



PRÉSENTÉ À
HABBA'S NGODJOU DOULAGA
8PRO408 – OUTILS DE PROGRAMATION
POUR LA SCIENCE DES DONNÉES

PAR
CÉRINA AWASHISH
AWAC18510500

MINI PROJET

10 ÉCEMBRE 2025

1. Introduction

Ce rapport propose une étude exploratoire du dataset Netflix, comprenant plus de 8000 titres de films et de séries provenant de divers pays. Le but est de comprendre la composition du catalogue, de déterminer les genres et les pays les plus présents, d'analyser les tendances au fil du temps, ainsi que de représenter graphiquement les changements du contenu que Netflix propose.

2. Méthodologie

L'analyse du dataset a été réalisée dans un notebook Jupyter, en utilisant plusieurs bibliothèques comme Pandas, Matplotlib, Seaborn et Plotly. L'exploration initiale du dataset concernait l'analyse de la distribution des types de contenus, l'étude des genres, y compris la séparation entre plusieurs genres, ainsi que l'exploration des types, des valeurs manquantes et des statistiques descriptives.

Ensuite, l'analyse de chaque colonne du dataset a été effectuée, c'est-à-dire l'analyse des genres, de la provenance des titres, des années de sortie et des dates d'ajout sur la plateforme. Pour finir, des graphiques ont été créés pour une visualisation statique et interactive.

Pendant l'exécution du code pour convertir les dates, une erreur est survenue. Il y a donc fallu convertir les dates pour avoir une analyse au fil du temps plus juste.

3. Principaux résultats

Pour ce qui concerne les types de contenus, l'analyse montre que Netflix suggère plus de films que de séries. Cela indique une tendance plus orientée vers la cinématographie, bien que les séries constituent une part significative du catalogue. Les trois genres qui ressortent le plus sont les drames, les comédies et les documentaires. Ces résultats démontrent que Netflix met plus d'énergie sur les contenus dramatiques et humoristiques, tout en proposant une certaine diversité dans ce qu'elle propose.

La grande majorité des titres proviennent des États-Unis, de l'Inde, et du Royaume-Uni, ce qui démontre que leurs productions, en termes de film ou de série, dominent le catalogue. Toutefois, Netflix offre aussi des contenus provenant d'autres pays, bien que leur nombre soit moins élevé.

Les années de sortie indiquent une forte concentration entre 2010 et 2020, ce qui indique que la plateforme se concentre sur les productions plus récentes. L'étude de la colonne `date_added` démontre qu'après 2015, Netflix a ajouté des titres de plus en plus rapidement, et que cette année correspond à une période de croissance importante pour la plateforme. Cette tendance illustre une approche d'élargissement et de diversification de son catalogue.

4. Conclusion

L'étude révèle une plateforme qui est en constante évolution, qui favorise les contenus plus récents, en forte demande, et qui sont principalement produits dans certains pays dominants. Netflix élargie de plus en plus et diversifie son contenu pour répondre à un large public. Les images parlent d'eux même, donc les visualisations interactives et statistiques ont permis d'approfondir la compréhension du dataset et tout ce qui en découle.





