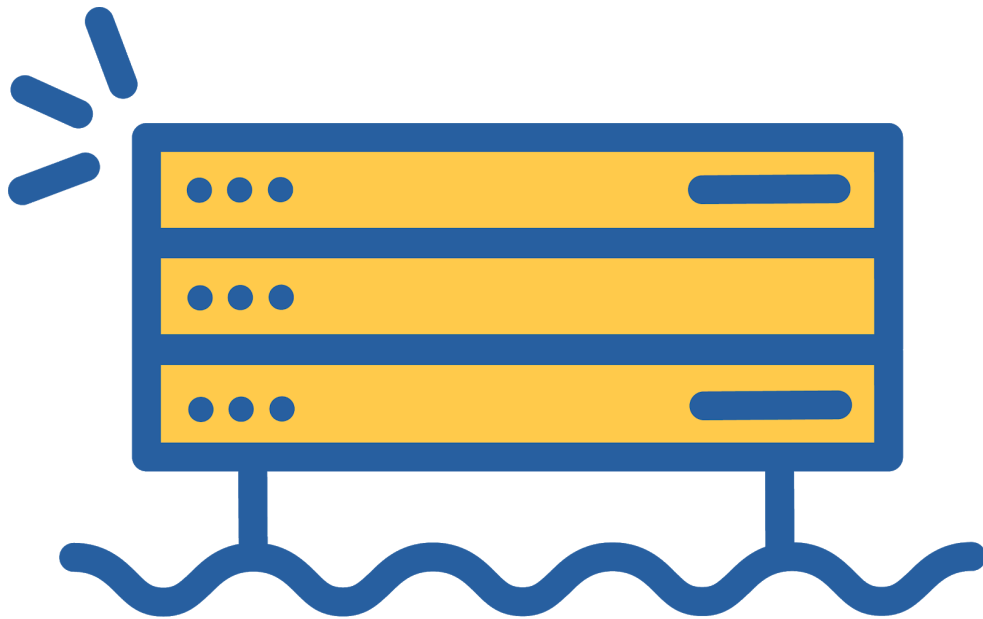

La Plateforme_

Bases de données relationnelle



La Plateforme_

Réaliser une veille sur les différents SGBD. Proposer un exemple concret d'utilisation pour argumenter vos explications.

- **Descriptions des SGBD relationnels classiques (forces des SGBD relationnels classiques et les propriétés ACID, faiblesses des SGBD relationnels classiques)**
- **Quelles sont les spécificités des systèmes NoSQL**
- **Quelles sont les spécificités des systèmes NewSQL**

1. Dataset
2. Récupérer les 2 datasets suivant :

2.1.

<https://www.kaggle.com/shivamb/netflix-shows>

netflix_titles.csv : 12 colonnes (9 string, 1 DateTime, 1 Integer, 1 other)
(show_id, type, title, director, cast, country, date_added, release_year, rating, duration, listed_in, description)

TV Shows and Movies listed on Netflix

This dataset consists of tv shows and movies available on Netflix as of 2019. The dataset is collected from Flixable which is a third-party Netflix search engine.

In 2018, they released an interesting report which shows that the number of TV shows on Netflix has nearly tripled since 2010. The streaming service's number of movies has decreased by more than 2,000 titles since 2010, while its number of TV shows has nearly tripled. It will be interesting to explore what all other insights can be obtained from the same dataset.

Integrating this dataset with other external datasets such as IMDB ratings, rotten tomatoes can also provide many interesting findings.

2.2

<https://www.kaggle.com/chasewillden/netflix-shows?select=Netflix+Shows.csv>

netflix_titles.csv : 7 colonnes (6 String, 1 Integer)

(title, rating, ratingLevel, ratingDescription, release year, user rating score, user rating size)

Context

Netflix in the past 5-10 years has captured a large population of viewers. With more viewers, there most likely an increase of show variety. However, do people understand the distribution of ratings on Netflix shows?

Because of the vast amount of time it would take to gather 1,000 shows one by one, the gathering method took advantage of Netflix's suggestion engine. The suggestion engine recommends shows similar to the selected show. As part of this data set, I took 4 videos from 4 ratings (totaling 16 unique shows), then pulled 53 suggested shows per video. The ratings include: G, PG, TV-14, TV-MA. I chose not to pull from every rating (e.g. TV-G, TV-Y, etc.). Acknowledgements

The data set and the research article can be found at The Concept Center

3. Installer une base de données MySQL ou MariaDB
4. Créer une base de données 'netflix'
5. Créer une table appelée 'netflix_title', importer les données provenant du fichier netflix_titles.csv
6. Créer une table appelée 'netflix_shows' provenant du fichier Netflix Shows.csv Avec la table netflix_titles

7. Afficher tous les titres de films de la table netflix_titles dont l'ID est inférieur strict à 80000000
8. 7. Afficher toutes les durée des TV Show (colonne duration)
9. Réaliser une veille sur ces notions MySQL (<https://sql.sh/fonctions/right>)
 - a. Tri des données
 - b. Renommage
 - c. Agrégation
 - d. Jointures
 - e. Opération
10. Afficher tous les noms de films communs aux 2 tables (netflix_titles et netflix_shows)
11. Calculer la durée totale de tous les films (Movies) de votre table netflix_titles
12. Compter le nombre de TV Shows de votre table 'netflix_shows' dont le 'ratingLevel' est renseigné.
13. Compter les films et TV Shows pour lesquels les noms (title) sont les mêmes sur les 2 tables et dont le 'release year' est supérieur à 2016.
14. Supprimer la colonne 'rating' de votre table 'netflix_shows'
15. Supprimer les 100 dernières lignes de la table 'netflix_shows'
16. Le champ "ratingLevel" pour le TV show "Marvel's Iron Fist" de la table 'netflix_shows' est vide, pouvez-vous ajouter un commentaire ?
17. Réaliser une veille sur le modèle d'analyse et de conception Merise
18. Modéliser (Merise) votre Base de données netflix. Que pouvez-vous dire de cette modélisation ?