Immo Eliza Scraping Project

Jacques Declercq

Heba Elabrak

Leonor Drummond

Table of content

- How we did it
- Who did what
- What went wrong
- How we solved it

How we did it

- With selenium we built a script to gather the links with the pages of houses and apartments;

- With beautifulSoup we tried to collect the information needed from each property page;

- Information was saved in a cvs file.

Who did what

- Jacques: Selenium;

- Leonor: Scraping with BeautifulSoup, Powerpoint;

- Heba: Scraping with BeautifulSoup; GitHub; Readme file.

What went wrong

- **Captcha:** Access to the soup was denied.

- **Understanding BeautifulSoup vs Selenium:** First we had the idea of only using BeautifulSoup!

- **Scraping a property page:** Lack of knowledge and experience;

- **Making selenium and BeautifulSoup collect the information:** We only got collected the information from the last property.

- **Not able to make the cvs file**

How we solved it

Captcha problem - Not solved

- We built a dictionary with the "Google cache" and referrer;

- Slowing down request time.

Understanding BeautifulSoup vs Selenium – Solved

- Each have their functionalities;

- The advice from the coaches was important to understand the steps.

Scraping a property page – Partly solved

- Tried to go to deep on the information;

- No being able to segment things properly;

- Elements on the top of the page were easily scraped.

Making selenium and BeautifulSoup collect the information - Not solved

Thank you for your attention!