**Projet Scientifique : Expérience en Laboratoire de Chimie Générale en Python**

**Contexte du Projet**

Ce projet vise à introduire les étudiants de première session en sciences de la nature aux bases de la programmation en Python à travers l’analyse de données chimiques. Le projet est divisé en quatre semaines, chacune se concentrant sur des compétences spécifiques en Python et en analyse de données.

**Objectifs du Projet**

* Apprendre à manipuler des données avec Pandas.
* Visualiser des données avec Matplotlib.
* Effectuer des analyses statistiques de base en utilisant des boucles et des listes.
* Comprendre et appliquer des concepts de dérivées et de régression linéaire.

**Semaine 1 : Introduction à Pandas et Extraction de Données**

1. **Installation et Configuration** :
   * Installer Python et les bibliothèques nécessaires : Pandas et Matplotlib.
   * Télécharger les fichiers .csv fournis.
2. **Exploration des Données** :
   * Charger les fichiers .csv avec Pandas.
   * Utiliser head(), info(), et describe() pour explorer les données.
   * Filtrer les données selon des critères spécifiques (ex : valeurs manquantes, plages de valeurs).
3. import pandas as pd
4. # Charger les données
5. df = pd.read\_csv('data1.csv')
6. # Exploration des données
7. print(df.head())
8. print(df.info())
9. print(df.describe())
10. # Filtrer les données
11. df\_filtered = df[df['Valeur'] > 50]

**Semaine 2 : Visualisation des Données avec Matplotlib**

1. **Création de Graphiques Simples** :
   * Histogrammes, boîtes à moustaches, et graphiques en nuage de points.
   * Personnalisation des graphiques (titres, légendes, couleurs).
2. import matplotlib.pyplot as plt
3. # Histogramme
4. plt.hist(df['Valeur'], bins=20, color='blue', edgecolor='black')
5. plt.title('Distribution des Valeurs')
6. plt.xlabel('Valeur')
7. plt.ylabel('Fréquence')
8. plt.show()
9. # Nuage de points
10. plt.scatter(df['Principe'], df['Valeur'], color='red')
11. plt.title('Relation entre Principe et Valeur')
12. plt.xlabel('Principe')
13. plt.ylabel('Valeur')
14. plt.show()

**Semaines 3 et 4 : Analyse Statistique et Régression Linéaire**

1. **Calculs Statistiques de Base** :
   * Moyennes, médianes, et écarts-types à l’aide de boucles et de listes.
2. # Calcul de la moyenne
3. valeurs = df['Valeur'].tolist()
4. somme = 0
5. for valeur in valeurs:
6. somme += valeur
7. moyenne = somme / len(valeurs)
8. print(f'Moyenne: {moyenne}')
9. **Régression Linéaire et Dérivées** :
   * Implémentation de la régression linéaire à partir de zéro.
   * Calcul des dérivées pour analyser les tendances.
10. import numpy as np
11. # Régression linéaire
12. x = df['Valeur'].values
13. y = df['Valeur'].values # Utiliser une autre colonne pour y si nécessaire
14. n = len(x)
15. xy = sum(x[i] \* y[i] for i in range(n))
16. x\_squared = sum(x[i] \*\* 2 for i in range(n))
17. a = (n \* xy - sum(x) \* sum(y)) / (n \* x\_squared - sum(x) \*\* 2)
18. b = (sum(y) - a \* sum(x)) / n
19. print(f'Equation de la droite: y = {a}x + {b}')
20. # Visualisation de la régression linéaire
21. plt.scatter(x, y, color='blue')
22. plt.plot(x, a \* x + b, color='red')
23. plt.title('Régression Linéaire')
24. plt.xlabel('x')
25. plt.ylabel('y')
26. plt.show()

**Fichiers .csv**

* **data1.csv** : Données sur les principes, lois et théories de base en chimie générale.
* **data2.csv** : Données sur les grandes découvertes chimiques.
* **data3.csv** : Données sur les composés les plus courants de notre environnement.
* **data4.csv** : Données sur l’analyse des données expérimentales.

Chaque fichier contient entre 50 et 100 lignes de données, avec des colonnes représentant différentes variables chimiques.

**Liens pour Télécharger les Fichiers .csv**

1. [data1.csv](blob:https://outlook.office.com/83c80e01-8fc6-4e2f-8d28-1874d292df52)
2. [data2.csv](blob:https://outlook.office.com/8394a455-ab08-473e-8e2e-4d87b75791f6)
3. [data3.csv](blob:https://outlook.office.com/2e77645f-d764-45c4-91df-efff0544e102)
4. [data4.csv](blob:https://outlook.office.com/0e845445-9253-4695-b89e-4cb7f46434fe)

**Conclusion**

Ce projet permettra aux étudiants de se familiariser avec les bases de la programmation en Python tout en appliquant ces compétences à des données réelles dans le domaine de la chimie générale. Les étudiants apprendront à extraire, visualiser et analyser des données, tout en développant une compréhension des concepts de dérivées et de régression linéaire.

Si vous avez besoin de plus d’aide ou d’informations supplémentaires, n’hésitez pas à demander !

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Une image contenant texte, capture d’écran, Police

Description générée automatiquement

Une image contenant texte, capture d’écran, Police, nombre

Description générée automatiquement

Une image contenant texte, capture d’écran, Police

Description générée automatiquement