

Analyse Exploratoire

Rapport : le financement participatif

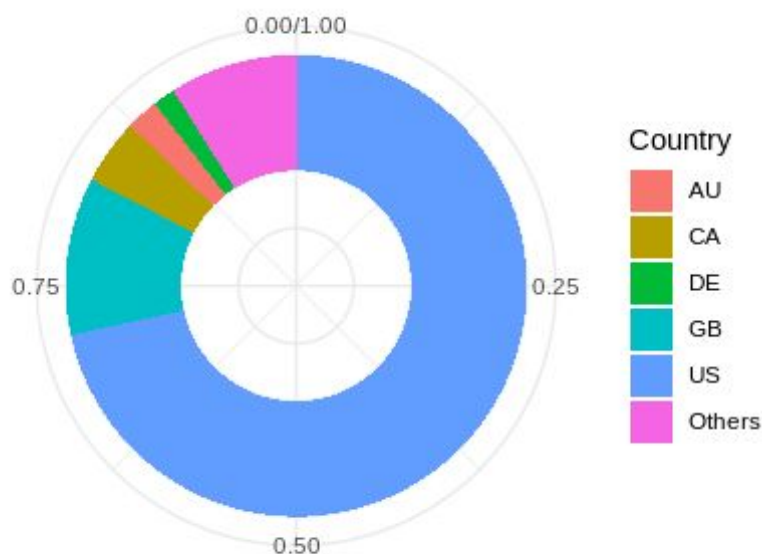
Quels sont les facteurs de succès d'une campagne participative ?

I. Introduction

Dans le cadre du cours d'analyse exploratoire, nous devons trouver un jeu de données et l'analyser en utilisant le langage R que nous avons appris ce semestre.

Nous avons choisi de mener diverses analyses sur le jeu de données suivant : <https://www.kaggle.com/yashkantharia/kickstarter-campaigns>

Il contient des informations sur les projets initiés sur la plateforme KickStarter, une entreprise américaine de financement participatif créée en 2009. Les données ont été récoltées entre avril 2009 et février 2019. Pour chaque projet, nous disposons d'informations sur son nom, les dates de début et de fin de la campagne, l'argent demandé et obtenu en fin de campagne, ainsi que le pays d'origine.



On note que les données concernent majoritairement des projets américains, représentant presque 75% des projets. Cela est dû au fait que KickStarter est une plateforme américaine, nous risquons donc d'avoir des résultats peu représentatifs pour des projets hors-US.

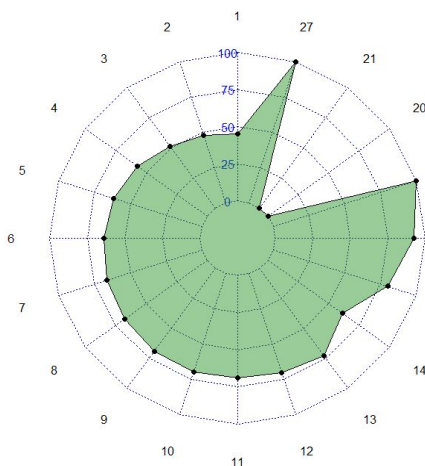
Nous cherchons alors à déterminer quels sont les facteurs de succès d'une campagne participative pour estimer les chances de réussite d'un éventuel nouveau projet. Nous allons donc analyser ce jeu de données et visualiser ces résultats via un script R. Le script utilisé pour produire toutes les analyses exposées ici, ainsi que le jeu de données, sont accessibles depuis le dépôt git suivant : <https://github.com/rizziemma/analyse-exp>

II. Analyse du titre

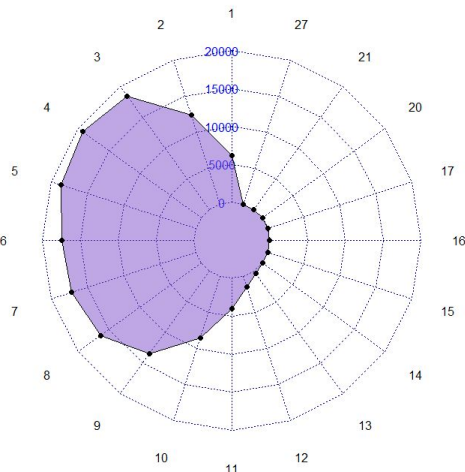
Le titre étant la première information que l'on lit, nous nous sommes demandé s'il avait un impact sur le succès du projet. Plus précisément, nous avons cherché l'impact du nombre de mots du titre. Dans notre jeu de données, la composition des titres varie de 1 mot à 27 mots, alors nous avons comparé le pourcentage de succès des projets en fonction du nombre de mots.

Nous avons remarqué des pourcentages étonnants (comme 0% et 100% de succès à 21 et 27 mots), nous avons donc décidé de visualiser également le nombre de projets en fonction du nombre de mots dans le titre. Cela nous a permis de constater qu'il était pertinent de regarder seulement les projets ayant un titre composé de 1 mot à 10 mots, car représentant au moins 5000 projets.

Pourcentage de succès selon le nombre de mots dans le titre



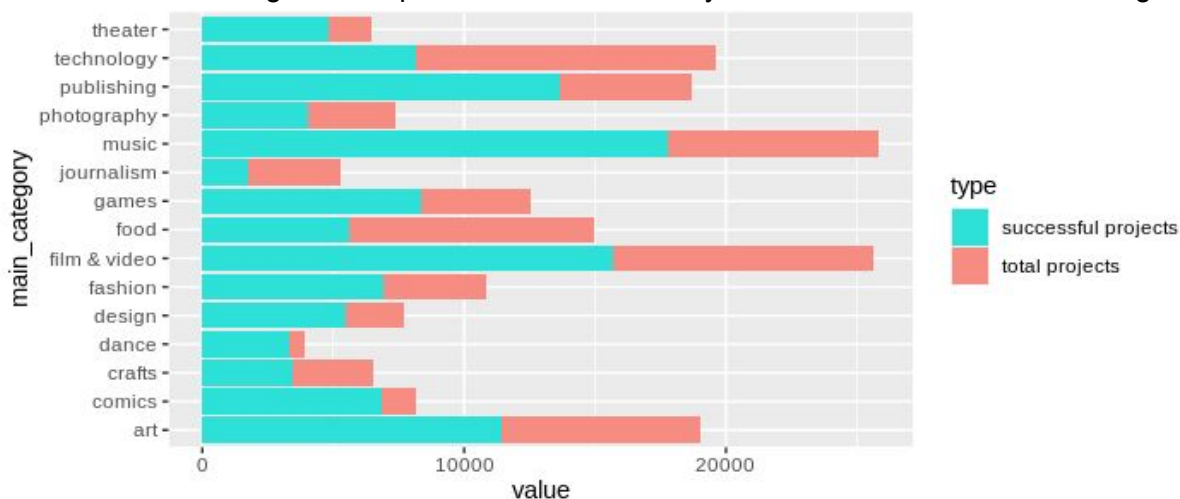
Nombre de projets selon le nombre de mots dans le titre



Nous constatons que le pourcentage de succès est croissant de 1 mot à 7 mots, puis qu'il se stabilise au-delà. En conclusion, nous pouvons dire que le titre doit être assez précis, d'un minimum de 7 mots si l'on veut augmenter ses chances de succès. On note également qu'il ne faut pas avoir peur des titres très long, puisque les projets ayant un titre de 10 mots ont la même chance de succès que ceux ayant 7 mots.

III. Analyse des catégories

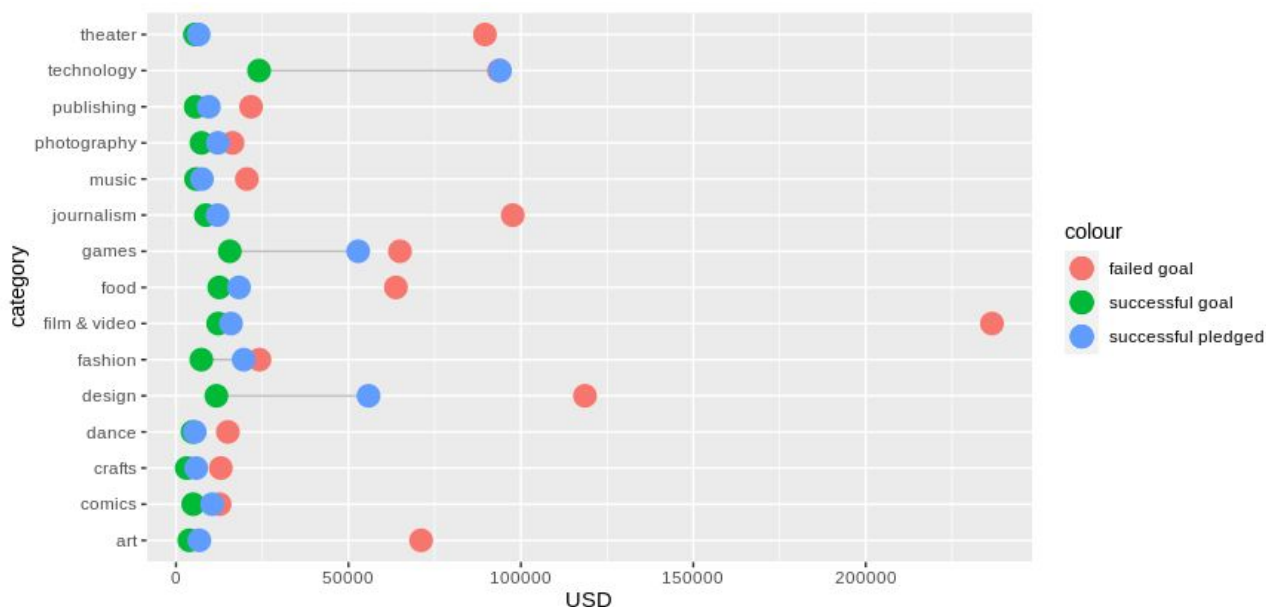
Lorsque l'on parcourt le site Kickstarter pour voir les campagnes en cours, il faut sélectionner quelle catégorie explorer. La catégorie d'un projet influe alors beaucoup sur sa visibilité sur le site. Quelles sont les catégories les plus fructueuses ? Analysons les données selon les catégories.



Nous avons observé que certaines catégories ont de très bons taux de réussite, avec très peu de projets échoués (en rouge), mais également avec un nombre total de projets plutôt bas. C'est le cas des catégories **theater**, **design**, **dance**, **comics**. On remarque tout de même que les catégories **photography**, **journalism** et **crafts** ont environ le même nombre de projets (moins de 10 000, mais ont beaucoup moins de succès. On en déduit que ces catégories intéressent globalement moins les participants, et qu'il sera donc difficile de réussir une campagne dans ces catégories.

Pour les catégories ayant plus de 10 000 projets, nous voyons que le taux de réussite est globalement moyen, autour de 50 à 70%, ce qui ne garantit donc pas la réussite. On peut en déduire que le nombre de projets dans ces catégories impact fortement le taux de succès : plus il y a de projets, moins la visibilité d'un nouveau projet sera bonne sur le site internet. Une nouvelle campagne risque donc de passer inaperçue aux côtés des milliers d'autres en cours. Pour lancer un projet dans ces catégories (**technology**, **publishing**, **music**, **games**, **food**, **film&video**, **fashion**, **art**), il faut s'attendre à une concurrence rude et savoir se démarquer, car les chances de succès sont plutôt faibles.

Pour détailler un peu plus ces résultats nous avons observé l'impact du montant demandé par les campagnes selon les catégories. Ici, nous représentons les sommes moyennes demandées par les projets réussis et échoués, ainsi que la somme moyenne obtenue par les projets réussis.



Nous observons que globalement, les projets échoués demandaient beaucoup plus d'argent que les projets réussis. L'écart est d'autant plus important dans les catégories **theater**, **technology**, **journalism**, **film&video**, **design** et **art**. Pour les autres catégories, l'écart des sommes demandées entre les projets réussis et échoués est très faible. De manière générale, pour toutes les catégories, il sera donc très difficile de réussir une campagne demandant plusieurs milliers de dollars. Il est très difficile d'obtenir de grandes sommes via le financement participatif.

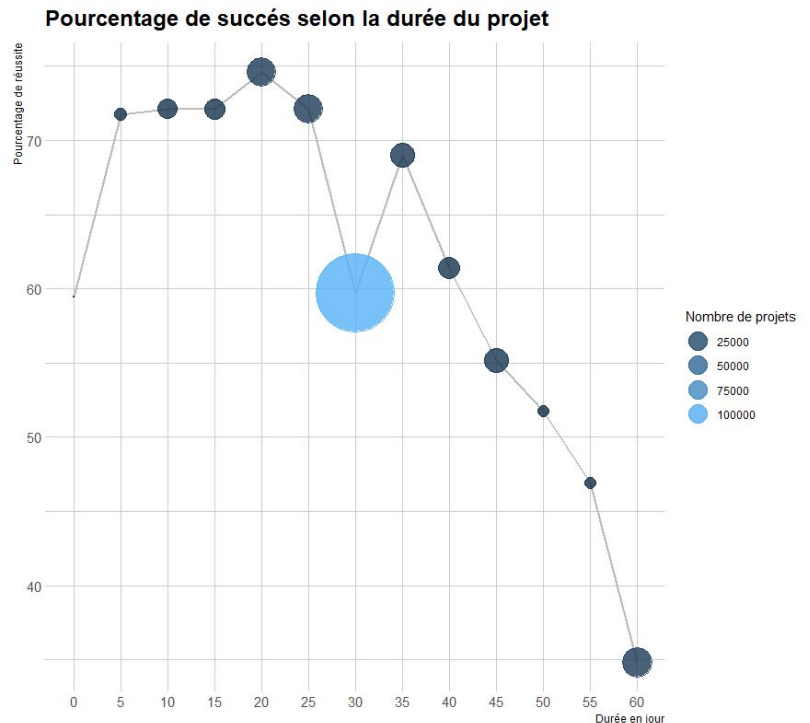
En revanche, lorsque l'on observe les sommes moyennes obtenues par les projets réussis, nous voyons que celles-ci dépassent légèrement la somme demandée, sauf pour les catégories **technology**, **games**, et **design** où la somme moyenne obtenue atteint jusqu'à 4 fois l'objectif. Cet écart nous montre que pour ces catégories, un projet qui plaît au public peut espérer obtenir beaucoup plus que son objectif. Nous expliquons cet écart notamment par la définition de plusieurs paliers objectifs, lorsque le premier palier est atteint le projet est considéré réussi, mais les dons peuvent continuer à affluer pour atteindre les objectifs suivants.

IV. Analyse de la durée des projets

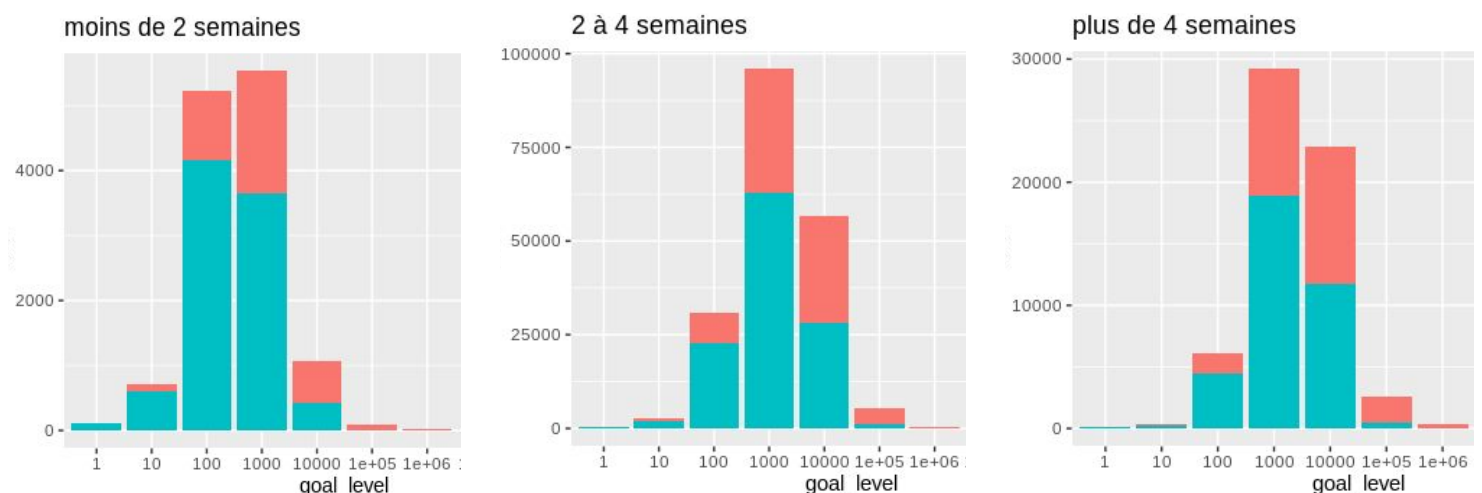
Enfin, nous nous sommes demandé si la durée des projets avait un impact sur la probabilité de succès. Cela pourrait nous permettre de savoir, par exemple, quelle est la durée optimale d'un projet ou quelles sont les chances de succès pour une durée donnée.

Pour ce faire, nous avons calculé et visualisé le pourcentage du succès des projets selon la durée, tout en prenant en compte le nombre de projets pour chaque durée. Nous avons décidé de prendre en compte les projets allant de 0 à 60 jours (durée maximale d'une campagne), et de faire nos calculs par palier de 5 jours.

Dans un premier temps, on peut alors penser que les projets de courtes durées sont plus propice au succès.



Cependant, nous nous sommes demandé si ce résultat était pertinent, et nous avons décidé de prendre également en compte l'argent demandé et obtenu en fonction de la durée des projets. Nous représentons ici le nombre de projets totaux ainsi que les projets réussis selon l'objectif. Les paliers sont représentés comme suit : 1 représente les objectifs de 1 à 9 dollars, 10 de 10 à 99 dollars 1000 de 1000 à 9999 dollars, ainsi de suite.



De manière générale, les projets demandant moins de 1000 dollars ont de très bonnes chances de réussite, peu importe leur durée. Pour les projets de plus de 1000 dollars, le taux de succès semble constant peu importe la durée, mais le nombre de projets dans ces paliers est bien plus important parmi les projets de plus de 2 semaines. Pour les projets de plus de 10000 dollars, le taux de réussite est légèrement meilleur à partir de 2 semaines (moins de 50% pour moins de 2 semaines, approche les 50% pour plus de 2 semaines au palier des 10 000 dollars).

Nous concluons que peu importe la durée du projet, les chances de réussite sont plutôt faibles lorsque l'on demande de grandes sommes (plus de 10 000 dollars), mais que dans ce cas il vaut mieux choisir une durée d'au moins 2 semaines. Pour les campagnes demandant moins, nous voyons que choisir une durée plus longue n'influe pas sur le taux de succès. Ainsi, pour un projet

demandant moins de 1000 dollars, on peut espérer atteindre l'objectif très rapidement (moins de 2 semaines). On peut estimer que l'objectif soit ainsi dépassé, et si l'on choisit une durée plus grande cela permettrait peut être d'atteindre d'éventuels paliers supérieur de l'objectif.

Parmi les curiosités observées, nous remarquons un certain nombre de projets demandant moins de 10 dollars. Nous avons remarqué que ces projets ont en fait un premier palier très bas, mais espèrent en réalité des montants plus grands. (un des exemples observés : <https://www.kickstarter.com/projects/tattoofixingbook/time-for-art-2-waterloo-bridge/description> ayant obtenu 6 dollars au total, mais espérait plus d'après les dernières mises à jour)

V. Conclusion

Nous avons réussi à déterminer plusieurs facteurs de succès d'une campagne participative en se basant sur les projets des 10 dernières années. Pour estimer les chances de réussite d'un projet, voici les différents facteurs observés :

- **titre** : il doit être détaillé, avec au minimum 7 mots il y a plus de chances de réussites
- **catégorie** : certaines ont plus de succès que d'autres. Theater, design, dance et comics ont généralement de meilleures chances de succès. Les projets technology, games, et design réussis obtiennent en moyenne 3 à 4 fois plus que leur objectifs.
- **objectif** : globalement un projet a plus de chances de réussir en demandant une petite somme (moins de 1000 dollars). Au-delà, il devient très difficile d'obtenir la somme voulue.
- **durée** : elle impacte peu les chances de succès, mais pour un projet demandant une somme importante (plus de 1000 dollars), il est préférable de choisir au moins 2 semaines.

Nous rappelons tout de même que ces résultats sont à relativiser car les données proviennent d'une seule plateforme américaine. Il pourrait donc exister des différences entre les plateformes et les pays. Malgré tout, nous n'avons pas observé de différence selon les pays d'origine concernant les résultats d'analyse sur le titre et les catégories, les tendances sont restées les mêmes pour les 5 pays les plus représentés. De plus, nous avons utilisé le pourcentage de réussite, mais c'est une notion qui reste vague. En effet, beaucoup de projets se font par palier et obtiennent le statut réussi dès que le premier palier est atteint.

Ce projet nous a permis d'apprendre à manipuler des jeux de données en utilisant le langage R. Il nous a fallu chercher quelles étaient les données intéressantes à analyser. En effet, ce rapport présente les résultats qui nous ont paru les plus pertinents. Nous avons réalisé d'autres calculs comme l'impact de la longueur du résumé ou la date de lancement d'un projet, mais ces résultats n'ont pas révélé de facteurs de succès.

Nous avons également appris qu'il était important de toujours vérifier si la quantité de données est assez importante pour pouvoir tirer des conclusions. En effet, bien que le jeu de données soit grand, lorsque l'on mène une analyse selon certains critères, la taille des données analysées peut devenir trop petite pour en tirer des conclusions pertinentes. Par exemple, certains paliers d'objectifs (au-delà de un million de dollars) contenaient moins d'une centaine de projets, ce qui est trop peu représentatif pour décrire une tendance globale par rapport aux milliers de projets présents dans le jeu de données.

Enfin, nous avons appris à prendre du recul sur le jeu de données choisi. Le nôtre, par exemple, ne concerne qu'une seule plateforme et contient principalement des projets américains. Bien qu'étant la plateforme de financement participatif la plus connue, nous pouvons donc obtenir des résultats très différents pour d'autres plateformes locales et ciblant certains types de projets ou pays (exemple : le site Zeste permet le financement de projets écologiques et solidaires).