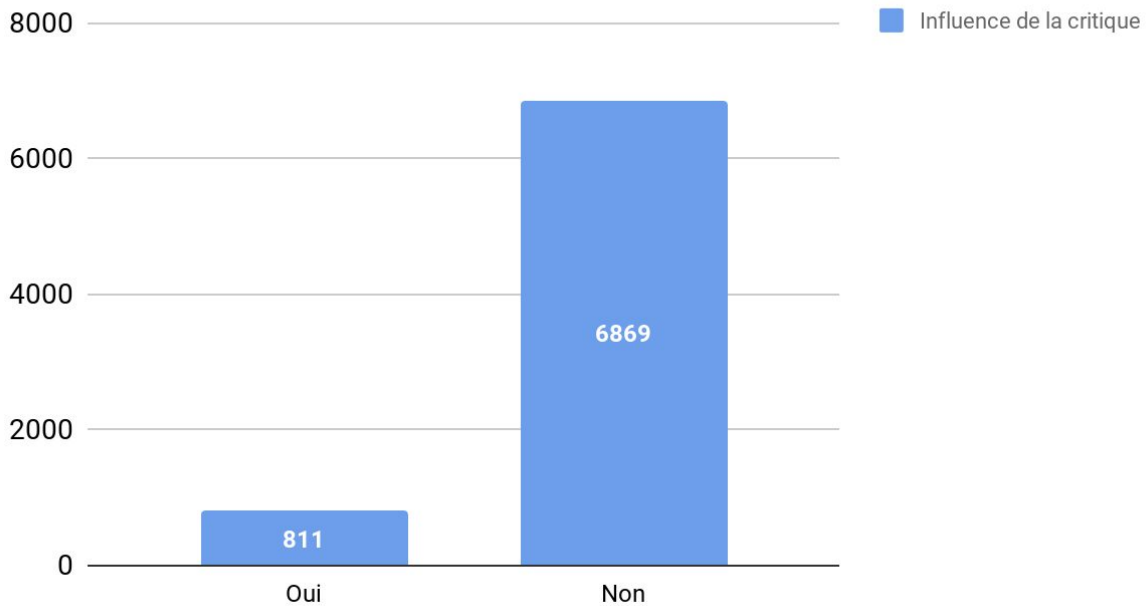
Le ratio du premier jeu pour chaque console référencé suite au second MapReduce sert alors de témoins pour ladite console.

Le quatrième MapReduce parcourt le jeu de données obtenu et vérifie si la vente de chaque jeu vidéo a été influencé ou non par la critique. Pour cela, il vérifie si l'écart entre le ratio témoin de la même plateforme et le ratio du jeu est inférieur à 20%, sinon, on considère que la note critique n'a eu aucune ou très peu d'influence sur le succès de la vente du jeu.

Le 5ème et dernier MapReduce sur le jeu de données permet de faire le pourcentage de vente de jeux ayant été influencés par la note de la critique. Le résultat affichée montre le pourcentage de ventes influencés (OUI) ou non (NON) par la note de la presse spécialisée.

5) Résultats et interprétation

Influence de la critique sur la vente ?



Au vue des résultats et de la méthodologie que nous avons mis en place, nous pouvons en déduire que les critiques spécialisées n'influencent généralement pas la vente des jeux de manière linéaire.