

# Données de panel sur League of Legend

Etude économétrique

Projet de recherche réalisé dans le cadre du master 1 MECEN



Université de Tours

Par Alexis VINCENT, Iruomachi IRUOMAH, Yann BROCHET

2024-2025

# 1 Introduction

Pour notre étude nous sommes partis sur quelque chose d'original, nous avons voulu faire des régressions sur un jeu vidéo que l'on affectionne depuis l'enfance. (Ne vous inquiétez pas nous avons fait en sorte que cela soit compréhensible et reste censé au niveau de l'interprétation pour vous).

League of Legends est un jeu vidéo où deux équipes de 5 joueurs s'affrontent sur une carte appelée la "Faille de l'Invocateur". L'objectif principal est de détruire la base de l'équipe adverse. La carte est divisée en trois voies : haute, centrale et basse. Chaque équipe commence aux coins opposés de la carte. Pour détruire la base ennemie, il faut d'abord traverser une voie, détruire les trois tourelles qui la composent, ainsi que l'inhibiteur associé. Ensuite, chaque équipe doit détruire les deux tourelles protégeant le cœur de la base ennemie. Une fois toutes ces structures détruites, la base elle-même devient vulnérable et peut être détruite pour remporter la victoire.

Les parties durent généralement entre 20 et 40 minutes, et accomplir des objectifs sur la carte permet de gagner des récompenses, offrant ainsi des avantages pour l'équipe qui les réalise. Dans cette étude, nous allons analyser quelles variables influencent la victoire d'une équipe, qu'elles aient un impact positif ou négatif.

Si vous souhaitez mieux comprendre le but du jeu, ce n'est pas obligatoire pour saisir l'étude, mais vous pouvez regarder cette petite vidéo sympathique qui explique les bases du jeu : [ici](#)

## 2 Variables et modèle

Nous avons étudié 5 variables en analysant les performances de 20 joueurs très compétitifs évoluant au plus haut niveau de League of Legends avec une partie tout les mois sur une années (qui correspond à une saison de compétition sur ce jeu):

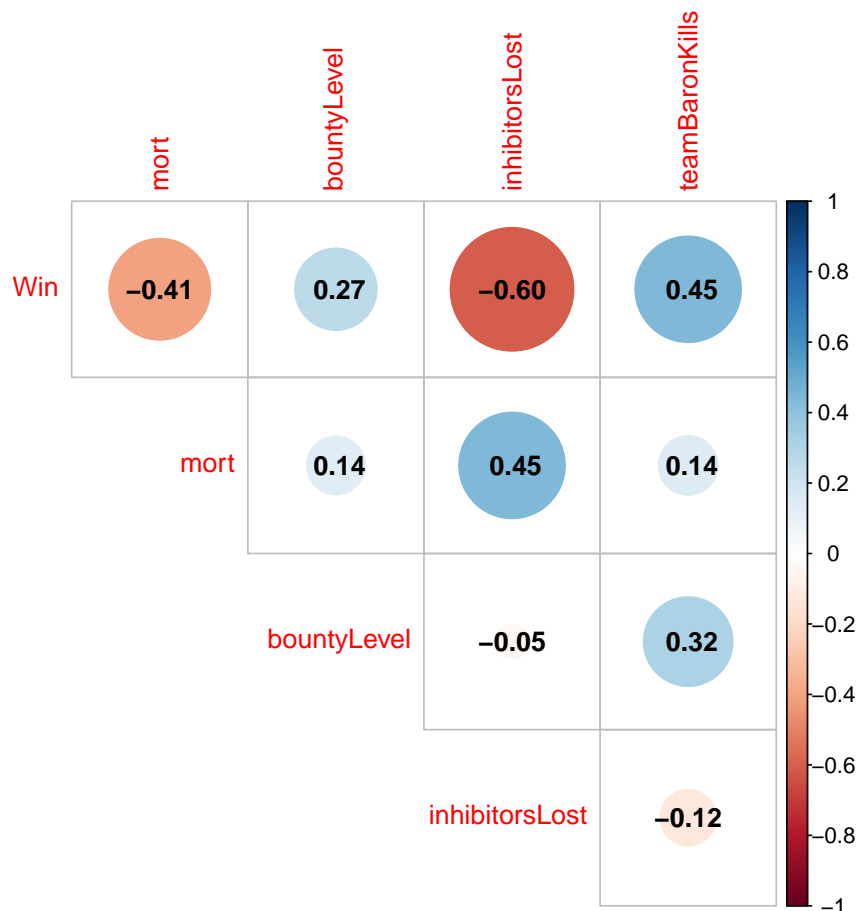
- teamBaronKills : Cette variable représente le nombre de barons tués par une équipe. Le baron est l'un des plus gros objectifs du jeu et procure un des plus grands avantages (augmentant ainsi les chances de gagner).
- mort : Cette variable représente le nombre de morts d'un joueur. Se faire tuer par un adversaire donne des ressources à ces derniers, ce qui leur permet de gagner plus facilement.
- inhibitorLost : Cette variable représente le nombre d'inhibiteurs perdus. Les inhibiteurs sont des bâtiments proches du cœur de la base des joueurs ; les perdre offre un très gros avantage à l'équipe adverse.
- bountyLevel : Cette variable représente la prime qu'un adversaire récupère lorsqu'il tue un joueur. Cette prime augmente si le joueur réalise une très bonne partie (c'est-à-dire, s'il obtient un très bon score).
- knowledge : Il existe 5 rôles dans League of Legends : Top, Jungle, Mid, ADC et Support. Cette variable, qui est fixe, indique le rôle de prédilection d'un joueur (expertise au Top, Jungle, Mid, ADC ou Support).

## 2.1 Voici la distribution statistique de nos variables

Win	mort	bountyLevel	inhibitorsLost	teamBaronKills	knowledge
Min. :0.000	Min. : 0.000	Min. : 0.0	Min. :0.0000	Min. :0.0000	adc :36
1st Qu.:0.000	1st Qu.: 3.000	1st Qu.: 3.0	1st Qu.:0.0000	1st Qu.:0.0000	jgl :48
Median :1.000	Median : 5.000	Median : 5.0	Median :0.0000	Median :0.0000	mid :48
Mean :0.575	Mean : 4.971	Mean : 6.1	Mean :0.5708	Mean :0.5333	supp:36
3rd Qu.:1.000	3rd Qu.: 7.000	3rd Qu.: 9.0	3rd Qu.:1.0000	3rd Qu.:1.0000	top :72
Max. :1.000	Max. :15.000	Max. :24.0	Max. :4.0000	Max. :3.0000	NA

- On peut constater que plus le nombre de parties est grand, plus la probabilité de faire une victoire semble se rapprocher en moyenne de 0,5. En effet, il existe un a priori selon lequel les développeurs du jeu feraient en sorte que l'on gagne au moins la moitié de nos parties (par exemple, en influençant nos équipiers au début de la partie choisie aléatoirement, je le rappelle).
- Pour la variable de Nashor tués, la moyenne est assez faible en raison du fait que c'est un objectif de fin de partie, donc souvent la partie se finit soit avant (donc sans Nashor tué), soit avec un ou deux grand max. On peut voir que le max ici est de 4, qui est un cas extrême et peu observable dans la réalité.
- Pour ce qui est de la variable knowledge, on note une sur-représentation du rôle de toplaner dans notre jeu de données.

Voici un aperçu des corrélations de nos variables (sans la variable knowledge)



Nous allons étudier ce modèle:

$$\text{Win}_{it} = \beta_0 + \beta_1 \cdot \text{teamBaronKills}_{it} + \beta_2 \cdot \text{mort}_{it} + \beta_3 \cdot \text{bountyLevel}_{it} + \beta_4 \cdot \text{inhibitorsLost}_{it} + \beta_5 \cdot \text{factor(knowledge)}_i + c_i + \epsilon_{it}$$

Table 1: Model with standard error correction

	<i>Dependent variable:</i>		
		Win	
	Poolé	Within	Random
teamBaronKills	0.264*** (0.029)	0.256*** (0.033)	0.262*** (0.030)
mort	-0.055*** (0.008)	-0.053*** (0.009)	-0.055*** (0.008)
bountyLevel	0.026*** (0.007)	0.029*** (0.009)	0.027*** (0.008)
inhibitorsLost	-0.225*** (0.030)	-0.221*** (0.033)	-0.224*** (0.031)
factor(knowledge)jgl	-0.049 (0.040)		-0.048 (0.041)
factor(knowledge)mid	-0.044 (0.060)		-0.042 (0.061)
factor(knowledge)supp	0.186*** (0.064)		0.190*** (0.067)
factor(knowledge)top	0.057 (0.045)		0.057 (0.046)
Constant	0.651*** (0.071)		0.643*** (0.075)
Observations	240	240	240
R <sup>2</sup>	0.611	0.591	0.605
Adjusted R <sup>2</sup>	0.597	0.547	0.592
F Statistic	45.289*** (df = 8; 231)	77.994*** (df = 4; 216)	354.476***

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Pour revenir sur nos corrélations, on avait bien un a priori sur les signes de nos variables : en effet, par exemple, avoir un nombre de morts important baisserait nos chances de victoire (plus expliqué dans la section suivante), tuer un Nashor augmenterait les chances de victoire, perdre un inhibiteur diminuerait nos chances de victoire et enfin avoir un score de recherche élevé (bounty level) augmenterait nos chances de victoire. On observe donc que les modèles suivent bien cette logique, pouvant donc induire peut-être un vrai lien de cause à effet entre ces variables et la chance de gagner une partie.

### 3 Interprétation

Nous avons réalisé une régressions sur 3 modèles : poolés, within et effets aléatoires afin d'observer lequel serait le meilleur pour notre étude.

- Premièrement, nous pouvons observer que toutes nos variables sont significatives au seuil de 1 %, hormis pour quelques indicatrices de notre variable knowledge.
- Éliminer un baron augmente, en moyenne, la chance de victoire de 26 %, toutes choses égales par ailleurs (comme expliqué dans l'introduction des variables, éliminer un baron offre de gros avantages à son équipe).
- Mourir une fois de plus diminue en moyenne la chance de victoire de 5 %, toutes choses égales par ailleurs (c'est dû au fait que mourir donne des ressources à l'adversaire).
- Si l'indice du bountyLevel augmente de 1 pour un joueur, sa chance de gagner la partie augmente en moyenne de 2,7 %, toutes choses égales par ailleurs (cela s'explique par le fait qu'un joueur avec un très gros bounty signifie qu'il a un bon score. Il a donc plus de chances de gagner, mais attention : s'il meurt, il offrira beaucoup de ressources à l'adversaire).
- Perdre un inhibiteur diminue en moyenne les chances de remporter une partie d'environ 22 %, toutes choses égales par ailleurs. En effet, un inhibiteur est un bâtiment très proche du noyau de la base d'une équipe. Si on le perd, cela offre tout d'abord des avantages à l'équipe adverse et signifie que l'adversaire est très proche de gagner !
- Il existe un a priori très fort dans League of Legends qui dit que le rôle support est le meilleur pour remporter le plus de parties. Dans notre régression, après avoir mis le rôle adc en référence afin d'éviter tout problème de multicollinéarité parfaite, on peut effectivement observer que la connaissance du rôle support est celui qui augmente le plus la chance de gagner en moyenne.
- De plus, notre R2 dépasse les 0,5, ce qui est plutôt satisfaisant.

Il est important de noter que la variable expliquée, "Win", est une variable binaire. Or, pour réaliser une régression avec une variable binaire et obtenir des résultats optimaux, il est nécessaire d'étudier la corrélation entre les variables explicatives ainsi que la distribution de la variable dépendante. Cela permet de déterminer s'il est préférable d'appliquer un modèle logit ou probit pour prédire la probabilité de faire une "Win". Dans notre cas, la méthode la plus optimale serait d'utiliser un modèle logit. Cependant, nous avons choisi de ne pas l'appliquer afin de favoriser une interprétation rapide des variables, au détriment de l'optimalité du modèle.

### 4 Conclusion

Nos coefficients sont très similaires, mais les erreurs types dans le modèle à effets aléatoires sont légèrement plus grandes que dans le modèle poolé. Cependant, c'est le modèle within qui possède les erreurs types les plus grandes. On aurait donc pu se diriger vers ce modèle, mais on perd l'information des caractéristiques individuelles, car elles ne varient pas dans le temps. On aurait donc pu conclure sur le choix d'un modèle à effets aléatoires.

Cependant, nous sommes dans un cas où nos effets individuels (knowledge d'un rôle, dans ce cas) sont corrélés à nos variables explicatives. En effet, la connaissance parfaite d'un rôle permet d'avoir moins de morts, un meilleur score, et donc un bounty plus élevé, ainsi que de faire plus d'objectifs, ce qui permet de remporter plus de parties. De par cette corrélation, nous devons nous diriger vers le modèle within (à effets fixes).

## 5 Annexe

Pour récupérer les données, nous avons conçu un script Python qui interagit avec l'API de Riot Games. Ce script est conçu pour effectuer des requêtes HTML sur les serveurs de Riot Games afin d'extraire les informations relatives aux parties des joueurs. Concrètement, le processus fonctionne de la manière suivante : le script envoie des requêtes via l'API de Riot Games, qui permet d'accéder à des bases de données contenant des informations sur les matchs joués par les joueurs. Ensuite, ces données sont récupérées et analysées, incluant des informations sur les différentes variables associées à chaque partie (comme le nombre de kills, de morts, la durée de la partie, etc.). Les informations sont ensuite extraites et stockées sous une forme structurée pour pouvoir être utilisées dans notre analyse.

Il faut noter que cette analyse reste très rudimentaire. En effet, League of Legends est un jeu comportant de nombreuses mécaniques complexes, tant explicites qu'implicites. Il existe donc probablement des méthodes plus rigoureuses pour analyser les chances de victoire. Toutefois, notre objectif était d'obtenir un premier aperçu des facteurs qui peuvent influencer une victoire.

De plus, étant donné que cette étude se concentre sur l'élite des joueurs de League of Legends, il est important de préciser que ces derniers jouent exclusivement à leur rôle de prédilection dans toutes leurs parties.