

WEB SCRAPPING PROJECT ON ALIEXPRESS



Fuffi
Note 14 Lite

6.26 inch
1+16GB
Google play

8:20

Shopping cart icon

FUFFI Note 14 Lite Mobile phones 6.26 inch...
3 sold



Fuffi

Note 12 Pro

6.8 inch
6+128GB
5000mAh

Shopping cart icon

FUFFI Note 12 Pro Smartphone Android 6.8...
Rs. **10,321.2**



Fuffi

P5

6.6 inch
4GB 128GB
4000 mAh

Shopping cart icon

Original FUFFI P5 Mobile phones Android 4...
Rs. **8,536.34**

DESCRIPTIONS:

The AliExpress Smartphone Data Scraper is a web scraping project aimed at extracting valuable information about smartphones available on the AliExpress e-commerce platform. The project focuses on collecting data such as description of smartphones, pricing to create a comprehensive dataset for analysis and comparison.

OUTLINE:

From this site, we are going to grab the following information:

- Description
- Price

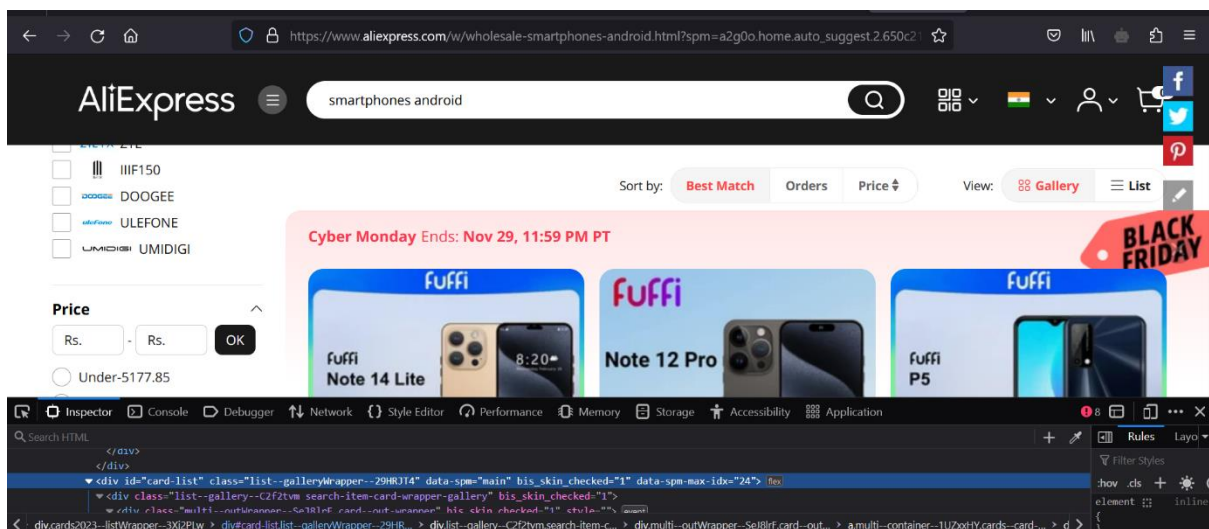
STEPS:

Choose the Website and Webpage URL

The first step is to select the website you want to scrape. We will try to extract data of smartphones from the website.

Inspect the website

Now the next step is to understand the website structure. Understand what the attributes of the elements that are of your interest are. Right click on the website to select "Inspect". This will open HTML code. Use the inspector tool to see the class name of all the elements to use in the code.



Installing the important libraries

Python has several web scrapping libraries. We will use the following libraries:

- **Requests** – for making HTTP requests to website
- **BeautifulSoup** – for parsing the HTML code
- **Pandas** – for storing the scraped data in data frame

Write the Python source code

We'll write the main python code. The code will perform the following steps:

- Using requests to send an HTTP GET requests
- Using BeautifulSoup to parse the HTML code
- Extracting the required data from the HTML code
- Store the information in a pandas dataframe

Exporting the extracted data

We'll export the data as a CSV file. We will use the pandas library. We'll use the pandas library.

BENEFITS:

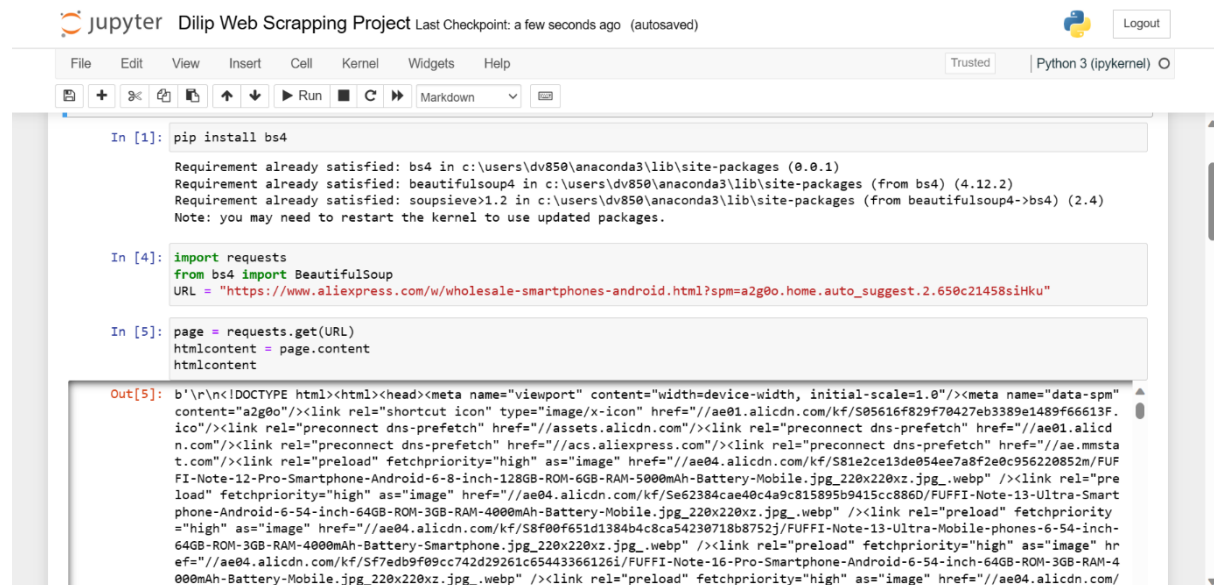
- Access to valuable data for analysis or research.
- Automation of data collection, saving time and effort.
- Stay up to date with change changes on the target websites.

RISK:

- Legal issues related to web scraping.
- Technical challenges due to website changes

AliExpress Smartphones Web Scrapping Coding:

1. Access AliExpress website by using import requests URL and BeautifulSoup library.



Jupyter Dilip Web Scrapping Project Last Checkpoint: a few seconds ago (autosaved)

```
In [1]: pip install bs4

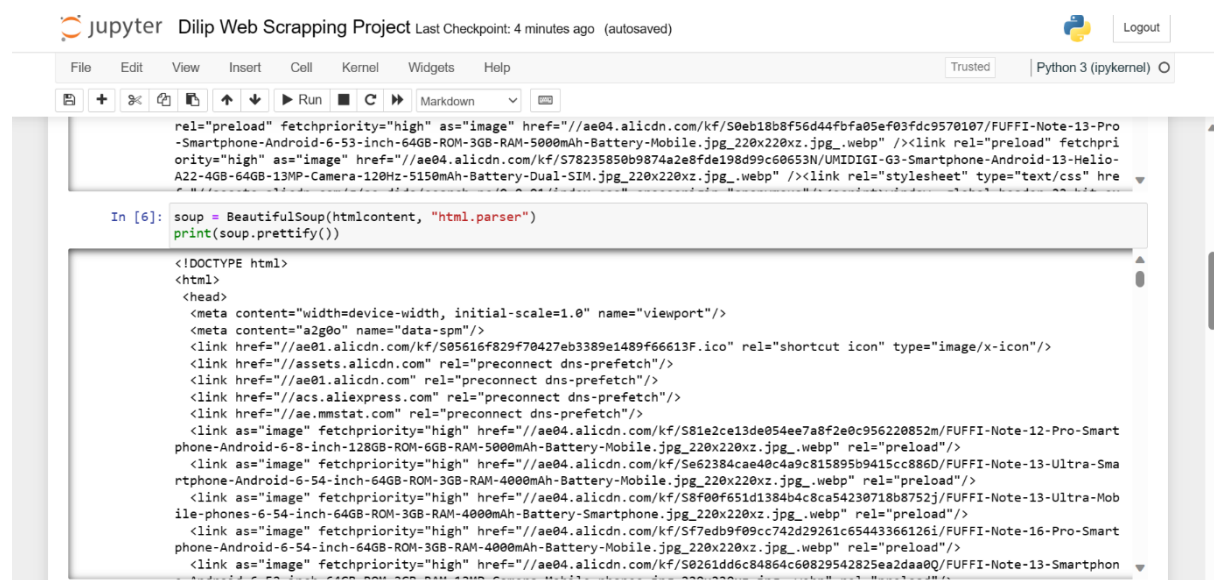
Requirement already satisfied: bs4 in c:\users\dv850\anaconda3\lib\site-packages (0.0.1)
Requirement already satisfied: beautifulsoup4 in c:\users\dv850\anaconda3\lib\site-packages (from bs4) (4.12.2)
Requirement already satisfied: soupsieve>1.2 in c:\users\dv850\anaconda3\lib\site-packages (from beautifulsoup4->bs4) (2.4)
Note: you may need to restart the kernel to use updated packages.

In [4]: import requests
from bs4 import BeautifulSoup
URL = "https://www.aliexpress.com/w/wholesale-smartphones-android.html?spm=a2g0o.home_auto_suggest.2.650c21458siHku"

In [5]: page = requests.get(URL)
htmlcontent = page.content
htmlcontent

Out[5]: b'\r\n<!DOCTYPE html><html><head><meta name="viewport" content="width=device-width, initial-scale=1.0"/><meta name="data-spm" content="a2g0o"/><link rel="shortcut icon" type="image/x-icon" href="//ae01.alicdn.com/kf/S05616f829f70427eb3389e1489f66613F.ico"/><link rel="preconnect dns-prefetch" href="//assets.alicdn.com"/><link rel="preconnect dns-prefetch" href="//ae01.alicdn.com"/><link rel="preconnect dns-prefetch" href="//acs.aliexpress.com"/><link rel="preconnect dns-prefetch" href="//ae01.alicdn.com"/><link rel="preconnect dns-prefetch" href="//ae04.alicdn.com/kf/S81e2ce13de054ee7a8f2e0c956220852m/FUFFI-Note-12-Pro-Smartphone-Android-6-8-inch-128GB-ROM-6GB-RAM-5000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp"/><link rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/kf/S62384cae40c4a9c815895b9415cc886D/FUFFI-Note-13-Ultra-Smartphone-Android-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp"/><link rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/kf/S8f00f651d1384b4c8ca54230718b8752j/FUFFI-Note-13-Ultra-Mobile-phones-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Smartphone.jpg_220x220xz.jpg_webp"/><link rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/kf/Sf7edb9f09cc742d29261c65443366126i/FUFFI-Note-16-Pro-Smartphone-Android-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp"/><link rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/
```

2. Parse the HTML:



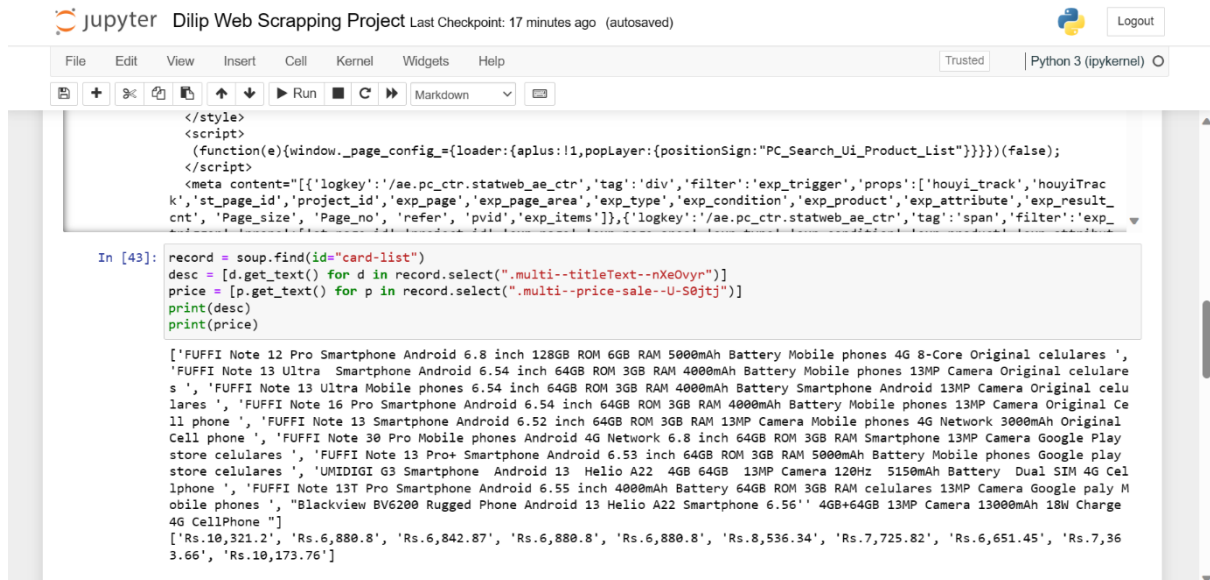
Jupyter Dilip Web Scrapping Project Last Checkpoint: 4 minutes ago (autosaved)

```
rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/kf/S0eb18b8f56d44fbfa05ef03fcd9570107/FUFFI-Note-13-Pro-Smartphone-Android-6-53-inch-64GB-ROM-3GB-RAM-5000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp"/><link rel="preload" fetchpriority="high" as="image" href="//ae04.alicdn.com/kf/S78235850b9874a2e8fde198d99c60653N/UMIDIIGI-G3-Smartphone-Android-13-Helio-A22-4GB-64GB-13MP-Camera-120Hz-5150mAh-Battery-Dual-SIM.jpg_220x220xz.jpg_webp"/><link rel="stylesheet" type="text/css" href="//ae04.alicdn.com/kf/S0eb18b8f56d44fbfa05ef03fcd9570107/FUFFI-Note-13-Pro-Smartphone-Android-6-53-inch-64GB-ROM-3GB-RAM-5000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp"/><link rel="stylesheet" type="text/css" href="//ae04.alicdn.com/kf/S78235850b9874a2e8fde198d99c60653N/UMIDIIGI-G3-Smartphone-Android-13-Helio-A22-4GB-64GB-13MP-Camera-120Hz-5150mAh-Battery-Dual-SIM.jpg_220x220xz.jpg_webp"/></html>

In [6]: soup = BeautifulSoup(htmlcontent, "html.parser")
print(soup.prettify())

<!DOCTYPE html>
<html>
<head>
<meta content="width=device-width, initial-scale=1.0" name="viewport"/>
<meta content="a2g0o" name="data-spm"/>
<link href="//ae01.alicdn.com/kf/S05616f829f70427eb3389e1489f66613F.ico" rel="shortcut icon" type="image/x-icon"/>
<link href="//assets.alicdn.com" rel="preconnect dns-prefetch"/>
<link href="//ae01.alicdn.com" rel="preconnect dns-prefetch"/>
<link href="//acs.aliexpress.com" rel="preconnect dns-prefetch"/>
<link href="//ae01.alicdn.com" rel="preconnect dns-prefetch"/>
<link as="image" fetchpriority="high" href="//ae04.alicdn.com/kf/S81e2ce13de054ee7a8f2e0c956220852m/FUFFI-Note-12-Pro-Smartphone-Android-6-8-inch-128GB-ROM-6GB-RAM-5000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp" rel="preload"/>
<link as="image" fetchpriority="high" href="//ae04.alicdn.com/kf/S62384cae40c4a9c815895b9415cc886D/FUFFI-Note-13-Ultra-Smartphone-Android-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp" rel="preload"/>
<link as="image" fetchpriority="high" href="//ae04.alicdn.com/kf/S8f00f651d1384b4c8ca54230718b8752j/FUFFI-Note-13-Ultra-Mobile-phones-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Smartphone.jpg_220x220xz.jpg_webp" rel="preload"/>
<link as="image" fetchpriority="high" href="//ae04.alicdn.com/kf/Sf7edb9f09cc742d29261c65443366126i/FUFFI-Note-16-Pro-Smartphone-Android-6-54-inch-64GB-ROM-3GB-RAM-4000mAh-Battery-Mobile.jpg_220x220xz.jpg_webp" rel="preload"/>
<link as="image" fetchpriority="high" href="//ae04.alicdn.com/kf/S0261ddc84864c60829542825ea2daa0Q/FUFFI-Note-13-Smartphon
```

3. Access AliExpress smartphones description and price using this source code.



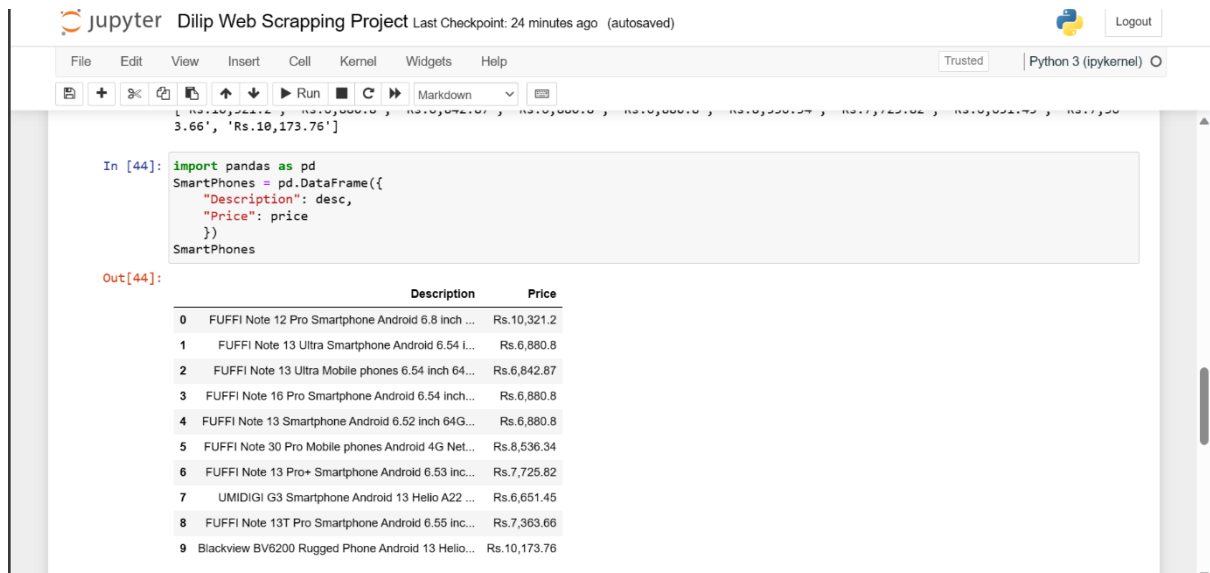
The Jupyter Notebook interface displays a code cell with HTML parsing logic using BeautifulSoup. The code finds a list of smartphone records and extracts their descriptions and prices. The output shows a list of strings, each containing a smartphone model and its price in Indian Rupees (Rs.).

```
</style>
<script>
(function(e){window._page_config={loader:{aplus:l1,popLayer:{positionSign:"PC_Search_Ui_Product_List"}}})(false);
</script>
<meta content="[{'logkey':'/ae.pc_ctr.statweb_ae_ctr','tag':'div','filter':'exp_trigger','props':['houyi_track','houyiTrac
k','st_page_id','project_id','exp_page','exp_page_area','exp_type','exp_condition','exp_product','exp_attribute','exp_result_
cnt','Page_size','Page_no','refer','pvid','exp_items']},{logkey:'/ae.pc_ctr.statweb_ae_ctr','tag':'span','filter':'exp_

In [43]: record = soup.find(id="card-list")
desc = [d.get_text() for d in record.select(".multi--titleText--nXeOvyr")]
price = [p.get_text() for p in record.select(".multi--price-sale--U-50jtj")]
print(desc)
print(price)

['FUFFI Note 12 Pro Smartphone Android 6.8 inch 128GB ROM 6GB RAM 5000mAh Battery Mobile phones 4G 8-Core Original celulares ',
'FUFFI Note 13 Ultra Smartphone Android 6.54 inch 64GB ROM 3GB RAM 4000mAh Battery Mobile phones 13MP Camera Original celulares ',
'FUFFI Note 13 Ultra Mobile phones 6.54 inch 64GB ROM 3GB RAM 4000mAh Battery Smartphone Android 13MP Camera Original celulares ',
'FUFFI Note 16 Pro Smartphone Android 6.54 inch 64GB ROM 3GB RAM 4000mAh Battery Mobile phones 13MP Camera Original Cell phone ',
'FUFFI Note 13 Smartphone Android 6.52 inch 64GB ROM 3GB RAM 13MP Camera Mobile phones 4G Network 3000mAh Original Cell phone ',
'FUFFI Note 30 Pro Mobile phones Android 4G Network 6.8 inch 64GB ROM 3GB RAM Smartphone 13MP Camera Google Play store celulares ',
'FUFFI Note 13 Pro+ Smartphone Android 6.53 inch 64GB ROM 3GB RAM 5000mAh Battery Mobile phones Google play store celulares ',
'UMIDIGI G3 Smartphone Android 13 Helio A22 4GB 64GB 13MP Camera 120Hz 5150mAh Battery Dual SIM 4G Cellphone ',
'FUFFI Note 13T Pro Smartphone Android 6.55 inch 4000mAh Battery 64GB ROM 3GB RAM celulares 13MP Camera Google paly Mobile phones ',
'Blackview BV6200 Rugged Phone Android 13 Helio A22 Smartphone 6.56'' 4GB+64GB 13MP Camera 13000mAh 18W Charge 4G CellPhone "]
['Rs.10,321.2', 'Rs.6,880.8', 'Rs.6,842.87', 'Rs.6,880.8', 'Rs.6,880.8', 'Rs.8,536.34', 'Rs.7,725.82', 'Rs.6,651.45', 'Rs.7,363.66', 'Rs.10,173.76']
```

4. Arrange all the data using import pandas library.



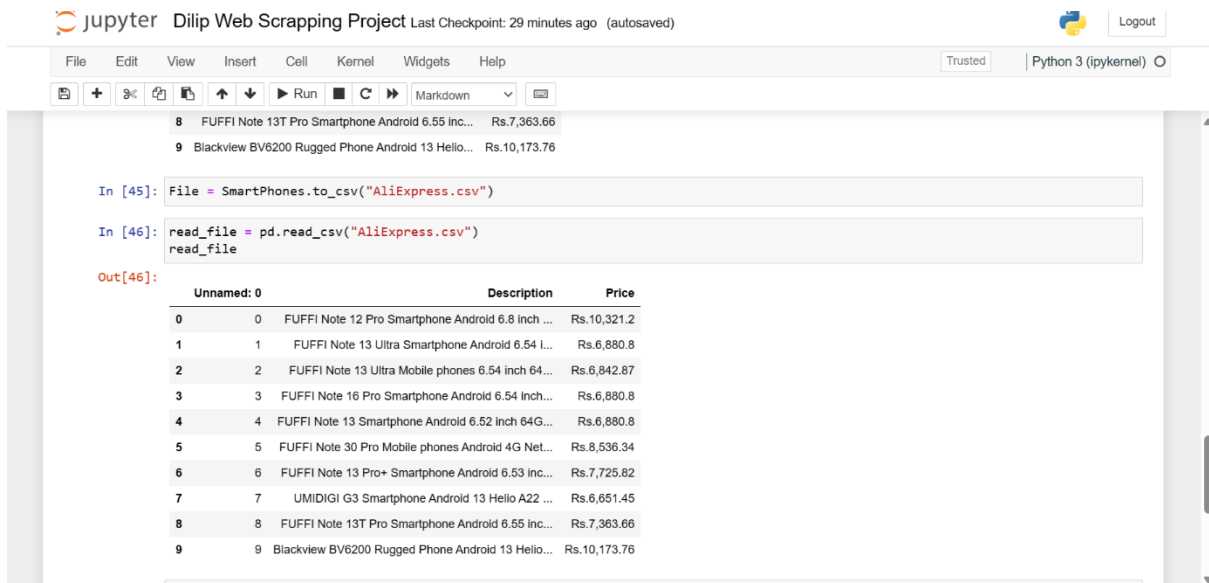
The Jupyter Notebook interface shows the next step in the data processing workflow. The code imports the pandas library and creates a DataFrame from the extracted descriptions and prices. The output is a structured table with two columns: 'Description' and 'Price'.

```
In [44]: import pandas as pd
SmartPhones = pd.DataFrame({
    "Description": desc,
    "Price": price
})
SmartPhones

Out[44]:
```

	Description	Price
0	FUFFI Note 12 Pro Smartphone Android 6.8 inch ...	Rs.10,321.2
1	FUFFI Note 13 Ultra Smartphone Android 6.54 L...	Rs.6,880.8
2	FUFFI Note 13 Ultra Mobile phones 6.54 inch 64...	Rs.6,842.87
3	FUFFI Note 16 Pro Smartphone Android 6.54 Inch...	Rs.6,880.8
4	FUFFI Note 13 Smartphone Android 6.52 inch 64G...	Rs.6,880.8
5	FUFFI Note 30 Pro Mobile phones Android 4G Net...	Rs.8,536.34
6	FUFFI Note 13 Pro+ Smartphone Android 6.53 inc...	Rs.7,725.82
7	UMIDIGI G3 Smartphone Android 13 Helio A22 ...	Rs.6,651.45
8	FUFFI Note 13T Pro Smartphone Android 6.55 inc...	Rs.7,363.66
9	Blackview BV6200 Rugged Phone Android 13 Helio...	Rs.10,173.76

5. Store above all data in CSV file and again open this file in application.



The image shows a Jupyter Notebook interface for a project titled "Dilip Web Scrapping Project". The notebook contains two code cells. The first cell, labeled "In [45]:", executes the command `File = SmartPhones.to_csv("AliExpress.csv")`. The second cell, labeled "In [46]:", executes `read_file = pd.read_csv("AliExpress.csv")` followed by `read_file`. The output of the second cell, labeled "Out[46]:", displays a DataFrame with three columns: "Unnamed: 0", "Description", and "Price". The DataFrame contains 10 rows of data, indexed from 0 to 9. The "Description" column contains truncated text of smartphone models, and the "Price" column contains their respective prices in Indian Rupees (Rs.).

```
In [45]: File = SmartPhones.to_csv("AliExpress.csv")

In [46]: read_file = pd.read_csv("AliExpress.csv")
read_file

Out[46]:
```

	Unnamed: 0	Description	Price
0	0	FUFFI Note 12 Pro Smartphone Android 6.8 Inch ...	Rs.10,321.2
1	1	FUFFI Note 13 Ultra Smartphone Android 6.54 I...	Rs.6,880.8
2	2	FUFFI Note 13 Ultra Mobile phones 6.54 inch 64...	Rs.6,842.87
3	3	FUFFI Note 16 Pro Smartphone Android 6.54 Inch...	Rs.6,880.8
4	4	FUFFI Note 13 Smartphone Android 6.52 Inch 64G...	Rs.6,880.8
5	5	FUFFI Note 30 Pro Mobile phones Android 4G Net...	Rs.8,536.34
6	6	FUFFI Note 13 Pro+ Smartphone Android 6.53 Inc...	Rs.7,725.82
7	7	UMIDIGI G3 Smartphone Android 13 Hello A22 ...	Rs.6,651.45
8	8	FUFFI Note 13T Pro Smartphone Android 6.55 inc...	Rs.7,363.66
9	9	Blackview BV6200 Rugged Phone Android 13 Hello...	Rs.10,173.76

CONCLUSION:

By doing the project we conclude that the data can then be easily read back into a DataFrame for further analysis or use in other data-related tasks. This type of workflow is commonly used for collecting and analyzing data from websites or external sources.

**THANK
YOU**
