

---

# Analyse Éditoriale des Animés

Une Approche Data-Driven (Version Étendue)

Science des Données & Curation

Environnement Python pour une Recommandation

"Sûre"



# Écosystème Technique

## Pandas

Manipulation et nettoyage des données tabulaires complexes.

## NumPy

Calculs numériques performants pour nos métriques propriétaires.

## Matplotlib

Visualisation des tendances et des distributions statistiques.



# Analyse Exploratoire (EDA)

## Biais de Notation

L'EDA a révélé que les séries longues affichent des notes plus fluctuantes, tandis que les formats courts maintiennent une stabilité supérieure.

## Corrélation Format/Stabilité

Identification des variables clés corrélées à la satisfaction spectateur pour poser les bases de notre segmentation.

# Nettoyage des Données

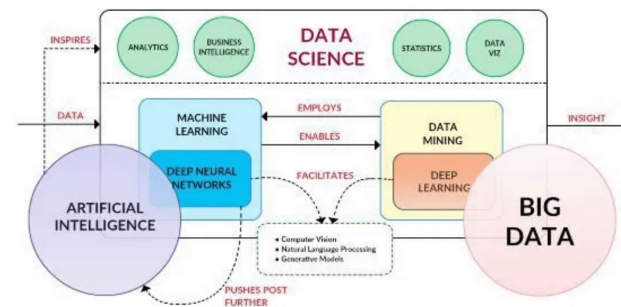
Le nettoyage ne s'est pas limité à la suppression des doublons. Nous avons développé une logique pour isoler le cœur narratif des œuvres.

## Isolation des "Fillers"

Identification des épisodes non-canoniques qui altèrent artificiellement la perception globale de la qualité.

## Traitement des "Recaps"

Suppression du bruit statistique généré par les épisodes récapitulatifs, souvent moins bien notés par le public.



# Le Concept de Régularité

## Définition Stratégique

La régularité est la clé de voûte de notre analyse. Elle mesure la capacité d'une production à maintenir un niveau d'excellence constant tout au long de sa diffusion.

Équation de Fiabilité

$$\text{RÉGULARITÉ} = 10 - (\text{MAX\_NOTE} - \text{MIN\_NOTE})$$

Une série magistrale avec un épisode catastrophique est jugée moins "sûre" qu'une série simplement excellente mais constante. Cette approche permet de quantifier la fiabilité d'un studio à maintenir son niveau de production sans décevoir le spectateur.

# Le Score de Qualité 10

Nous avons synthétisé nos découvertes dans une métrique propriétaire. Ce score composite est calculé selon une pondération précise pour valoriser l'excellence constante.

Facteur d'Analyse	Pondération	Objectif Éditorial
Note Globale	60%	Capture l'appréciation historique et l'impact public global.
Régularité	30%	Mesure la stabilité de l'expérience de visionnage par épisode.
Volume	10%	Valorise les œuvres capables de durer sans faiblir qualitativement.
OBJECTIF : HIÉRARCHISATION OBJECTIVE DE L'EXCELLENCE		



# Analyse des Studios

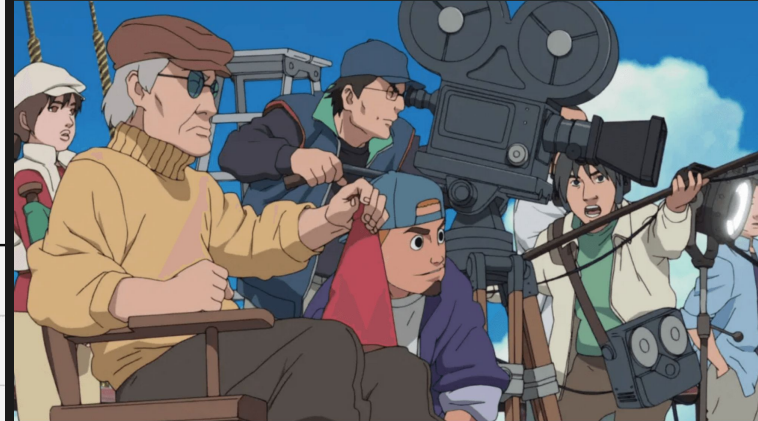
L'analyse met en lumière des disparités notables. Certains studios affichent une régularité impressionnante, tandis que d'autres montrent une plus grande instabilité liée aux cycles de production.

Studio	Régularité Moy.	Score Qualité
Madhouse	8.2 / 10	Élevé
Bones	7.9 / 10	Constant
Sunrise	7.5 / 10	Variable
Studio Pierrot	6.8 / 10	Inégal

# Distribution et Impact des Genres

L'analyse révèle que la popularité massive d'un genre ne garantit pas sa stabilité qualitative. Certains genres de niche offrent une expérience plus constante.

Genre	Régularité Moyenne	Observation Éditoriale
Slice of Life	8.2 / 10	Stabilité exceptionnelle, peu de "creux".
Seinen	7.8 / 10	Qualité élevée, public ciblé et fidèle.
Shonen	6.5 / 10	Forte popularité mais irrégularité marquée.





# Impact du Format

## Séries TV

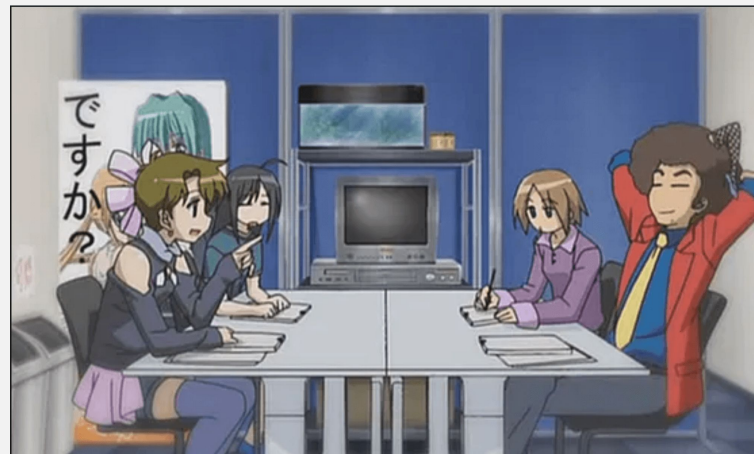
Soumises à des rythmes hebdomadaires intenses. Favorisent l'apparition de "creux" qualitatifs et d'épisodes de transition (fillers).

## Films de Cinéma

Bénéficient de budgets et de temps de production confortables. Scores de régularité frôlant souvent la perfection.

## OAV / Formats Courts

Compromis idéal offrant une liberté créative supérieure et une stabilité technique exemplaire sur une durée limitée.



**Observation :** Le rythme de production est le premier facteur d'instabilité. Plus le cycle est court, plus le risque de déviation qualitative augmente.

# Top 10 : L'Élite de la Régularité

Rang	Animé	Score Qualité	Note Globale	Régularité	Épisodes
01	Frieren: Beyond Journey's End	8.50	9.3	7.9	28
02	Steins;Gate	8.21	9.1	8.1	24
03	Monster	7.97	8.9	8.3	74
04	Vinland Saga	7.85	8.8	8.5	48
05	Hajime no Ippo	7.79	8.8	8.3	76
06	Les Carnets de l'apothicaire	7.79	8.8	8.4	24
07	Mushishi	7.69	8.7	8.6	26
08	Fullmetal Alchemist: Brotherhood	7.68	9.1	6.7	64
09	Your Lie in April	7.62	8.7	8.3	22
10	Cowboy Bebop	7.55	8.8	7.5	26

## Analyse

Ce classement valide notre approche : l'excellence éditoriale est indissociable de la constance. Les œuvres en tête maintiennent une qualité élevée sans déception majeure pour le spectateur.

# Segmentation Stratégique

---

Score  $\geq 8$

## Chef-d'œuvre

L'excellence absolue et constante. Ces œuvres représentent le sommet qualitatif du catalogue et garantissent une satisfaction totale.

Score  $\geq 6$

## Très bon mais inégal

Une qualité globale élevée malgré quelques passages plus faibles. Recommandé pour les spectateurs avertis acceptant quelques creux.

Score  $\geq 4$

## Culte mais risqué

Des œuvres marquantes qui peuvent diviser ou lasser. Nécessitent une curation humaine pour cibler le bon public.

## À éviter

Score  $< 4$

Contenus dont la qualité globale ou la régularité est insuffisante pour une recommandation standard.

# Similarité Thématique

Notre algorithme utilise l'**Indice de Jaccard** pour calculer la proximité entre deux œuvres.

En comparant les ensembles de tags de genres, nous identifions les séries qui partagent une "âme" commune.

## Indice de Jaccard

$$J(A, B) = |A \cap B| / |A \cup B|$$

Intersection des tags divisée par leur union.

# Filtre de Sécurité

## "Anti-Déception"

Le cœur de la recommandation "sûre" est son triple verrou de sécurité. Pour être recommandée, une œuvre doit impérativement valider trois critères de fiabilité.

### Indice de Régularité

SEUIL  $\geq 7.0$  / 10

### Note Plancher (Épisode)

AUCUN ÉPISODE  $< 6.0$  / 10

### Validation Segment

"CHEF-D'ŒUVRE" OU "TRÈS BON"

# Résultats et Validation

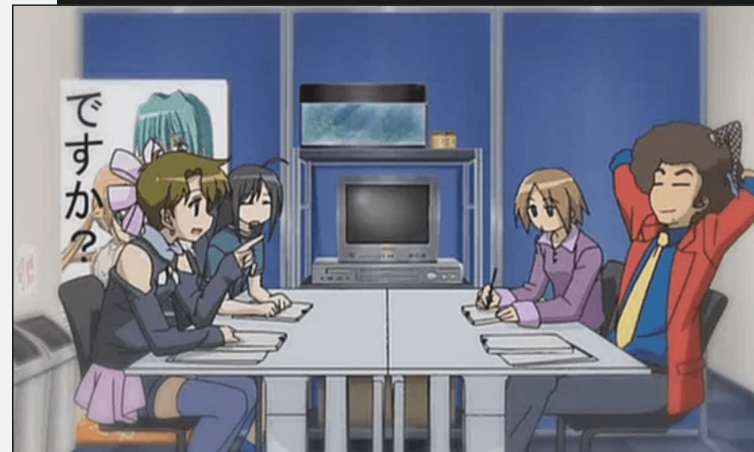
---

## Rétention Spectateur

Les tests montrent une corrélation forte entre nos recommandations et la rétention. En évitant les séries "instables", nous réduisons drastiquement le taux d'abandon.

## Pépites Cachées

Le système parvient à dénicher des œuvres qui, malgré une popularité moindre, offrent une qualité constante exemplaire, enrichissant ainsi le catalogue proposé.



# Conclusion et Perspectives

---

## Valeur Ajoutée

Ce projet prouve que la data science peut enrichir le travail éditorial sans le remplacer. En remplaçant l'intuition par des métriques de régularité, nous offrons un outil précieux pour fidéliser les audiences.

## Évolutions Futures

L'avenir réside dans l'intégration de l'analyse de sentiments sur les commentaires textuels pour capturer l'émotion des spectateurs et affiner la précision des recommandations.

