

```

from
request
s
import
get

from requests.exceptions import RequestException
from contextlib import closing
from bs4 import BeautifulSoup
import re
import csv
class DataObject:
    def __init__(self):
        self.name_company=""
        self.email=""
        self.link=""
        self.address=""
        self.directornames=""
    def __str__(self):
        return("Company name: "+str(self.name_company)+" Link:" +str(self.link)+ "
Email:"+str(self.email)+ " Address: "+str(self.address)+ " directornames"+
str(self.directornames))
    def csvrow(self):
        return
[ str(self.name_company),str(self.link),str(self.email),str(self.address),str(self.directornames
)]
def simple_get(url):
    """
    Attempts to get the content at `url` by making an HTTP GET request.
    If the content-type of response is some kind of HTML/XML, return the
    text content, otherwise return None.
    """
    try:
        with closing(get(url, stream=True)) as resp:
            if is_good_response(resp):
                return resp.content
            else:
                return None
    except RequestException as e:
        log_error('Error during requests to {0} : {1}'.format(url, str(e)))
        return None
def is_good_response(resp):
    """
    Returns True if the response seems to be HTML, False otherwise.
    """
    content_type = resp.headers['Content-Type'].lower()
    return (resp.status_code == 200
            and content_type is not None

```

```

        and content_type.find('html') > -1)
def log_error(e):
    """
    It is always a good idea to log errors.
    This function just prints them, but you can
    make it do anything.
    """
    print(e)
searchterm = input("Enter search term:")
count=0
data_array=[]
getstr='https://www.zaubacorp.com/companysearchresults/'+searchterm
print(getstr)
html=simple_get(getstr)
html = BeautifulSoup(html, 'html.parser')
for tr in html.select('tr'):
    for a in tr.select('a'):
        tempObj = DataObject()
        tempObj.link=a['href']
        tempObj.name_company = a.text
        data_array.append(tempObj)
    count+=1;
csvData=[]
for tempObj in data_array:
    companypage = simple_get(tempObj.link)
    companypage = BeautifulSoup(companypage, 'html.parser')
    tempObj.email =companypage.find(text=re.compile('^([a-zA-Z0-9_.-]+)@[a-zA-Z0-9-]+\.[a-zA-Z0-9-.-.]+$'))
    tempObj.address = companypage.find('p',text=re.compile('Address:')).findNext('p').text
    dirnames = companypage.findAll('tr', id=re.compile('^package'))
    dirnames = [dir.findNext('a').text for dir in dirnames]
    tempObj.directornames = ",".join(dirnames)
    print(tempObj.name_company,tempObj.directornames)
    csvData.append(tempObj.csvrow())
with open(searchterm+'.csv', 'w') as csvFile:
    writer = csv.writer(csvFile)
    writer.writerows(csvData)
csvFile.close()

```