Mardi 26 Mai 2020

Daniel Kaddous

Joshua Harris

Deveulopper.se Simplon.co

**Rapport de projet**

**Projet de graduation PokeStat:**

**Présentation de dashboards et d’analyse graphique**

**eSports.**



**Table des matières**

1. INTRODUCTION
2. OBJECTIF DU PROJET
   1. Définition
   2. Motivation
   3. Note de cadrage
3. TECHNOLOGIES
   1. SQL
   2. WorkBench
   3. Python
   4. Panda
   5. Plotly
   6. Dash
   7. Flask
   8. BeautifulSoup
   9.  Grafana





1. **INTRODUCTION**

Dans le cadre de nos 7 mois de formation à Simplon.co en la qualité de développer data, il nous a été demander de travailler sur un projet qui ferai office de projet de graduation.   
  
  
 Ce projet a pour objectif de regrouper et représenter toutes ou majoritairement les qualifications que nous avons développé tout au long de cette formation. Au cours de son développement le projet pris plusieurs mises à jour majeures. Dans un premier temps l’application devait mettre en scène deux entité dans un combat à l’instar d’un jeux vidéo Pokémon. Daniel KADDOUS et Joshua HARRIS étant passionné de jeux e-sportif (jeux compétitifs à niveau professionnel) nous avons voulu recentrer le projet pour qu’il englobe plus de technologie et nous nous somme projeter sur des graphiques Analytics qui avait pour thème de desceller des tendances au sein de jeux comme Fortnite, Rocket League et finalement Pokémon.

1. **OBJECTIF & NOTE DE CADRAGE**
   1. **Définition**

Avant toute chose ils nous à sembler pertinent de destiner une partie à la définition du support de nos analyses l’eSport.

L’eSport se définit par des évènements mettant en scène des compétiteurs sportif au sein d’un jeu vidéo. Ces rencontres en ligne ou en présentiel (LAN-party) prennent de plus en plus de places dans nos sociétés jusqu’à en acquérir la légitimité de sport dit « classique ». Le milieu devient de plus en plus lucratif en adéquation d’un nombre accrue de supporter, de sponsors ainsi que de représentations.   
 Ces mêmes représentations sous format de championnat partagent les mêmes codes auxquelles on a l’habitude et sont disponibles sur des flux-vidéo en direct sur des plates-forme de diffusion spécialisé dans le jeux-vidéo ont a pour exemple les deux mastodontes Twitch.tv et Mixer.com.



* 1. **Motivation**

Vis-à-vis du projet amener l’utilisateur à une meilleure compréhension de l’eSport mais dans une plus grande mesure une certaine sensibilisation niveau données font partie de nos principales motivations.

* 1. **Note de cadrage**

Les objectifs du projet peuvent être ainsi résumés par priorités décroissantes :

* Référencements de jeux de données pertinent et intéressant.
* Nettoyage de la donnée si nécessaire et modélisation de la base de données.
* Création des premiers plots graphiques.
* Hébergement en local de nos analyses avec Flask ainsi qu’une conception légère d’une interface utilisateur.
* Création d’une documentation ReadTheDocs pour l’application



Objectifs de l’équipe de projet

L’équipe doit travailler suivant des principes généraux visant à valider nos compétences acquises au cours de la formation Deveulopper.se Data à Simplon dont les principales caractéristiques sont :

* Concevoir / maintenir une base de données avec un langage de requête SQL
* Analysé / nettoyer la donnée avec un langage de programmation Python

Début du projet le 21 avril 2020 en parallèle du cursus de formation qui suit son cours en distanciel. La fin du projet doit prendre fin et remplir ses objectifs avant le 24 mai pour pouvoir basculer sur la rédaction d’un rapport de projet. Préparation de slides ou d’un support équivalant afin de pouvoir présenter le projet.



1. **TECHNOLOGIES**
   1. **SQL**

Pour notre base de données l’outil principal pour sa maintenance, son organisation et sa structure à été le langage de requête SQL. Dans notre cas nos données ne sont pas liées entre les tables et n’a pas eu besoin d’une modélisation comme la méthode MERISE que nous avions apprise pendant la formation.  
   
 Pour notre base de données on se retrouve avec 4 tables indépendantes où elles sont tout simplement stocker. Nous avons du créé les tables avec les colonnes avec un a data-type adéquat de sorte qu’il match avec nos métadonnées.   
  
Notre base de données est donc composée de 4 tables, qui référencie nos 3 jeux de données : un data-set pour Fortnite, pour Pokémon, Rocket League et une dernière table qui stocke des données calculées depuis la base de données Rocket League.

* 1. **Workbench**

Quand on parle de langage de requête et base de données on est dans la majorité des cas amener à utiliser un Système de gestion de base de données (SGBD). Dans notre cas nous avons utilisé WorkBench qui avait été présenté au début de notre formation.   
  
 Ce qui va varier d’un SGBD à l’autre sont les features qui lui sont intégré et dans notre cas WorkBench était amplement suffisant, gratuit et partiellement open-source sont ses principaux atouts.

* 1. **Python**

Quand on se réfère au top trending Python est LE langage de programmation du moment, la plupart du temps au coude à coude avec Java C/C++. Facile à apprendre et à utilisé si on se fie au avis des utilisateurs il est donc le langage de choix pour les développer.Junior que nous sommes.



  
  
 Nous avons pu développer diverse compétence au cours de notre formation et notamment l’ors de notre stage en AI à Cerema où nous devions utiliser des réseaux neuronaux pré-entrainer comme NEAT-EO afin de lui faire identifier des objets sur des photo panoramique.

Ce sont les librairies utilisées sous Python qui sont majoritairement les technologies les plus présente dans notre projet.

* 1. **Pandas**

Nous avons utilisé Pandas dans nos script python pour concevoir avec nos data-set des tableaux de données auquel on lui a imposé des calculs, à le nettoyer pour exemple on a ajouté des colonnes qui référencer des pourcentages en fonction du nombre d’entrée et d’un motif d’une colonne. On a utilisé Panda également pour importé nos data-set directement dans nos base de données SQL et finalement pour nettoyer les data-set c’est-à-dire supprimer ou ajouter un élément ou alors changer le format d’une métadonnée ect..

Dans le panel de nos compétence acquise ou en cours d’acquisition nous estimons que Panda est le meilleur outil pour manipuler la donnée avant même de l’importer.

* 1. **Plotly**

Plotly est une librairie Python développer par Plotly , c’est un outil de visualisation de données j’ai utilisé Plotly car c’est très complet on peut annoter très facilement ses graphiques pour les faire comprendre.

Dans le projet j’ai utilisé majoritairement des graphiques en Histrogramme , j’ai trouvé ça plus adapté et plus compréhensible à la comparaisons de Statistique entre plusieurs Pokemon, et un Scatter étais très inadapté parce qu’on comparé plusieurs statistiques du Pokemon en même temps donc cela devenais très vite illisible

Je suis satisfait d’avoir réussis mais je ne suis pas très satisfait de mon analyse car elle ne révèle rien de nouveau c’est plus un outil de comparaison que de l’analyse qui donne une vraie conclusion, on peux certes y observer des choses intéressante que ne voit pas forcément sans s’y pencher et composer plus vite une équipe Pokemon.

Je n’ai pas rencontré de grosses difficultés avec Plotly hormis sur les layout des graphiques et la

* 1. **Dash**

Dash est une librairie python qui permet de créer une UI de visualisations de données, d’analyse, et d’exploration de données dans une App Web sans utiliser de langages Web Comme le Javascript , elle est aussi développer par Plotly.

Ma première difficulté à été d’insérer différents tableau Dash sur la même app Web. J’ai suivi un tutoriel trouver sur internet qui m’as mener pas à pas vers ce que je voulais réaliser.

Une autre difficulté et que j’ai utiliser un composant web de Dash

Appeler le Dropdown (dcc.dropdown). Le problème avec se Dropdown c’est que suivant le navigateur il n’agissait pas pareil , j’ai du donc travailler sur le Responsive de ce dernier, j’ai fais en effet un post sur le forum Dash et il fonctionner très bien chez les autres forumeur mais pas chez moi , j’ai dû donc creuser pour trouver d’où venais le problème.

Je suis quand même satisfait de ce que j’ai réussis à en faire car ce n’étais pas une mince affaire d’intégrer Dash dans Flask pour me laisser une liberté dans Front de l’application. Mais sur internet on trouve tout pour nous aider !

* + 1. **Flask**

Flask est un framework open-source de développement web en Python. Il reste léger afin de garder le souplesse du Python. Il utilise un système dis de templates, c’est-à-dire que l’on peut rappeler plusieurs fois la View d’une page pour la mettre sur d’autre pages

Je n’ai pas rencontré de grosses difficulté dû fait de mon background de Développeur Full-Stack je comprends déjà le fonctionnement d’un Framework Web , mais ça quand même était un défi , il fonctionne

Pas comme Rails , mais j’ai tout de même pas eu de grande difficulté sur le Back-end définir les Routes etc…

La features nouvelle que j’ai pu rencontrer est Jinja2 le moteur de templates Python. J’ai réussi à le prendre en main mais j’ai encore beaucoup de choses à explorer de ce côté-là mais c’est quand même très similaire au ERB de Ruby.

Je suis assez satisfait de mon travail sur Flask de plus ça m’as permis de beaucoup m’entrainer en Python et de pratiquer autres choses que la manipulation des données.

 **Conclusion**

Vis-à-vis du projet j’aurais aimé trouver un moyen de récolté des données en direct sur mon Ordinateur sur des jeux au quel je joue ou sur mes composants.

Je creuse encore ce côté-là mais c’est quand même compliqué et quelques choses d’intéressant à la fois.

Dans l’ensemble je suis assez satisfait surtout de la partie Flask/Dash qui m’as permis de beaucoup m’entrainer sur le volet web du Python et sur le langages dans sa globalité .

J’aurais aimé aussi avoir le temps de tout Conteneuriser avec Docker pour pouvoir l’héberger sur Heroku mais je n’ai pas eu le temps de m’y pencher pour faire un déploiement stable.

Avant de commencer le projet je voulais en faire un jeu ou l’ont récoltes les données des parties mais vite confronté a la réalité du temps , le jeu aurait été très limité et donc aurait eu un nombres de données très faibles et n’aurais pas été intéressant à partager.































