

Vous travaillez chez La poule qui chante, une entreprise française d'agroalimentaire. Son activité principale est l'élevage et la vente de poulets sous le label "Poulet Agriculture Biologique".

Son activité actuelle est franco-française mais Patrick, le PDG de l'entreprise souhaite évaluer la possibilité de se développer à l'international. Pour l'instant, le champ des possibles est bien large : aucun pays particulier ni aucun continent n'est pour le moment choisi. Tous les pays sont envisageables !

Comme on en a déjà parlé, je te fais ce mail rapide pour te briefer sur notre mission d'expansion à l'international.

Ton objectif sera de proposer une **analyse des groupements de pays** que l'on peut **cibler** pour exporter nos poulets. Nous approfondirons ensuite l'étude de marché.

Tu seras en autonomie sur ce projet, notamment sur le choix des données à analyser et du langage (R ou Python).

1. **Pars** des données de la **FAO** (*Food and Agriculture Organization*) en pièce jointe pour commencer ton analyse puis :
  - utilise l'analyse PESTEL pour trouver des idées de nouvelles données à ajouter (nous voulons au minimum 8 variables) ;
  - récupère et utilise toutes les données en *open data* que tu souhaites sur [le site de la FAO](#), [de la banque mondiale](#) ou encore [sur données mondiales](#).
2. **Prépare** et **nettoie** les données :
  - Si tu utilises plusieurs sources de données, regroupe-les dans un même fichier.
  - L'idéal serait d'avoir au minimum 100 pays dans notre analyse (qui couvrent au moins 60% de la population mondiale).
3. Passe ensuite à l'**exploration des données** (en Python ou en R).

Je ne vais pas t'apprendre ton travail, je crois que tu as suffisamment fait cela !

Pour la partie analytique, réalise une **ACP** avec réductions des dimensions et un **clustering** (dans un notebook séparé de l'exploration des données) :

- Analyse le **cercle des corrélations** et la **projection des individus**.
- Regroupe les pays en utilisant les données de ton ACP ou les données brutes.
- Commence par la **classification ascendante hiérarchique** puis un **k-means**.

Juste pour te rappeler que tu présenteras tes résultats dans un mois devant le COMEX.

Nous attendons des préconisations de pays où nous étendre, et puis tu feras un point sur ta méthodologie de collecte, les traitements et les analyses. Mais pas trop technique n'est-ce-pas ? La plupart des directeurs n'ont pas ton niveau en data !

Pendant la soutenance, l'évaluateur jouera le rôle de Patrick, votre PDG :

- **Présentation des livrables** (15 minutes)
  - Vous présenterez vos livrables à l'aide du support de présentation réalisé, vous structurerez votre présentation en rappelant :
    - le contexte du projet ;
    - votre démarche et méthodologie d'analyse ;
    - vos résultats et recommandations de pays pour l'entreprise.
- **Discussion** (10 minutes)
  - Toujours dans le rôle du PDG, l'évaluateur vous posera des questions sur votre travail, sur vos choix et sur vos préconisations, telles que :
    - Est-ce qu'il y a un pays qui ressort plus que les autres ?
    - En quoi est-ce intéressant d'utiliser une ACP ?
    - Est-ce que les deux méthodes de clustering ont donné les mêmes résultats ?