



臺灣銀行外匯爬蟲

一、前置作業

1. 預先安裝：

```
pip install selenium pandas concurrent-log-handler
```

2. 下載 Microsoft Edge WebDriver：

前往官方下載頁面：<https://developer.microsoft.com/zh-tw/microsoft-edge/tools/webdriver/>

3. 驅動程式配置：

```
# Windows
set EDGE_DRIVER_PATH=C:\python\msedgedriver.exe

# Linux/macOS
export EDGE_DRIVER_PATH=~/.bin/msedgedriver
```

二.程式碼說明

1.瀏覽器初始化模組

```
# 使用 raw string 處理Windows路徑
service = Service(r"C:\python\msedgedriver.exe")

# 實務技巧：隱藏瀏覽器自動化標記
options = webdriver.EdgeOptions()
options.add_argument("--disable-blink-features=AutomationControlled")
driver = webdriver.Edge(service=service, options=options)
```

2.分頁處理機制

```
try:
    # 點擊第N個歷史查詢按鈕
    current_buttons[index].click()

    # 實務技巧：等待新分頁開啟
    WebDriverWait(driver, 10).until(
        lambda d: len(d.window_handles) > 1
    )

    # 切換到最新分頁
    driver.switch_to.window(driver.window_handles[-1])

except IndexError:
    print(f"⚠ 第 {index+1} 個按鈕無法點擊")
```

3.定位貨幣名稱**

```
# 使用CSS選擇器精準定位
currency_element = driver.find_element(
    By.CSS_SELECTOR,
    "div.visible-phone.print_hide > h2"
)
currency = currency_element.text.split()[0]
```

4. 表格迭代處理

```
for row in table.find_elements(By.TAG_NAME, "tr"):
    # 跳過表頭
    if "header" in row.get_attribute("class"):
        continue

    # 動態等待單行數據
    WebDriverWait(driver, 5).until(
        lambda _: len(row.find_elements(By.TAG_NAME, "td")) >= 4
    )
    # 數據提取...
```

擴充接口設計

```
class BankCrawler:
    def __init__(self, driver_path):
        # 初始化邏輯...

    def fetch_currency_list(self):
        """取得可用幣別清單"""
        # 實作細節...

    def get_history(self, currency_code, start_date, end_date):
        """取得指定幣別歷史數據"""
        # 實作細節...

    def export_csv(self, filename):
        """匯出CSV檔案"""
        # 實作細節...

# 使用範例
crawler = BankCrawler(r"C:\python\msedgedriver.exe")
crawler.get_history("USD", "2023-01-01", "2023-12-31")
crawler.export_csv("usd_2023.csv")
```