

UQAC

Université du Québec à Chicoutimi

Département Informatique et Mathématiques – DIM

8PRO408 – Outils de programmation pour la science des données

Mini-Projet : Analyse des contenus Netflix

Samuel BRASSARD

David NDIMINA

Ornela SAKOUE

Présenté à

Habba's NGODJOU DOUKAGA

Date de remise : 10 déc.2025

[Automne 2025]

[David NDIMINA]
[Samuel BRASSARD]
[Ornela SAKOUE]

[10 déc.2025]

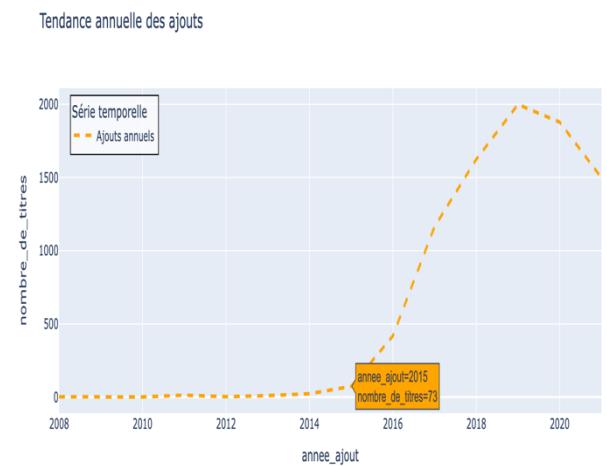
I- ANALYSE GLOBALE :

Dans l'ensemble, l'analyse a d'abord permis de nettoyer et structurer le jeu de données : les types de variables ont été harmonisés, les dates converties dans un format exploitable et les valeurs manquantes traitées, ce qui garantit une base fiable pour les analyses suivantes. Sur cette base, le catalogue a été décrit selon plusieurs angles complémentaires. On observe d'abord une forte présence des films par rapport aux séries, ce qui confirme que l'offre reste majoritairement centrée sur le format "Movie". La répartition géographique met ensuite en évidence une domination claire de certains pays producteurs, en particulier les États-Unis et l'Inde, qui concentrent une grande partie des titres disponibles, tandis que les autres pays sont représentés de manière plus marginale. Du point de vue thématique, l'offre est très concentrée autour de quelques genres majeurs : les drames, les films internationaux et les comédies occupent les premières places et structurent une grande part du catalogue. Enfin, les analyses temporelles montrent que cette offre n'est pas stable dans le temps : le volume de sorties et d'ajouts varie fortement d'une période à l'autre, avec des phases de croissance rapide suivies de ralentissements, ce qui reflète l'évolution de la stratégie de contenu de la plateforme et les contraintes du contexte extérieur.

II- ANALYSE DES TENDANCES :

a- Tendance annuelle d'ajouts :

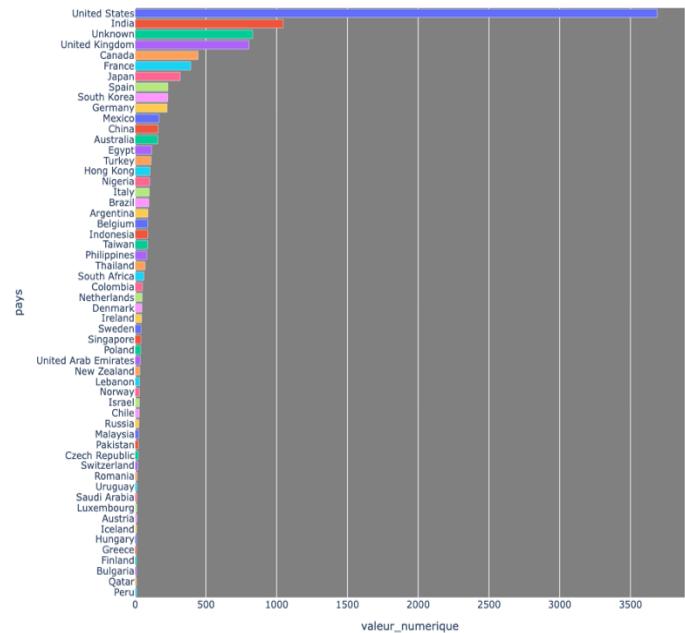
On observe une forte augmentation du nombre de contenus ajoutés à partir de 2015, avec un pic autour de 2019. Après cette date, le volume de nouveaux titres diminue nettement, ce qui pourrait être lié à des facteurs externes comme la pandémie de COVID-19.



b- Pays dominant dans le catalogue :

Cette analyse montre que le pays dominant en nombre de contenus est les États-Unis, suivis de l'Inde. Le graphique présenté peut toutefois être légèrement biaisé : il ne retient que les pays ayant au moins 10 titres. Dans ce cadre, la catégorie « Unknown » apparaît en troisième position, alors que, dans l'analyse textuelle globale, elle occupe la deuxième place. Ce décalage s'explique par le filtrage appliqué, qui exclut une partie des contenus « Unknown ». Cela n'affecte cependant pas la validité générale de notre interprétation.

Répartition géographique des contenus (≥ 10 titres)



c- Genre de contenus dominant :

L'analyse des genres révèle que trois catégories dominent nettement le catalogue : « Drama », « International Movies » et « Comédies ». Ces genres arrivent en tête en nombre de titres, ce qui indique que la plateforme privilégie des contenus narratifs (dramatique ou comique) à forte portée internationale. Cela confirme une stratégie axée sur des histoires accessibles à un large public, souvent issues de cinémas variés à travers le monde.

Répartition des genres

