



Target Brilliance

cahier des charges

projet d'analyse

2024

Introduction

CONTEXTE DU PROJET

Le projet "Target Brilliance" se concentre sur l'analyse des données des commandes de Target au Brésil, couvrant une période de trois ans, de 2016 à 2018. Ce projet s'inscrit dans un effort plus large visant à comprendre et à améliorer les performances commerciales de Target dans une région clé. Le marché brésilien est particulièrement intéressant en raison de sa taille, de sa diversité géographique, et des habitudes de consommation qui peuvent varier considérablement d'une région à l'autre.



Portée du projet (Scope) :

La portée du projet "Target Brilliance" englobe l'analyse détaillée des données des opérations commerciales de Target au Brésil sur la période de 2016 à 2018. Ce projet vise à examiner divers aspects des ventes, y compris le comportement d'achat des clients.



Sommaire



INTRODUCTION	02
OBJECTIFS	04
LES DONNÉES	06
MÉTHODOLOGIE	09
RESSOURCES	12
LIVRABLES	14
PLANNING	15
RISQUES	16



2. Objectifs

QUESTION PRINCIPAL DU PROJET

Comment Target peut-elle améliorer ses opérations commerciales au Brésil en utilisant l'analyse des données pour comprendre le comportement des clients, les tendances de vente et les performances des produits entre 2016 et 2018 ?

Les objectifs principaux de ce projet «Target Brilliance» sont les suivants :

1. Quelles sont les catégories de produits qui vendent le plus ? Y a-t-il des différences entre les villes ?
2. Comment le nombre de commandes change-t-il chaque mois ? Y a-t-il des saisons où les ventes sont plus élevées ?
3. Quelles sont les méthodes de paiement que les clients utilisent le plus ? Le paiement en plusieurs fois influence-t-il le nombre ou la valeur des commandes ?
4. Quels facteurs influencent le temps de livraison ? Est-ce qu'il y a un lien entre ces facteurs et la satisfaction des clients ?
5. Quelles villes ou régions ont une forte croissance des ventes ? Y a-t-il des opportunités pour ouvrir plus de magasins ou améliorer les processus ?



Les Données

RÉSUMÉ DES DONNÉES



1. Contexte : Dans le projet "Target Brilliance", nous allons travailler avec plusieurs jeux de données qui contiennent des informations essentielles sur les opérations de Target au Brésil. Voici un aperçu des principales sources de données

2. Contenu des données : le contenu des données est constitué de plusieurs fichiers CSV qui contiennent des informations clés sur les opérations commerciales de Target au Brésil entre 2016 et 2018. Voici un aperçu du contenu de chaque fichier

Les Données

DESCRIPTION DES TABLES :

customers



- **customer_id** : Identifiant unique pour chaque client.
- **customer_unique_id** : Identifiant unique reflétant l'identité du client.
- **customer_zip_code_prefix** : Code postal du client.
- **customer_city** : Ville de résidence du client.
- **customer_state** : État de résidence du client.

- **geolocation_zip_code_prefix** : Code postal associé à une localisation.
- **geolocation_lat** : Latitude de la localisation.
- **geolocation_lng** : Longitude de la localisation.
- **geolocation_city** : Ville associée à la localisation.
- **geolocation_state** : État associé à la localisation.

geolocation



order_items



- **order_id** : Identifiant de la commande à laquelle l'article appartient.
- **order_item_id** : Identifiant unique de chaque article dans la commande.
- **product_id** : Identifiant du produit associé à l'article.
- **seller_id** : Identifiant du vendeur.
- **shipping_limit_date** : Date limite d'expédition pour l'article.
- **price** : Prix de l'article.
- **freight_value** : Valeur du fret associée à l'article.

Les Données

DESCRIPTION DES TABLES :

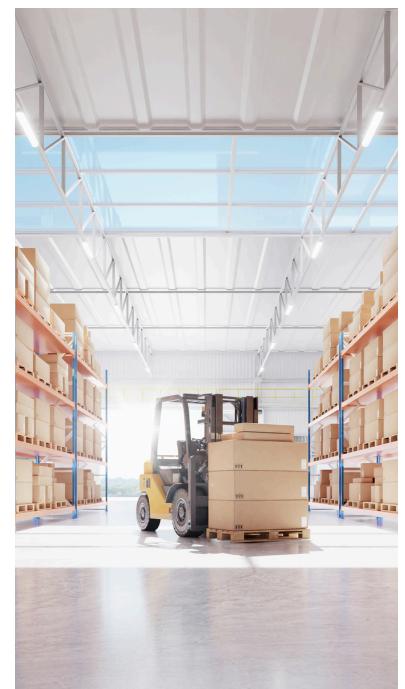
Orders



- **order_id** : Identifiant unique de la commande.
- **customer_id** : Identifiant du client ayant passé la commande.
- **order_status** : Statut actuel de la commande (par exemple, "livré", "en attente").
- **order_purchase_timestamp** : Date et heure d'achat.
- **order_approved_at** : Date et heure d'approbation de la commande.
- **order_delivered_carrier_date** : Date à laquelle la commande a été livrée au transporteur.
- **order_delivered_customer_date** : Date à laquelle le client a reçu la commande.
- **order_estimated_delivery_date** : Date estimée de livraison.

- **product_id** : Identifiant unique pour chaque produit.
- **product_category** : Catégorie à laquelle le produit appartient.
- **product_name_length** : Longueur du nom du produit.
- **product_description_length** : Longueur de la description du produit.
- **product_photos_qty** : Nombre de photos disponibles pour le produit.
- **product_weight_g** : Poids du produit en grammes.
- **product_length_cm** : Longueur du produit en centimètres.
- **product_height_cm** : Hauteur du produit en centimètres.
- **product_width_cm** : Largeur du produit en centimètres.

products



Les Données

DESCRIPTION DES TABLES :

sellers



- **product_id** : Identifiant du produit associé au vendeur.
- **product_seller_id** : Identifiant unique du vendeur.
- **seller_zip_code_prefix** : Code postal associé au vendeur.
- **seller_city** : Ville où se trouve le vendeur.
- **seller_state** : État où se trouve le vendeur.

NETTOYAGE DES DONNÉES

Avant d'analyser les données du projet "Target Brilliance", il est essentiel de faire un traitement préalable pour les nettoyer et les préparer correctement. Cela inclut la gestion des valeurs manquantes, la suppression des doublons, et l'uniformisation des formats de donnée

Ensuite, il faudra établir des relations clés primaires et clés étrangères entre les différentes tables de données. Par exemple, la colonne `customer_id` dans la table `orders` peut être reliée à la même colonne dans `customers`, tandis que `product_id` dans `order_items` peut être relié à `products`. Ces relations sont cruciales pour pouvoir analyser les données de manière cohérente et tirer des conclusions précises à partir des différents jeux de données.



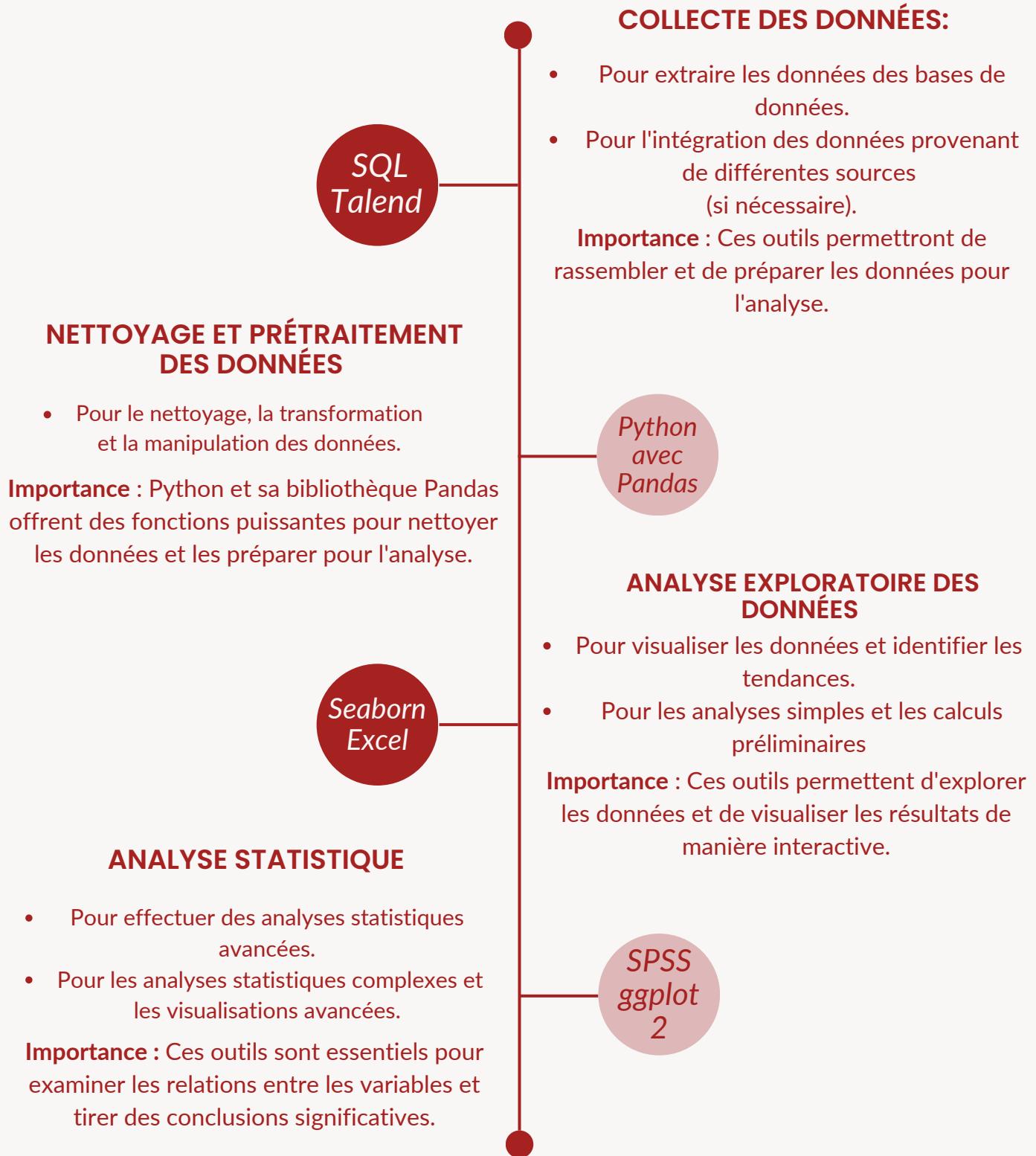
Méthodologie

Utiliser des outils comme SQL et Python pour analyser les données, identifier des tendances et des patterns, et répondre aux principales questions du projet (par exemple, quelles catégories de produits génèrent le plus de ventes, ou comment les ventes varient-elles selon les saisons et les régions).

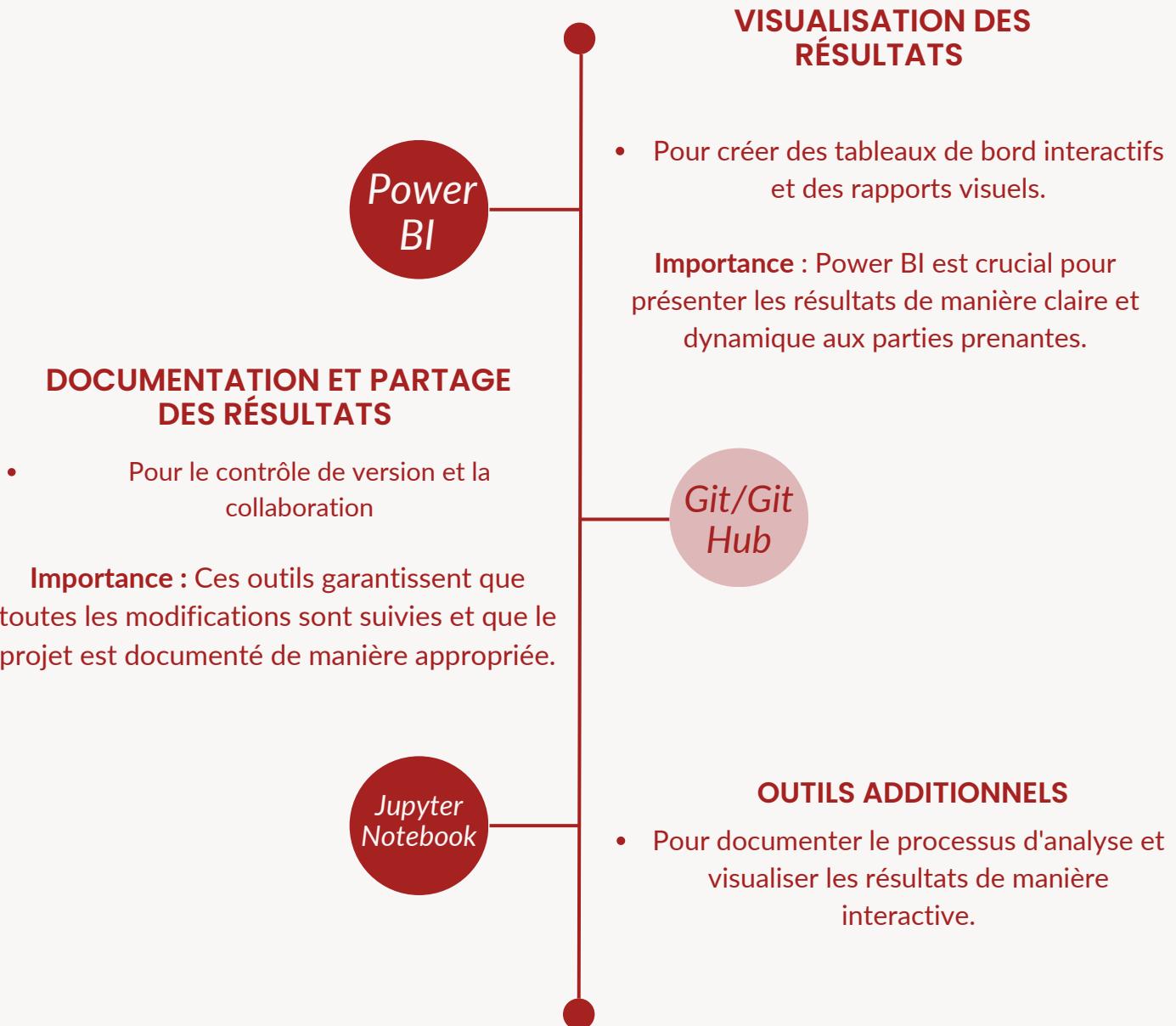
Sources de données : Rassembler les différents jeux de données disponibles (customers.csv, orders.csv, products.csv, etc.) pour avoir une vue globale sur les opérations commerciales de Target au Brésil entre 2016 et 2018

Création des relations entre les tables : Établir les relations clés primaires et clés étrangères entre les tables. Par exemple, relier la table orders.csv avec customers.csv via la clé customer_id et lier order_items.csv à products.csv via product_id. Cela permettra d'analyser les données de manière cohérente et interconnectée.

Méthodologie



Méthodologie



Outil Principal

L'outil central tout au long du projet sera **Python**, principalement avec la bibliothèque Pandas pour la manipulation des données et Seaborn pour la visualisation. Cela nous permettra de mener une analyse complète et approfondie des performances académiques, tout en étant flexible et adaptable aux besoins spécifiques du projet.

Ressources Nécessaires

Pour assurer le bon déroulement du projet « Target Brilliance », plusieurs ressources seront nécessaires, regroupées en différentes catégories :

Ressources Humaines :

- **Analystes de données** : Pour l'extraction, le nettoyage et l'analyse des données.
- **Statisticiens** : Pour effectuer des analyses statistiques avancées et interpréter les résultats.
- **Développeurs** : Pour automatiser les processus d'extraction et de transformation des données.
- **Experts en visualisation** : Pour créer des rapports et des tableaux de bord interactifs.



Ressources Techniques :

- Matériel informatique :
 - Ordinateurs avec des spécifications adéquates (processeur puissant, RAM suffisante) pour exécuter des analyses lourdes.

Logiciels :

- **Systèmes de gestion de base de données (DBMS)** : SQL Server, MySQL ou PostgreSQL.
- **Outils d'analyse et de visualisation** : SPSS, R, Power BI, Talend, Excel.
- **Environnement de développement** : Anaconda pour Python, Jupyter Notebook pour la documentation et la visualisation interactive.
- **Plateforme de gestion de version** : Git et GitHub pour le suivi des modifications et la collaboration.

Ressources Nécessaires

- **Ressources de Données** : Les données locales de Kaggle seront utilisées pour analyser les commandes, clients, produits et paiements.
- **Données Externes** : Des données externes, comme les campagnes marketing ou les tendances du marché, peuvent être ajoutées pour enrichir l'analyse.

Ressources Pédagogiques :

- **Documentation** : Manuels et guides sur les outils utilisés (Python, SPSS, Power BI).
- **Formations** : Sessions de formation pour l'équipe sur l'utilisation des outils d'analyse et de visualisation, ainsi que sur les méthodes d'analyse statistique.

Conclusion

La mise à disposition de ces ressources est essentielle pour garantir le succès du projet « Target Brilliance ». Cela permettra de réaliser une analyse approfondie des performances commerciales et de proposer des recommandations pratiques pour l'optimisation continue des ventes et de la satisfaction client.



Livrables

LES PRINCIPAUX LIVRABLES DU PROJET «TARGET BRILLIANCE» SONT :



Rapport d'analyse des données : Incluant les résultats, les conclusions, et des recommandations pour améliorer les performances académiques.



Tableaux de bord interactifs : Crées avec Power BI et Excel pour visualiser les données et suivre les performances des étudiants et enseignants.



Modèles statistiques et prédictifs : Analyses réalisées dans SPSS et R pour identifier les facteurs clés et prédire les tendances.



Documentation technique : Code commenté, guides d'utilisation, et documentation pour faciliter la maintenance et l'utilisation des outils.



Présentation finale : Résumé visuel des résultats et recommandations, avec une discussion des conclusions.

Planning

VOICI UNE PROPOSITION SIMPLIFIÉE DU PLANNING POUR LE PROJET « TARGET BRILLIANCE »

1

PRÉPARATION (JOUR 1-3)

- Définir les objectifs et créer le cahier des charges.
- Collecter et préparer les données.

2

ANALYSE DES DONNÉES (JOUR 4-10)

- Nettoyage et exploration des données avec Python et Talend.
- Analyse statistique de base avec Excel et SPSS.

3

MODÉLISATION ET VISUALISATION (JOUR 11-18)

- Analyse approfondie avec R et création de modèles statistiques.
- Visualisation des résultats avec Power BI.

4

VALIDATION ET DOCUMENTATION (JOUR 19-22)

- Documenter les analyses et valider les résultats.
- Préparer les livrables finaux et les rapports.

5

PRÉSENTATION FINALE (JOUR 23-25)

- Présenter les conclusions et les recommandations.



Risques

RISQUES POTENTIELS



QUALITÉ DES DONNÉES

Données incomplètes ou erronées

Solution : Nettoyage et validation anticipée.



COMPLEXITÉ DE L'ANALYSE

Analyse complexe

Solution : Utiliser des outils adaptés aux compétences de l'équipe.



RESSOURCES TECHNIQUES

Outils non disponibles

Solution : Prévoir l'accès aux outils ou utiliser des alternatives.



GESTION DU TEMPS

Retards

Solution : Suivi rigoureux avec des jalons clairs.



COMMUNICATION

Mauvaise collaboration

Solution : Utiliser des plateformes de gestion de projet et communication.



DONNÉES HISTORIQUES

Données biaisées

Solution : Analyser la validité des conclusions avec des données alternatives.



CE PROJET **TARGET BRILLIANCE**, RÉALISÉ PAR **AYMEN RAMI**, VISE À ANALYSER LES PERFORMANCES ACADEMIQUES DES ÉTUDIANTS ET DES ENSEIGNANTS POUR IDENTIFIER LES FACTEURS INFLUENÇANT CES PERFORMANCES. L'OBJECTIF PRINCIPAL EST DE PERMETTRE AUX PARTICIPANTS D'APPLIQUER DES COMPÉTENCES PRATIQUES, DE DÉVELOPPER LEURS CAPACITÉS DE COMMUNICATION ET DE RENFORCER LE TRAVAIL EN ÉQUIPE.

JE TIENS À EXPRIMER MES REMERCIEMENTS À TOUS CEUX QUI PARTICIPERONT À LA CONSTRUCTION DE CE PROJET, CAR LEUR CONTRIBUTION EST ESSENTIELLE À NOTRE SUCCÈS COLLECTIF.

Merci.

Theramitkarface