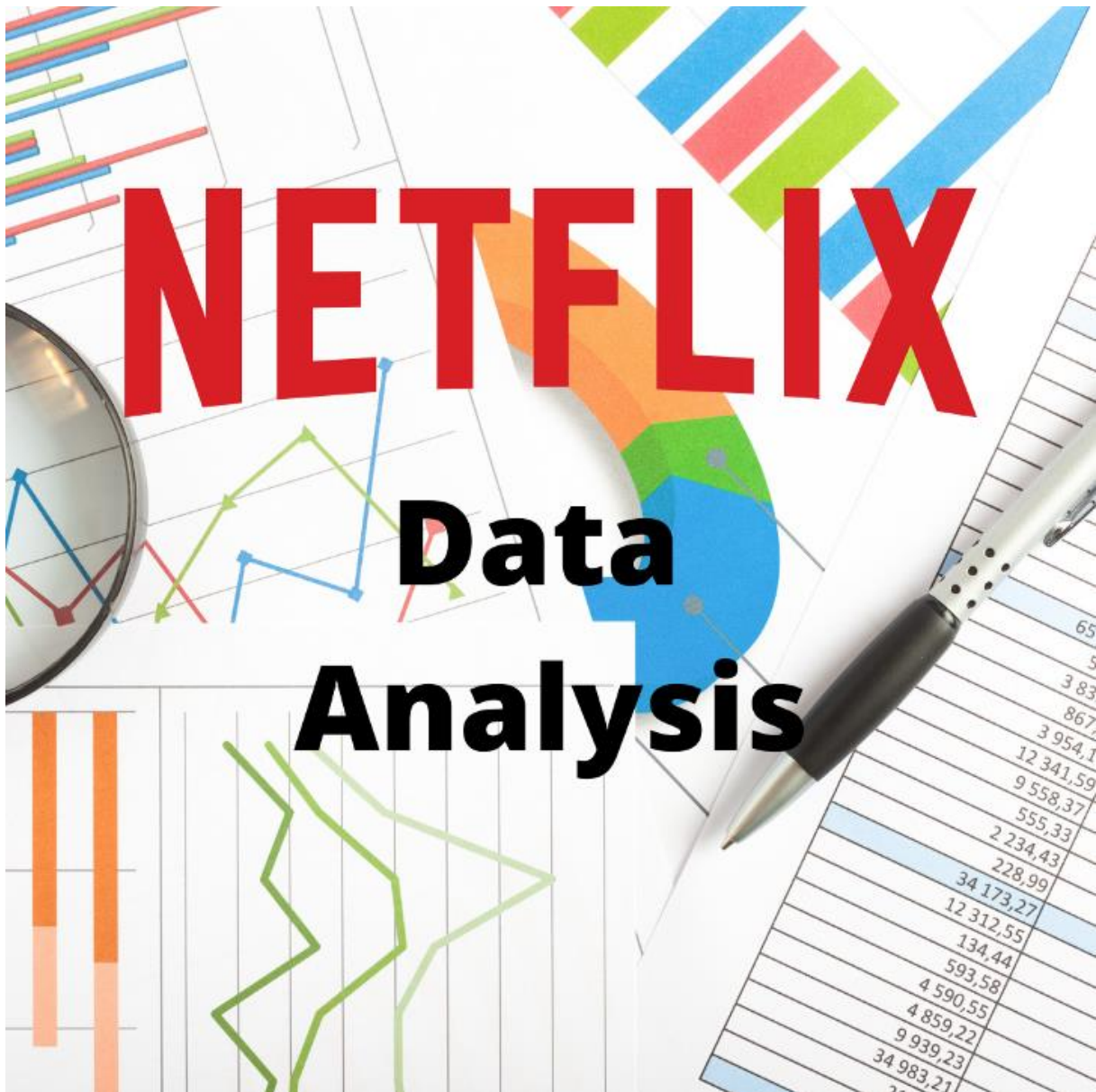


PROJET PYTHON

Moteur de recherche **NETFLIX**



Brief

Contexte

Vous appartenez à l'équipe de développement de la société « NETFLIX ».

En naviguant sur Netflix ou Prime Vidéo, vous devriez avoir rencontré une section qui vous permet de trouver des films et des émissions de télévision en fonction de votre choix. Malheureusement les choix sont limités et vous ne pouvez qu'effectuer une recherche en fonction du genre.

Vous allez devoir travailler pour construire un tout nouveau moteur de recherche. On va vous demander de produire un POC (Proof of concept, ou simplement prouver que votre concept fonctionne), il n'aura pas besoin d'interface graphique mais il devra s'appuyer sur la console Python.

Bien entendu, vous allez devoir gérer ce projet avec une méthodologie précise :

- Découpage des tâches
- Planification
- Présentation du POC
- Rédaction de rapports

Vous trouverez plus bas les spécifications techniques.

Prérequis

Pour ne pas travailler directement sur le catalogue Netflix, vous allez devoir utiliser les données provenant de Kaggle : « The Movies Dataset » ; téléchargez simplement les fichiers « movies_metadata.csv » et « credits.csv », les IDs sont corrélés entre les fichiers. Ces fichiers contiennent une gamme d'informations sur plus de 45k films tels que le nom du film, la classification, le genre, la vue d'ensemble, les acteurs et bien plus encore.

Vous allez devoir comprendre la structure d'un fichier CSV, et l'intégrer dans votre programme en python. (cf. ressources)

Prenez bien le temps de comprendre la structure du fichier, il pourrait vous être utile pour enregistrer des informations.

Cahier des charges

4 parties fonctionnelles pour le programme

I. Fonctionnalités globales

- Le programme est entièrement interfaçable en CLI
 - Le programme est construit de manière hiérarchisée
 - Il est possible de naviguer entre les différentes options, back return, sélection de menu, erreur
 - Des contrôles de données sont effectués lors des entrées utilisateur (ex : si on propose un menu 1 2 3, avoir uniquement ces possibilités)
 - Il ne présente pas de bug majeur (Ex : sortie non prévue du programme)
- Il doit présenter les statistiques globales des données
 - Nombre totale de film
 - Durée totale
 - Ratio acteur / film
 - Pourcentage en fonction du pays
 - *Vous pouvez imaginer d'autres statistiques*

II. Système de recherche

- Nous devons pouvoir effectuer une recherche par :
 - Genre / Durée / Acteur / Langue
 - Limité à 5 réponses
 - Si pas de réponses affichées, on applique un critère en moins et ainsi de suite (message de non-correspondance si rien de trouvé)
- La recherche doit être rapide et ne pas dépasser quelques secondes
- La recherche doit être efficace (ex : au plus près des durées demandées)

Pour aller plus loin, vous pouvez définir des nouveaux critères de recherche comme les genres croisés (exemple : comédie d'horreur)

Système de filtre

Nous pouvons afficher la liste complète des films en prévoyant un défilement à l'aide d'une touche et un choix de filtre à l'aide d'options.

- Une possibilité de filtres devra être proposé
 - Genre
 - Durée
 - Pays
- Plusieurs filtres pourront être actifs simultanément

III. Fonctionnalités utilisateur

- L'utilisateur aura la possibilité d'enregistrer ses recherches
- L'utilisateur pourra consulter ses statistiques
 - Nombre de connexion au moteur de recherches
 - Ses tendances préférés (genre, pays, durée moyenne, acteurs)
- L'utilisateur pourra noter un film de 1 à 5
- L'utilisateur pourra supprimer ses données personnelles

1 partie **bonus** pour le programme

Gestion d'administration du programme :

- Nombre d'utilisateur ayant utilisé le moteur de recherche
- Tendance des recherches
- Suppression des données utilisateurs
- Ajout / Modification / Suppression d'un film

Gestion de projet

Vous allez devoir organiser une gestion de projet qui prévoit :

- Des réunions
- Des prises de décisions
- Une planification
- La rédaction d'un rapport
- La préparation d'une soutenance

Attention il sera important de désigner un chef de projet. Il aura pour rôle de s'assurer que les tâches sont bien distribuées et que l'avancement est assuré.

Tips

- Effectuez une réunion de lancement
 - Recontextualiser le projet
 - Prévoyez un plan d'action
 - Découper et planifier vos tâches
- Prévoyez des réunions régulières pour vérifier l'état d'avancement
- Mettez en place un outil pour vérifier l'état d'avancement du projet (Excel, Trello, Gant)

Livrable

Jeudi 07 décembre 2023

Rapport de projet :

1. Contexte
2. Planification
3. Documentation technique
4. Guide utilisateur
5. Difficultés rencontrées
6. Perspectives d'évolutions
7. Conclusion
8. Bilan de groupe
9. Bilans personnels
10. Glossaire

Vendredi 08 décembre 2023 Jour de la soutenance

Programme en Python (lien GitHub):

- Il doit répondre aux spécifications techniques
- Être codé avec de bonnes pratiques (code clair et commenté)

Soutenance de Projet

- 20 minutes maximum de présentation avec Slides
 - Intro
 - Planification / Répartition des tâches
 - Démonstration
 - Difficultés rencontrées
 - Perspectives d'évolution
 - Conclusion
 - Bilan de groupe
 - **Bilan perso**
- 10 minutes de question

Ressources

Liens Movies Dataset:

[The Movies Dataset \(kaggle.com\)](https://www.kaggle.com/datasets/tommot/the-movies-dataset)

CSV

https://pixees.fr/informatiquelycee/n_site/nsi_prem_csv.html

<https://docs.python.org/fr/3/library/csv.html>

Gestion de projet

<https://www.journaldunet.fr/web-tech/guide-de-l-entreprise-digitale/1443834-scrum-guide-de-la-methode-agile-star/>

<https://www.journaldunet.fr/web-tech/guide-de-l-entreprise-digitale/1443832-kanban-une-methode-agile-de-gestion-de-projet-visuelle-et-continue/>

Un repo git qui regroupe des exercices utiles pour manipuler les données avec pandas !

[GitHub - guipsamora/pandas_exercises: Practice your pandas skills!](https://github.com/guipsamora/pandas_exercises)

Modalités d'évaluations

Passage par groupe de 4

Tous les membres du groupe doivent participer oralement

La soutenance doit être accompagnée d'une présentation visuelle

La soutenance ne doit pas dépasser 20 minutes

La présentation est claire et bien préparée

Les questions ne dépasseront pas 10 minutes

Le jury sera composé du formateur et du responsable pédagogique (il pourra être assisté par une personne extérieure à la formation)

Chaque participant intervient

Le rapport regroupe tous les éléments demandés (il est bien structuré)

La planification montre que chaque membre s'est vu attribuer une tâche

Objectifs visés d'apprentissage

Savoir écrire un code en Python

Savoir structurer son code de manière efficace

Savoir répondre aux spécificités du cahier des charges

Savoir commenter et rendre son code clair

Savoir (en python) :

- Utiliser des structures de contrôle
- Utiliser des boucles
- Utiliser des tableaux et des dictionnaires
- Lire et écrire dans un fichier
- Manipuler des DataFrames