

### Case Study

Hello,

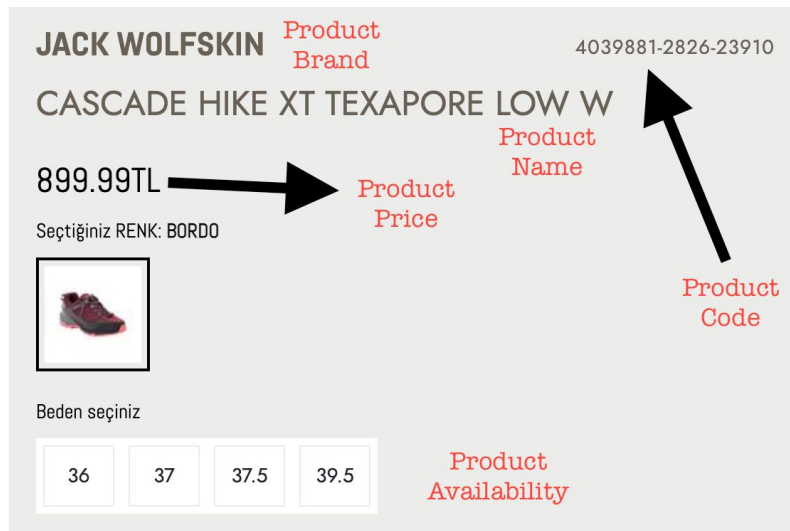
Dear candidate, in this project you will build a simple scraper for one of our customer.

### Description of Assignment

As you know AnalyticaHouse is a consultant on digital marketing area. We have bunch of customers taking consulting services from us. Lets say one of our customers demanded a Product Report which consisting of product code, product name, product brand, product availability( as a percent ) , product price for all of products they have.

Your mission is build a crawling system which will go to those URLs (There is an xlsx file at the attachment), scrape necessary metrics and write to Excel as seperate columns. At the below picture all metrics can be seen.

Note: You should append [www.spx.com.tr](http://www.spx.com.tr) to every path for reaching page.



The screenshot shows a product page for Jack Wolfskin shoes. Annotations with arrows point to specific elements:

- Product Brand:** JACK WOLFSKIN
- Product Name:** CASCADE HIKE XT TEXAPORE LOW W
- Product Price:** 899.99TL
- Product Code:** 4039881-2826-23910
- Product Availability:** Indicated by the size selection buttons (36, 37, 37.5, 39.5) at the bottom.

Other visible text includes: "Seçtiğiniz RENK: BORDO" and "Beden seçiniz".

## Additional Questions

1. If I'd have 10000 urls that I should visit, then it takes hours to finish. What can we make to fasten this process?
2. Please briefly explain what is API and how it works.

## Sending the project

To demonstrate your work, you should follow these steps.

1. Use git and upload your source code(Python) on to your GitHub page as a public repository.
2. Add a README.md file with the project description into your repo. Project description should include following items.
  - a. Used tech-stack ( library or etc.)
  - b. A brief description of the challenges you face.
  - c. What did you learn from this project?
  - d. Answers of the additional questions.

**Deadline is 12 February 2021, 23:59.**

That's it. Good Luck and Take Care! 🍀