

# Forward School

**Program Code: J620-002-4:2020**

**Program Name: FRONT-END SOFTWARE DEVELOPMENT**

**Title : Case Study - Webscrapping on Trip Advisor**

**Name: Phua Yan Han**

**IC Number: 050824070059**

**Date : 4/7/23**

**Introduction : learning webScrapping**

**Conclusion : learned more efficient ways to do webScrapping**

Instruction: Use Selenium and Beautiful Soup to extract the first 5 comments and their titles from the trip advisor link.

Url: [https://www.tripadvisor.com/Airline\\_Review-d8729220-Reviews-AirAsia-AirAsia-Berhad-Malaysia](https://www.tripadvisor.com/Airline_Review-d8729220-Reviews-AirAsia-AirAsia-Berhad-Malaysia) ([https://www.tripadvisor.com/Airline\\_Review-d8729220-Reviews-AirAsia-AirAsia-Berhad-Malaysia](https://www.tripadvisor.com/Airline_Review-d8729220-Reviews-AirAsia-AirAsia-Berhad-Malaysia))

```
In [13]: from selenium import webdriver
         from bs4 import BeautifulSoup

         driver = webdriver.Chrome('C:\\Users\\Asus\\Documents\\ChromeDriver\\chromedriver\\c
```

```

In [21]: import time
from selenium import webdriver
from bs4 import BeautifulSoup
import time
import pandas as pd
from datetime import datetime
driver.get("https://www.tripadvisor.com/Airline_Review-d8729220-Reviews-AirAsia")
data = []
ammountLooped = 0
while ammountLooped <=4:
    tit=""
    desc=""
    time.sleep(5)
    soup = BeautifulSoup(driver.page_source, "html.parser")
    for card in soup.find_all('div',attrs={'class': 'WAllg _T'}):
        for title in card.find_all('div',attrs={'class': 'KgQgP MC _S b S6 H5 _a
#         print(f"title {title.text.rstrip()}")
        tit=f"title {title.text.rstrip()}"
        for description in card.find_all('a',attrs={'class': 'Qwuub'}):
#         print(f"description {description.text.rstrip()}")
        desc =f"description {description.text.rstrip()}"
        data.append([tit,desc])
    button = driver.find_element_by_class_name('ui_button.nav.next.primary')
    button.click()
    time.sleep(5)
    ammountLooped+=1
for item in data:
    print(item[0])
    print(item[1])
    print()
print(len(data))

```

title No refund from Aiaasia for flight cancellation more than 8 months  
description No refund from Aiaasia for flight cancellation more than 8 months

title Not good  
description Not good

title Don't use this airline  
description Don't use this airline

title The staffs are very very rude  
description The staffs are very very rude

title No in-flight entertainment  
description No in-flight entertainment

title Nicole and dime airline company  
description Nicole and dime airline company

title mr  
description mr

title Great experience, plane needs some refurbishing though.  
description Great experience, plane needs some refurbishing though.

title One of the worst customer service  
description One of the worst customer service

title Terrible Booking Process and Waste of Time  
description Terrible Booking Process and Waste of Time

title AirAsia is a FRAUD  
description AirAsia is a FRAUD

title Air Asia (& Hutchgo Wholeseller) is Organised Crime, Taking Advantage of Medical Emergency  
description Air Asia (& Hutchgo Wholeseller) is Organised Crime, Taking Advantage of Medical Emergency

title Horrendous app navigation that leads to 0 checked in baggage  
description Horrendous app navigation that leads to 0 checked in baggage

title They dont refund and their app or website keep breakdown, cant even book flight under airasia.  
description They dont refund and their app or website keep breakdown, cant even book flight under airasia.

title Booking nightmare  
description Booking nightmare

title Damaged luggage, dirty broken chair, broken armrest, unhelpful ground and air staff  
description Damaged luggage, dirty broken chair, broken armrest, unhelpful ground and air staff

title They wont refund  
description They wont refund

title Worse customer service ever and do not refund after canceling a flight.  
description Worse customer service ever and do not refund after canceling a flight.

title Worst airline ever  
description Worst airline ever

title Only fly with AirAsia if you have no alternative  
description Only fly with AirAsia if you have no alternative

title I got my money back. SG residents, contact CASE, it refunded my money without a Single-Word.  
description I got my money back. SG residents, contact CASE, it refunded my money without a Single-Word.

title Scammers who reschedule and refuse refunds  
description Scammers who reschedule and refuse refunds

title Bad experience in Air asia  
description Bad experience in Air asia

title Must be the worst airline  
description Must be the worst airline

title Scam artists!  
description Scam artists!

25

Reference: <https://medium.com/ymedialabs-innovation/web-scraping-using-beautiful-soup-and-selenium-for-dynamic-page-2f8ad15efe25> (<https://medium.com/ymedialabs-innovation/web-scraping-using-beautiful-soup-and-selenium-for-dynamic-page-2f8ad15efe25>)