# Analyse des Tendances Économiques et Sociales via le Web Scraping et la Visualisation de Données

Ce projet vise à exploiter ces données pour fournir des analyses approfondies sur des thèmes d'actualité, comme l'évolution des prix, les tendances du marché de l'emploi ou encore les comportements des consommateurs. L'objectif est de combiner le **web scraping**, l'**analyse statistique avancée**, et la **visualisation interactive** pour mieux comprendre et interpréter les données.

#### **Objectifs**

1. **Extraction des données** :
   * Collecter des données économiques et sociales à partir de sources fiables telles que des sites gouvernementaux, des portails statistiques, des marketplaces, ou encore des médias.
   * Utiliser des techniques de web scraping pour automatiser la collecte d'informations.
2. **Analyse des données** :
   * Nettoyer, organiser et analyser les données collectées.
   * Appliquer des modèles statistiques pour détecter des tendances, identifier des corrélations ou prédire des comportements futurs.
3. **Visualisation et communication** :
   * Présenter les résultats sous forme de tableaux de bord interactifs, graphiques dynamiques ou infographies.
   * Faciliter la prise de décision grâce à une communication visuelle claire.

#### **Thématiques ouvertes possibles(Vous pourez proposez des thematiques ouvertes)**

1. **Analyse des prix des produits de consommation** :
   * Collecter les prix des produits dans plusieurs marketplaces (Amazon, eBay, etc.) et comparer les tendances.
   * Étudier l'effet des promotions et des saisons sur les variations de prix.
2. **Emploi et économie** :
   * Scraper des offres d'emploi pour analyser les compétences les plus recherchées dans différents secteurs.
   * Identifier les régions avec les opportunités les plus nombreuses.
3. **Analyse des tendances de consommation** :
   * Étudier les avis des consommateurs pour comprendre les préférences par secteur (mode, électronique, alimentation, etc.).
   * Analyser les ventes et les popularités de produits spécifiques.
4. **Impact des politiques économiques** :
   * Scraper des publications ou rapports économiques pour évaluer l’impact des politiques économiques (hausse des taxes, subventions, etc.) sur la vie quotidienne.

#### **Méthodologie**

1. **Étape 1 : Collecte des données**
   * Choisir des sources fiables et pertinentes pour le thème choisi.
   * Utiliser des outils comme **BeautifulSoup**, **Scrapy**, ou **Selenium** pour le web scraping.
2. **Étape 2 : Nettoyage et pré-traitement**
   * Gérer les données manquantes, les doublons et les incohérences.
   * Utiliser Python (pandas, NumPy) ou R pour structurer les données.
3. **Étape 3 : Analyse statistique**
   * Effectuer des analyses descriptives (moyenne, médiane, écart-type).
   * Identifier les tendances à l’aide de modèles statistiques (régression, séries temporelles).
   * Explorer des relations causales ou prédictives à l’aide d’algorithmes de machine learning simples (régression logistique, K-means).
4. **Étape 4 : Visualisation**
   * Utiliser des outils comme **Matplotlib**, **Plotly**, ou **Tableau** pour créer des visualisations interactives.
   * Développer un tableau de bord ou **Power BI** ou tout autre outils de visualisation
5. **Étape 5 : Interprétation et rapport**
   * Résumer les résultats et proposer des interprétations claires pour une audience technique et non technique.
   * Fournir des recommandations basées sur les analyses.

#### **Livrables**

* Code source du script de scraping et d’analyse.
* Ensemble de données collectées et nettoyées.
* Rapport analytique détaillé avec interprétations.
* Tableau de bord interactif ou visualisation finale.

#### **Outils et technologies suggérés**

* **Langages** : Python, R
* **Web scraping** : BeautifulSoup, Scrapy, Selenium
* **Traitement des données** : pandas, NumPy
* **Analyse statistique** : scikit-learn, Statsmodels
* **Visualisation** : Matplotlib, Plotly, Tableau, Power BI

Deadline : 06 Janvier 2024

Github : ingngolo@gmail.com