TravelNest Exercise

# Airbnb Web Scraping

Code to scrape Airbnb links to fetch property details using python and json.

#Problem Statement

Please write some code that scrapes property name, property type (e.g Apartment), number of bedrooms, bathrooms and list of the amenities for the following 3 properties:

URLs:

-- https://www.airbnb.co.uk/rooms/33571268

-- https://www.airbnb.co.uk/rooms/33090114

-- https://www.airbnb.co.uk/rooms/50633275

# Approach solving the problem:

- Research and learn about web scraping

- Fetch all the data/contents from each link using html response and requests

- Finding the site contents and <div> tags in the scarped data

- Splitting data in lines

- From the scraped data, its seen that the first index of data is always the property name and thus storing the name of the property in variable.

- Then enumerating the specific strings in the site data to find the matching requirements such as number of bedrooms, bathrooms and list of amenities

- Saving all the required data in a dictionary

- Output data is displayed as console output and also saved as json file with the file name as name of the property

- Showing error message for unsuccessful scraping of url or specific data

- Making methods static and using try-exception block for better handling of exceptions

# How to Run:

Create virtual environment

Requirements: pip install -r requirements.txt

Run main.py

Outputs in the Result folder and on console.

References

https://requests.readthedocs.io/projects/requests-html/en/latest/

https://www.w3schools.com/python/default.asp

https://www.codecademy.com/courses/learn-web-scraping