財報爬蟲程式碼

#C:\Users\lll12\Desktop  chrom driver site

from ast import keyword

from selenium import webdriver

import time

from selenium.webdriver.common.keys import Keys

import pandas as pd

from selenium import webdriver

from selenium.webdriver.common.by import By

from selenium.webdriver.support.ui import WebDriverWait

from selenium.webdriver.support import expected\_conditions as EC

from selenium.webdriver.common.keys import Keys

from re import T

path="C:/Users/lll12/Desktop/chromedriver.exe"#電腦中chromedriver的位置

from selenium import webdriver

#打開瀏覽器的用途

def myfun(df,item):

    df[item[0]] = item[1:]

    return df

driver=webdriver.Chrome(path)

driver.get("https://www.cmoney.tw/finance/f00041.aspx?s=2330")

time.sleep(5)

items=['2849','2891','2832','2886','2809','2324','2382','2357','3231','2365']

searcher = WebDriverWait(driver, 10).until(

        EC.presence\_of\_element\_located((By.XPATH, '//\*[@id="qStockId"]')))

for item in items:

    keyword=item

    searcher.clear()

    searcher.send\_keys(keyword)

    searcher.send\_keys(Keys.RETURN)

    item1=['日期']

    item2=['營業毛利']

    item3=['營業利益']

    item4=['稅後純益']

    for i in range(2,10):

        dates=driver.find\_elements\_by\_xpath('//\*[@id="MainContent"]/ul/li/article/div[2]/div/table/tbody/tr[1]/th[%s]'%i)

        for date in dates:

            print("日期",date.text)

            item1.append(date.text)

        gpms=driver.find\_elements\_by\_xpath('//\*[@id="MainContent"]/ul/li/article/div[2]/div/table/tbody/tr[4]/td[%s]'%i)

        for gpm in gpms:

            print("營業毛利",gpm.text)

            item2.append(gpm.text)

        profits=driver.find\_elements\_by\_xpath('//\*[@id="MainContent"]/ul/li/article/div[2]/div/table/tbody/tr[6]/td[%s]'%i)

        for profit in profits:

            print("營業利益",profit.text)

            item3.append(profit.text)

        atss=driver .find\_elements\_by\_xpath('//\*[@id="MainContent"]/ul/li/article/div[2]/div/table/tbody/tr[10]/td[%s]'%i)

        for ats in atss:

            print("稅後純益",ats.text)

            item4.append(ats.text)

    df = pd.DataFrame()

    for i in [item1,item2,item3,item4]:

        df = myfun(df,i)

        print(df)

        df.to\_csv(keyword+"財報.csv",encoding='Big5',index=False)