

QUESTION 4.2, EXERCISE 2

INTRODUCTION TO PARSEHUB:

ParseHub is an opensource, user-friendly web scraping tool. It allows the user to extract data from websites. It doesn't require coding and provides GUI which is simple to learn.

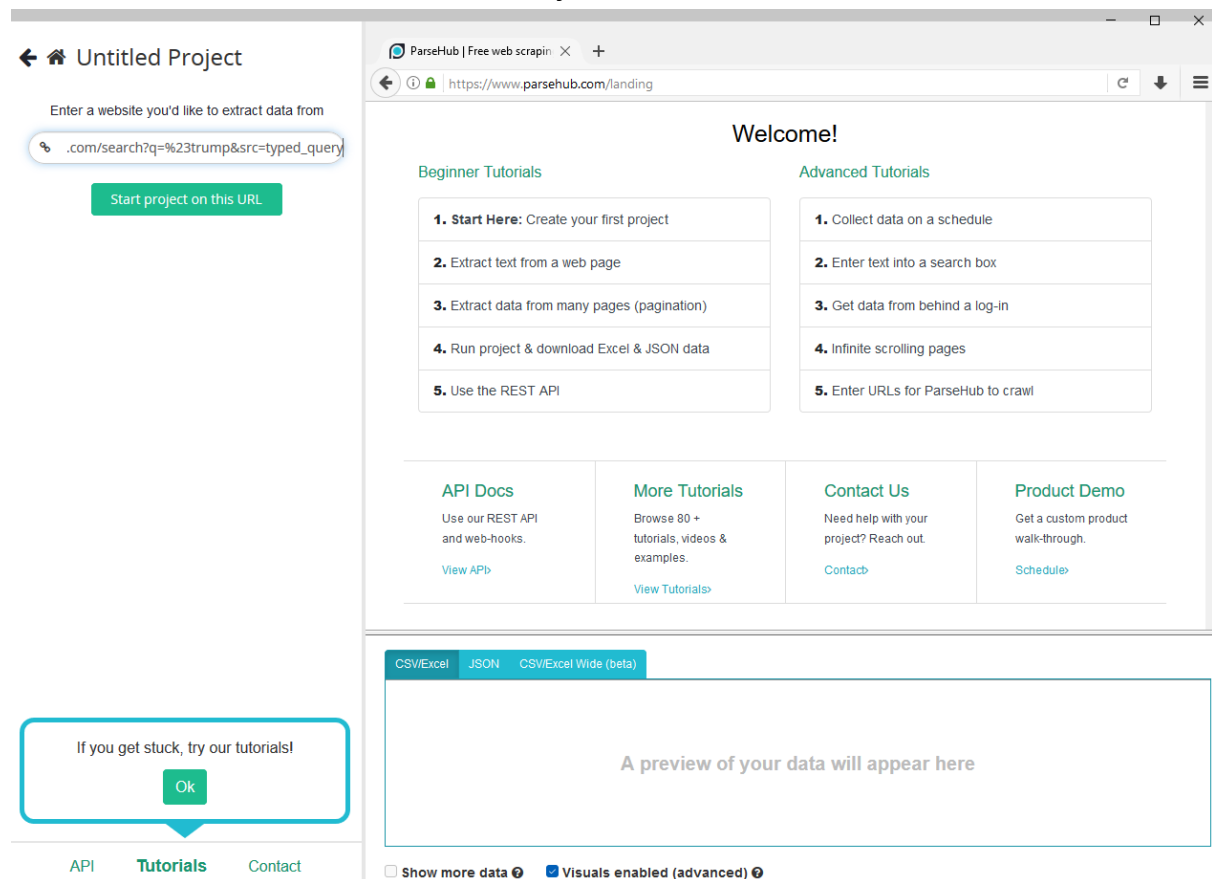
STEPS TO WEB SCRAPE WITH PARSEHUB:

Step 01: Install ParseHub

We can download and install it for free from www.parsehub.com

Step 02: Create new project:

Open the app after installation and click on "New Project" in home screen. Now enter the URL and click "Start Project".



Step 03: Select Data to Scrape:

Using the interface to select data to scrape from the webpage. Create "Select" and "Extract" commands to scrape.

en.wikip...

main_template

Select page

Select & Extract Selection1 (1)

Get Data

Selection Node: Edit

> All element s with class mw-page-container

☐ Wait up to 60 seconds for elements to appear.

API Tutorials Contact

W Indigenous peoples of the Americas

https://en.wikipedia.org/wiki/Indigenous_peoples_of_the_Americas

Demography [edit] Select Mode

Further information: Population history of Indigenous peoples of the Americas

The following table provides estimates for each country in the Americas of the populations of Indigenous people and those with partial Indigenous ancestry, each expressed as a percentage of the overall population. The total percentage obtained by adding both of these categories is also given.

Note: these categories are inconsistently defined and measured differently from country to country. Some figures are based on the results of population-wide genetic surveys while others are based on self-identification or observational estimation.

Indigenous populations of the Americas
as estimated percentage of total country's population

Country	Indigenous	Ref.	Part Indigenous	Ref.	Combined total	Ref.
North America						
Greenland	89%		%		89%	[179]
Canada	1.8%		3.6%		5.4%	[180]
Mexico	7%		83%		90%	[181]
United States	1.1%		1.8%		2.9%	[182]
Dominican Republic	%		%		%	
Grenada	~0.4%		~0%		~0.4%	[183]

CSV/Excel JSON CSV/Excel Wide (beta)

selection1

Contents

move to sidebar hide

(Top)

1Terminology

Toggle Terminology subsection

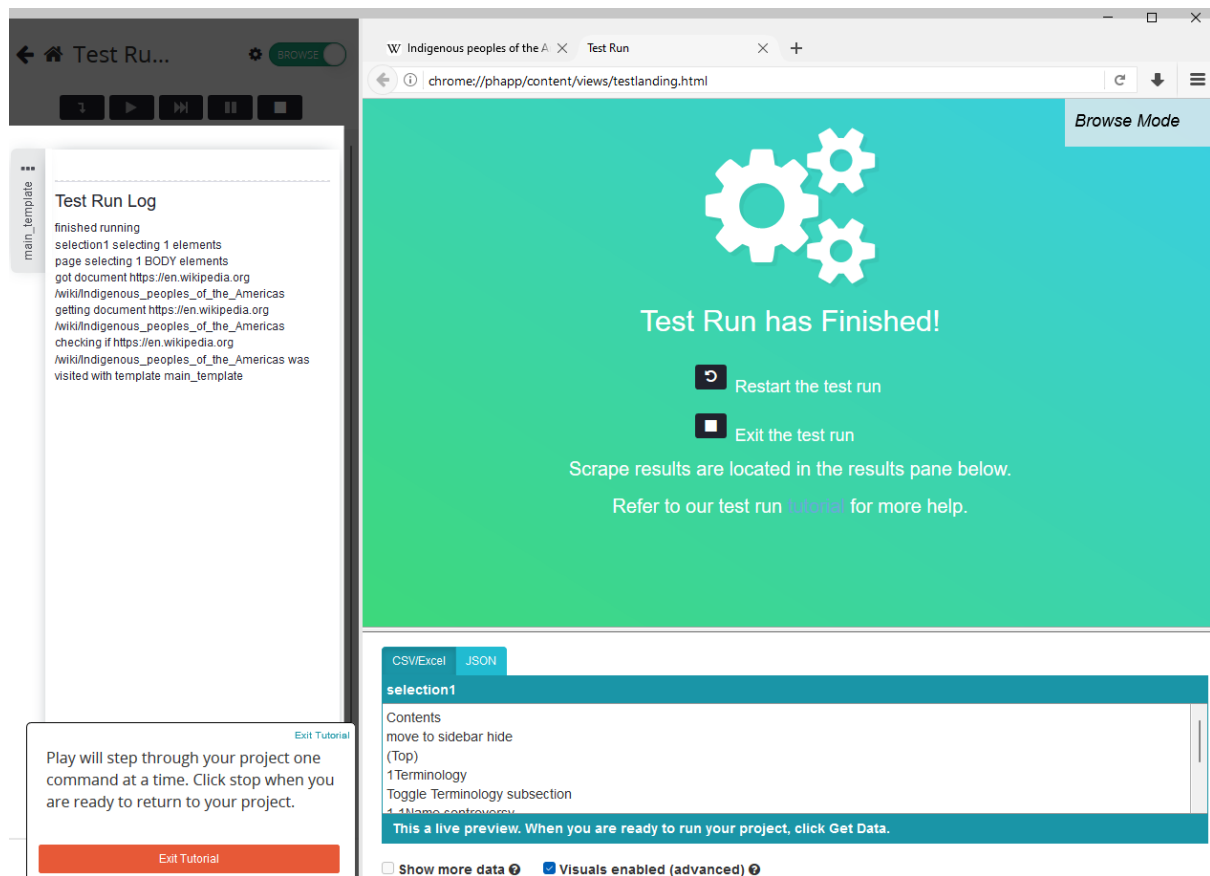
1.1Name controversy

This a live preview. When you are ready to run your project, click Get Data.

☐ Show more data ☒ Visuals enabled (advanced)

Step 04: Test your Project:

Click the run button once done selecting and test your scraping.



Step 05: Refine and Extract Data:

Click on “Get Data” and click “Run” to start scraping data.

en.wikipedia.org Project

Edit project

Your data is ready! Click on the green buttons to download.

Download Data

CSV/Excel

JSON

API

Template Name

Pages Scraped

main_template

1

All dates and times are in UTC +0000.

Empty file with no results? [Click here to fix.](#)

CSV file too big? Save the JSON file and [click here to convert to CSV.](#)

Run Details

Status

Pages

Initialized

Start Time

Finished

API Key

Project Token

Run Token

complete

1 collected

2023-09-19T07:38:14

2023-09-19T07:38:14

2023-09-19T07:38:40

tn1Supb1BeQn

tihZ7db_j1QE

toEU04T0sjYP

Settings

URL

Starting Template

Starting Value

Load Javascript

Rotate IPs

https://en.wikipedia.org/wiki/Indigenous_peoples_of_the_Americas

main_template

{}

true

false

Step 06: Export Data:

We can export the data to csv/excel or Json or APIs. Click “Download” button to save the file in local computer.

File Home Insert Page Layout Formulas Data Review View Help Tell						
Clipboard		Font			Alignment	
F69						
	A	B	C	D	E	F
61	United States 1.1% 1.8% 2.9% [182]					
62	Dominican Republic % % %					
63	Grenada ~0.4% ~0% ~0.4% [183]					
64	Haiti % % % [184]					
65	Jamaica % % %					
66	Puerto Rico 0.4% [185] 84% [186][187] 84.4%					
67	Saint Kitts and Nevis % % %					
68	Saint Lucia % % %					
69	Saint Vincent and					
70	the Grenadines 2% % % [188]					
71	Trinidad and Tobago 0.8% 88% 88.8%					
72	Country Indigenous Ref. Part Indigenous Ref. Combined total Ref.					
73	South America					
74	Argentina 2.38% [189] 27% [190][191] 27.38%					
75	Bolivia 20% 68% 88% [192]					
76	Brazil 0.4% 12% 12.4% [193]					
77	Chile 10.9% % % [194]					
78	Colombia 4.4% [195] 49% [196] 53.4%					
79	Ecuador 25% 65% 90% [197]					
80	French Guiana % % %					
81	Guyana 10.5% [198] % %					
82	Paraguay 1.7% 95% 96.7% [199]					
83	Peru 25.8% 60.2% 86% [200]					
84	Suriname 2% [201] % %					
run_results						
Ready Accessibility: Unavailable						

Overall, ParseHub simplifies the task of scraping without using coding. With it's user-friendly GUI it makes extracting data much easier. The ability to save the extracted data in various file formats is another plus point of using ParseHub.