

Épreuve pratique BAC NSI

Sujet 2 - Analyse d'écart de salaires

Cette situation d'évaluation comporte ce document ainsi que des fichiers de codes et de données présents sur l'ordinateur à la disposition du candidat. Le candidat doit restituer ce document avant de sortir de la salle d'examen. Le candidat doit agir en autonomie et faire preuve d'initiative tout au long de l'épreuve.

En cas de difficulté, le candidat peut solliciter l'examineur afin de lui permettre de continuer la tâche. Des moments privilégiés pour solliciter l'examineur sont indiqués dans le document sous la forme d'appels professeur.

L'examineur peut intervenir à tout moment, s'il le juge utile.

Des écarts de salaires subsistent entre les femmes et les hommes, même à poste équivalent. En France, l'écart de salaire moyen est encore d'environ 15 % en 2023 selon l'INSEE.

Afin d'observer cet écart et ces conséquences, on dispose de jeux de données représentant les employés d'une entreprise. Chaque jeu de données est une liste de dictionnaires où chaque dictionnaire représente un employé avec les champs suivants et le type des valeurs associées :

- 'experience' (int, en années, indiquant l'expérience professionnelle) ;
- 'etudes' (int en années, indiquant le nombre d'années d'étude après l'obtention du baccalauréat) ;
- 'sexe' (str, 'F' ou 'M') ;
- 'salaire' (int, en euros).

Deux jeux de données vous sont fournis, un premier jeu de données de test dans le fichier `donnees.py` dont le contenu est reproduit ci-dessous et un jeu de données plus complet de 2000 employés dans le fichier `donnees_completes.py`.

```
employees = [  
    {'experience': 5, 'etudes': 3, 'sexe': 'F', 'salaire': 2400},  
    {'experience': 3, 'etudes': 3, 'sexe': 'M', 'salaire': 2550},  
    {'experience': 5, 'etudes': 5, 'sexe': 'F', 'salaire': 2500},  
    {'experience': 3, 'etudes': 5, 'sexe': 'M', 'salaire': 2800},  
    {'experience': 2, 'etudes': 5, 'sexe': 'F', 'salaire': 2300},  
    {'experience': 2, 'etudes': 3, 'sexe': 'M', 'salaire': 2700}  
]
```

Le fichier `analyse.py` présente des éléments d'analyse de ces jeux de données qui vont être complétés et améliorés dans les questions qui suivent.

1. Écrire le code de la fonction `salaires_moyen_condition` qui prend en paramètres un tableau d'employés dans le format décrit ci-dessus, le nom d'un des trois champs 'experience', 'etudes' ou 'sexe' ainsi qu'une valeur et qui renvoie un nombre flottant (type `float`) correspondant au salaire moyen des employés dont la valeur associée au champ est la valeur fournie.

La fonction doit renvoyer `None` s'il n'y a pas d'employés ayant la valeur recherchée pour le champ donné.

Ainsi, un appel à `salaires_moyen_condition(employes, 'sexe', 'F')` permettra de renvoyer le salaire moyen des femmes employées.

Une fonction de test sur le jeu de données de test est fournie. Déterminer le salaire moyen des femmes et des hommes pour le jeu de données complet.

Appel 1	Appeler le professeur pour lui présenter vos réponses et votre fonction ou en cas de difficultés de compréhension de la représentation.
----------------	---

2. Écrire en Python une fonction nommée `effectif_par_sexe` qui prend en paramètre un tableau non vide d'employés et qui renvoie un dictionnaire ayant deux clés 'F' et 'H' associées respectivement à l'effectif des femmes et à l'effectif des hommes employés.

Par exemple, avec le tableau `employes` précédent :

```
>>> effectif_par_sexe(employes)
{'F' :3, 'M' :3}
```

Appel 2	Appeler le professeur pour lui présenter votre fonction et son fonctionnement ou en cas de difficultés.
----------------	---

On définit l'écart de salaire moyen en pourcentage comme :

$$\text{écart} = \frac{\text{salaire moyen hommes} - \text{salaire moyen femmes}}{\text{salaire moyen hommes}} \times 100$$

Une fonction `calcule_ecart_sexe` est écrite dans le fichier `analyse.py`

3. Expliquer pourquoi le code de cette fonction est incorrect et proposer quelques tests simples sous forme d'assertions qui permettent de mettre ce ou ces problèmes en évidence :
 - vérifier que le résultat est `None` dans le cas où un seul sexe est présent ;
 - vérifier qu'un écart de salaires exprimé en pourcentage soit toujours compris entre 0 et 100.

Proposer une version corrigée de la fonction `ecart_salaire` qui valide ces tests et renvoie le bon écart. En déduire l'écart de salaire moyen dans les données complètes.

Appel 3	Appeler le professeur pour lui présenter vos tests et la correction proposée.
----------------	---

4. Afin de proposer un salaire d'embauche à un nouvel employé, le service informatique de l'entreprise a proposé d'utiliser l'algorithme des k-plus proches voisins et renvoyer la moyenne des salaires des trois employés aux caractéristiques les plus proches.

La fonction `salaire_par_proximite` effectue ce calcul pour faire une proposition.

Tester et comparer les salaires proposés aux deux futurs employés suivants :

```
{'experience': 3, 'etudes': 3, 'sexe': 'F'}  
{'experience': 3, 'etudes': 3, 'sexe': 'M'}
```

Identifier la source des écarts entre les deux propositions de salaire dans le programme et la corriger.

Appel 4	Appeler le professeur pour lui présenter vos tests, votre analyse des écarts et la correction proposée.
----------------	---

Description du dossier fourni

Le dossier fourni au candidat sur l'ordinateur comporte les éléments suivants :

- Une version PDF de l'énoncé ;
- Un code source à analyser et compléter `analyse.py` ;
- un jeu de données de tests `donnees.py` ;
- un jeu de données complet de 2000 employés `donnees_completes.py`.