**Modul 3 (Scraping Tripadvisor dan IMDB Movie)**

**Scraping Tripadvisor**

Kali ini kita akan melakukan *scraping* situs Tripadvisor untuk mendapatkan data-data seperti daftar hotel yang ada di Yogyakarta dan kegiatan atau wisata yang ada di Yogyakarta. Scraping ini akan menggunakan *library* python Beautiful Soup. Yang pertama akan dilakukan adalah melakukan scraping daftar hotel yang ada di Yogyakarta.

Yang pertama adalah, buka situs https://www.tripadvisor.com kemudian pilih Hotel dan masukkan tempat “Yogyakarta Region”, atau dapat menggunakan link berikut <https://www.tripadvisor.com/Hotels-g294230-Yogyakarta_Region_Java-Hotels.html>. Maka tampilan web tripadvisor adalah seperti gambar 1.0 di bawah ini.

Graphical user interface, website

Description automatically generated

**Gambar 1. 1** tampilan website tripadvisor

Kemudian buka IDE, dapat menggunakan *Google Collab* atau menggunakan *Visual Studio Code*. Di sini saya menggunakan *Visual Studio Code*. Kemudian setelah itu *import* beberapa *library* yang akan digunakan seperti requests, dan BeautifulSoup.

import requests

from bs4 import BeautifulSoup

import csv

kemudian, buatlah variabel yang berfungsi untuk menampung situs url, requests, objek soup, dan list kosong untuk menampung data. Di sini saya menggunakan varianel

* url = untuk menampung url
* req = untuk menampung requests
* soup = untuk menampung objek soup
* data\_hotel = list kosong untuk menampung data

dan kemudian tuliskan kode seperti di bawah ini:

url = "https://www.tripadvisor.com/Hotels-g294230-Yogyakarta\_Region\_Java-Hotels.html"

req = requests.get(url)

soup = BeautifulSoup(req.text, "html.parser")

data\_hotel = []

S

Graphical user interface, text

Description automatically generated

**Gambar 1. 2** inspect element interface website

as