

RALPH LAUREN

Ralph Lauren is a traded fashion company, that produces products ranging from the mid-range to the luxury segments. You will be asked to scrape images from this website of Ralph Lauren.

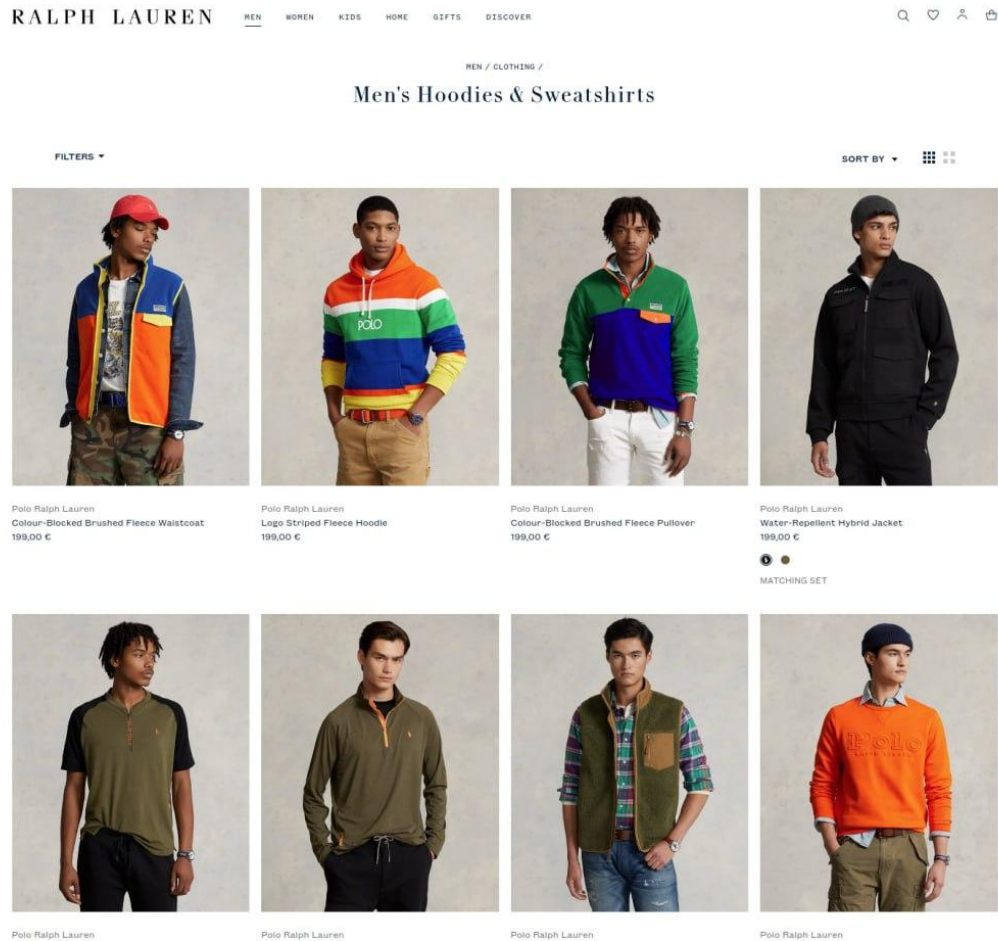
Check out this link below.

<https://www.ralphlauren.nl/en/men/clothing/hoodies-sweatshirts/10204?webcat=men%7Cclothing%7Cmen-clothing-hoodies-sweatshirts>

You will find the brands of certain men cloths from Raulph Lauren firm. Under the subtitle “Polo Ralph Lauren” you will see several items. Each item represents a particular clothing, consisting of images of a person wearing a certain cloth and the cloth itself. Check out, as well, the specific items of cloths clicking on each cloths and being rendered to that specific cloth page.

Imagine yourself as a data engineer, whose first task is to scrape all the items and later do some image processing operations on them.

Part 1



1. Scrap all the images of the people wearing a particular cloth.



2. Scrap all the images of the cloths (not worn on somebody)



3. Scrap all the colors of the particular image of a person and its corresponding cloths. (This is a bonus, totally optional)



and other colors (e.g. all colors that are available for a particular item)

Once you finish task 1, separate the scrapped images into 2 different folders

1. a person in particular cloths
2. the cloths (not worn on people).

Your program should look like this

```
def scrape_images(url)
    # saves the images in particular folders
```

Scrape all images and upload them to github/gdrive, share with us. There should be around 154 items each containing person and the separate cloth.

Note: if you could not accomplish task 1 by means of coding, you still can do the part 2. Instead download some images without code (click download button on the website) and proceed to part 2. However you will lose points from part 1 in this case. Though this gives you a chance to proceed with part 2, anyway.

Part 2

Write a python/any other code that converts all the scrapped images into .jpg format

Go to

<https://colab.research.google.com/drive/18RenTYhuPVip9SHdMLn-vnK0K57B--um#scrollTo=D0h2Y-oOCnXJ>

And use your cloths images to extract the cloth mask

1. Randomly choose an image of a cloth (not worn on a person) in your dataset. The model on the colab is pretrained to generate cloth masks. Browse your images and download the generated mask (the black and white output).
2. Write a program that given the image of the cloth (original) and the mask of the cloth (taken from the website), outputs the original image of the cloth with the background of the cloth in blue color.

Program should look like

```
def preprocess(cloth, cloth_mask):
    # returns/saves blue_background_cloth
```

This program should work for all the cloths in your dataset, however you don't have to separately do this for all items. Just check if it works for a couple of images and that's it, you are done!

Rules:

1. After you manage to do the task, upload it in a github/any other repository and share it in a public link with us.
2. Upload the dataset that you scrapped with 2 different folders described above.
3. Create report file and describe in it, which parts you could and couldn't accomplish with

the task. Explain your logic.