FLIGHT PRICE PREDICTION

Submitted by:

RUPAMANANDA NANDI

ACKNOWLEDGMENT

I am over helmed in all humbleness and gratefulness to acknowledge my depth to all those who have helped me to put these ideas, well above the level of simplicity and into something concrete. I would like to express my special thanks of gratitude to my mentor who gave me the golden opportunity to do this wonderful project on the topic (Ratings Prediction), which also helped me in doing a lot of Research and I came to know about so many new things. I am really thankful to them. Any attempt at any level can ‘t be satisfactorily completed without the support and guidance of my parents and friends. I would like to thank Datatrained helped me a lot in gathering different information, collecting data and guiding me from time to time in making this project, despite of their busy schedules, they gave me different ideas in making this project unique

INTRODUCTION

- ## Business Problem Framing

  The goal of this project is to build a web scraper that will run and perform searches on flight prices with flexible dates (up to 3 days before and after the dates you select first), for a particular destination. It saves an excel with the results and sends an email with the quick stats. Obviously, the objective is to help us find the best deals!

- ## Conceptual Background of the Domain Problem

  The cheapest available ticket on a given flight gets more and less expensive over time. This usually happens as
  an attempt to maximize revenue based on -
  1. Time of purchase patterns (making sure last-minute purchases are expensive)
  2. Keeping the flight as full as they want it (raising prices on a flight which is filling up in order
  to reduce sales and hold back inventory for those expensive last-minute expensive
  purchases)

- ## Review of Literature

Firstly, to web scrap a website where Flights with price and their location is to be mentioned alongside date and time.

Then making a dataset and then Data Collection, Data Analysis, Model Building is to be done.

- ## Motivation for the Problem Undertaken

Often it comes back to better goal setting, making it more compelling (to both the conscious

and unconscious mind), being clearer about why you want to do something. NLP, is about the cause-and-effect equation.

Analytical Problem Framing

jupyter flight price 11 Last Checkpoint: 8 hours ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Trusted Python 3

Run Code

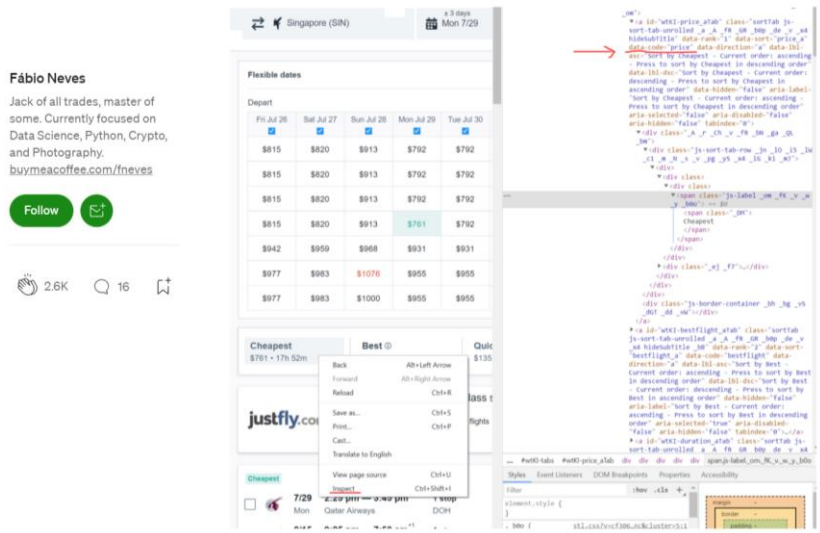
```
In [ ]: kayak = 'https://www.kayak.com/flights/LIS-SIN/2019-07-29-flexible/2019-08-15-flexible?sort=bestflight_a'
driver.get(kayak)
sleep(3)
```

```
In [ ]: def load_more():
    try:
        more_results = '//a[@class = "moreButton"]'
        driver.find_element_by_xpath(more_results).click()
        print('sleeping.....')
        sleep(randint(25,35))
    except:
        pass
```

```
In [ ]: def start_kayak(city_from, city_to, date_start, date_end):
    """City codes - it's the IATA codes!
    Date format -  YYYY-MM-DD"""

    kayak = ('https://www.kayak.com/flights/' + city_from + '-' + city_to +
             '/' + date_start + '-flexible/' + date_end + '-flexible?sort=bestflight_a')
    driver.get(kayak)
    sleep(randint(8,10))

    try:
        xp_popup_close = '//button[contains(@id,"dialog-close") and contains(@class,"Button-No-Standard-Style close ")]'
        driver.find_elements_by_xpath(xp_popup_close)[5].click()
    except Exception as e:
```



The easiest function to code is to load more results, so let's start with that. I want to maximize the amount of flights I get, without triggering the security check, so I will click once in the "load more results" button every time a page is displayed. The only thing new is the try statement, which I added because sometimes the button was not loading properly.

If it acts up with you too, simply comment it out in the start kayak function that I will show ahead.

After all this, we might as well come up with a simple loop to start using the functions we just created and keep them busy. Completed with four "fancy" prompts for you to actually write the cities and dates (the inputs). Since when we're testing, we don't want to type these variables every time, alternate it with the explicit way below them when needed.

CONCLUSION

- Key Findings and Conclusions of the Study

Creating a Dataset and finding the best possible model