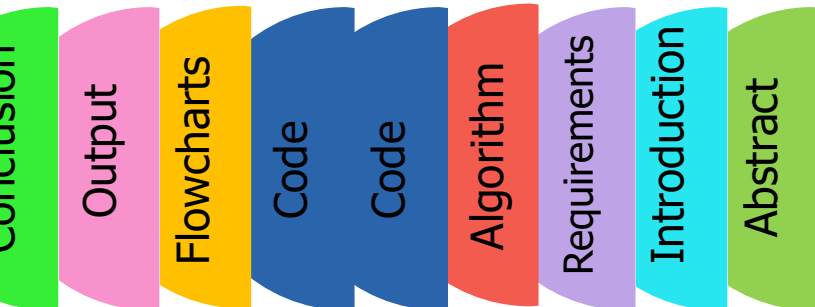


# WEB SCRAPING





## ➤ ABOUT THE PROJECT

- ABSTRACT

## ➤ AGENDA

- INTRODUCTION
- ADVANTAGES
- PROBLEM DEFINITION
- SYSTEM REQUIREMENTS
- TECHNOLOGIES USED

## ➤ IMPLEMENTATION

- TOPICS REQUIRED
- FLOWCHARTS
- ALGORITHMS

## ➤ PROCESSING

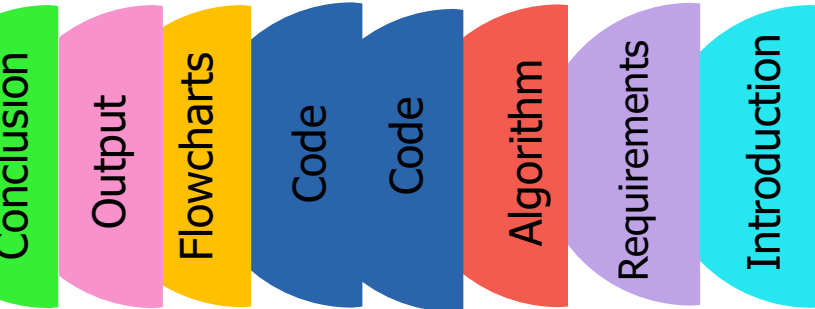
- INPUT
- OUTPUT

## ➤ CONCLUSION

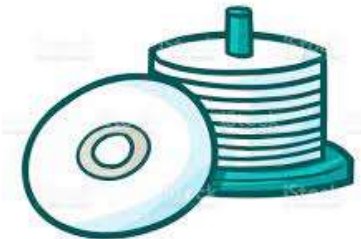


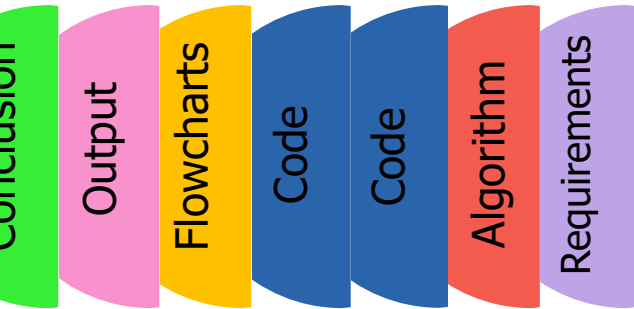
Contents



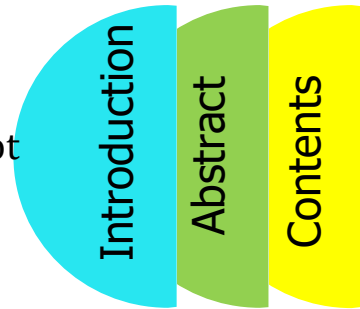


Main objective of Web Scraping is to extract information from one or many websites and process it into simple structures such as spreadsheets, database or CSV file. web scraping extracts HTML code and, with it, data stored in a database. The scraper can then replicate entire website content elsewhere.





Our main motive in this project is to scrap data from some job searching websites like Flexjobs and creating a website for job searching. In today's time of data science & engineering, it is entirely expected to gather information from sites for examination purposes, job purposes. The main reason to choose this topic is to bring different job availability details into one frame as the websites mentioned above has different prioritized job availability information but we can make a place where a person who is searching for a job can get in touch with all the available job offers being offered, it is mostly useful for the people who are looking for a job but not a particular one and that can be done by their capabilities. In this project we are going to use Beautiful soup library for scraping and parsing. First, we are going to collect the data from different job sites using Beautiful soup library and then we will store that data and using that we will create our own job website.





Programming language:

**PYTHON**

Platform used:

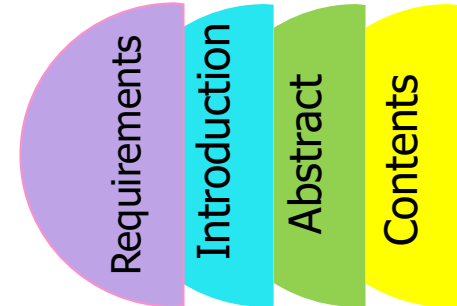
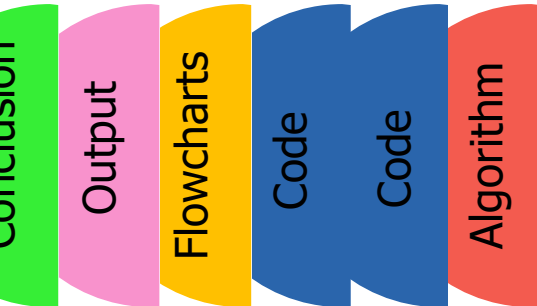
**GOOGLE COLABORATORY**

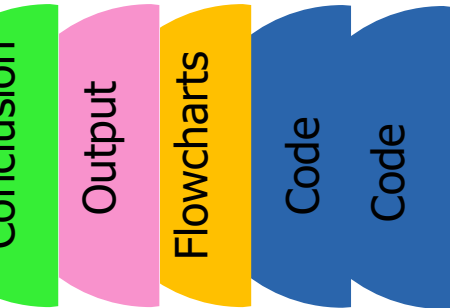
### **SOFTWARE REQUIREMENTS :**

- Language used : Python
- Operating system : Windows 11

### **HARDWARE REQUIREMENTS :**

- Hard disk : 10 MB
- Processor : Intel(IR) Core i5-5300 CP





**STEP 1:** Start.

**STEP 2:** Import Modules.

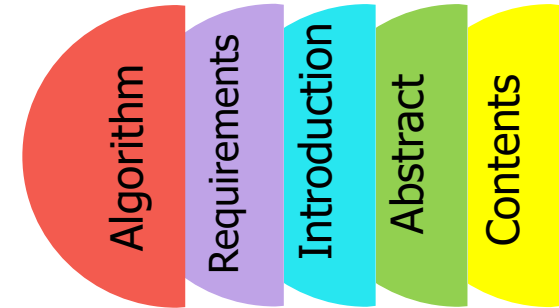
**STEP 3:** Get url using requests.

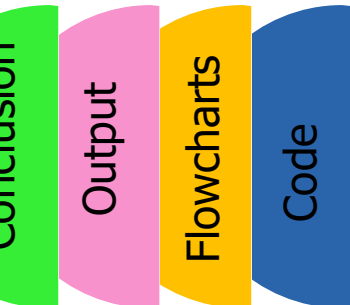
**STEP 4:** Using “beautiful soup” scrap data from given url.

**STEP 5:** From that scraped data we will collect required information.

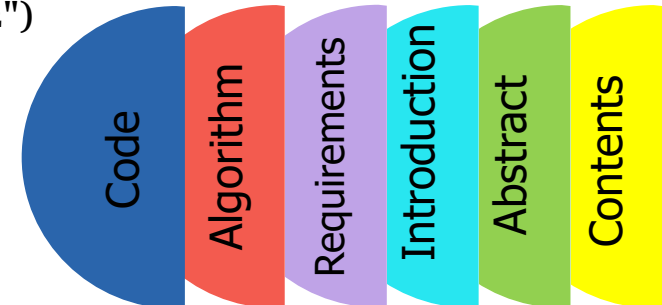
**STEP 6:** Then we convert our collected information into CSV file.

**STEP 7:** End .





```
#!/pip install beautifulsoup4
import requests
import csv
from bs4 import BeautifulSoup
def getdata(job):
    job_title=job.a.text.strip()
    job_url=job.a.get('href')
    posted_time=job.find('div',class_="job-age").text.strip()
    job_details=job.find('div',class_="row align-items-center mb-2")
    job_place=job.find('div',class_="col pe-o job-locations text-
truncate").text.strip()
    job_desc=job.find('div',class_="job-description")
    #print(job_title)
    #print(job_url)
    #print(posted_time)
    #print(job_desc)
    #print(job_place)
    information=[job_title,posted_time,job_desc.get_text(),job_place,"
https://www.flexjobs.com"+job_url]
    return information
def main():  total_information=[]
url='https://www.flexjobs.com/search?search=work+from+home
+part+time&location='
```





```
# print(len(jobs))
# while True:
    r=requests.get(url)  soup=BeautifulSoup(r.text,'html.parser')
    jobs=soup.find_all('div',class_="col-md-12 col-12")
    for job in jobs:
        information=getdata(job)
    total_information.append(information)
    with open('results.csv','w',encoding='utf-8')as f:
        writer=csv.writer(f)
        writer.writerow(['job_title','posted_time','job_desc','job_place','j
        ob_url'])
        writer.writerows(total_information)
    main()
```

Conclusion

Output

Flowcharts



Code

Code

Algorithm

Requirements

Introduction

Abstract

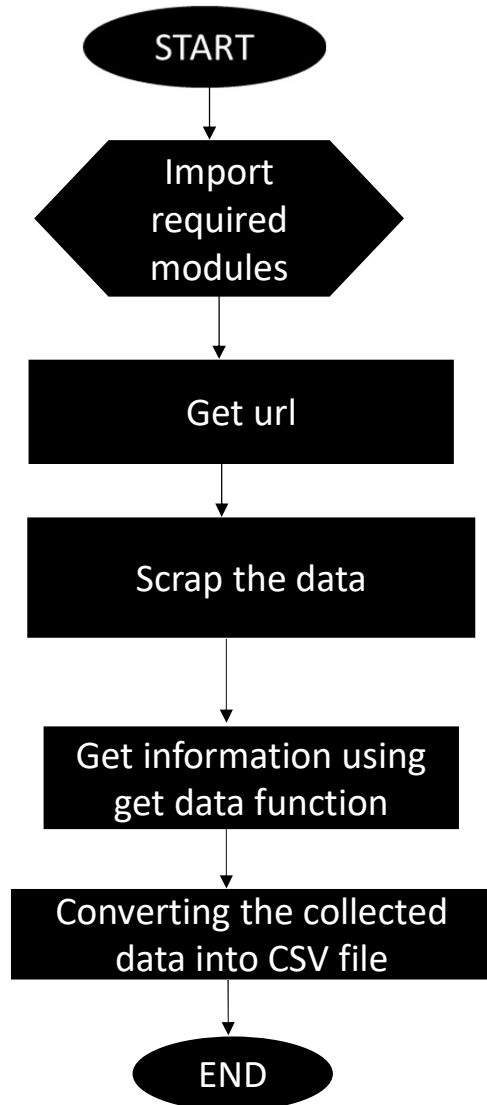
Contents







Conclusion  
Output



Flowcharts  
Code  
Algorithm  
Requirements  
Introduction  
Abstract  
Contents





Conclusion

results.xlsx

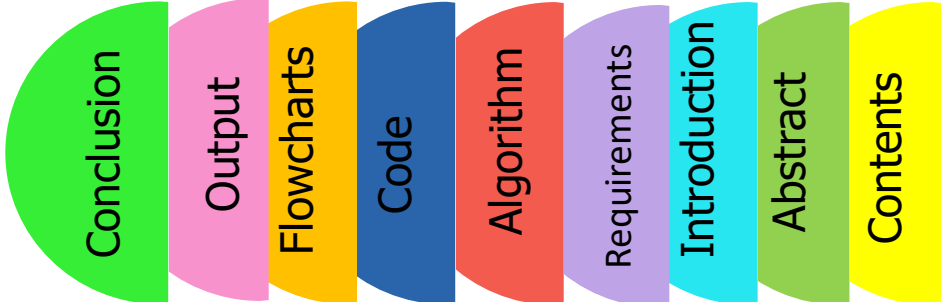
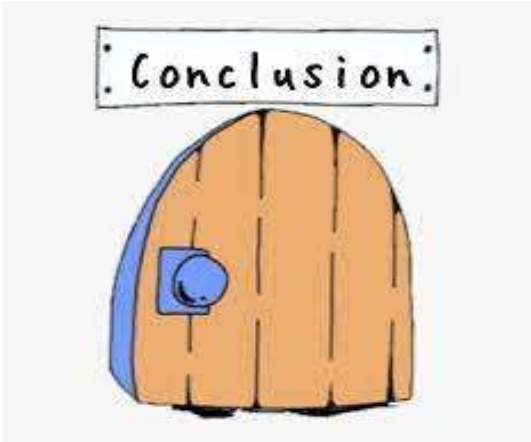
Output	Flowcharts	Code	Algorithm	Requirements	Introduction	Abstract	Contents
--------	------------	------	-----------	--------------	--------------	----------	----------

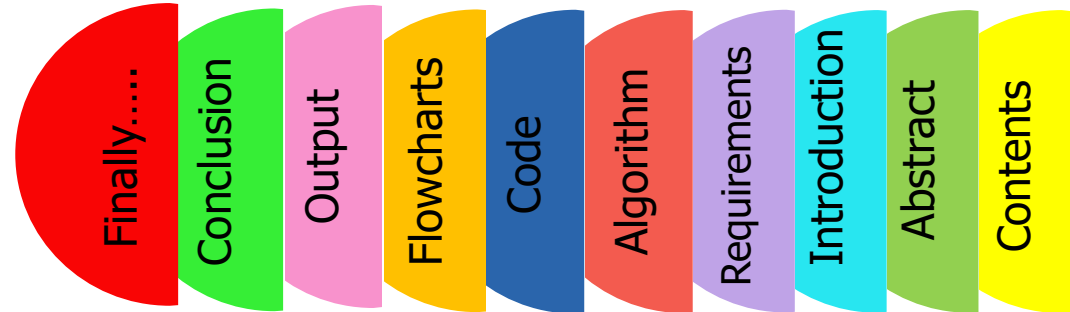




Finally.....

We now successfully scraped the information of work from home job using beautiful soup and created a csv containing the job information such as job name , job url, place from where job has been offered and job description. From the above information the user can choose the job they are interested in and can get access to the main job page by the provide url.





**Under Guidance:**

**Dr. Poonam Yadav**

**REGARDS:**

Sravya Chepuri	-	AP21110010977
Navya Kotha	-	AP21110010992
Manohar Kandi	-	AP21110010994
Renuka Kondapaneni	-	AP21110011250