Requests

Question 1:

Your challenge is to create a Python program that performs HTTP information gathering by making an HTTP GET request to a specified URL. The gathered information, including the HTML content of the web page, will be saved to a JSON file. Additionally, handle exceptions gracefully.

Expected Output:

Enter a target URL (Ctrl + C to exit): https://google.com
Information gathered from https://google.com.
Enter a target URL (Ctrl + C to exit): https://wikipedia.org
Information gathered from https://wikipedia.org.
Enter a target URL (Ctrl + C to exit):
Program terminated.

JSON file


```
{"url": "https://google.com", "html_content": "<!doctype html><html itemscope=\"\" itemtype=\"http://schema.org/WebPage\"
lang=\"en-IN\"><head><meta content=\"text/html; charset=UTF-8\" http-equiv=\"Content-Type\"><meta content=\"/images/branding/googleg/1x/
googleg_standard_color_128dp.png\" itemprop=\"image\"><title>Google</title><script nonce=\"AZa4uQwPhCy9Ufb_uduLnA\">(function(){var _g=
{kEI:'wQ6cZZTvA_m7vr0Pnb6_0AY',kEXPI:'0,798226,3,562889,4349,207,4804,1132070,870537,327244,620,380090,16114,28684,23792,12320,2815,14764,
4998,17075,38444,2872,2891,4139,7615,606,50058,10632,2614,13491,230,1014,1,16916,2652,4,17215,42402,27050,6624,7596,1,42154,2,16737,23024,
5679,1021,31121,4569,6258,23418,1249,33067,2,2,1,26632,8155,8861,14490,20506,7,1922,9779,8213,34246,20198,20137,14,82,13332,15660,18832,5122,
3030,15816,1804,7734,25,6601,2,1706,6769,12431,2414,3073,7755,87,120,8840,31542,1883,1195,549,12350,5210139,2,365,1104,117,16,5993363,2806666,
7475460,5,20540004,1008,13289,2376,43886,3,1603,3,262,3,234,3,2121276,2585,23029351,8163,10336,2709,8027,872,5800,1967,13021,4428,10576,5874,
2,2,19038,4349,1297,421,1,5886,9210,2,1299,2052,27,6,5,15,265,1280,110,4145,663,209,438,664,1767,1403,1068,1551,1157,2252,161,1110,5,2269,
2666,381,4469,207,427,8,2421,3,3622,961,149,2383,664,1590,2,20,551,1,6,2602,214,825,7,4,3299,1543,2456,3,308,250,657,128,716,92,57,107,40,270,
2271,210,1130,3366,5,132,1033,2,64,211,719,273,1974,322,403,4,1159,790,6,371,1309,14,283,300,2,137,1566,3,4,2,2,2,256,24,505,3,999,724,359,94,
273,606,591,283,247,8,647,1079,437,1,6,1988,342,2646,310,960,282,700,923,44,1121,38,600,122,12,78,1598,1,582,225,1,1,56,826,1596,133,215,170,
4,757,1217,2,2,422,238,307,1214,531,1231,833,563,83,270,2,6,2,591,5,14,218,63,3057,57,97,2,86,66,251,1270,191,328,448,1,6,431,1164,3,2,253,
2312,3,953,483,2113,711,3,21727928,218,4005,5,274',kBL:'Qwmp',kOPI:89978449};(function(){var a;(null==(a=window.google)?0:a.stvsc)?google.
kEI=_g.kEI:window.google=_g;}).call(this);})();(function(){google.sn='webhp';google.kHL='en-IN';})();(function(){\nvar h=this||self;function l
(){return void 0!==window.google&&void 0!==window.google.kOPI&&0!==window.google.kOPI?window.google.kOPI:null};var m,n=[];function p(a){for
(var b;a&&(!a.getAttribute||!(b=a.getAttribute(\"eid\")));)a=a.parentNode;return b||m}function q(a){for(var b=null;a&&(!a.getAttribute||!(b=a.
getAttribute(\"leid\")));)a=a.parentNode;return b}function r(a){/^http:/i.test(a)&&\"https:\"===window.location.protocol&&(google.ml&&google.
ml(Error(\"a\"),!1,{src:a,glmm:1}),a=\"\");return a}\nfunction t(a,b,c,d,k){var e=\"\";-1===b.search(\"&ei=\")&&(e=\"&ei=\"+p(d),-1===b.search
(\"&lei=\")&&(d=q(d))&&(e+=\"&lei=\"+d));d=\"\";var g=-1===b.search(\"&cshid=\")&&\"slh\"!==a,f=[];f.push([\"zx\",Date.now().toString()]);h.
_cshid&&g&&f.push([\"cshid\",h._cshid]);c=c();null!=c&&f.push([\"opi\",c.toString()]);for(c=0;c<f.length;c++){if(0===c||0<c)d+=\"&\";d+=f[c]
[0]+\"=\"+f[c][1]}return\"/\"+(k||\"gen_204\")+\"?atyp=i&ct=\"+String(a)+\"&cad=\"+(b+e+d)};m=google.kEI;google.getEI=p;google.getLEI=q;
```


Question 2:

Your challenge is to create a Python program that utilizes a class to perform an HTTP GET request to a specified URL, retrieve the HTTP headers, and analyze or display specific information from the headers.

Hint:

Example Headers
print(f"Content-Type: {headers.get('Content-Type', 'N/A')}")
print(f"Server: {headers.get('Server', 'N/A')}")
print(f"Date: {headers.get('Date', 'N/ A')}")

Expected Output:

Enter a target URL (Ctrl + C to exit): https://wikipedia.fr
HTTP Headers Information:
Content-Type: text/html; charset=UTF-8
Server: nginx
Date: Mon, 08 Jan 2024 15:15:47 GMT
Enter a target URL (Ctrl + C to exit):
Program terminated.

Question 3:

Your challenge is to create a Python program that utilizes the requests library to perform a simple IP geolocation lookup. The program should prompt the user to enter an IP address, and it will then make an HTTP GET request to the "ipinfo.io" API to retrieve and display basic geolocation information.

Expected Output:

Enter an IP address (Ctrl + C to exit): 8.0.0.8
IP Geolocation Information:
IP Address: 8.0.0.8
City: Broomfield
Region: Colorado
Country: US
Location: 39.8854,-105.1139
Organization: AS3356 Level 3 Parent, LLC
Timezone: America/Denver
AS (Autonomous System): N/A
Enter an IP address (Ctrl + C to exit):
Program terminated.

Question 4:

The challenge is to create a Python program with web scraping using the requests library and BeautifulSoup. To analyze a webpage, extract specific content, and perform basic data manipulation. Extract count of all headings (h1, h2, etc.). Display the first 3 paragraphs.
Expected Output:

Example output for this url: `https://en.wikipedia.org/wiki/Computer_security`

```
Total Number of Headings: 81
First 3 Paragraphs:
1.

2. Computer security, cybersecurity, digital security or information technology security (IT security) is the protectio
n of computer systems and networks from attacks by malicious actors that may result in unauthorized information disclos
ure, theft of, or damage to hardware, software, or data, as well as from the disruption or misdirection of the services
 they provide.[1][2]

3. The field is significant due to the expanded reliance on computer systems, the Internet,[3] and wireless network sta
ndards such as Bluetooth and Wi-Fi. Also, due to the growth of smart devices, including smartphones, televisions, and t
he various devices that constitute the Internet of things (IoT). Cybersecurity is one of the most significant challenge
s of the contemporary world, due to both the complexity of information systems and the societies they support. Security
 is of especially high importance for systems that govern large-scale systems with far-reaching physical effects, such
as power distribution, elections, and finance.[4][5]
```

Question 5:

The challenge is to create a Python program that utilizes the requests library to download files from the web. Students are required to download multiple files of different types and save them locally**.**

Expected Output:

File 'amazon-01.jpg' downloaded successfully.

File 'downloaded_document.pdf' downloaded successfully.

Downloaded files