



HMIN122M Entrepôts de données et big-data

Mini Projet - Winamax

Auteur:

Gracia-Moulis Kévin (21604392) Canta Thomas (21607288)

Master 1 - AIGLE/DECOL Faculté des sciences de Montpellier Année universitaire 2020/2021

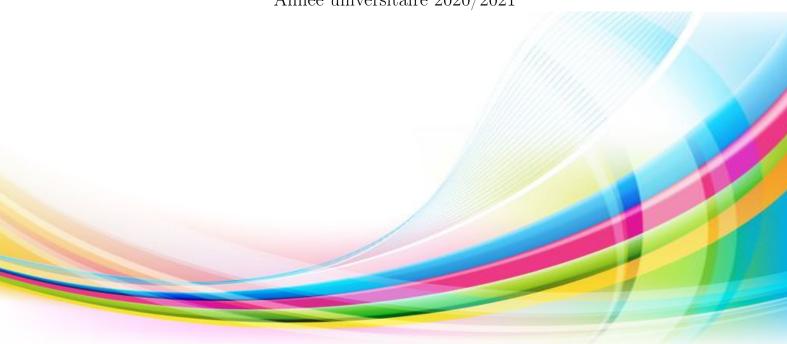


Table des matières

Analyse des traitements	2
Conception	3

Analyse des traitements

Les jeux d'argents et les paris font partis intégrante de notre société depuis des milliers d'années. De notre ère ou le monde est en ligne tout as donc été dématérialisé sur de multiples plate-formes web. Nous allons nous intéresser au cas de Winamax qui est un site de poker, de pari sportif et de Ligue Fantasy. Il permet au bookmaker du monde entier de pouvoir parier de l'argent, réel, sur les différents services vu précédemment. L'objectif de cette entreprise est donc d'amasser un maximum d'utilisateur, notamment par l'organisation d'évènement et/ou compétition ou même encore en sponsorisant des joueurs professionnel.

Winamax est donc un site d'affluence quotidienne, chaque utilisateur ayant son propre compte et son propre soldes. Chaque jour des milliers d'utilisateurs affiche des pertes ou des gains durant leurs divers partis. On constate donc qu'il devient très intéressant d'analyser la base de données de ce site, en se focalisant notamment sur la gestion de l'afflux d'argent. L'erreur n'est pas permise ici, la perte ou gain d'argent venant d'une erreur dans la base de donnée est problématique, elle se doit donc d'être bien construite et réfléchis.

Les services du site sont accessible à toutes personnes majeurs, ceci offre une très forte audience potentielle. Malgré une forte concurrence comme la Française des Jeux, le PMU en France ou encore Pokerstars, Unibet dans le monde, Winamax à su captiver son public en étant l'un des premiers site à allier un maximum de services possibles sur une même plate-forme, notamment le poker et les paris sportifs. Cette multitude de service complexifie donc la base de données afin de gérer des milliers de mises ou paris différent au même moment, mais permet aussi sur le long terme d'observer et analyser les statistiques de ces données.

Winamax a quant à lui pour plus gros paris de conserver en toute sécurité chaque solde de ses utilisateurs, ceci implique que la base de donnée du site se doit, malgré sa complexité, d'être parfaitement organisé afin d'offrir une sécurité maximal et rapidité. Les principales sources de revenus du site proviennent de "taxes" en ponctionnant une partie des gains des joueurs, lors de dépôts ou retraits fait par l'utilisateur, lors de partenariat, comme en France avec TF1 en mars 2000 ou même lors d'inscriptions de bookmakers pour un tournois et bien d'autres.

Nous avons listés les actions/opérations qui nous semblent judicieux de tracer afin de récolter les informations énoncés durant l'analyse précédente. Pour chaque action, il lui succède trois traitements possibles.

- 1. Inscription des joueurs lors d'un tournoi, table et pari
 - Sélection du nombre de joueur total inscrit sur un événement(tournoi, table, pari) précis.
 - Moyenne du nombre de joueur inscrit à un des événements possibles.
 - Nombre total des inscriptions aux événements du mois.
- 2. Gain des joueurs
 - Somme d'argent total reversé d'un événement précis.
 - Nombre de joueur touchant un cashprize sur un événement.
 - → Joueur ayant touché le plus cashprize sur le mois.
- 3. Dépôt d'argent d'un joueur sur son compte Winamax
 - Somme des dépôts ayant été effectué par un joueur précis sur un mois précis.
 - ➤ Moyenne des dépôts annuels effectué par les joueurs de 20ans.
 - Fréquence des dépôts lors d'une promotion.
- 4. Retrait d'argent d'un joueur sur son compte personnel (banque, paypal...)
 - Quantité d'argent retiré en moyenne par retrait.
 - ➤ Quantité d'argent retiré sur le mois.
 - Somme d'argent total retiré par un joueur.

Par ordre de rentabilité le plus important, analytiquement parlant, à traiter est le dépôt d'argent (3) et les inscriptions des joueurs (1). Étant les principales sources de revenus de l'entreprise, elles possèdent énormément de données exploitables. Ensuite vient le gain des joueurs (2) et le retrait d'argent (4) qui possèdent elles aussi un potentielle d'analyse mais moins importante que les précédentes.

Conception

Les deux actions / opérations qui nous semblent le plus pertinents à analyser sont les dépôt d'argent sur le compte utilisateur (joueur) et les différents inscriptions (tournoi, table ou pari). Nous avons représenté ci-dessous nos deux actions principales sous la forme d'un data-mart (ou modèle étoile) :

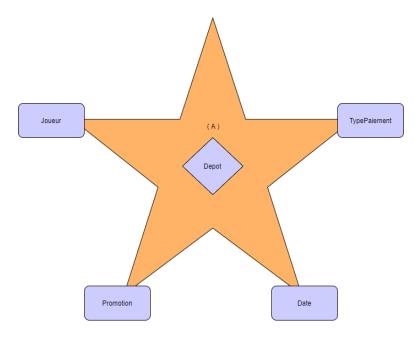


FIGURE 1 – Data-mart de nos Dépôt

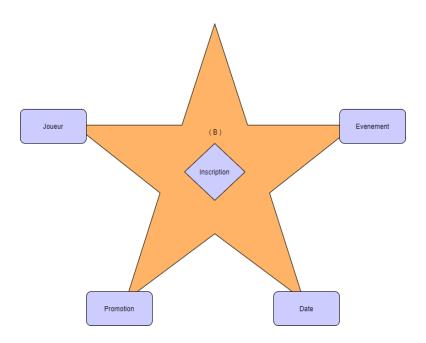


Figure 2 – Data-mart de nos inscriptions

Pour chaque actions/opérations la liste des faits à ajouter dans l'entrepôt sont les suivants :

1. Un fait (j, t, p, d, x) existe lorsqu'un joueur j effectue un paiement le jour d avec un certains type de paiement et une promotion.

La mesure x correspond au montant du paiement.

Ce fait est transactionnel et est additive.

1. Un fait (j, e, d, x) existe pour lorsque qu'un joueur j est inscrit le jour d à un évènement e.

La mesure x correspond au nombre total de joueur inscrits le jour d. Ce fait est snapshot et est non-additive.

Voici une représentation des dimensions nécessaires pour chacun de nos modèles :

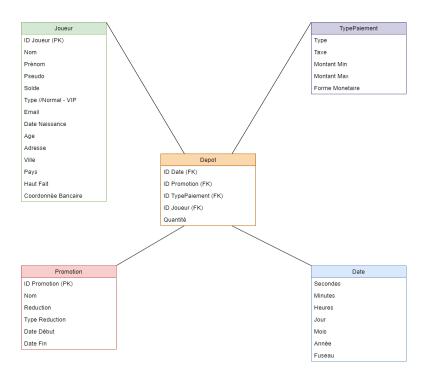


FIGURE 3 – Représentation de notre table depot

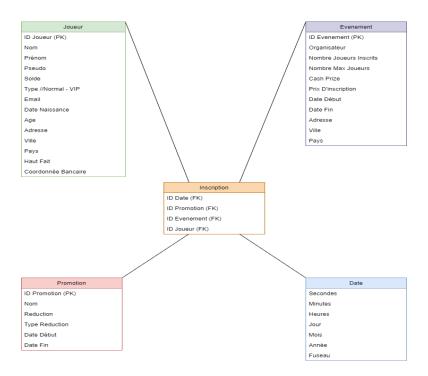


FIGURE 4 – Représentation de notre table inscription

Grâce au modèle que nous avons mis en place, il est possible de répondre aux différents traitements car la représentation de notre data-mart à été faite de sorte qu'elle puisse y répondre. Nous avons mis pour rappel le traitement que nous avions énoncés plus haut, la réponse possible la succède.

♠ Depot :

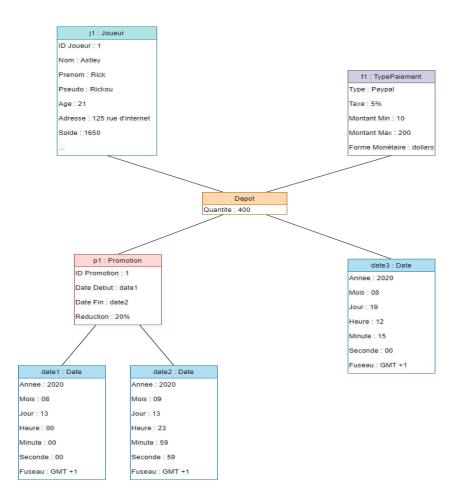
- Somme des dépôts ayant été effectué par un joueur précis sur un mois précis : On additionne la quantité de chaque dépôt d'un joueur (ID Joueur) sur un mois précis (Date Annee, Date Mois).
- Moyenne des dépôts annuels effectué par les joueurs de 20ans :
 On récupère la moyenne des quantité déposé par an (Date Annee) par les joueurs (ID Joueur) avec un age (Joueur age) égal à 20.
- Fréquence des dépôts lors d'une promotion :

 On regarde le nombre de dépôt ayant eu lieu entre la date de début (Promotion Date Début) et de fin (Promotion Date Fin) de la promotion (ID Promotion).

Solution Solution :

- Sélection du nombre de joueur total inscrit sur un événement (tournoi, table, pari) précis : On récupère le nombre d'inscription à un événement (ID Evenement).
- Moyenne du nombre de joueur inscrit à un des événement :
 On fait la moyenne du nombre de joueurs inscrits (Evenement Nombre Joueur Inscrit)
 d'une évenement (ID Evenement) donnée.
- Nombre total des inscriptions aux événements du mois :
 On additionne le nombre de joueurs inscrits (Evenement Nombre Joueur Inscrit) sur un mois précis (Date Annee, Date Mois).

Voici un exemple d'instance pour chaque modèle afin de démontrer leur pertinence :



 $FIGURE\ 5-Instance\ d'un\ depot$

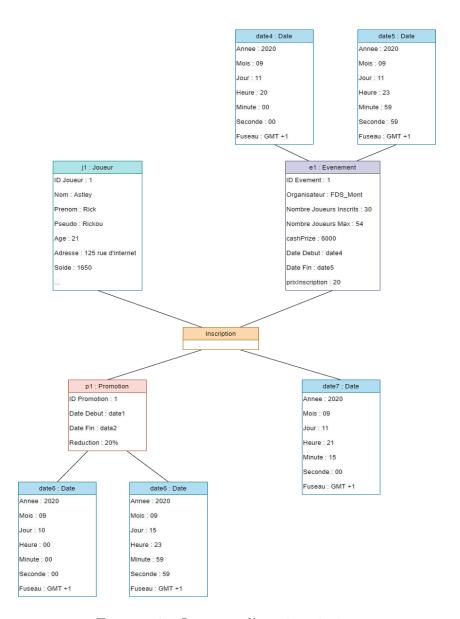


FIGURE 6 – Instance d'une inscription