

Exercice 1 :

- 1)
- 2) Rappel du code que vous allez compléter dans les questions suivantes.

```
1 import fonctions_fichiers_csv_json as fic
2
3
4 def extrait_data(data:list[list[str]]) -> tuple[list[str],list[str],list[str],list[st
5     """Cette fonction extrait les données
6     et renvoie 4 listes de même longueur"""
7     sexes : list = []
8     prenoms :list = []
9     annees :list = []
10    nb      :list = []
11
12    for k in range(1,len(data)): # on retire l'entête.
13        ligne=data[k]
14        sexes.append(ligne[0])
15        prenoms.append(ligne[1 ])
16        annees.append(ligne[2])
17        nb.append(ligne[3])
18
19    return sexes, prenoms, annees, nb
```

- 3) La fonction qui renvoie le nombre total de naissances de garçons et de filles de 1900 à 2020 avec les années XXXX et les prénoms rares qui ne sont pas comptés est la suivante :

```

1 def compteGouF(sexes : list [str] , prenoms : list[str] , annees : list [str] ,nb
  : list [str] ) -> tuple[int,int] :
2     """Compte le total de Garçons ou Fille né de 1900 à 2020
3     Cette fonction renvoie 2 valeurs : le nombre total de garçons
4     et le nombre total de filles."""
5     nbG : int = 0
6     nbF : int = 0
7     for k in range(len(annees)) :
8         if annees[k]!="XXXX" :
9             if 1900<= int(annees[k]) and int(annees[k]) <=2020 :
10                if prenoms!='_PRENOMS_RARES':
11                    if sexes[k]=='1' :
12                        nbG = nbG + int(nb[k])
13                    else :
14                        nbF = nbF + int(nb[k])
15     return nbF,nbG

```

Après la fonction `compteGouF` on ajoute les lignes d'appel de la fonction et d'affichage des résultats intermédiaires (lignes 7 et 9 à la fin du code dans le cadre ci-dessous). Vous noterez qu'avant d'aller plus loin il est rassurant d'afficher les résultats que la fonction renvoie.

Vous devriez ainsi voir s'afficher *"Entre 1900 et 2021 il y a eu 42541437 naissances de filles et 43278958 naissances de garçons."*

```

1 nom_fichier="nat2021.csv"          # code déjà présent
2 data : list[str] =fic.chargeFichierCsv(nom_fichier) # code déjà présent
3 sexes, prenoms, annees, nb = extrait_data(data)    # code déjà présent
4
5 # on appelle la fonction et on regarde ce qu'elle renvoie
6 # avec un affichage.
7 nbF,nbG =compteGouF(sexes , prenoms , annees,nb    )
8
9 print(f"Entre {annees[0]} et {annees[-2]} il y a eu {nbF} "
10      f"naissances de filles et {nbG} naissances de garçons.")

```

- 4) La fonction `calculPourcentages()` qui reçoit le nombre de naissance de garçons et le nombre de naissance de filles et renvoie les pourcentages correspondants est la suivante. Vous noterez qu'avant d'aller plus loin il est rassurant d'afficher les résultats que la fonction renvoie.

Vous devriez ainsi voir s'afficher *Entre 1900 et 2021 il y a eu 49.57 pourcents de filles et 50.43 pourcents de garçons.*

```

1 def calculPourcentages(nbF : int ,nbG : int ) -> tuple[float, float] :
2     """Cette fonction calcule le pourcentage de naissance G ou F.
3         Cette fonction renvoie 2 valeurs, le pourcentage de F et
4         le pourcentage de G. """
5     pG : float = 0
6     pF : float = 0
7
8     pG = nbG / (nbF + nbG)
9     pF = nbF / (nbF + nbG)
10    # arrondis à 2 chiffres.
11    pG =100*round(pG,4)
12    pF =100*round(pF,4)
13    return pG , pF
14
15 ## ci-dessus le programme principal.
16 nom_fichier="nat2021.csv"
17 data : list[str] =fic.chargeFichierCsv(nom_fichier)
18 sexes, prenom, annees, nb = extrait_data(data)
19
20 # on appelle la fonction compteGouF et on regarde ce qu'elle renvoie
21 # avec un affichage.
22 nbF,nbG =compteGouF(sexes , prenom , annees,nb )
23
24 print(f"Entre {annees[0]} et {annees[-2]} il y a eu {nbF} "
25       f"naissances de filles et {nbG} naissances de garçons.")
26
27 # on appelle la fonction calculPourcentages et on regarde ce qu'elle renvoie
28 # avec un affichage.
29 (pG , pF )=calculPourcentages(nbF ,nbG )
30
31 print(f"Entre {annees[0]} et {annees[-2]} il y a eu {pF} pourcents de filles "
32       f" et {pG} pourcents de garçons.")

```

5) La procédure qui affiche les résultats dans des phrases d'explications peut s'écrire :

```

1 def affichageJoliDesResultats(annees, nbF, nbG, pG , pF) :
2     print(f"Entre {annees[0]} et {annees[-3]} :")
3     print(f" Le nombre total de naissance est de {nbF+nbG}. ")
4     print(f"Il y a eu {nbF} naissances de filles.")
5     print(f"Il y a eu{nbG} naissances de garçons.")
6     print(f"Il y a eu {pF} pourcents de filles et"
7         f" et {pG} pourcents de garçons.")

```

et produit l'affichage suivant :

```

1 Entre 1900 et 2020 :
2 Le nombre total de naissance est de 85820395.
3 Il y a eu 42541437 naissances de filles.
4 Il y a eu43278958 naissances de garçons.
5 Il y a eu 49.57 pourcents de filles et et 50.43 pourcents de garçons.

```

6) La fonction *naissancesParAns* complétée est la suivante :

```

1 def naissancesParAns(sexes : list[str] ,prenoms :
2     list[str], annees: list[str], nb : list [str], G_ou_F : int ) -> dict[int,int] :
3     """Cette fonction lit les données chargées depuis le csv puis
4     renvoie un dictionnaire dont les clés sont les années
5     et les valeurs l'effectif cumulé des naissances."""
6     dicoAnNb : dict={}
7     for k in range(1900,2021,1):      # on s'arrête en 2020.
8         dicoAnNb[k]=0
9
10    for k in range(len(annees)) :
11        if annees[k]!="XXXX" and int(annees[k])<=2020 :
12            # on s'arrête en 2020.
13            if prenoms[k] != "_PRENOMS_RARES" and sexes[k]==str(G_ou_F) :
14                dicoAnNb[int(annees[k])] = dicoAnNb[int(annees[k])] + int(nb[k])
15    return dicoAnNb

```

Cette fonction s'appelle avec les commandes :

```

1 dicoG = naissancesParAns(sexes,prenoms , annees , nb , G_ou_F=1 )
2 dicoF = naissancesParAns(sexes,prenoms , annees , nb , G_ou_F=2 )

```

Il est possible de consulter le dictionnaire en tapant `dicoG[2000]+dicoF[2000]` 773055 et de comparer à une rapide recherche internet qui donne 779000 par exemple. Ce qui compte c'est l'ordre de grandeur ici l'écart est faible. Il est donc légitime que la fonction renvoie un résultat correct.

7)

8) La procédure de tracé graphique *traceCleValeurDoubleDico* est la suivante :

Pour le tracé graphique vous devez en début de code importer les modules :

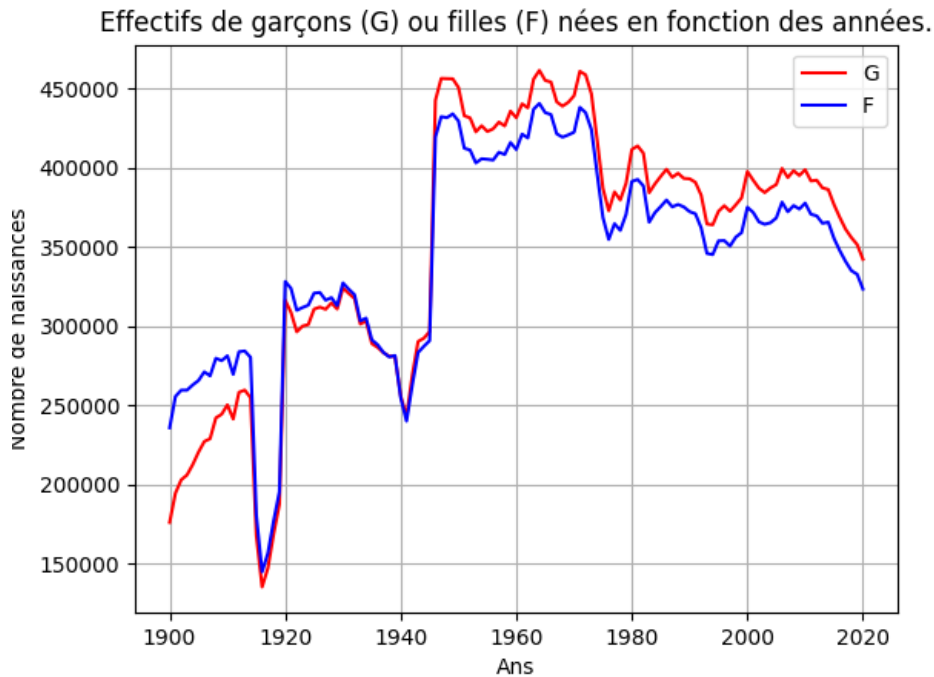
```
1 import matplotlib.pyplot as plt
```

```
1 def traceCleValeurDoubleDico( dico1 : dict[int,int],
2                               dico2 : dict[int,int],
3                               descriptionAxeX : str ,
4                               descriptionAxeY : str,
5                               titre : str,
6                               etiquette1:str,
7                               etiquette2:str ) -> None :
8     """Cette fonction trace les valeurs de chaque dictionnaire
9     en fonction des clés correspondantes"""
10    abs1 : list[int]= list(dico1.keys())
11    ord1 : list[int]= list(dico1.values())
12    abs2 : list[int]= list(dico2.keys())
13    ord2 : list[int] =list(dico2.values())
14
15    plt.plot(abs1,ord1, label=etiquette1, color='red')
16    plt.plot(abs2,ord2, label=etiquette2, color='blue')
17    plt.xlabel(descriptionAxeX)
18    plt.ylabel(descriptionAxeY)
19    plt.title(titre)
20    plt.legend()
21    plt.grid()
22    plt.show()
```

Vous pourrez l'appeler en définissant les paramètres avant l'appel de la fonction :

```
1 titre = "Effectifs de garçons (G) ou filles (F) nées en fonction des années."
2 descriptionAxeX="Ans"
3 descriptionAxeY="Nombre de naissances"
4 etiquette1='G'
5 etiquette2='F'
6 traceCleValeurDoubleDico( dicoG ,dicoF , descriptionAxeX ,
7                             descriptionAxeY ,titre , etiquette1, etiquette2 )
```

La courbe que vous devriez obtenir est la suivante :



9) Au debut de votre code, importez le module numpy :

```
1 import numpy as np
```

La fonction qui convertit un dictionnaire en 2 tableaux numpy est la suivante :

```
1 def convertitDictionnaireEnTableau (dico : dict[int,int]) -> np.ndarray :
2     """Convertit les clés et valeurs du dictionnaire en 2 tableaux
3     numpy"""
4     annees      : np.ndarray = np.array(list(dico.keys()))
5     nbNaissance : np.ndarray = np.array(list(dico.values()))
6     return annees, nbNaissance
7
8 # appel et test "léger" de la fonction.
9 anneesNpG, nbNaissNpG = convertitDictionnaireEnTableau (dicoG)
10 anneesNpF, nbNaissNpF = convertitDictionnaireEnTableau (dicoF)
11
12 assert( anneesNpF.size == 121)
13 assert( nbNaissNpF.size == 121)
14 assert( anneesNpG.size == 121)
15 assert( nbNaissNpG.size == 121)
```

10) Nous allons maintenant calculer les proportions.

Rappel : La proportion de naissance de filles s'écrit

$$p_F = \frac{\text{nombre}_F}{\text{nombre}_F + \text{nombre}_G}$$

et de garçons :

$$p_G = \frac{\text{nombre}_G}{\text{nombre}_F + \text{nombre}_G}$$

Le code de calculs sur les tableaux numpy pour obtenir le quotient ou proportion année par année des naissances s'écrit :

```
1 quotientG = nbNaissNpG / (nbNaissNpG+nbNaissNpF)
2 quotientF = nbNaissNpF / (nbNaissNpG+nbNaissNpF)
```

11) Pour visualiser les résultats sous forme de courbe, La procédure qui reçoit les tableaux numpy et affiche une courbe en rouge et l'autre en bleu est la suivante :

code

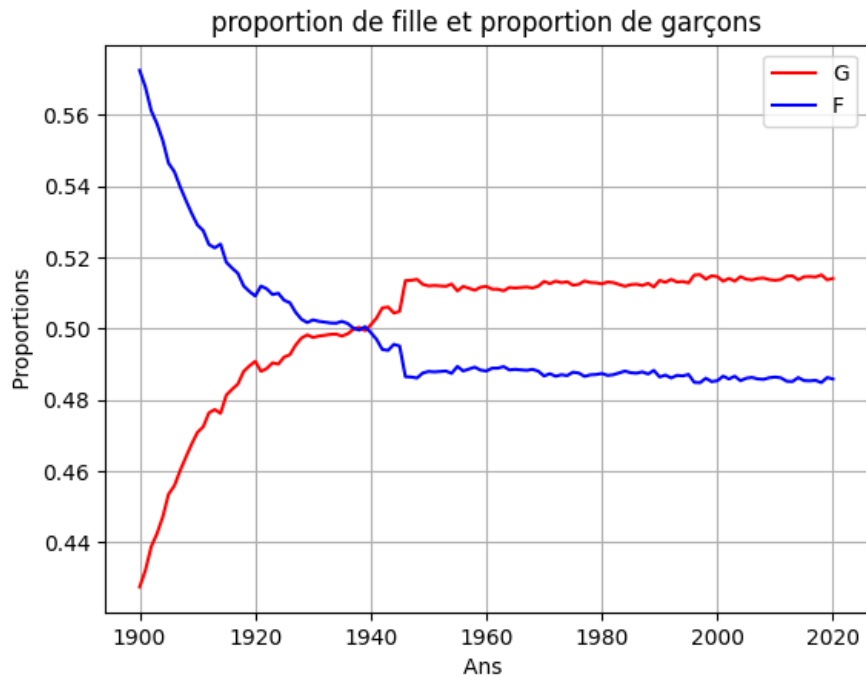
```
1 def traceDeuxTableaux(xtab1 : np.ndarray ,
2                       ytab1 : np.ndarray ,
3                       xtab2 : np.ndarray ,
4                       ytab2 : np.ndarray ,
5                       descriptionAxeX : str ,
6                       descriptionAxeY : str, titre : str,
7                       etiquette1:str,
8                       etiquette2:str ) -> None :
9     """procédure qui trace 2 courbes de tableau numpy"""
10    plt.plot(xtab1,ytab1, label=etiquette1, color='red')
11    plt.plot(xtab2,ytab2, label=etiquette2, color='blue')
12    plt.xlabel(descriptionAxeX)
13    plt.ylabel(descriptionAxeY)
14    plt.title(titre)
15    plt.legend()
16    plt.grid()
17    plt.show()
```

Pour l'appel :

code

```
1 descriptionAxeX= 'Ans '
2 descriptionAxeY = ' Proportions '
3 titre='proportion de fille et proportion de garçons'
4 etiquette1='G'
5 etiquette2='F'
6 traceDeuxTableaux( anneesNpG , quotientG ,anneesNpF ,quotientF
7   , descriptionAxeX ,
8   , descriptionAxeY ,titre , etiquette1, etiquette2 )
```

La courbe que vous devriez obtenir est la suivante :



Il est étonnant de voir le croisement des courbes avant 1940. On constate qu'à partir de 1945, les courbes sont quasi horizontale. Les données avant 1945 semblent entachées d'erreur. En regardant le site et en allant dans l'onglet Documentation. <https://www.insee.fr/fr/statistiques/2540004#consulter>.

Il est possible de lire :

Avertissement : Le fichier des prénoms est établi à partir des seuls bulletins de naissance des personnes nées en France (métropole et départements d'outre-mer). En conséquence, l'exhaustivité n'est pas garantie sur toute la période, notamment pour les années antérieures à 1946.