TP Thème 2 – Analyse de logs SSH simulés (auth.log)

Objectifs pédagogiques :

- Lire et parcourir un fichier de log
- Extraire des informations pertinentes (IPs, utilisateurs, erreurs)
- Compter et classer les tentatives de connexion suspectes
- Visualiser les IPs les plus actives (bonus)

Consignes

Contexte : Vous êtes analyste sécurité dans une entreprise. Vous recevez une copie d'un fichier de log SSH et devez identifier des comportements suspects : IPs avec de nombreuses tentatives d'accès échouées.

Fichier fourni

Un fichier **auth.log** est mis à disposition. Il contient un mélange de tentatives échouées et réussies de connexion SSH.

Partie 1 – Analyse textuelle (script simple)

1. Ouvrir le fichier auth.log
2. Extraire toutes les lignes contenant **"Failed password"**
3. Extraire les adresses IP de ces lignes à l'aide d'une **expression régulière**
4. Compter le nombre d'occurrences de chaque IP
5. Afficher les **5 IPs** ayant généré le plus d'échecs

Partie 2 – Visualisation (script avancé)

1. Utiliser la bibliothèque *matplotlib* (utilisez **pip install matplotlib** si nécessaire)
2. Créer un **graphique de barres** représentant les IPs avec le plus grand nombre d'échecs
3. Comparer les IPs ayant échoué et celles ayant réussi (bonus)
4. Ajouter une **légende, un titre**, et des **axes lisibles**

Bonus – Analyse avancée

- Filtrer les tentatives réussies (Accepted password) et les comparer aux échecs
- Exporter les résultats dans un fichier CSV ou JSON

- Créer une interface utilisateur simple (avec input()) pour explorer les résultats

Conseils pratiques

- Travaillez par étapes, testez chaque partie du code
- Utilisez des **expressions régulières simples**, testables dans regex101.com
- Pensez à **normaliser vos données** (IP, formats de ligne)