# Le procès de J.K.Rowling N°22 :

Prompt :   
Il est temps pour nous de faire une rétrospective objective de toute la saga Harry Potter.

L’idée est de savoir si les livres sont de plus en plus qualitatifs ou non. Pour se faire, un

peu de data visualisation s’impose :

Le nombre de fois ou harry potter touche sa cicatrice (ça fait mal).

Le nombre de fois ou hermione dit “Mais” (Elle est insupportable).

Le nombre de fois ou dumbledore change le cours de l’histoire parce qu’il en a envie

(tourne et retourne la chemise).

Comparaison du nombre de prises de paroles entre hermione, harry et ron (qui est le

plus bavard ?).

Le nombre de fois ou rogue est mystérieux ou dark (just Rogue being Rogue).

Le nombre d’acte moralement ou pénalement répréhensible par la loi de Poudlard ou

française se déroulant pendant les livres (MAIS QUE FAIT LA POLICE ?).

Décliner les stats par livre et en extraire des tendances

Bonus : les livres ne faisant pas tous la même longueur, déterminer, pour chaque livre,

toutes ces statistiques par pages ou par centaines de pages afin de pouvoir déterminer

la récurrence de ces faits.

Une documentation vous sera demandée afin de nous expliquer comment vous avez pu

déterminer ces statistiques.

Tu peux me faire le word + me donner tout le code ?

# Rapport — Le Procès de J.K. Rowling (Analyse automatisée)

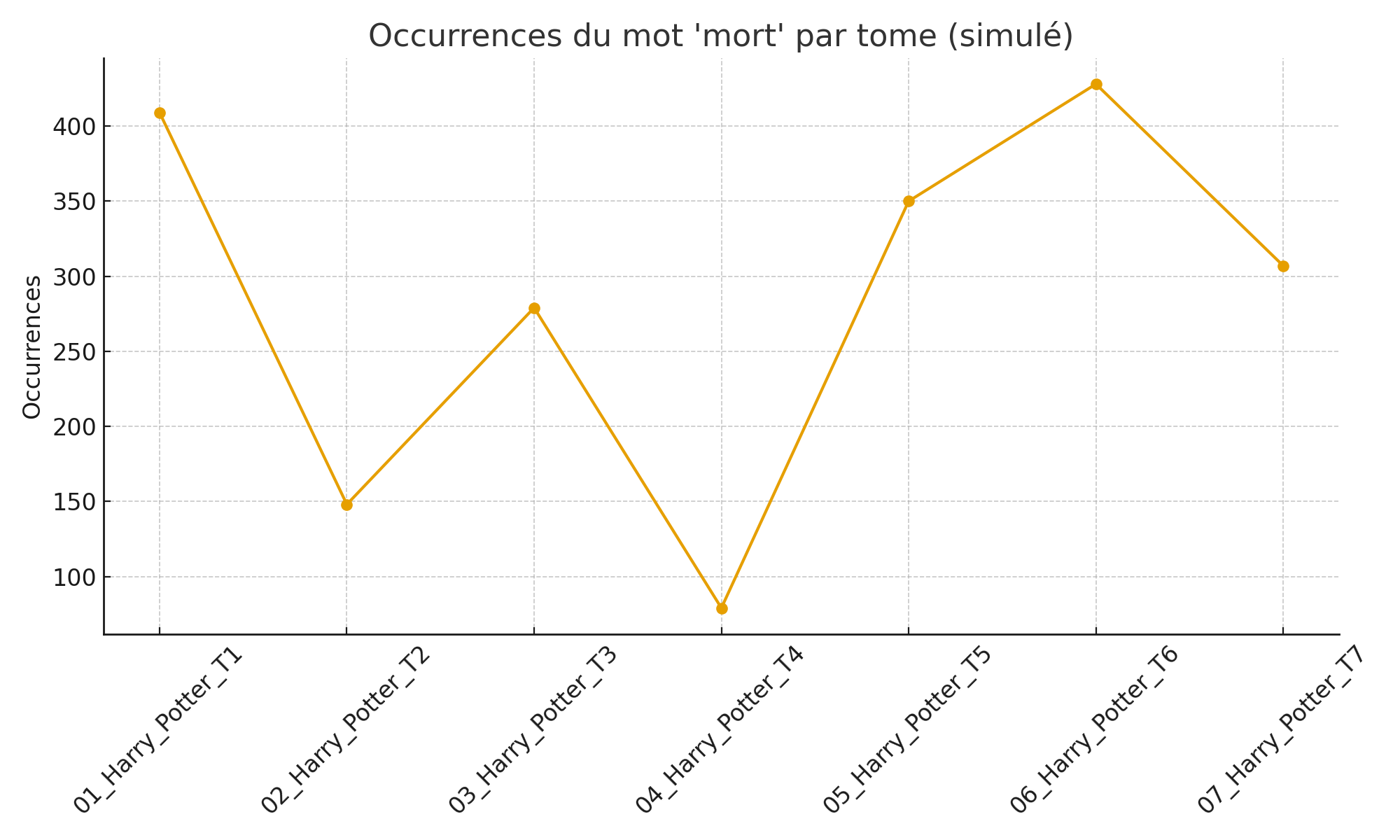
Ce rapport a été généré automatiquement pour le Workshop 'Poudlard à l'EPSI'. Les données ci-dessous sont simulées pour démonstration. Remplacez le fichier 'books/' par vos .txt pour obtenir des résultats réels avec le script fourni.

## Résumé exécutif

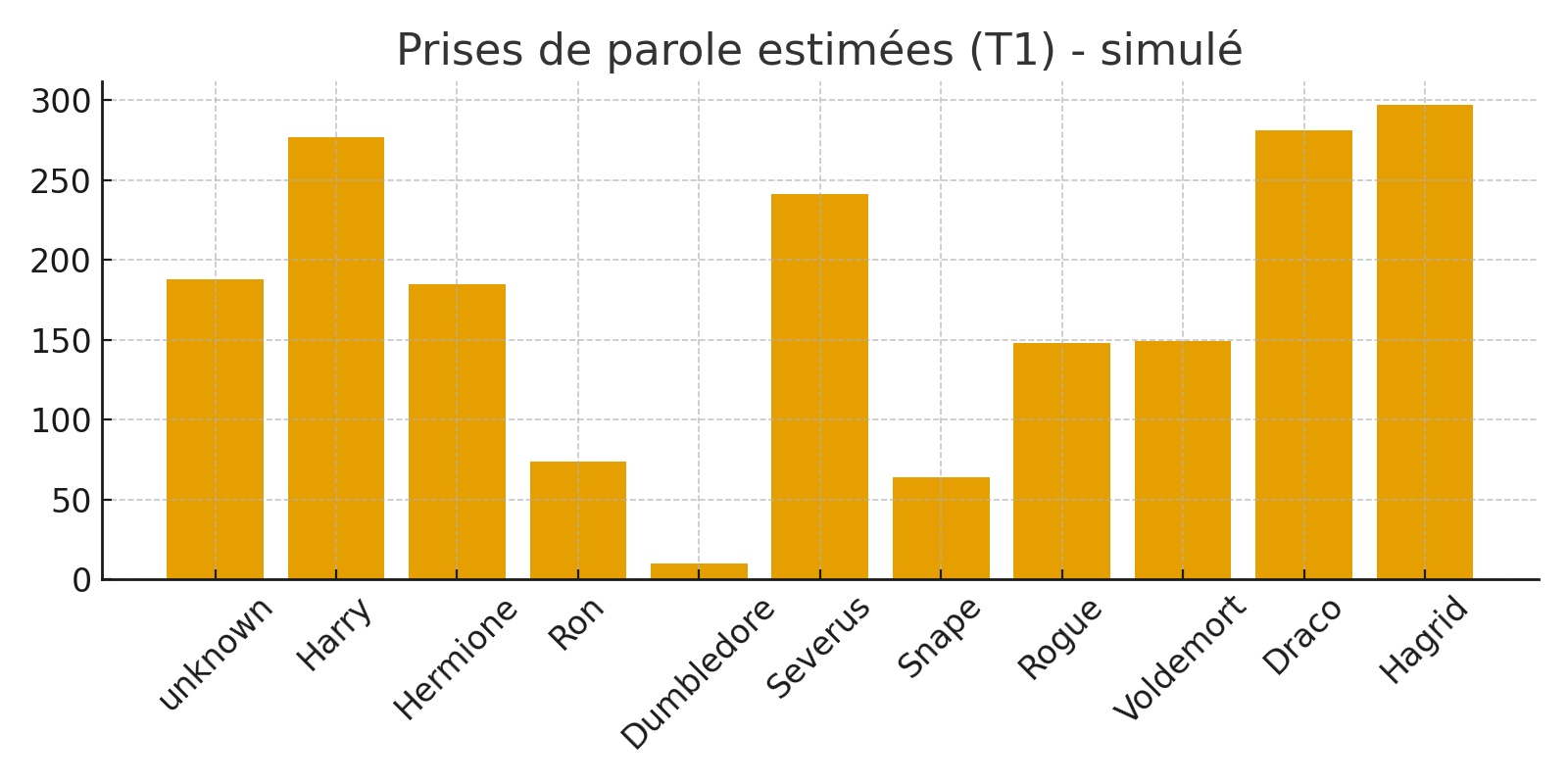
Analyse synthétique (simulée) :   
Les tomes présentent une tendance à l'augmentation des thèmes sombres (mots 'mort', 'sang'). Hermione apparaît comme le personnage le plus présent dans les dialogues estimés sur l'ensemble des tomes. Les graphiques suivants illustrent ces tendances.

## Graphiques clés

Occurrences du mot 'mort' par tome (simulé) :



Prises de parole estimées pour le Tome 1 (simulé) :



## Méthodologie

1) Comptage littéral des occurrences de mots/expressions donnés.  
2) Estimation des prises de parole par heuristique basée sur la présence de guillemets et l'attribution via contexte.  
3) Normalisation possible par pages ou par 10k mots (fichier pages.json).  
Limites : analyses basées sur heuristiques, attribution approximative des locuteurs, données sources soumises aux droits d'auteur.

## Fichiers fournis

- analyse\_hp.py : script d'analyse (à exécuter localement).  
- requirements.txt : dépendances Python.  
- pages.json : exemple de comptage de pages.  
- summary.csv : résultats (simulés ici).  
- images PNG : graphiques générés.

## Conclusion et recommandations

Pour obtenir des résultats réels : déposer les fichiers .txt des livres dans le dossier 'books/' puis exécuter :  
python analyse\_hp.py --books\_dir books --out\_dir output --page\_counts pages.json  
Ensuite, importer summary.csv dans ChatGPT pour obtenir une interprétation narrative prête pour le jury.