

Résumé du Projet : Analyse des Contenus Netflix

1. Introduction : Comprendre le "Pourquoi"

L'objectif principal de ce projet est de structurer et d'analyser les données des contenus Netflix pour en extraire des tendances clés et des insights sur l'évolution et la diversité de son offre. Pour y parvenir, l'analyse a dû surmonter deux défis majeurs liés à la préparation des données brutes.

- **Normalisation des données** : *Cette étape était cruciale pour éviter les répétitions d'informations et organiser la base de données de manière efficace en reliant différentes tables (films, acteurs, réalisateurs).*
- **Nettoyage des données** : *Cette étape était cruciale pour corriger les incohérences, les valeurs manquantes et les erreurs de formatage afin de garantir la fiabilité des résultats de l'analyse.*

Pour relever ces défis, une méthodologie rigoureuse en plusieurs étapes a été mise en place.

2. La Méthodologie : Comment les Données ont été Préparées

La préparation des données s'est déroulée en trois grandes phases, transformant un simple fichier Excel en une base de données relationnelle propre et prête pour l'analyse.

1. **Importation et Inspection** Le fichier Excel initial (netflix_titles.xlsx) a été importé dans l'outil d'analyse (SAS). Une première inspection a ensuite été menée pour vérifier la structure globale et la qualité des données brutes, permettant d'identifier en amont les éventuels problèmes à corriger.
2. **Normalisation et Structuration** Pour organiser les données de manière logique et efficace, un entrepôt de données relationnel a été créé. L'information, initialement contenue dans une seule table, a été répartie dans plusieurs tables spécialisées et interconnectées, comme illustré ci-dessous :

Type de Table	Rôle et Exemple Concret
Table Principale	Contient les informations centrales sur chaque titre. (Exemple : table titles avec title_id, type, release_year).
Tables de Dimensions	Stockent les informations uniques pour éviter la redondance. (Exemples : directors, actors, genres, countries).
Tables de Liaison	Gèrent les relations "plusieurs-à-plusieurs". (Exemples : title_directors pour lier un film à ses réalisateurs, title_genres pour lier un titre à ses genres).

1. **Nettoyage et Insertion** Des actions de nettoyage clés ont été effectuées, notamment la suppression des espaces inutiles dans les champs de texte et la conversion des dates en un format numérique standard. Les données ainsi nettoyées ont ensuite été insérées dans la nouvelle structure de base de données normalisée, prêtes pour l'analyse.

Cette base de données propre et structurée a permis de mener des analyses révélatrices sur le catalogue de la plateforme.

3. Analyses et Découvertes Clés : Ce que les Données Révèlent

Cette section présente les principaux enseignements tirés de l'analyse des données préparées.

3.1. Le Catalogue Netflix : Volume, Durée et Nouveautés

L'analyse de l'évolution du catalogue montre une augmentation constante du nombre de films et de séries ajoutés au fil des ans, avec une accélération marquée ces dernières années. Cette croissance reflète une stratégie d'investissement massive visant à enrichir continuellement l'offre pour fidéliser les abonnés.

Concernant la durée des films, deux points ressortent :

- **La norme :** La durée moyenne d'un film sur la plateforme est d'environ **99,32 minutes**, ce qui correspond à un standard de l'industrie cinématographique.
- **L'exception :** Le film le plus long est Black Mirror: Bandersnatch (312 minutes). Sa longueur exceptionnelle s'explique par sa nature **interactive**. Cette expérimentation avec le récit interactif est un différentiateur clé sur un marché saturé, conçu pour augmenter l'engagement de l'utilisateur et créer des expériences de visionnage uniques que la télévision linéaire traditionnelle ne peut reproduire.

L'analyse des ajouts récents, comme les 8 nouveaux titres prévus pour le 1er mai 2025, confirme la stratégie de diversification continue du contenu. L'inclusion simultanée de films, de documentaires et de spectacles de stand-up indique une volonté de répondre aux goûts d'un large éventail d'abonnés et de maintenir un flux constant de nouveautés.

3.2. Les Acteurs Clés : Réaliseurs, Acteurs et Pays

L'analyse met en lumière les créateurs et les pays qui façonnent le plus le catalogue Netflix.

La présence de spécialistes du stand-up comme Raúl Campos et Marcus Raboy en tête des réalisateurs les plus prolifiques est un puissant indicateur de la stratégie de contenu de Netflix, qui valorise des productions à haut volume et au coût relativement faible pour satisfaire une audience dédiée.

Réalisateur(s)	Nombre de Productions
Raúl Campos, Jan Suter	324
Marcus Raboy	256
Martin Scorsese	144
Steven Spielberg	100

L'analyse de l'origine des productions révèle une domination écrasante des États-Unis, suivis de très loin par l'Inde. Le Royaume-Uni complète le trio de tête, soulignant la prépondérance du contenu anglophone tout en montrant l'importance stratégique du marché indien.

Enfin, l'analyse des acteurs les plus présents souligne la prédominance des stars de Bollywood, comme **Shah Rukh Khan** (729 titres), aux côtés de figures populaires d'Hollywood comme **Adam Sandler** (361 titres). Ce constat illustre une stratégie de casting à deux volets, visant à la fois le marché indien, massif et en pleine croissance, et le public mondial familier des têtes d'affiche américaines.

La connexion entre ces points de données est éclairante : la prédominance de réalisateurs spécialisés dans le stand-up, un genre très populaire aux États-Unis, est directement liée à la domination de ce

pays dans la production globale. Cela suggère une stratégie d'acquisition de contenu fortement axée sur le marché américain, utilisant des formats efficaces pour générer un volume élevé de programmes.

3.3. La Diversité au Coeur de l'Offre

L'analyse des genres confirme que Netflix propose une offre très variée pour s'adresser à des publics multiples. Les **3 genres les plus populaires** sont les Drames, les Comédies et les Documentaires. Cependant, le catalogue est enrichi par de nombreuses catégories plus spécifiques comme Anime Series, Stand-Up Comedy ou encore LGBTQ Movies. Cette stratégie à deux niveaux permet à la fois de capter un large public avec des genres universels et de fidéliser des communautés de niche avec des contenus spécialisés.

Certains créateurs se distinguent par leur grande polyvalence. Le réalisateur **Justin G. Dyck** a travaillé sur 5 genres différents, tandis que des acteurs comme **Nicolas Cage** ont joué dans 6 genres distincts, démontrant leur capacité à naviguer entre des univers cinématographiques très variés. Pour Netflix, s'associer à des talents aussi flexibles est un atout stratégique, leur permettant de produire une grande variété de contenus sans cesse renouvelés.

Ensemble, ces analyses révèlent une double stratégie : alors que Netflix s'appuie sur des formats standardisés et des créateurs prolifiques pour construire son volume, son véritable avantage concurrentiel réside dans une diversification délibérée et granulaire des genres et des talents, garantissant ainsi qu'il s'adresse à la fois aux marchés de masse et aux publics de niche.

4. Conclusion : Bilan et Prochaines Étapes

4.1. Synthèse des Apports du Projet

Ce projet a permis de dégager trois conclusions fondamentales sur la stratégie de contenu de Netflix :

1. L'analyse confirme **l'expansion continue et l'enrichissement constant** du catalogue Netflix au fil des ans.
2. Bien qu'il existe des standards (comme la durée moyenne des films), Netflix se distingue par **l'exploration de formats innovants et une grande diversité de genres** pour satisfaire tous les publics.
3. Le contenu est porté par des **créateurs et acteurs extrêmement prolifiques**, avec une forte présence de la production américaine et indienne, ainsi que des réalisateurs spécialisés dans des genres populaires comme le Stand-Up Comedy.

4.2. Limites et Perspectives d'Avenir

Malgré la richesse des enseignements, l'étude présente des limites qui ouvrent des pistes pour de futures analyses.

Limitations de l'Étude	Pistes d'Amélioration Futures
Manque de certaines données (ex: durée des épisodes).	Enrichir la base de données avec de nouvelles dimensions (langue, nombre d'épisodes).
Analyse purement quantitative, sans aspects qualitatifs.	Intégrer des données qualitatives comme les évaluations ou les critiques des utilisateurs.

Présentation des résultats sous forme de tableaux bruts.	Créer des visualisations de données (graphiques, cartes) pour rendre les résultats plus accessibles et interactifs.
--	---