

Scraping Coding Project

With this project we will assess your ability to solve problems on your own and present a fully working solution for a real use case.

You will be given a chance to show your skills in scraping development: code, design, autonomy, pragmatism, rigor, etc.

The project is about implementing a scraper for the website **<https://www.transavia.com/fr-FR>**.

Specifications

Your goal is to write a script that given a departure date, a return date, an origin and a destination returns all the possible flights and their pricing. You can limit origin destination to **PARIS & AMSTERDAM or any available route**.

For instance, for the query: PARIS | AMSTERDAM | today() + 15 days | today() + 22 days the script will have to return something like the following list:

- HV5194 | ORY | AMS | 21:45 | 23:00 | basic | 98€
- HV5193 | AMS | ORY | 19:45 | 21:05 | basic | 88€

See the attached image to see the original landing page on which theses prices have been scraped.

Optional

Besides the features described above, you may add the following improvements:

- Use a proxy to obfuscate your IP: <https://www.xroxy.com/proxy-country/>, tor browser ...
- Allow the user to enter any origin / destination tuple as input
- Try running you script on a large amount of departure dates – are you banned ? Can you prevent this ban ? Can you speed up the process with parallel jobs ?

Technical specifications

We expect a script that take command line arguments as input but you may use whatever language/ framework you want.

Project presentation

You have been given a certain amount of days to complete this project. We expect you to send us a link to a public repository (github, gitlab, bitbucket...) before the agreed limit date. We will then review your code and, if this first internal review is positive, will ask you to come present the results of your work to us.

The presentation will allow you to explain your technical decisions, and the problems you encountered.

It is essential that your script works and return the correct prices, focus on this aspect first.

← → ↻ 🔒 https://www.transavia.com/fr-FR/reservez-un-vol/vols/rechercher/ 🔍 ☆ 📄 ⋮

🏠

Réservez un vol

EUR

Modifier votre recherche ▼

✈️ Vol aller

Paris (Orly) - Amsterdam (Schiphol)

Panorama du mois 📅

dim. 7

à partir de 163 €

lun. 8

Pas de vols disponibles

mar. 9

Pas de vols disponibles

juil. 2019

mer. 10

Pas de vols disponibles

jeu. 11

à partir de 98 €

ven. 12

Pas de vols disponibles

sam. 13

Pas de vols disponibles

←

→

Thu 11 Jul 2019

🕒 21:45 ✈️ 23:00

Numéro de vol HV5194

2 sièges disponibles à ce prix

98 €

Sélectionner

Évaluation de TripAdvisor pour Paris (Orly) - Amsterdam (Schiphol) avec Transavia: 🏠 🌳 🌳 🌳 🌳 🌳

Pas de frais de réservation ! Les horaires sont indiqués en heures locales. Les prix affichés peuvent varier en fonction de la disponibilité du tarif. Le prix définitif du billet s'affiche dès que vous commencez la procédure de paiement.

✈️ Vol retour

Amsterdam (Schiphol) - Paris (Orly)

Panorama du mois 📅

mer. 17

Pas de vols disponibles

jeu. 18

à partir de 98 €

ven. 19

Pas de vols disponibles

juil. 2019

sam. 20

Pas de vols disponibles

dim. 21

à partir de 88 €

lun. 22

Pas de vols disponibles

mar. 23

Pas de vols disponibles

←

→

Thu 18 Jul 2019

🕒 19:45 ✈️ 21:05

Numéro de vol HV5193

3 sièges disponibles à ce prix

98 €

Sélectionner

Évaluation de TripAdvisor pour Amsterdam (Schiphol) - Paris (Orly) avec Transavia: 🏠 🌳 🌳 🌳 🌳 🌳

Pas de frais de réservation ! Les horaires sont indiqués en heures locales. Les prix affichés peuvent varier en fonction de la disponibilité du tarif. Le prix définitif du billet s'affiche dès que vous commencez la procédure de paiement.

Paris (Orly)
Pas de vol sélectionné

↔ Amsterdam (Schiphol)
Pas de vol sélectionné

👤 1

Détails ^

Total 0,00 €

Suivant →

Mon avis