

Problem Statement: -

To Scrape the Flipkart Website Searching for 'TV'. To retrieve all the data from the first page. Save the data in a CSV file. Analyze it a bit based on Price and Rating.

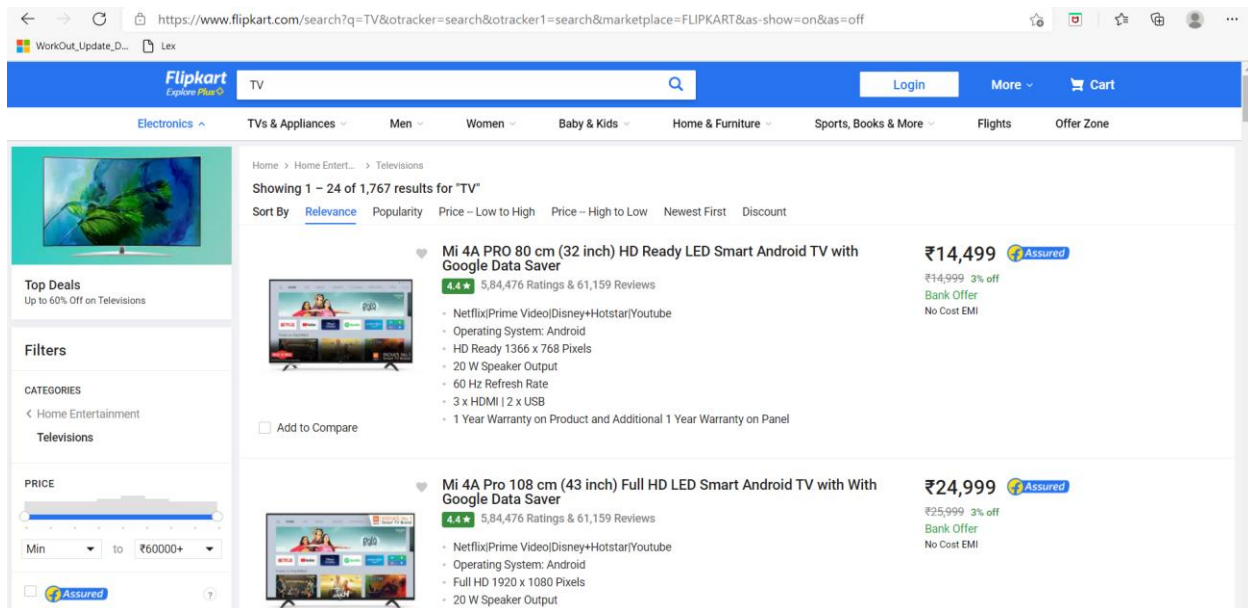
After that select a product with the following desired features: -

- Should support Netflix
- Should have a rating of 4.4 or above
- Should have discount on original price of over 30%
- Satisfying all the above condition the product should be cheapest among the bunch

Solution: -

We will go to Flipkart Website and search for a string 'TV'

We came across the below Page: -



We will be scraping this website.

Open Jupiter notebook and import the below Libraries:

Panda

Numpy

Beautiful Soup

Requests

Warnings (Optional)

Commands:-

```
import pandas as pd
```

```
import numpy as np
```

```
from bs4 import BeautifulSoup as bs
```

```
import requests
```

```
#library to deal with warning
```

```
import warnings
```

```
warnings.filterwarnings('ignore')
```

Snap:

```
In [233]: # Web Scrapping
import pandas as pd
import numpy as np
from bs4 import BeautifulSoup as bs
import requests
#library to deal with warning
import warnings
warnings.filterwarnings('ignore')
```

Next, we will define a variable 'link' with the value of the URL.

Then we will try to get the access to the link through requests

We will write the command to check if we can get

```
<Response [200]>
```

, which is a confirmation that we have access to the web page through our Jupiter Notebook.

In case it doesn't give the desired output try refreshing and rerunning the command.

If in case the Website is not provided for scraping, we cannot get

```
<Response [200]>.
```

Commands:-

```
link = "https://www.flipkart.com/search?q=TV&otracker=search&otracker1=search
&marketplace=FLIPKART&as-show=on&as=off"
```

```
page = requests.get(link)
```

```
page
```

Snap:-

```
In [52]: link = "https://www.flipkart.com/search?q=TV&otracker=search&otracker1=search&marketplace=FLIPKART&as-show=on&as=off"

In [53]: page = requests.get(link)

In [54]: page

Out[54]: <Response [200]>
```

We can check the content of the page now.

Command:-

page.content

Snap:-

```
In [55]: page.content

Out[55]: b'<!doctype html><html lang="en"><head><link href="https://rukminim1.flixcart.com" rel="preconnect"/><link rel="stylesheet" h
ref="//static-assets-web.flixcart.com/www/linchpin/fk-cp-zion/css/app.chunk.07395f.css"/><meta http-equiv="Content-type" cont
ent="text/html; charset=utf-8"/><meta http-equiv="X-UA-Compatible" content="IE=Edge"/><meta property="fb:page_id" content="10
2988293558"/><meta property="fb:admins" content="658873552,624500995,100000233612389"/><meta name="robots" content="noodp"/><
link rel="shortcut icon" href="https://static-assets-web.flixcart.com/www/promos/new/20150528-140547-favicon-retina.ico"/><li
nk type="application/opensearchdescription+xml" rel="search" href="/osdd.xml?v=2"/><meta property="og:type" content="websit
e"/><meta name="og_site_name" property="og:site_name" content="Flipkart.com"/><link rel="apple-touch-icon" sizes="57x57" href
="/apple-touch-icon-57x57.png"/><link rel="apple-touch-icon" sizes="72x72" href="/apple-touch-icon-72x72.png"/><link rel="app
le-touch-icon" sizes="114x114" href="/apple-touch-icon-114x114.png"/><link rel="apple-touch-icon" sizes="144x144" href="/appl
e-touch-icon-144x144.png"/><link rel="apple-touch-icon" href="/apple-touch-icon-57x57.png"/><meta name="twitter:card" content
="app"><meta name="twitter:site" content="@flipkart"><meta name="twitter:creator" content="@flipkart"><meta name="twitter:tit
le" content="TV- Buy Products Online at Best Price in India - All Categories | Flipkart.com"><meta name="twitter:description"
content="Shop for electronics, apparels & more using our Flipkart app Free shipping & COD."/><meta name="twitter:app:country"
content="in"><meta name="al:ios:app_name" content="Flipkart"><meta name="al:ios:app_store_id" content="742044692"><meta name
="twitter:app:name:iphone" content="Flipkart"><meta name="twitter:app:id:iphone" content="742044692"><meta name="twitter:app:
url:iphone" content="http://dl.flipkart.com/dl/home?"/><meta name="twitter:app:name:ipad" content="Flipkart"><meta name="twitt
er:app:id:ipad" content="742044692"><meta name="twitter:app:url:ipad" content="http://dl.flipkart.com/dl/home?"/><meta name="t
witter:app:name:googleplay" content="Flipkart"><meta name="twitter:app:id:googleplay" content="com.flipkart.android"><meta na
me="twitter:app:url:googleplay" content="http://dl.flipkart.com/dl/home?"/><style>#container {<br>\n\t\t\ttheight: 100%;<br>\n\t\t}</st
yle></script></html>
```

Clearly this value is in crude form.

We will beautify this using HTML parser

Command:-

soup=bs(page.content,'html.parser')

Snap:-

```
In [56]: soup=bs(page.content, 'html.parser')

In [44]: soup

Out[44]: <!DOCTYPE html>
<html lang="en"><head><link href="https://rukminim1.flixcart.com" rel="preconnect"/><link href="//static-assets-web.flixcart.com/www/linchpin/fk-cp-zion/css/app.chunk.07395f.css" rel="stylesheet"/><meta content="text/html; charset=utf-8" http-equiv="Content-type"/><meta content="IE=Edge" http-equiv="X-UA-Compatible"/><meta content="102988293558" property="fb:page_id"/><meta content="658873552,624500995,100000233612389" property="fb:admins"/><meta content="nooodp" name="robots"/><link href="https://static-assets-web.flixcart.com/www/promos/new/20150528-140547-favicon-retina.ico" rel="shortcut icon"/><link href="/osdd.xml?v=2" rel="search" type="application/opensearchdescription+xml"/><meta content="website" property="og:type"/><meta content="Flipkart.com" name="og_site_name" property="og:site_name"/><link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes="57x57"/><link href="/apple-touch-icon-72x72.png" rel="apple-touch-icon" sizes="72x72"/><link href="/apple-touch-icon-144x144.png" rel="apple-touch-icon" sizes="144x144"/><link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon"/><meta content="app" name="twitter:card"/><meta content="@flipkart" name="twitter:site"/><meta content="@flipkart" name="twitter:creator"/><meta content="TV- Buy Products Online at Best Price in India - All Categories | Flipkart.com" name="twitter:title"/><meta content="Shop for electronics, appliances & more using our Flipkart app Free shipping & COD." name="twitter:description"/><meta content="in" name="twitter:app:country"/><meta content="Flipkart" name="al:ios:app_name"/><meta content="742044692" name="al:ios:app_store_id"/><meta content="Flipkart" name="twitter:app:name:iphone"/><meta content="742044692" name="twitter:app:id:iphone"/><meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:iphone"/><meta content="Flipkart" name="twitter:app:name:ipad"/><meta content="742044692" name="twitter:app:id:ipad"/><meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:ipad"/><meta content="Flipkart" name="twitter:app:name:googleplay"/><meta content="com.flipkart.android" name="twitter:app:id:googleplay"/><meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:googleplay"/></style></head><body>
```

We can take this to a level up, by using prettify

Command:-

```
print(soup.prettify())
```

Snap:-

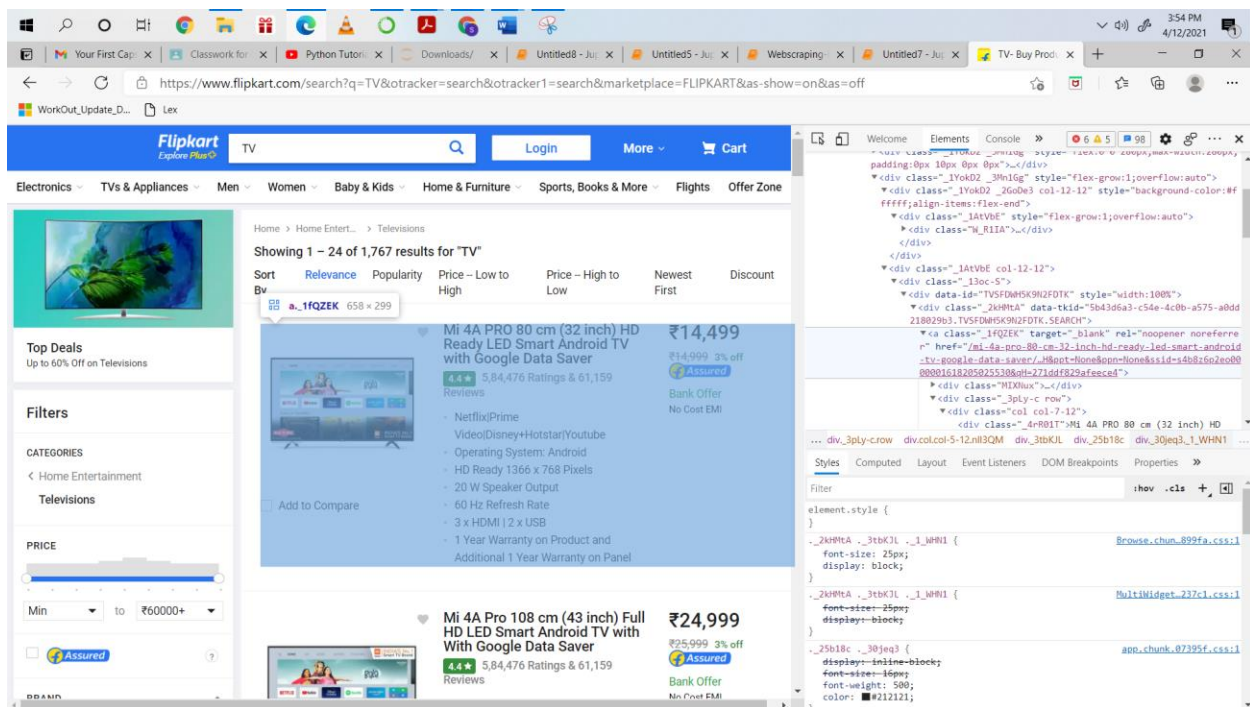
```
In [57]: print(soup.prettify())

<!DOCTYPE html>
<html lang="en">
<head>
  <link href="https://rukminim1.flixcart.com" rel="preconnect"/>
  <link href="//static-assets-web.flixcart.com/www/linchpin/fk-cp-zion/css/app.chunk.07395f.css" rel="stylesheet"/>
  <meta content="text/html; charset=utf-8" http-equiv="Content-type"/>
  <meta content="IE=Edge" http-equiv="X-UA-Compatible"/>
  <meta content="102988293558" property="fb:page_id"/>
  <meta content="658873552,624500995,100000233612389" property="fb:admins"/>
  <meta content="nooodp" name="robots"/>
  <link href="https://static-assets-web.flixcart.com/www/promos/new/20150528-140547-favicon-retina.ico" rel="shortcut icon"/>
  <link href="/osdd.xml?v=2" rel="search" type="application/opensearchdescription+xml"/>
  <meta content="website" property="og:type"/>
  <meta content="Flipkart.com" name="og_site_name" property="og:site_name"/>
  <link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes="57x57"/>
  <link href="/apple-touch-icon-72x72.png" rel="apple-touch-icon" sizes="72x72"/>
  <link href="/apple-touch-icon-144x144.png" rel="apple-touch-icon" sizes="144x144"/>
  <link href="/apple-touch-icon-57x57.png" rel="apple-touch-icon" sizes="144x144"/>
  <meta content="app" name="twitter:card"/>
  <meta content="@flipkart" name="twitter:site"/>
  <meta content="@flipkart" name="twitter:creator"/>
  <meta content="TV- Buy Products Online at Best Price in India - All Categories | Flipkart.com" name="twitter:title"/>
  <meta content="Shop for electronics, appliances & more using our Flipkart app Free shipping & COD." name="twitter:description"/>
  <meta content="in" name="twitter:app:country"/>
  <meta content="Flipkart" name="al:ios:app_name"/>
  <meta content="742044692" name="al:ios:app_store_id"/>
  <meta content="Flipkart" name="twitter:app:name:iphone"/>
  <meta content="742044692" name="twitter:app:id:iphone"/>
  <meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:iphone"/>
  <meta content="Flipkart" name="twitter:app:name:ipad"/>
  <meta content="742044692" name="twitter:app:id:ipad"/>
  <meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:ipad"/>
  <meta content="Flipkart" name="twitter:app:name:googleplay"/>
  <meta content="com.flipkart.android" name="twitter:app:id:googleplay"/>
  <meta content="http://dl.flipkart.com/dl/home?" name="twitter:app:url:googleplay"/>
</style></head>
<body>
```

Next, we will analyze the website.

We see the website has 24 products on display. We will do an inspect on the page and try to identify a block/Class which identifies, each product block uniquely.

We are doing this in order to formulate a loop approach to get the data of each product in one go.



The class I got for my execution time: -

`_2kHMTa`

We will save this as a variable `mainbox`. As to me it is kind of mainbox having information about various product in different box. Like `Box[0]`, `Box[1]` so on...

We will execute `len(mainbox)` to ensure the number of products matches.

Command:-

```
mainbox = soup.find_all("div", {"class": "_2kHMTa"})
```

```
len(mainbox)
```

Snap:-

```
In [58]: mainbox = soup.find_all("div", {"class": "_2kHMTa"})

In [62]: len(mainbox)
Out[62]: 24
```

Next we will check if we are able to retrieve the value of the fields with the class name with `mainbox[0]`.

If it works we can proceed with the For Loop.

Command:-

```
box = mainbox[0]

box.find("div",{"class": "_4rR01T"}).text

box.find("div",{"class": "_4rR01T"}).text.strip()

name = box.find("div",{"class": "_4rR01T"}).text.strip()

discounted_price= box.find("div",{"class": "_30jeq3 _1_WHN1"}).text.strip()

rating= box.find("div",{"class": "_3LWZlK"}).text.strip()

original_price = box.find("div",{"class": "_3I9_wc _27UcVY"}).text.strip()

discount_percent = box.find("div",{"class": "_3Ay6Sb"}).text.strip()

Feature= box.find("div",{"class": "fMghEO"}).text.strip()

print(name)

print(original_price)

print(discount_percent)

print(discounted_price)

print(Feature)
```

Snap:-

```
In [63]: box = mainbox[0]

In [64]: box.find("div",{"class": "_4rR01T"}).text
Out[64]: 'iFFALCON by TCL AI Powered K31 108 cm (43 inch) Ultra HD (4K) LED Smart Android TV with HDR 10'

In [50]: box.find("div",{"class": "_4rR01T"}).text.strip()
Out[50]: 'iFFALCON by TCL AI Powered K31 108 cm (43 inch) Ultra HD (4K) LED Smart Android TV with HDR 10'

In [65]: name = box.find("div",{"class": "_4rR01T"}).text.strip()

In [66]: discounted_price= box.find("div",{"class": "_30jeq3 _1_WHN1"}).text.strip()
rating= box.find("div",{"class": "_3LWZlK"}).text.strip()
original_price = box.find("div",{"class": "_3I9_wc _27UcVY"}).text.strip()
discount_percent = box.find("div",{"class": "_3Ay6Sb"}).text.strip()
Feature= box.find("div",{"class": "fMghEO"}).text.strip()

In [68]: print(name)
print(original_price)
print(discount_percent)
print(discounted_price)
print(Feature)

iFFALCON by TCL AI Powered K31 108 cm (43 inch) Ultra HD (4K) LED Smart Android TV with HDR 10
₹47,990
49% off
₹23,999
Netflix|Prime Video|Disney+Hotstar|YoutubeOperating System: AndroidUltra HD (4K) 3840 x 2160 Pixels20 W Speaker Output60 Hz Ref
resh Rate2 x HDMI | 1 x USB-A+ Grade UHD 10-bit DLED Panel1 Year Warranty on Product
```

Value fetching are correct.

Next, we will create a loop.

We will fetch the below values for each product:-

Product Name,

Price,

Ratings,

Original Price

Discount Percent

Discounted Price

Feature

Command:-

```
data_list=[]
```

```
for box in mainbox:
```

```
    temp_dict={}
```

```
    temp_dict['Product_name']=box.find("div",{"class": "_4rR01T"}).text.strip()
```

```
    if box.find("div",{"class": "_3I9_wc _27UcVY"}):
```

```
        temp_dict['Original_price']= box.find("div",{"class": "_3I9_wc _27UcVY"}).text.replace('₹', '').strip()
```

```
    else:
```

```
        temp_dict['Original_price']= box.find("div",{"class": "_30jeq3 _1_WHN1"}).text.replace('₹', '').strip()
```

```
    temp_dict['Discounted_price']=box.find("div",{"class": "_30jeq3 _1_WHN1"}).text.replace('₹', '').strip()
```

```
    if box.find("div",{"class": "_3Ay6Sb"}):
```

```
        temp_dict['Discount_percent']= box.find("div",{"class": "_3Ay6Sb"}).text.replace('% off', '').strip()
```

```
    else:
```

```
        temp_dict['Discount_percent']= '0'
```

```
    temp_dict['rating']=box.find("div",{"class": "_3LWZlK"}).text.strip()
```

```
    temp_dict['Feature']=box.find("div",{"class": "fMghEO"}).text.strip()
```

```
    data_list.append(temp_dict)
```

Snap:-

```
In [182]: data_list=[]
for box in mainbox:
    temp_dict={}
    temp_dict['Product_name']=box.find("div",{ "class": "_4rR01T" }).text.strip()
    if box.find("div",{ "class": "_3I9_wc _27UcVY" }):
        temp_dict['Original_price']= box.find("div",{ "class": "_3I9_wc _27UcVY" }).text.replace('₹', '').strip()
    else:
        temp_dict['Original_price']= box.find("div",{ "class": "_30jeq3 _1WHN1" }).text.replace('₹', '').strip()
    temp_dict['Discounted_price']=box.find("div",{ "class": "_30jeq3 _1WHN1" }).text.replace('₹', '').strip()
    if box.find("div",{ "class": "_3Ay6Sb" }):
        temp_dict['Discount_percent']= box.find("div",{ "class": "_3Ay6Sb" }).text.replace('% off', '').strip()
    else:
        temp_dict['Discount_percent']= '0'
    temp_dict['rating']=box.find("div",{ "class": "_3LWZ1K" }).text.strip()
    temp_dict['Feature']=box.find("div",{ "class": "fMghEO" }).text.strip()
    data_list.append(temp_dict)
```

We have added an if an else condition in the loop.

This is to handle the null.

We see for some scenario there is no discount.

In that the else part will populate 0% by default.

For few scenarios we do not have the Original Price Mentioned.

In that case our Discounted Price Becomes Original Price.

In the loop above we create a Data_List

And temporary a dictionary to store the value.

We are appending the Data_list with the dictionary after the loop.

This can be done in many different ways, I have gone with an approach which I found easy.

Snap:-

```
In [183]: data_list
Out[183]: [{'Product_name': 'iFFALCON by TCL AI Powered K31 108 cm (43 inch) Ultra HD (4K) LED Smart Android TV with HDR 10',
  'Original_price': '47,990',
  'Discounted_price': '23,999',
  'Discount_percent': '49',
  'rating': '4.4',
  'Feature': 'Netflix|Prime Video|Disney+Hotstar|YoutubeOperating System: AndroidUltra HD (4K) 3840 x 2160 Pixels20 W Speaker Output60 Hz Refresh Rate2 x HDMI | 1 x USB4+ Grade UHD 10-bit DLED Panel1 Year Warranty on Product'},
 {'Product_name': 'SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV 2020 Edition with Voice Search',
  'Original_price': '20,900',
  'Discounted_price': '17,999',
  'Discount_percent': '13',
  'rating': '4.4',
  'Feature': 'Netflix|Disney+Hotstar|YoutubeOperating System: TizenHD Ready 1366 x 768 Pixels20 W Speaker Output60 Hz Refresh Rate2 x HDMI | 1 x USB1 Year Comprehensive Warranty on Product and 1 Year Additional warranty on Panel'},
 {'Product_name': 'Mi 4A PRO 80 cm (32 inch) HD Ready LED Smart Android TV with Google Data Saver',
  'Original_price': '14,999',
  'Discounted_price': '14,499',
  'Discount_percent': '3',
  'rating': '4.4',
  'Feature': 'Netflix|Prime Video|Disney+Hotstar|YoutubeOperating System: AndroidHD Ready 1366 x 768 Pixels20 W Speaker Output60 Hz Refresh Rate2 x HDMI | 1 x USB1 Year Comprehensive Warranty on Product and 1 Year Additional warranty on Panel'}
```


Next we will convert the Data_list to Data Frame.

Command:-

```
df=pd.DataFrame(data_list)
```

df

Snap:-

```
In [184]: df=pd.DataFrame(data_list)
```

```
In [185]: df
```

Out[185]:

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	IFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23,999	49	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
1	SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV ...	20,900	17,999	13	4.4	Netflix Disney+ Hotstar YoutubeOperating System...
2	Mi 4A PRO 80 cm (32 inch) HD Ready LED Smart A...	14,999	14,499	3	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
3	LG 80 cm (32 inch) HD Ready LED Smart TV 2020 ...	21,990	16,999	22	4.3	Netflix Prime Video Disney+ Hotstar YoutubeOper...
4	Mi 4A Pro 108 cm (43 inch) Full HD LED Smart A...	25,999	24,999	3	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
5	Mi 4X 108 cm (43 inch) Ultra HD (4K) LED Smart...	28,999	27,999	3	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
6	Mi 4X 125.7 cm (50 inch) Ultra HD (4K) LED Sma...	34,999	34,999	0	4.4	Netflix Disney+ Hotstar YoutubeOperating System...
7	OnePlus Y Series 108 cm (43 inch) Full HD LED ...	29,999	25,999	13	4.3	Netflix Prime Video Disney+ Hotstar YoutubeOper...
8	Mi 4A 100 cm (40 inch) Full HD LED Smart Andro...	24,999	21,999	12	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
9	LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV	52,990	34,999	33	4.4	Netflix Disney+ Hotstar YoutubeOperating System...
10	Mi 4A Horizon Edition 80 cm (32 inch) HD Ready...	15,999	15,999	0	4.4	Netflix Prime Video Disney+ Hotstar YoutubeOper...
11	realme 80 cm (32 inch) HD Ready LED Smart Andr...	17,999	15,999	11	4.3	Netflix Prime Video Disney+ Hotstar YoutubeOper...
12	realme 108 cm (43 inch) Full HD LED Smart Andr...	25,999	24,999	3	4.3	Netflix Prime Video Disney+ Hotstar YoutubeOper...
13	SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma...	1,99,900	1,29,999	34	4.5	Netflix Disney+ Hotstar YoutubeOperating System...

We have highlighted the two scenario that we have handles.

Alternatively, we can import the whole data directly without if else, and perform a drop.

However, dropping a product just because it doesn't have Rating/Discount Percent may not work for other scenarios and is not recommended.

We will analyze the data frame:-

```
df.info()
```

```
In [148]: df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 24 entries, 0 to 23
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  --
0   Product_name     24 non-null     object
1   Original_price    22 non-null     object
2   Discounted_price  24 non-null     object
3   Discount_percent  22 non-null     object
4   rating           24 non-null     object
5   Feature          24 non-null     object
dtypes: object(6)
memory usage: 1.2+ KB
```

We will copy this info in a csv File using Panda:-

Command:-

```
df.to_csv('WebScrapingTV.csv', index=False)
```

After this command the csv file will generated in the same folder where the ipynb file is saved.

We can read/view the df by following command:-

```
df1=pd.read_csv('WebScrapingTV.csv')
```

```
df1
```

Snap:

```
In [235]: df1=pd.read_csv('WebScrapingTV.csv')
```

```
In [236]: df1
```

```
Out[236]:
```

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	IFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23,999	49	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
1	SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV ...	20,900	17,999	13	4.4	Netflix Disney+Hotstar YoutubeOper...
2	Mi 4A PRO 80 cm (32 inch) HD Ready LED Smart A...	14,999	14,499	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
3	LG 80 cm (32 inch) HD Ready LED Smart TV 2020 ...	21,990	16,999	22	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
4	Mi 4A Pro 108 cm (43 inch) Full HD LED Smart A...	25,999	24,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
5	Mi 4X 108 cm (43 inch) Ultra HD (4K) LED Smart...	28,999	27,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
6	Mi 4X 125.7 cm (50 inch) Ultra HD (4K) LED Sma...	34,999	34,999	0	4.4	Netflix Disney+Hotstar YoutubeOperating System...
7	OnePlus Y Series 108 cm (43 inch) Full HD LED ...	29,999	25,999	13	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
8	Mi 4A 100 cm (40 inch) Full HD LED Smart Andro...	24,999	21,999	12	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
9	LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV	52,990	34,999	33	4.4	Netflix Disney+Hotstar YoutubeOperating System...
10	Mi 4A Horizon Edition 80 cm (32 inch) HD Ready...	15,999	15,999	0	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
11	realme 80 cm (32 inch) HD Ready LED Smart Andr...	17,999	15,999	11	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
12	realme 108 cm (43 inch) Full HD LED Smart Andr...	25,999	24,999	3	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
13	SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma...	1,99,900	1,29,999	34	4.5	Netflix Disney+Hotstar YoutubeOperating System...

We can do some Univariate analysis through below commands:-

Plot Histograms of Price and Rating

```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

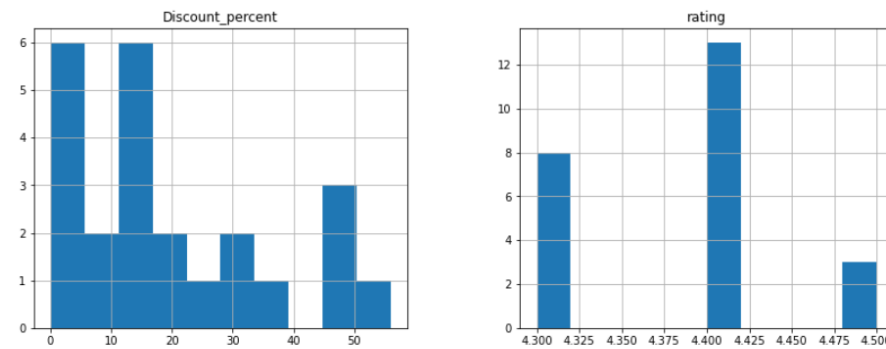
```
df1.hist(figsize=(14,5))
```

```
plt.show()
```

Snap:-

```
In [237]: # Plot Histograms of Price and Rating
import seaborn as sns
import matplotlib.pyplot as plt

df1.hist(figsize=(14,5))
plt.show()
```



Here we see in the first graph,

No of products against the discount.

Example: Total 6 products are having discounts from 0-5(percent) so on.

And in the 2nd Graph we see the no of products against the rating.

Example: Total 8 products are having rating in-between 4.3 to 4.325.

We can do more univariate/bivariate analysis but let's focus on identifying the desired product.

Featured Desired as per the problem statement: -

- Should support Netflix
- Should have a rating of 4.4 or above
- Should have discount on original price of over 30%
- Satisfying all the above condition the product should be cheapest among the bunch

To filter out we can apply all the filter together, but lets apply filter one by one to understand.

Command: -

```
Filter1_df=df[df['Feature'].str.contains('Netflix',na=False)]
```

Filter1_df

Explanation: -

This will filter records where feature contains string 'Netflix'.

Since pretty much all the records has Netflix, Filter one doesn't helps much and all the records are moved to data frame Filter_df

Snap:-

```
In [242]: Filter1_df=df[df['Feature'].str.contains('Netflix',na=False)]
          Filter1_df
```

Out[242]:

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	IFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23,999	49	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
1	SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV ...	20,900	17,999	13	4.4	Netflix Disney+Hotstar YoutubeOperating System...
2	Mi 4A PRO 80 cm (32 inch) HD Ready LED Smart A...	14,999	14,499	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
3	LG 80 cm (32 inch) HD Ready LED Smart TV 2020 ...	21,990	16,999	22	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
4	Mi 4A Pro 108 cm (43 inch) Full HD LED Smart A...	25,999	24,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
5	Mi 4X 108 cm (43 inch) Ultra HD (4K) LED Smart...	28,999	27,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
6	Mi 4X 125.7 cm (50 inch) Ultra HD (4K) LED Sma...	34,999	34,999	0	4.4	Netflix Disney+Hotstar YoutubeOperating System...
7	OnePlus Y Series 108 cm (43 inch) Full HD LED ...	29,999	25,999	13	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
8	Mi 4A 100 cm (40 inch) Full HD LED Smart Andro...	24,999	21,999	12	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
9	LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV	52,990	34,999	33	4.4	Netflix Disney+Hotstar YoutubeOperating System...
10	Mi 4A Horizon Edition 80 cm (32 inch) HD Ready...	15,999	15,999	0	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
11	realme 80 cm (32 inch) HD Ready LED Smart Andr...	17,999	15,999	11	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
12	realme 108 cm (43 inch) Full HD LED Smart Andr...	25,999	24,999	3	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
13	SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma...	1,99,900	1,29,999	34	4.5	Netflix Disney+Hotstar YoutubeOperating System...
14	OnePlus Y Series 80 cm (32 inch) HD Ready LED ...	19,999	15,499	22	4.3	Netflix Prime Video Disney+Hotstar YoutubeOper...
15	LG 108 cm (43 inch) Full HD LED Smart TV 2020 ...	40,990	30,999	24	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...

Applying 2nd Filter:-

Command:-

```
Filter2_df=Filter1_df[Filter1_df['rating']>='4.4']
```

Filter2_df

Explanation:

This filter applied will take records where rating are greater or equal to 4.4

Snap:-

```
In [243]: Filter2_df=Filter1_df[Filter1_df['rating']>='4.4']
          Filter2_df
```

Out[243]:

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	IFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23,999	49	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
1	SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV ...	20,900	17,999	13	4.4	Netflix Disney+Hotstar YoutubeOperating System...
2	MI 4A PRO 80 cm (32 inch) HD Ready LED Smart A...	14,999	14,499	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
4	MI 4A Pro 108 cm (43 inch) Full HD LED Smart A...	25,999	24,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
5	MI 4X 108 cm (43 inch) Ultra HD (4K) LED Smart...	28,999	27,999	3	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
6	MI 4X 125.7 cm (50 inch) Ultra HD (4K) LED Sma...	34,999	34,999	0	4.4	Netflix Disney+Hotstar YoutubeOperating System...
8	MI 4A 100 cm (40 inch) Full HD LED Smart Andro...	24,999	21,999	12	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
9	LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV	52,990	34,999	33	4.4	Netflix Disney+Hotstar YoutubeOperating System...
10	MI 4A Horizon Edition 80 cm (32 inch) HD Ready...	15,999	15,999	0	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
13	SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma...	1,99,900	1,29,999	34	4.5	Netflix Disney+Hotstar YoutubeOperating System...
15	LG 108 cm (43 inch) Full HD LED Smart TV 2020 ...	40,990	30,999	24	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
17	SAMSUNG 80 cm (32 inch) HD Ready LED Smart TV ...	19,900	17,499	12	4.4	Netflix Disney+Hotstar YoutubeOperating System...
18	MI 4X 138.8 cm (55 inch) Ultra HD (4K) LED Sma...	44,999	39,999	11	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
19	Iffalcon 138.6 cm (55 inch) QLED Ultra HD (4K)...	1,26,990	54,999	56	4.5	Netflix Disney+Hotstar YoutubeOperating System...
22	SAMSUNG 108 cm (43 inch) Full HD LED Smart TV	38,900	32,811	15	4.4	Netflix Disney+Hotstar YoutubeOperating System...

Filter 3

Command:-

```
Filter3_df=Filter2_df[Filter2_df['Discount_percent']>'30']
```

Filter3_df

Explanation:

This filter gives product which have Discount_percent>30

Snap:-

```
In [244]: Filter3_df=Filter2_df[Filter2_df['Discount_percent']>'30']
          Filter3_df
```

Out[244]:

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	IFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23,999	49	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...
9	LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV	52,990	34,999	33	4.4	Netflix Disney+Hotstar YoutubeOperating System...
13	SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma...	1,99,900	1,29,999	34	4.5	Netflix Disney+Hotstar YoutubeOperating System...
19	Iffalcon 138.6 cm (55 inch) QLED Ultra HD (4K)...	1,26,990	54,999	56	4.5	Netflix Disney+Hotstar YoutubeOperating System...

We now have 4 records; among them we need to select the record with lowest price.

Command:-

```
Filter4_df=Filter3_df
```

```
Filter4_df['Discounted_price']=Filter4_df['Discounted_price'].str.replace(',','')
```

```
Filter4_df["Discounted_price"]=pd.to_numeric(Filter4_df["Discounted_price"])
```

```
Filter4_df.info
```

```
Filter4_df=Filter4_df.sort_values(by='Discounted_price')
```

```
Filter4_df.head(1)
```

Explanation-

So the main idea here is we will sort the DF based on discounted_price.

Before that we need to replace “,”, and do a typecasting.

Else without type casting we may see a scenario where

```
Value 129999
```

May be placed before

```
Value 23999
```

Hence above Codes.

Snap:-

```
In [246]: Filter4_df=Filter3_df
Filter4_df['Discounted_price']=Filter4_df['Discounted_price'].str.replace(',','')

Filter4_df["Discounted_price"]=pd.to_numeric(Filter4_df["Discounted_price"])
Filter4_df.info
```

```
Out[246]: <bound method DataFrame.info of
0 iFFALCON by TCL AI Powered K31 108 cm (43 inch... 47,990
9 LG 108 cm (43 inch) Ultra HD (4K) LED Smart TV 52,990
13 SAMSUNG 163 cm (65 inch) Ultra HD (4K) LED Sma... 1,99,900
19 Iffalcon 138.6 cm (55 inch) QLED Ultra HD (4K)... 1,26,990

Discounted_price Discount_percent rating \
0 23999 49 4.4
9 34999 33 4.4
13 129999 34 4.5
19 54999 56 4.5

Feature
0 Netflix|Prime Video|Disney+Hotstar|YoutubeOper...
9 Netflix|Disney+Hotstar|YoutubeOperating System...
13 Netflix|Disney+Hotstar|YoutubeOperating System...
19 Netflix|Disney+Hotstar|YoutubeOperating System... >
```

```
In [250]: Filter4_df=Filter4_df.sort_values(by='Discounted_price')
Filter4_df.head(1)
```

```
Out[250]:
```

	Product_name	Original_price	Discounted_price	Discount_percent	rating	Feature
0	iFFALCON by TCL AI Powered K31 108 cm (43 inch...	47,990	23999	49	4.4	Netflix Prime Video Disney+Hotstar YoutubeOper...

We got the product:-

iFFALCON by TCL AI Powered K31 108 cm

This product

- Supports Netflix
- Has a rating of 4.4
- Is having a discount of 49% over the Marked Price.
- Is the Cheapest among all options, INR 23999