



Data Analysis : La vérité sur les notes d'animés?

Création d'un algorithme de score basé sur la régularité des épisodes.

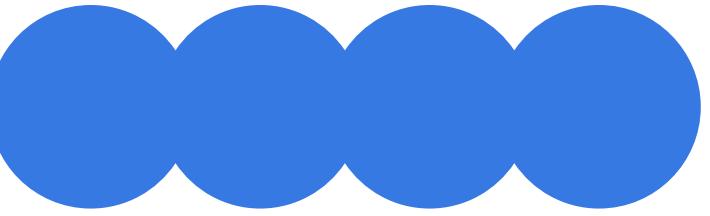


Présenté par Aaron
Boti



Le Problème

- Sur les sites de streaming, on se fie à la note globale (ex: 8.5/10).
- Problème : Cette note masque la vérité. Une série peut avoir 8.5 en ayant des épisodes géniaux et des épisodes catastrophiques.
- Objectif : Détecter les séries "instables" pour ne recommander que des valeurs sûres.



Outils & Environnement utilisés



Python : Langage principal.



Jupyter Notebook : Pour le développement itératif.



Pandas : Pour la manipulation du Dataset.



Matplotlib : Pour la Data Visualization



GitHub : Pour le versioning.



Les données de départ

- Source : Dataset anime.csv.
- Problèmes rencontrés : Valeurs manquantes (NaN), formats de dates incorrects, colonnes inutiles.
- Défi : Rendre ces données lisibles par l'ordinateur.

```
Anime,Genre_Tags,Source,Nb_Episodes,Status,Studio,Date_Pub>Note_Globale,Note_Meilleur_Ep,Meilleur_Ep_
" Naruto ", " Action / Shonen ", " Manga ", 220, " Fini ", " Studio Pierrot ", " 2002-10-03 ", 8.4, 9.6, " Ep 133: Le cri de la bête "
" Trigun ", " Sci-Fi / Action ", " Manga ", 26, " Fini ", " Madhouse ", " 1998-04-01 ", 8.2, 9.1, " Ep 24: Le péché (Sin) "
" Dragon Ball Z ", " Action / Shonen ", " Manga ", 291, " Fini ", " Toei Animation ", " 1989-04-26 ", 8.7, 9.4, " Ep 237: La mort de Freeza "
" Fullmetal Alchemist: Brotherhood ", " Aventure / Military ", " Manga ", 64, " Fini ", " Bones ", " 2009-04-05 ", 9.1, 9.5, " Ep 15: La guerre des corps "
" Cowboy Bebop ", " Sci-Fi / Noir ", " Original ", 26, " Fini ", " Sunrise ", " 1998-04-03 ", 8.8, 9.7, " Ep 26: The Real Flame "
" Sword Art Online ", " Isekai / Romance ", " Light Novel ", 25, " Fini ", " A-1 Pictures ", " 2012-07-08 ", 7.2, 8.9, " Ep 15: L'Amour au combat "
" Re:Zero ", " Isekai / Psychologique ", " Light Novel ", 50, " Fini ", " White Fox ", " 2016-04-04 ", 8.2, 9.8, " Ep 15: Le destin de la dernière personne "
" Dr. Stone ", " Sci-Fi / Adventure ", " Manga ", 57, " En cours ", " TMS Entertainment ", " 2019-07-05 ", 8.3, 9.1, " Ep 2: La révolution "
" One Piece ", " Adventure / Shonen ", " Manga ", 1090, " En cours ", " Toei Animation ", " 1999-10-20 ", 8.7, 9.8, " Ep 10: La guerre mondiale "
" Blue Lock ", " Sport / Thriller ", " Manga ", 36, " En cours ", " 8bit ", " 2022-10-09 ", 8.2, 9.3, " Ep 11: La dernière partie "
" Haikyuu!! ", " Sport / Volley ", " Manga ", 85, " Fini ", " Production I.G ", " 2014-04-06 ", 8.7, 9.9, " Ep 10 (S3): La finale "
" One Punch Man ", " Action / Comedy ", " Web Manga ", 24, " En cours ", " Madhouse ", " 2015-10-05 ", 8.5, 9.6, " Ep 12: Le combat "
" Solo Leveling ", " Action / Fantasy ", " Manhwa ", 12, " En cours ", " A-1 Pictures ", " 2024-01-07 ", 8.3, 9.5, " Ep 12: Le défi "
" Les Carnets de l'apothicaire ", " Mystery / Historical ", " Light Novel ", 24, " En cours ", " OLM ", " 2023-10-22 ", 8.0, 8.2, " Ep 1: Coach et mentor "
" Medalist ", " Sport / Figure Skating ", " Manga ", 3, " En cours ", " ENGI ", " 2025-01-05 ", 8.0, 8.2, " Ep 1: Coach et mentor "
" Du mouvement de la Terre ", " Historical / Drama ", " Manga ", 8, " Fini ", " Madhouse ", " 2024-10-05 ", 8.4, 8.9, " Ep 1: Le mouvement de la Terre "
" Samurai Champloo ", " Adventure / Chanbara ", " Original ", 26, " Fini ", " Manglobe ", " 2004-05-20 ", 8.5, 9.2, " Ep 26: La mort de Jinsei "
" L'Attaque des Titans ", " Action / Mystery ", " Manga ", 89, " Fini ", " Wit Studio / MAPPA ", " 2013-04-07 ", 8.5, 9.9, " Ep 25: Silence "
" Death Note ", " Thriller / Supernatural ", " Manga ", 37, " Fini ", " Madhouse ", " 2006-10-04 ", 8.6, 9.5, " Ep 25: Silence "
" Hunter x Hunter (2011) ", " Adventure / Shonen ", " Manga ", 148, " Fini ", " Madhouse ", " 2011-10-02 ", 9.0, 9.7, " Ep 50: Requiem "
" Code Geass ", " Sci-Fi / Mecha ", " Original ", 50, " Fini ", " Sunrise ", " 2006-10-06 ", 8.7, 9.8, " Ep 50: Requiem "
" Steins;Gate ", " Sci-Fi / Thriller ", " Visual Novel ", 24, " Fini ", " White Fox ", " 2011-04-06 ", 9.1, 9.7, " Ep 22: L'Échange "
" GTO ", " Comedy / Slice of Life ", " Manga ", 43, " Fini ", " Studio Pierrot ", " 1999-06-30 ", 8.7, 9.2, " Ep 43: La dernière partie "
" Naruto Shippuden ", " Action / Shonen ", " Manga ", 500, " Fini ", " Studio Pierrot ", " 2007-02-15 ", 8.2, 9.7, " Ep 476: La mort de Pain "
" Psycho-Pass ", " Sci-Fi / Police ", " Original ", 41, " Fini ", " Production I.G ", " 2012-10-12 ", 8.3, 9.2, " Ep 11: Saï "
" My Hero Academia ", " Action / Super Power ", " Manga ", 138, " En cours ", " Bones ", " 2016-04-03 ", 7.9, 9.6, " Ep 49: La mort de All For One "
" Bleach ", " Action / Supernatural ", " Manga ", 366, " Fini ", " Studio Pierrot ", " 2004-10-05 ", 7.9, 9.7, " Ep 366 (Tome 65) "
" Parasite: La Maxime ", " Sci-Fi / Horror ", " Manga ", 24, " Fini ", " Madhouse ", " 2014-10-09 ", 8.3, 9.3, " Ep 12: Coévolué "
" Demon Slayer ", " Action / Historical ", " Manga ", 55, " En cours ", " Ufotable ", " 2019-04-06 ", 8.5, 9.9, " Ep 19: Le sacrifice "
" Erased ", " Mystery / Supernatural ", " Manga ", 12, " Fini ", " A-1 Pictures ", " 2016-01-08 ", 8.3, 9.4, " Ep 8: Spiral "
" Tokyo Ghoul ", " Horror / Action ", " Manga ", 48, " Fini ", " Studio Pierrot ", " 2014-07-04 ", 7.5, 9.6, " Ep 12: Ghoul "
" Mushishi ", " Slice of Life / Supernatural ", " Manga ", 26, " Fini ", " Artland ", " 2005-10-23 ", 8.7, 9.4, " Ep 12: Orphelin "
" JoJo's Bizarre Adventure ", " Action / Supernatural ", " Manga ", 26, " Fini ", " David Production ", " 2012-10-06 ", 8.7, 9.4, " Ep 12: Orphelin "
```

Nettoyage des données avec Pandas

```
df_clean['Note_Pire_Ep'] = df_clean['Note_Pire_Ep'].fillna(0)
df_clean['Nb_Episodes'] = pd.to_numeric(df_clean['Nb_Episodes'], errors='coerce').fillna(12)
```

- Traitement des valeurs nulles (remplacement par la médiane ou 0).
- Conversion des types (Textes → Nombres).
- Création d'un DataFrame propre : `animes_clean.csv`.

Création du "Score de Régularité"

C'est le cœur de l'algorithme.

Formule:

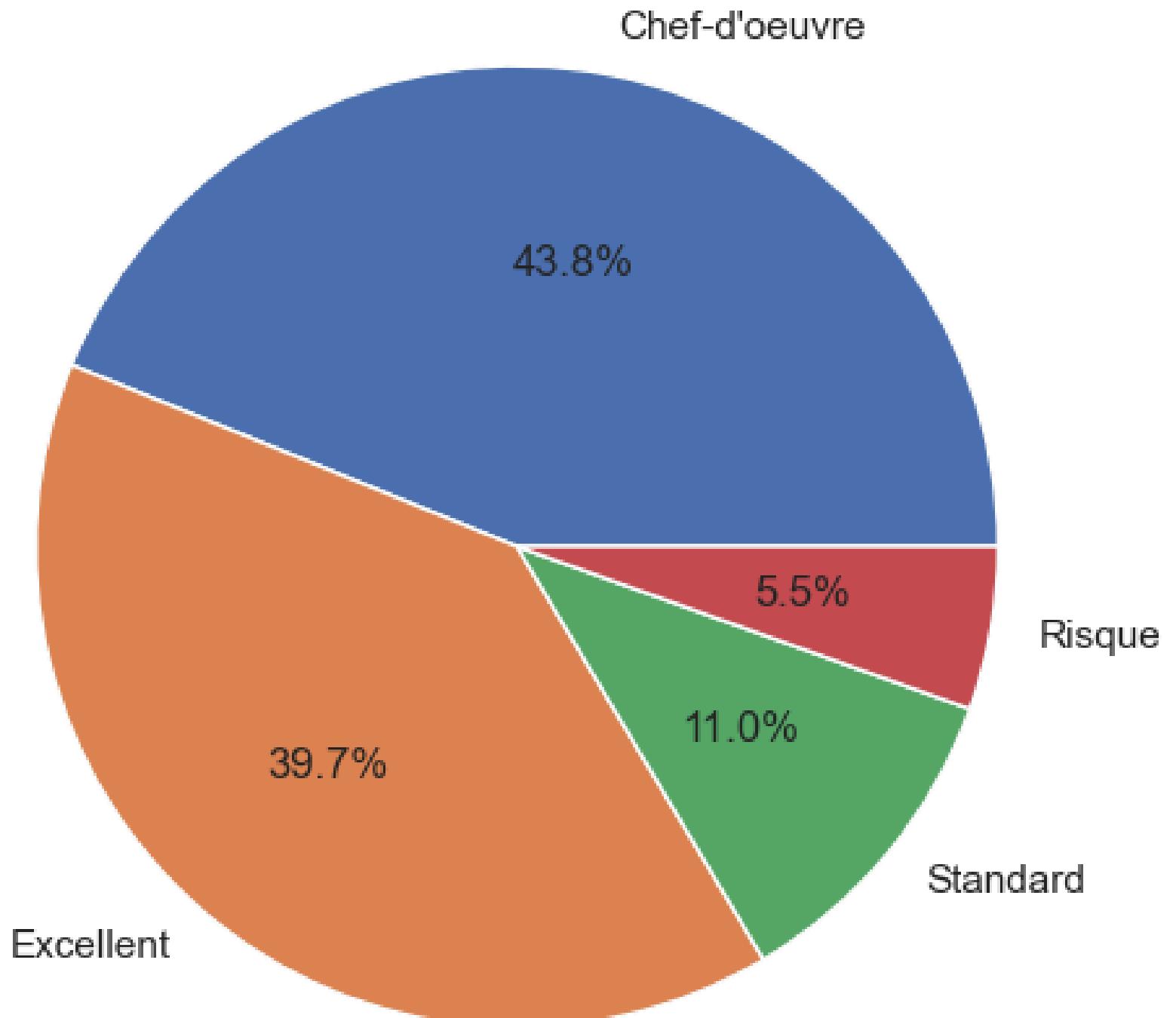
$$\text{Régularité} = 10 - (\text{Meilleur Episode} - \text{Pire Episode})$$

Si l'écart est grand → Pénalité (Série risquée).

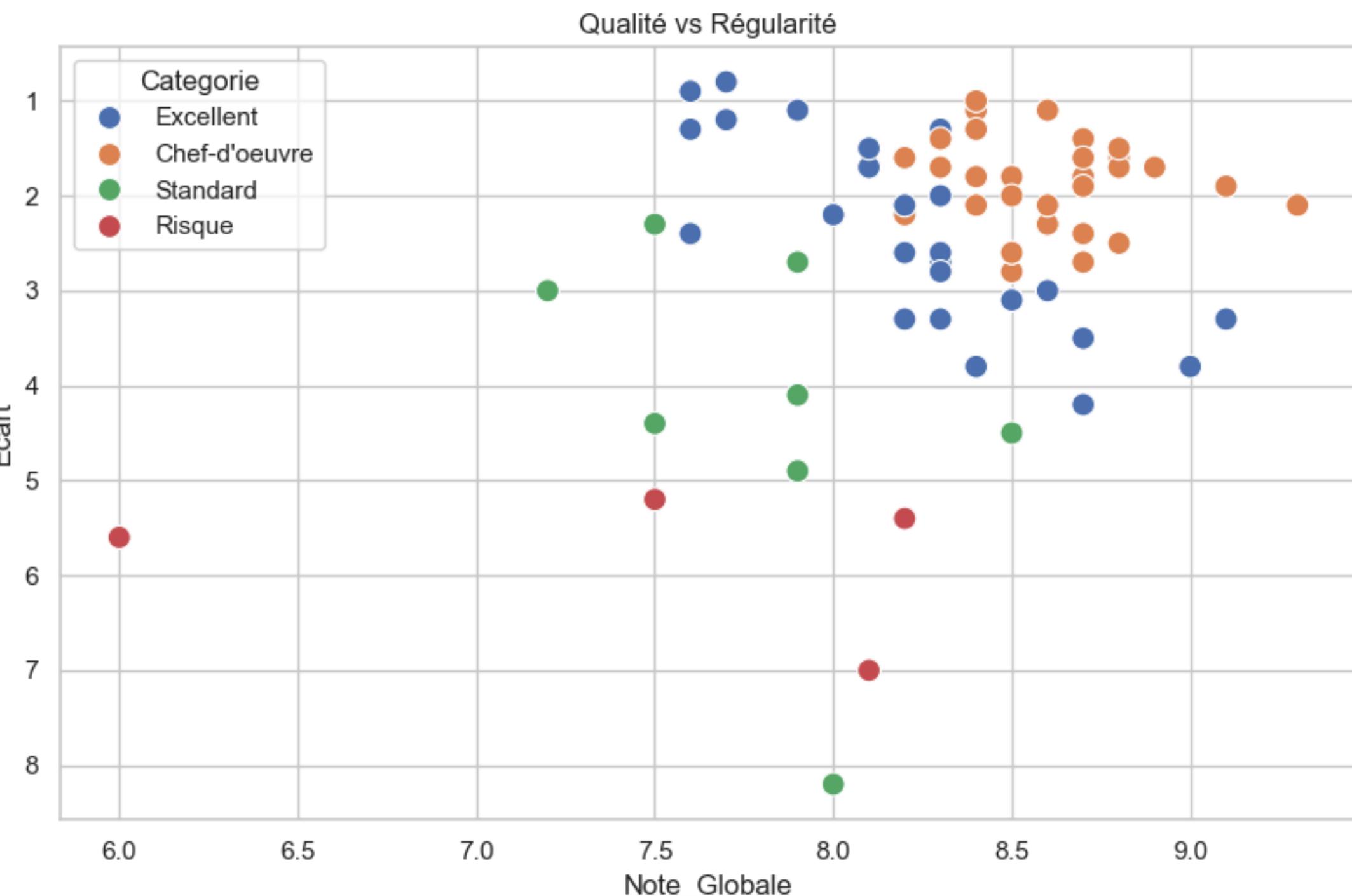
Si l'écart est petit → Bonus (Série stable).

Vue d'ensemble du marché

- Les genres Action et Shonen sont les plus représentés.
- Il y a très peu de "Chefs-d'œuvre" (catégorie verte), la majorité des séries sont "Standards".



Mapping: Qualité vs Régularité



Axe X: Note Globale (La popularité)

En rouge : Les séries "risquées" (instables)

Axe Y: Régularité (La stabilité)

En vert : Les "Chefs-d'œuvre" (stables)

Les résultats de l'algorithme

Rang	Animé	Score	Mention
1	Steins;Gate	8.98	Chef-d'œuvre
2	Your Lie in April	8.96	Chef-d'œuvre
3	Frieren	8.91	Chef-d'œuvre

Affiche le Top 3 ou Top 5 des animés que ton code
a trouvés (ceux avec le meilleur Score Final).

Ce que j'ai appris

1. Importance du Data Cleaning (80% du travail).
2. Puissance de Python pour automatiser les analyses.
3. Importance de la Visualisation pour prendre des décisions.

"Merci de votre lecture ! Code complet dispo sur ce repo."