
2IA-TC-8.1 PROJET COLLECTE ET STOCKAGE DES DONNÉES

**Collecte et stockage de données issues de League of
Legends pour l'analyse et la détermination des
facteurs décisifs menant à la victoire**

Étudiants :

Kawtar MOULAHID
Johanna SAOUD
Hicham TALAOUBRID
Amine TORKI

Enseignant :

L. BAYARD
S. HARISPE
N. SUTTON-CHARANI
A. TCHEMEDJIEV

Table des matières

1 Présentation de League of Legends	2
1.1 Introduction	2
1.2 Éléments du jeu	2
2 Design of Experiment	7
2.1 Définition de la question et de la population	7
2.2 Taille d'échantillon et plan d'expérience	7
2.3 Organisation et déroulement de la collecte	8
2.4 Stockage de données sous une base de donnée MySQL	10
3 Conclusion	10

1 Présentation de League of Legends

1.1 Introduction

League of Legends est un jeu vidéo sorti en 2009 de type MOBA (*multiplayer online battle arena*), développé et édité par *Riot Games*. Il est très rapidement apprécié de la communauté et reçoit des critiques majoritairement positives. Il gagne des prix pour son accessibilité, le design de ses personnages et sa compétitivité. En juillet 2012, il est le jeu le plus joué sur ordinateur en nombre d'heures en Europe et aux États-Unis, et maintient cette position très longtemps. Encore aujourd'hui, il reste l'un des jeux les plus joués et récupère cette première place de temps à autres. Il est souvent considéré comme le jeu avec la scène compétitive la plus développée, notamment par son championnat du monde ayant lieu chaque année qui touche un large audimat.

En 2019, le championnat du monde bat tous ses records d'audience : 44 millions de spectateurs simultanés aux pics d'audience et 99,6 millions de spectateurs uniques.

Pour le championnat du monde 2020, *Riot Games* parvient malgré la pandémie de COVID-19 à organiser une compétition en physique à Shanghai, toute la compétition est jouée dans un studio fermé au public. Seule la finale est ouverte et se déroule au stade de football de Pudong où les sud-coréen de l'équipe *DAMWON Gaming* s'impose face à l'équipe chinoise *Suning*.

1.2 Éléments du jeu



FIGURE 1 – Vue aérienne de la Faille de l'Invokeur

League of Legends est un jeu de stratégie en temps réel où deux équipes de 5 joueurs s'affrontent dans une arène virtuelle nommée *Faille de l'Invokeur*. Le but du jeu est de détruire la base adverse, appelée *Nexus*.

Selection des rôles Chaque joueur choisit un parmis les 5 rôles existant sur le jeu. Une équipe est nécessairement constituée de 5 joueurs, ayant chacun des rôles différents. Ces rôles sont : *Top laner*, *Jungler*, *Mid laner*, *Bot laner* et *Support*.

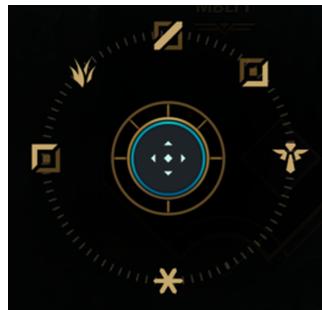


FIGURE 2 – Les 5 rôles sur League of Legends

Répartition dans l’arène La Faille de l’Invokeur est divisée en deux camps symétriques, séparés par une rivière : celui de l’équipe Bleue, et celui de l’équipe Rouge (cf fig 1). Elle est constituée de 3 *voies*, ou *lanes*, qui relient les bases de chaque équipe. Le reste de l’arène constitue la *jungle*. Les joueurs de chaque équipe se positionnent dans l’arène selon le rôle qu’ils ont sélectionné.

- Le *Bot laner* et le *Support* combattent sur la voie du bas
- Le *Mid laner* combat sur la voie du milieu
- Le *Top laner* combat sur la voie du haut
- Le *Jungler* combat les monstres de la jungle et aide ses coéquipiers sur toutes les autres voies

L’objectif pour chaque laner est de prendre du terrain, en détruisant les *tourelles* de défense ennemis. Pour cela, ils doivent d’abord éliminer leurs vis-à-vis adverse qui combat sur leur lane. L’arène est affectée par un *brouillard de guerre*, qui empêche les joueur de voir l’entièreté de la carte. Chaque joueur a un champ de vision restreint autour de lui et peut poser des balises (*ward*) qui donnent à l’équipe une vision tout autour. (cf. fig 5)

Selection des champions Avant le commencement de la partie, chaque joueur doit sélectionner un personnage (*champion*), qu’il incarnera pendant toute la durée de la partie. Les champions ont chacun des capacités de combat (*spells*) et des caractéristiques différentes. Dans l’idéal, il faut communiquer avec son équipe pour choisir des champions ayant de bonnes synergies entre-eux, ce qui facilitera les combats en équipe.

Bien qu’ils soient tous très différents, il existe plusieurs catégories de champions :

1. *Assassins* : Champions très mobiles et qui infligent de lourds dégâts dans une courte période.
2. *Combattants* : Champions à la fois offensifs et défensifs, ils sont très polyvalent, on les qualifie de champions de mêlée.
3. *Mages* : Champions qui lancent des sorts et causes des gros dégâts magiques à distance, il sont cependant très fragiles, ils ne peuvent pas se battre au corps à corps.

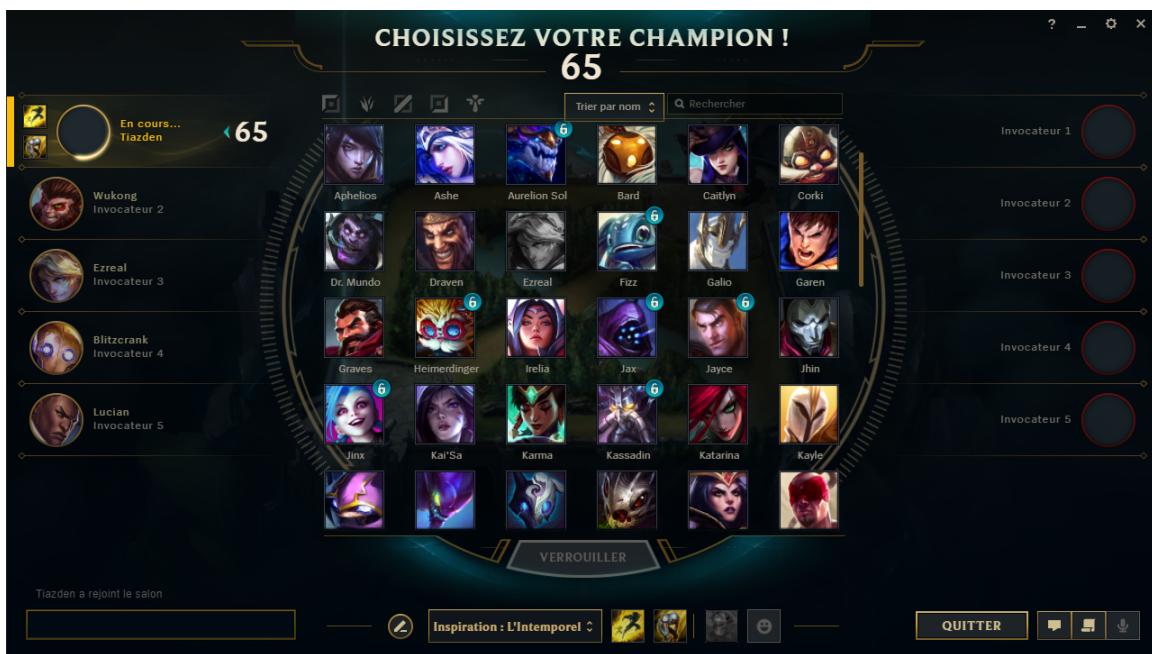


FIGURE 3 – Selection des champions

4. *Tirreurs* : Champions qui combattent à distance et causent de lourds dégâts. Ils sont aussi très fragiles et donc inefficace en combat corps à corps.
5. *Supports* : Champions utilitaire, ayant des compétences permettant de soutenir leurs alliés (ex : *soins*) ou d'affaiblir les ennemis (ex : *immobilisation*).
6. *Tanks* : Champions qui se battent au corps à corps, ils possèdent des capacités défensives et sont très résistants.

Combat Lorsqu'un joueur J1 en élimine un autre J2 au combat, J1 gagne de l'expérience (*XP*) et des pièces d'or (*PO*). Tandis que le joueur J2 est éliminé pendant 6 à 52 secondes suivant l'avancement de la partie. Il est ensuite renvoyé au point de réapparition proche du Nexus de sa base.

Pendant toute la période d'élimination, il est seulement spectateur de la partie et ne peut pas interagir avec ses coéquipiers. L'équipe perd donc momentanément l'un de ses membres, ce qui rend plus difficile la défense ou l'attaque. De plus, le joueur éliminé est retardé dans sa progression car il ne gagnera ni XP, ni PO pendant cette période.

Les joueurs combattent pour : - accumuler des ressources - empêcher l'adversaire d'en récupérer - sécuriser une prise d'objectif (dragon/baron) -

Entités Il y a différentes entités dans le jeu :

Les champions, qui sont des joueurs, ils possèdent des caractéristiques (vie, points d'attaque, points de défense, vitesse de déplacement etc.) qui leur sont propres et sont augmentées lorsque le joueur gagne un niveau d'expérience. Il est aussi possible de modifier ces caractéristiques au point de réapparition en achetant des objets (*items*) contre des PO. Lorsque le joueur passe un niveau d'expérience (*level up*) il peut améliorer une compétence de combat de son champion.

Les tourelles sont situées à intervalles réguliers sur les lanes (3 par lane + 2 devant le Nexus). Elles octroient des PO lorsqu'elles sont détruites. Avant 14 minutes de jeu, les

tourelles sont protégées par des boucliers de pierre, appelés *plates*. Les plates sont des segments de la barre de vie totale de la tourelle. La destruction d'une plate donne 160 PO répartis entre les joueurs ayant participé à sa destruction. La destruction de toutes les plates avant 14minutes de jeu amène à la destruction complète de la tourelle.



FIGURE 4 – Tourelle de l'équipe rouge ayant encore ses 5 plates

Les sbires (*minions, en anglais*) et les monstres de la jungle donnent des points d'XP et des PO au joueur qui les a tués. Les monstres de la jungle sont neutres, tandis que les sbires appartiennent soit à l'équipe bleue, soit à l'équipe rouge. Les sbires apparaissent sur les trois lanes et se déplacent automatiquement pour aider les joueurs à attaquer les tourelles adverses, ce qui leurs permet de se frayer un passage vers le Nexus.

Les monstres élites donnent un avantage stratégique aux joueurs de l'équipe qui l'ont éliminé. Il y a trois type de monstre élite différent.

1. Les *dragons*, au nombre de 5 actuellement, offrent un avantage de combat permanent à l'équipe.
2. Le *Héraut de la Faille*, il peut être capturé puis relâché sur une tourelle adverse qu'il détruira presque complètement. Il octroie aussi des PO et de l'XP lorsqu'il est capturé.
3. Le *Baron Nashor*, le plus puissant monstre élite. Il octroie 300 PO et de l'XP à toute l'équipe lorsqu'il est vaincu. Il augmente les dégâts des sbires et des joueurs de l'équipe pendant 90 secondes. Les sbires sont aussi beaucoup plus résistants.

Les balises, ou *wards*, elles sont placées par un joueur et donnent de la vision tout autour. Elles peuvent aussi être détruites par l'équipe adverse. Un joueur en possède en quantité limitée et doit donc les placer de manière rigoureuse, pour qu'elles aient un impact stratégique tout en restant discrète pour ne pas être facilement détruisible par un adversaire.

Éléments stratégiques Les wards constituent un élément stratégique important dans League of Legends. Elles permettent d'avoir de l'information sur l'emplacement des joueurs adverse dans l'arène lorsqu'ils passent à proximité. Elles permettent aussi de s'assurer qu'il n'y a pas d'adversaire aux alentours lors de la prise d'un objectif (un monstre d'élite par exemple). Si des adversaires sont à proximité, ils pourraient tenter de voler l'objectif, et donc récupérer les avantages qu'il procure.

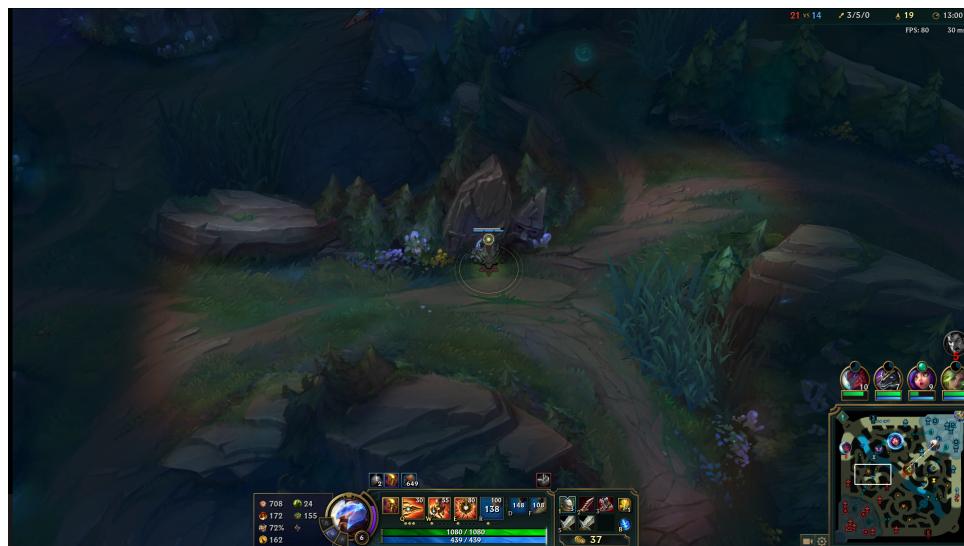


FIGURE 5 – Une ward donnant la vision dans la jungle adverse

Prendre des objectifs (monstre d'élite) permet de faciliter les combats, la destruction des tourelles et donc l'avancement d'une équipe vers le Nexus de la base ennemie. Les objectifs apparaissent à certains moments de la partie, et souvent les équipent se réunissent pour jouer ensemble la prise d'objectif. Ce sont souvent des combats d'équipe déterminants, qui peuvent retourner le déroulement d'une partie.

Il est important de détruire une tourelle avant 14 minutes de jeu, car elle procurera au total 1150 PO aux joueurs. En effet, chaque plate forme donne 160 PO, ce qui fait 800 PO au total, et la destruction d'une tourelle donne 350 PO. Cela donne un avantage considérable à ces joueurs qui pourront ensuite acheter des items pour augmenter leurs caractéristiques de combat.

Amasser des éliminations, ou *kills*, permet d'empêcher un adversaire de jouer pendant un laps de temps tout en octroyant des PO et de l'XP. Gagner les combats est donc primordial pour gagner en puissance rapidement. Plus l'on est puissant, plus les combats seront faciles par la suite, et plus on amassera d'éliminations. Il y a un effet boule de neige très important dans League of Legends et qui impacte fortement les parties.

Les rangs League of Legends est un jeu compétitif, à chaque joueur est attribué un rang correspondant à son Elo (analogique à l'Elo aux Échecs). Les rangs, du plus petit au plus grand, sont : *Fer*, *Bronze*, *Argent*, *Or*, *Platine*, *Diamant*, *Master*, *Grandmaster* et *Challenger*. Les Challengers sont les 300 meilleurs joueurs d'Europe.



FIGURE 6 – Les rangs

2 Design of Experiment

2.1 Définition de la question et de la population

Comme nous l'avons vu précédemment, il y a de nombreux éléments stratégiques qui influent une partie de League of Legends, et ce dès les premières minutes. Pour chercher à optimiser les chances de victoires d'une équipe, nous voulons donc répondre à la question suivante : Sur quels facteurs faut-il se concentrer pendant les dix premières minutes pour remporter une partie de League of Legends ?

La population étudiée sera constituée de joueurs Challenger tirés aléatoirement en Europe. Nous choisissons de nous restreindre aux 300 meilleurs joueurs d'Europe car nous pouvons être certains que ce sont de bons joueurs, qu'ils connaissent les règles et qu'ils essayent pleinement de remporter la partie. Les Challengers savent comment tirer profit de leur avantage de PO et d'XP, donc les impacts des facteurs décisifs seront plus marqués. De plus, les Challengers choisiront des champions ayant de bonnes synergies entre eux.

Il faudra prendre en compte cet aspect lors de l'interprétation, les conclusions tirées à partir de cette étude seront difficilement applicables à des joueurs ayant un rang assez bas, puisque ce seront des populations trop éloignées.

2.2 Taille d'échantillon et plan d'expérience

Nous faisons l'hypothèse que les choix de champions n'impactent pas les chances de victoires car nous observerons des parties de Challenger, qui savent choisir leurs champions. Si ce n'était pas le cas, une équipe pourrait être désavantagée par un choix de champion inadapté à un rôle par exemple. Nous ne prendrons donc pas en compte les choix de champions dans cette étude.

Nous négligeons aussi les impacts psychologiques liés à des séries de défaites ou des séries de victoires, ou tout simplement les différences de niveau entre les joueurs. C'est à dire que nous considérons la population de Challenger comme relativement homogène.

Nous voulons récolter les données d'un nombre p de facteurs sur un nombre N de parties de League of Legends, et ce pendant les 10 premières minutes. En effet, les 10 premières minutes semblent parfois conditionner la victoire dans les parties compétitives de haut niveau. Les p facteurs observés pour une équipe seront :

1. Les ward posées (ward_placed)
2. Les ward détruites (ward_kill)
3. Les montées de niveau (level_up)
4. Les éliminations de champions (champion_kill)
5. Les assistances aux éliminations de champions (champion_assist)
6. Les éliminations en série (champion_special_kill)
7. Les captures d'objectifs (elite_monster_kill)
8. Les tourelles détruites (building_kill)
9. Les plates détruites (turret_plate_destroyed)
10. Les PO récoltées (total_gold)
11. Les sbires et monstres de la jungle tués (minion_killed)

On a $p = 11$, ce qui nous fait une matrice de facteurs de taille $10 \times p$ puisqu'on observe ces facteurs chaque minute pendant les 10 premières minutes. Nous nous fixons comme objectif un ratio de 40 entre le nombre de facteurs total (10×11) et le nombre de parties N , ce qui fixe à 4400 la taille minimale d'échantillon à collecter. Cependant, il semble tout à fait possible de réduire la dimensionnalité en sommant ces facteurs sur les 10 premières minutes, ce qui nous donne $1 \times p$ facteurs à évaluer sur 4400 échantillons. Cela donne un ratio de 400, donc une faible variance des estimateurs. En effet, plus la taille de l'échantillon est grande devant le nombre de facteurs, plus la variance des estimateurs diminue.

Pour chaque partie, nous collectons aussi l'équipe gagnante.

Nous pourrons ensuite entraîner des modèles de machine learning de classification, qui tenteront de prédire la victoire ou non d'une équipe à partir des 11 facteurs que l'on a dégagé. Ensuite, une analyse du modèle permettra d'évaluer l'importance de chaque facteur dans les prédictions, et donc de déterminer quels sont les facteurs les plus importants pour mener à la victoire.

Cependant, si nous n'arrivons pas à établir de modèle de prédiction satisfaisant, trois cas se présentent à nous :

1. Soit les 10 premières minutes n'ont pas ou peu d'importance sur la victoire d'une équipe
2. Soit les modèles utilisés n'étaient pas adaptés
3. Soit les 11 paramètres/facteurs que l'on a choisi ne sont pas suffisant pour caractériser une partie

Il devrait être aussi possible de les évaluer en utilisant des matrices de Hadamard et et calculer les effets de chaque facteur, ou bien tout simplement une analyse en composantes principales (ACP).

2.3 Organisation et déroulement de la collecte

Nous avons utilisé l'API de League of Legends disponible ici : <https://developer.riotgames.com/apis>. Puisque nous ne sommes pas développeur d'une application reconnue par Riot, nous avions des limitations pour l'utilisation de l'API. (cf. fig 7)

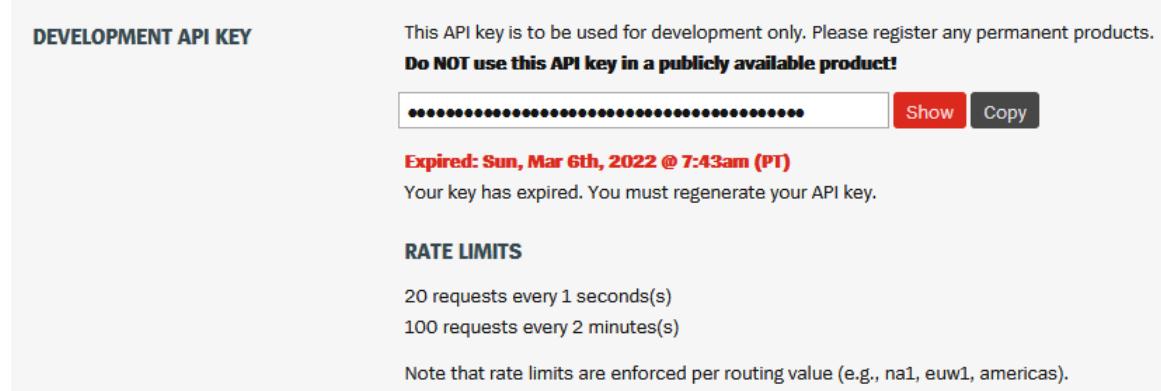


FIGURE 7 – Limitation de la clé d'API fournie par Riot

Les APIs que nous avons utilisé sont :

1. `/riot/account/v1/accounts/by-riot-id/gameName/tagLine` qui permet de récupérer le PUUID d'un joueur à partir de son pseudonyme et de sa tagLine. Pour les joueurs en europe de l'ouest, la tagline est "EUW". L'UUID (Player Universally Unique IDentifier) est un l'identifiant d'un joueur, unique à l'échelle mondiale.
2. `/lol/match/v5.matches/by-puuid/puuid/ids` qui permet de récupérer une liste d'identifiant de parties (match) pour un PUUID donné. Pour cela, il faut passer les paramètre de la requête dans l'URL. Par exemple, si nous voulons récupérer 10 matchs classés, nous devions écrire : "/ids?type=ranked&count=10" à la fin de l'URL.
3. `/lol/league/v4/challengerleagues/by-queue/RANKED_SOLO_5x5` qui permet de récupérer les pseudonymes et tagLines des Challengers en Europe.
4. `/lol/match/v5.matches/matchId/timeline` qui renvoie la timeline contenant tous les événements d'une partie, minute par minute.

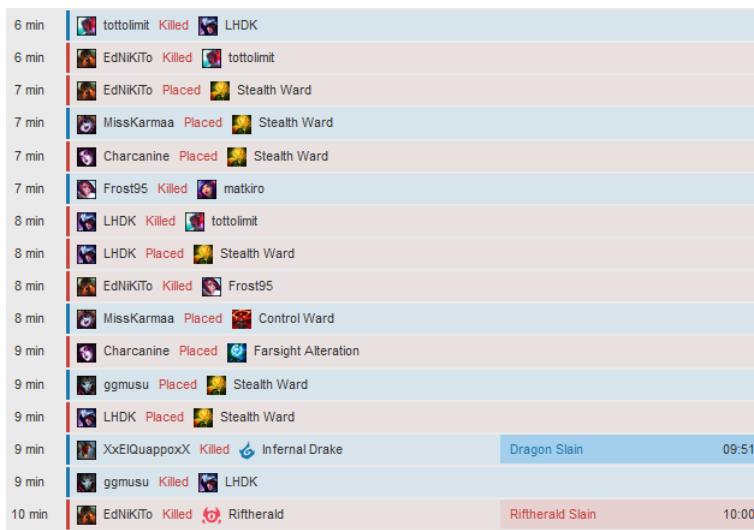


FIGURE 8 – Représentation d'un extrait de timeline

Grâce à ces 4 APIs, nous avons récolté les timelines de 4541 matchs de League of Legends. Tout d'abord, nous avons récolté les pseudonymes des 290 meilleurs joueurs d'Europe (Challengers), pour chaque joueur nous avons cherché les derniers matchs auxquel il a participé. Cependant, il était possible que deux joueurs aient participé au même match alors il fallait éliminer les doublons. Nous avons pu extraire 4541 matchId en approximativement 10 heures. En effet, il fallait ajouter artificiellement des pauses (sleep) dans notre programme pour respecter les limites de notre token d'API, qui avait un nombre limité de requêtes par minute. On a plusieurs fois rencontré des erreurs "429 - Too Many Requests", qui faisaient planter notre récolte et l'on devait la reprendre du début.

Une fois que l'on avait une liste de matchId, nous avons récupéré la timeline de chaque match et nous avons sélectionné les évènements du match qui suivaient les 11 facteurs que l'on a dégagé précédemment. Nous les avons récupéré pour l'équipe bleue et l'équipe rouge, ainsi que l'équipe gagnante.

Nous avons ensuite exporté les données dans deux CSV, l'un contenant les valeurs des facteurs à chaque minute d'une partie, l'autre renseignant sur l'équipe gagnante.

2.4 Stockage de données sous une base de donnée MySQL

Nous avons choisi d'utiliser un système de base de données MySQL permettant de manipuler des bases de données et d'accéder ainsi à leur contenu. Dans le cadre de notre projet, le jeu de donnée utilisé n'étant pas immense, il ne nécessitait pas forcément l'utilisation d'une base de données NOSQL par exemple. Afin de traiter nos données, nous avons réalisé trois scripts en python. Le premier, *DataImport.py*, sert à l'importation des données des deux fichiers *.csv* précédemment obtenue : *TLC4541.csv* et *WC4541.csv* recensant les matchs et les gagnants. Le second, *DataBaseConnectionCreation.py* quant à lui, sert à la connexion avec le serveur MySQL et en la création de notre base de données. Le dernier, *DataInsertion.py*, est le script qui va servir à l'insertion des données des deux fichiers *.csv* dans les tables de notre base de données qui a été précédemment créée.

Enfin, on écrit un script Shell Bash *excution.sh*, afin de lancer exécution automatique de ces trois fichiers l'un après l'autre permettant ainsi d'automatiser la tache.

3 Conclusion

Nous avons réussi à récupérer les données relatives aux éléments stratégiques mis en place par les équipes dans plus de 4500 matchs au niveau de Challenger. Nous les avons stockées dans une base de donnée SQL qui permettra, avec une simple fonction d'agrégation, de sélectionner seulement les dix premières minutes. Ces données pourront être par la suite analysées par différentes méthodes, qui montreront si les dix premières minutes ont un impact lors d'un match de League of Legends, et le cas échéant quels en sont les facteurs les plus important pour mener à la victoire.