2023.01.18 – verifying wiki page is the TW politician

pseudo code

input:

1. excel – politician name and personal information
2. folder – with 843 wiki pages

Step 1.

open excel – (Python 201) as a pandas DataFrame using read\_excel(…)

name, birthday, L1\_Division, Party

Step 2.

start with one wiki file

write a function to check if politician info is contained in that .html file

def function\_name(file\_name, politician info):

open file

get text

check if info is contained in the text – birthdate & political party

return TRUE/FALSE

save TRUE/FALSE to the pandas DataFrame as a new column

Step 3.

apply function to all 843 wiki files

Step 4.

compare Firai’s T/F with Tasha’s T/F - do they match or not?