

Rapport projet 7

Script du questionnaire :

Objectif : rassembler tous les fichiers csv contenant les résultats des questionnaires et les fusionner dans un seul csv afin d'avoir un data complet

```
import os    # pour manipuler les fichiers
import glob  # pour fusionner csv
import pandas as pd
os.chdir("C:\\Users\\utilisateur\\Documents\\microsoft_ia\\Cahier d'apprentissage\\Projets\\Projet_7_k-NN\\Quest
# on indique le chemin du dossier où sont contenus les csv
```

```
# on crée une variable qui servira à indiquer l'extension csv
extension = 'csv'

# pour chaque document contenu dans le fichier on crée une boucle qui ajoute les documents
all_filenames = [i for i in glob.glob('*.{}'.format(extension))]
```

```
# On combine tous les fichiers qu'on ajoute à une liste contenue dans la variable `combined_csv`
combined_csv = pd.concat([pd.read_csv(f) for f in all_filenames ])

# On peut exporter le csv
combined_csv.to_csv( "combined_csv.csv", index=False, encoding='utf-8-sig')
```

Description du code du questionnaire :

On procède aux importations nécessaires (numpy, pandas, os) :

```
import numpy as np
import pandas as pd
import os
```

On crée une classe *Question* et on initialise le constructeur avec *prompt* :

```
class Question:
    def __init__(self, prompt):
        self.prompt = prompt
        #self.answer = answer
```

On crée une variable *question_prompts* qui contient une liste. Cette liste contient toutes les questions, on peut accéder à chacune des questions par son indice dans la liste. Cette liste est donc composée de 10 questions en format str avec un saut de ligne (\n) et séparées par une virgule :

```
question_prompts = [
    "Question 1 :\nQuelle importance accordez-vous au succès ?\n\n(a)Une importance moyenne.\n(b)Une très grande importance.\n(c)Une importance moyenne.\n(d)Une importance moyenne.\n\nQuestion 2 :\nVous est-il déjà arrivé de prendre des jours de congé parce que vous étiez stressé(e) ?\n\n(a)Oui, très souvent.\n(b)Oui, parfois.\n(c)Non, jamais.\n\nQuestion 3 :\nVous considérez-vous comme une personne sachant conserver son sang froid en période de crise ?\n\n(a)Oui, très souvent.\n(b)Oui, parfois.\n(c)Non, jamais.\n\nQuestion 4 :\nParmi les propositions suivantes, laquelle est la plus à même de vous détendre et de réduire votre stress ?\n\n(a)Prendre un bain chaud.\n(b)Prendre un bain chaud.\n(c)Prendre un bain chaud.\n\nQuestion 5 :\nLes délais vous stimulent-ils ?\n\n(a)Non, mais les délais sont un mal nécessaire avec lequel il faut composer.\n(b)Non, mais les délais sont un mal nécessaire avec lequel il faut composer.\n(c)Non, mais les délais sont un mal nécessaire avec lequel il faut composer.\n\nQuestion 6 :\nPensez-vous que la vie actuelle génère plus de stress qu'il y a 40 ans ?\n\n(1)Peut-être.\n(2)Oui, beaucoup.\n(3)Non, pas du tout.\n\nQuestion 7 :\nVotre neveu vous demande de garder ses trois enfants un peu turbulents pendant le week-end ?\n\n(a)Oui, très souvent.\n(b)Oui, parfois.\n(c)Non, jamais.\n\nQuestion 8 :\nLe stress vous a-t-il déjà conduit(e) à endommager des choses ?\n\n(1)Non, je n'ai jamais eu de problème.\n(2)Oui, parfois.\n(3)Oui, très souvent.\n\nQuestion 9 :\nVous arrive-t-il d'être ennuyé(e) par des petits riens ?\n\n(1)Oui, parfois.\n(2)Assez souvent.\n(3)Non, jamais.\n\nQuestion 10 :\nQue ressentiriez-vous si vous deviez vous familiariser avec une nouvelle technologie ?\n\n(a)Très à l'aise.\n(b)Très à l'aise.\n(c)Très à l'aise.
```

Idem que précédemment : on crée une variable *Interpretation* qui contient une liste avec trois phrases détaillant le score.

```
Interpretation = [
    'Score entre 15 et 30 : \n\nVous savez très bien gérer votre stress. \nVous êtes très probablement perçu(e) comme une personne capable de gérer le stress.\n\nScore entre 11 et 14 : \n\nMême s'il vous arrive d'être tendu(e) et stressé(e) en certaines occasions, vous êtes capable de gérer le stress.\n\nScore moins de 10 : \n\nVotre score indique que le stress vous affecte de façon négative.',
]
```

Création d'une variable *questions* contenant également une liste. Chaque élément appelle la classe *Question* pour hériter d'elle et renvoie à l'attribut du constructeur *prompt*. On peut accéder à chaque question grâce à l'indice de la liste.

```
questions = [  
    Question(question_prompts[0]),  
    Question(question_prompts[1]),  
    Question(question_prompts[2]),  
    Question(question_prompts[3]),  
    Question(question_prompts[4]),  
    Question(question_prompts[5]),  
    Question(question_prompts[6]),  
    Question(question_prompts[7]),  
    Question(question_prompts[8]),  
    Question(question_prompts[9]),  
]
```

On crée une fonction *run_quiz* qui fait tourner le questionnaire avec comme attributs *questions* et *Interpretation* (qui sont des variables activées précédemment). La variable *score* est initialisée à 0. On crée également la variable *Data* qui contient une liste vide.

```
def run_quiz(questions, Interpretation):  
    score = 0  
    Data = []  
    for question in questions:  
        answer = input(question.prompt)  
        Data.append(answer)  
        if answer == 'a' or answer == '1':  
            score += 1  
        elif answer == 'b' or answer == '2':  
            score += 0  
        elif answer == 'c' or answer == '3':  
            score += 2  
  
    print("\nVotre Score est : ", score, '\n')
```

La boucle *for* va tourner sur chaque élément de la liste *questions*. Pour chaque question, on va demander à l'utilisateur d'entrer une réponse avec *input*. On va ajouter chaque réponse (contenue dans la variable *answer*) à la liste *Data* qu'on pourra sauvegarder par la suite.

On procède ensuite à des vérifications par plusieurs conditions (*if* et *elif*). Si la réponse vaut 'a' ou '1' la variable *score* est incrémentée de 1, si la réponse vaut 'b' ou '2' le score ne change pas et si la réponse vaut 'c' ou '3' la variable *score* augmente de 2.

Une fois sorti de la boucle, on affiche le score de l'utilisateur.

On peut alors analyser le score de l'utilisateur avec des conditions :

```
if score < 30 and score >=20:
    Label = 'A'
    print(Interpretation[0])
elif score < 20 and score >=10:
    Label = 'B'
    print(Interpretation[1])
elif score < 10:
    Label = 'C'
    print(Interpretation[2])
```

Si la variable *score* est inférieure à 30 et supérieure ou égale à 20, on lui donne la lettre 'A' dans la variable *Label*. On affiche ensuite le premier élément de la liste *Interpretation*.

Si le score est inférieur à 20 et supérieur ou égal à 10, son label est 'B' et on affiche le deuxième élément de la liste *Interpretation*.

Enfin, si le score est strictement inférieur à 10, on donne la lettre 'C' à la variable *Label* puis on affiche le troisième élément d'*Interpretation*.

```
Data.append(score)
Data.append(Label)
Data = np.array(Data)
name_generator = chr(np.random.randint(500))+chr(np.random.randint(500))+chr(np.random.randint(500))
```

On ajoute la variable *score* à la liste *Data* ainsi que la variable *Label*. On transforme ensuite la liste *Data* en tableau (*array*) numpy. La variable « *name-generator* » contient une concaténation de trois caractères aléatoires (fonction *random.randint*).

```
d = {'Q1' : Data[0],
     'Q2' : Data[1],
     'Q3' : Data[2],
     'Q4' : Data[3],
     'Q5' : Data[4],
     'Q6' : Data[5],
     'Q7' : Data[6],
     'Q8' : Data[7],
     'Q9' : Data[8],
     'Q10' : Data[9],
     'Score' : Data[10],
     'Interpretation' : Data[11],}
```

La variable *d* abrite dans un dictionnaire les réponses de la liste *Data*. Il récupère chaque position de la liste grâce à l'index.

```
if any(File.endswith(".csv") for File in os.listdir('./Dataset/')):  
    df = pd.read_csv('./DataSet/'+os.listdir('./DataSet/')[0])  
    df = df.append(d, ignore_index=True)  
    df.to_csv('./DataSet/'+os.listdir('./DataSet/')[0], index=False)  
else:  
    df = pd.DataFrame(d, index = ['1'])  
    df.to_csv('./DataSet/DataSet__'+name_generator+'__.csv', index=False)  
return print(df)  
  
run_quiz(questions, Interpretation)
```

On applique une condition qui vise à remplacer ou non un fichier csv déjà existant. Si le fichier n'existe pas, alors il est automatiquement créé avec comme nom un nom aléatoire grâce à la variable *name_generator*.

Enfin, on peut lancer le questionnaire en appelant *run_quiz*.