

# Template feedback Hackathon#1

Notions évaluées : Extraction, traitement et consolidation de données

---

## Feedback – Hackathon

**Nom des élèves :** Adil, Sabri

**Nom du projet :** The Ultimate Streaming Guide

**Nom du professeur évaluateur :** Clara Martinez

**Date :** 03/07/25

---

### Feedback global (2 phrases maximum)

#### Points positifs :

- **Excellente méthodologie et structuration :** Le projet est très bien organisé, suivant un workflow logique et facile à suivre (Chargement → Nettoyage → Analyse). L'utilisation de titres et de commentaires en Markdown rend le notebook particulièrement clair et professionnel.
- **Initiative et enrichissement des données :** La démarche de combiner deux datasets et surtout d'aller chercher des données manquantes (les genres) via scraping d'API (TheMovieDB) est une excellente initiative. Cela démontre une grande proactivité et une très bonne compréhension des enjeux d'un projet de data science.
- **Analyse statistique pertinente :** Le groupe ne s'est pas contenté de visualisations simples. L'identification d'un biais potentiel (la variabilité des notes des films anciens due au faible volume) et l'utilisation d'un test statistique comme l'ANOVA pour valider une hypothèse sur les décennies montrent une réelle profondeur d'analyse.
- **Qualité du code et interactivité :** Le code est propre et lisible. L'utilisation de widgets interactifs (**ipywidgets**) pour explorer la distribution des genres est une touche technique très appréciée qui améliore l'exploration.

#### Points à améliorer :

- **Approfondir l'analyse croisée :** L'analyse répond bien à des questions distinctes, mais pourrait gagner en profondeur en croisant davantage les axes. Par exemple : "Les genres les mieux notés (Drame, Comédie) sont-ils les mêmes sur toutes les plateformes ?" ou "La baisse des notes post-2010 affecte-t-elle plus les films ou les séries ?".
  - **Optimisation des visualisations :** La matrice de corrélation est très dense et difficile à lire. Une bonne pratique serait de la filtrer pour n'afficher que les corrélations supérieures à un certain seuil (ex:  $> 0.3$ ) afin de faire ressortir les relations les plus pertinentes. De même, pour le graphique sur l'évolution des notes, une moyenne mobile (**rolling mean**) aurait pu lisser la courbe et rendre la tendance plus visible.
- 

### Évaluation notée (notes sur 5)

(Ces notes servent à départager les groupes pour la sélection du meilleur projet)

Critère	Note /5	Commentaire
Clarté du code et commentaires pertinents	[ 5 ]	Code bien segmenté, clair, lisible et ordonné.
Utilisation efficace des notions vues pendant le bootcamp (Pandas, visualisation, EDA, etc.)	[ 3 ]	Les outils sont utilisés, mais l'analyse exploratoire pourrait être plus poussée pour générer des insights plus profonds.
Résultats corrects et visualisations cohérentes	[ 5 ]	Les graphiques répondent directement aux questions posées et les calculs semblent corrects.
Structuration du notebook / workflow clair et reproductible	[ 5 ]	Le notebook est exemplaire en termes de structure, de la lecture des données à l'analyse finale.
Pertinence de l'analyse exploratoire (EDA) et capacité à tirer des insights	[ 4 ]	Très bonne base d'EDA avec de bonnes questions. L'identification d'un biais est un vrai plus. Pour un 5, il aurait fallu croiser davantage les analyses pour des insights plus complexes.
Capacité à répondre aux questions posées via des analyses/statistiques simples (NumPy/SciPy)	[ 5 ]	Excellente utilisation de SciPy pour un test ANOVA, ce qui va au-delà d'une simple observation et démontre une rigueur statistique.

---

#### ❏ Commentaire final (optionnel)

→ Excellent travail d'équipe. Le projet est très solide, non seulement dans son exécution technique mais aussi dans sa démarche méthodologique. L'initiative d'enrichir les données

via une API est particulièrement remarquable et témoigne d'une grande maturité. Le groupe a démontré une très bonne maîtrise des concepts du bootcamp, en allant même jusqu'à utiliser des tests statistiques pour valider leurs hypothèses. C'est une base de très haute qualité qui, avec quelques analyses croisées supplémentaires, pourrait devenir un projet d'analyse de portfolio. Félicitations à l'équipe.