

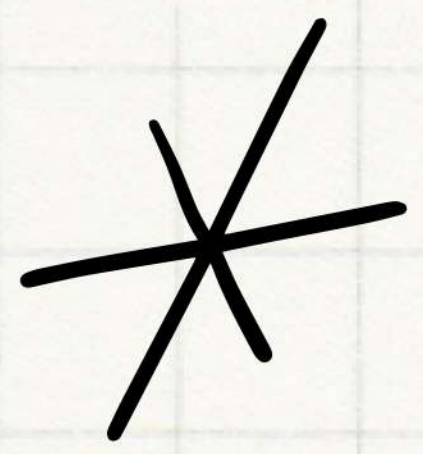


Web scraping

Anne & Lou



Kesako ?



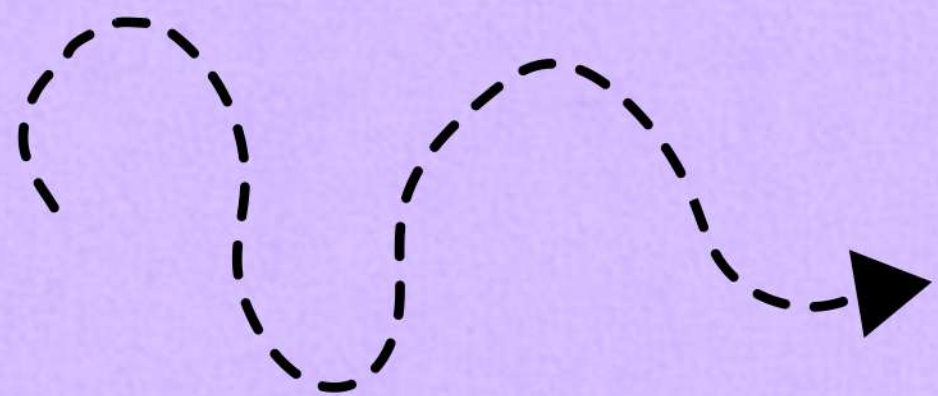
En tant que data Analyst et Développeur d'application IA, votre rôle sera de récupérer des données sur les offres d'emploi afin d'analyser les tendances de recrutement sur ce marché. Etapes:

1- Collecte de données : Développer des scripts en python pour récupérer le contenu HTML des pages web.

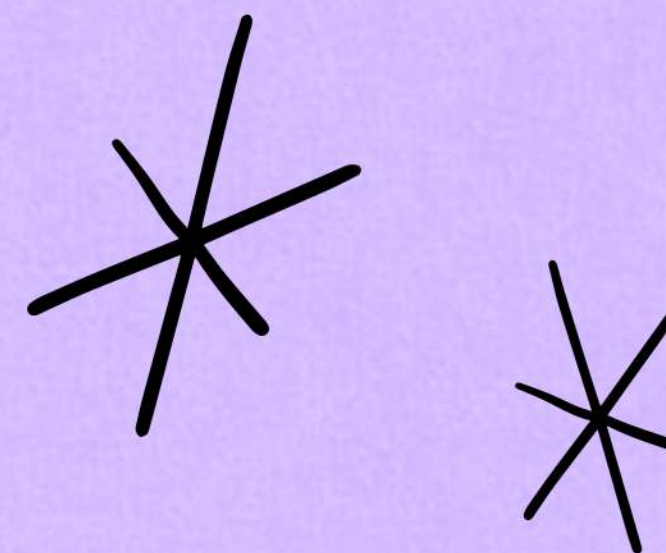
2- Nettoyage et préparation des données : nettoyer les valeurs, convertir les types de données si nécessaire, gérer les valeurs manquantes, etc.

3- Stockage des données : Proposer une méthode et un schéma de stockage adaptée.





Scraping



Indeed ayant décidé de nous faire honneur d'un 403.
Nous sommes partie sur le scraping de Météo Job.

Indeed Local

Récupération de toutes les
informations pertinentes

```
import pandas as pd
import requests
from bs4 import BeautifulSoup
```

```
def extract_job_data(url):
    response = requests.get(url)
    if response.status_code == 200:
        htmlData = response.content
        soup = BeautifulSoup(htmlData, 'html.parser')

        job_listings = soup.find_all('div', class_='truncated-text-fix')
        jobsdata = []
        for job in job_listings:
            title = job.find('h2').text.strip() if job.find('h2') else ''
            company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
            types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
            localisation = job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0').text.strip() if job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0') else ''
            date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
            salaire = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''

            jobsdata.append({"Titre de l'offre": title, "Société": company, "Lieu": localisation,
                            "Type d'offre": types, "Date d'actualisation": date,
                            "Salaires": salaire})

        df_jobs = pd.DataFrame(jobsdata)
        df_jobs = df_jobs.iloc[:,1]
        return df_jobs
    else:
        return None
```

Exemple d'utilisation de la fonction

```
url = "https://www.meteojob.com/jobs?what=data&where=Lyon"
```

```
df_jobs = extract_job_data(url)
```

```
if df_jobs is not None:
    display(df_jobs)
```

```
else:
    print("Une erreur s'est produite lors de la récupération des données.")
```

	Titre de l'offre	Société	Lieu	Type d'offre	Date d'actualisation	Salaires
0	Architecte Data (H/F)	Gif	Lyon 07 (69)	CDI	Il y a 8 jours	
1	Architecte Data (H/F)	Gif			Il y a 8 jours	
2	DATA Engineer Cloud	Le Mercato de l'Emploi	Lyon (69)	CDI	Il y a 3 jours Nouveau	
3	DATA Engineer Cloud	Le Mercato de l'Emploi			Il y a 3 jours Nouveau	
4	Data Scientist (h/f)	LHH Recruitment Solutions	Villeurbanne (69)	CDI	Il y a 13 jours	40 000 € - 46 000 €
5	Data Scientist (h/f)	LHH Recruitment Solutions			Il y a 13 jours	
6	Data Controller Comptable H/F	Talents AEC	Lyon 09 (69)	CDI	Il y a 4 jours Nouveau	25 000 € - 35 000 € par an
7	Data Controller Comptable H/F	Talents AEC			Il y a 4 jours Nouveau	
8	Data Ingénieur Talend H/F	Amiltone	Lyon (69)	CDI	Il y a 18 jours	28 000 € - 38 000 € par an
9	Data Ingénieur Talend H/F	Amiltone			Il y a 18 jours	
10	Data Engineer F/H	SYNERGIE	Chassieu (69)	CDD - 6 mois	Hier Nouveau	30 000 € - 35 000 € par an
11	Data Engineer F/H	SYNERGIE			Hier Nouveau	
12	(H/F) Gestionnaire Master Data SAP	ADSEARCH	Chassieu (69)	Intérim - 6 mois	Il y a 4 heures Nouveau	30 000 € - 50 000 €
13	(H/F) Gestionnaire Master Data SAP	ADSEARCH			Il y a 4 heures Nouveau	
14	Data scientist en alternance (H/F)	OpenClassrooms	Solaize (69)	Alternance / Apprentissage	Hier Nouveau	905 € - 1 600 € par mois
15	Data scientist en alternance (H/F)	OpenClassrooms			Hier Nouveau	
16	(H/F) Développeur Data BI	ADSEARCH	Saint-Priest (69)	CDI	Il y a 9 jours	40 000 € - 50 000 €
17	(H/F) Développeur Data BI	ADSEARCH			Il y a 9 jours	
18	GESTIONNAIRE MASTER DATA (F/H)	Randstad	Saint-Genis-Laval (69)	Intérim - 6 mois	Il y a 7 jours	À partir de 2 200 € par mois
19	GESTIONNAIRE MASTER DATA (F/H)	Randstad			Il y a 7 jours	
20	Alternance - Product Owner data F/H	OpenClassrooms	Lyon (69)	Alternance / Apprentissage	Hier Nouveau	905 € - 1 600 € par mois
21	Alternance - Product Owner data F/H	OpenClassrooms			Hier Nouveau	
22	PMO DATA STAGE (H-F)	CESI	Villeurbanne (69)	Stage	Il y a 13 jours	

MétéoJob Online

Récupération de toutes les
informations pertinentes

```
import pandas as pd
import requests
from bs4 import BeautifulSoup

def scrape_job_listings(url, pages):
    jobs_data = []

    for i in range(2, pages + 5):
        scroll_url = f"{url}&page={i}"
        response = requests.get(scroll_url)

        if response.status_code == 200:
            htmlData = response.content
            soup = BeautifulSoup(htmlData, 'html.parser')
            job_listings = soup.find_all('div', class_='truncated-text-fix')

            for job in job_listings:
                title = job.find('h2').text.strip() if job.find('h2') else ''
                company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
                types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
                date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
                salary = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''
                jobs_data.append({"Titre": title, "Company": company, "Type_de_contrat": types, "date_de_publication": date, "salaire": salary})

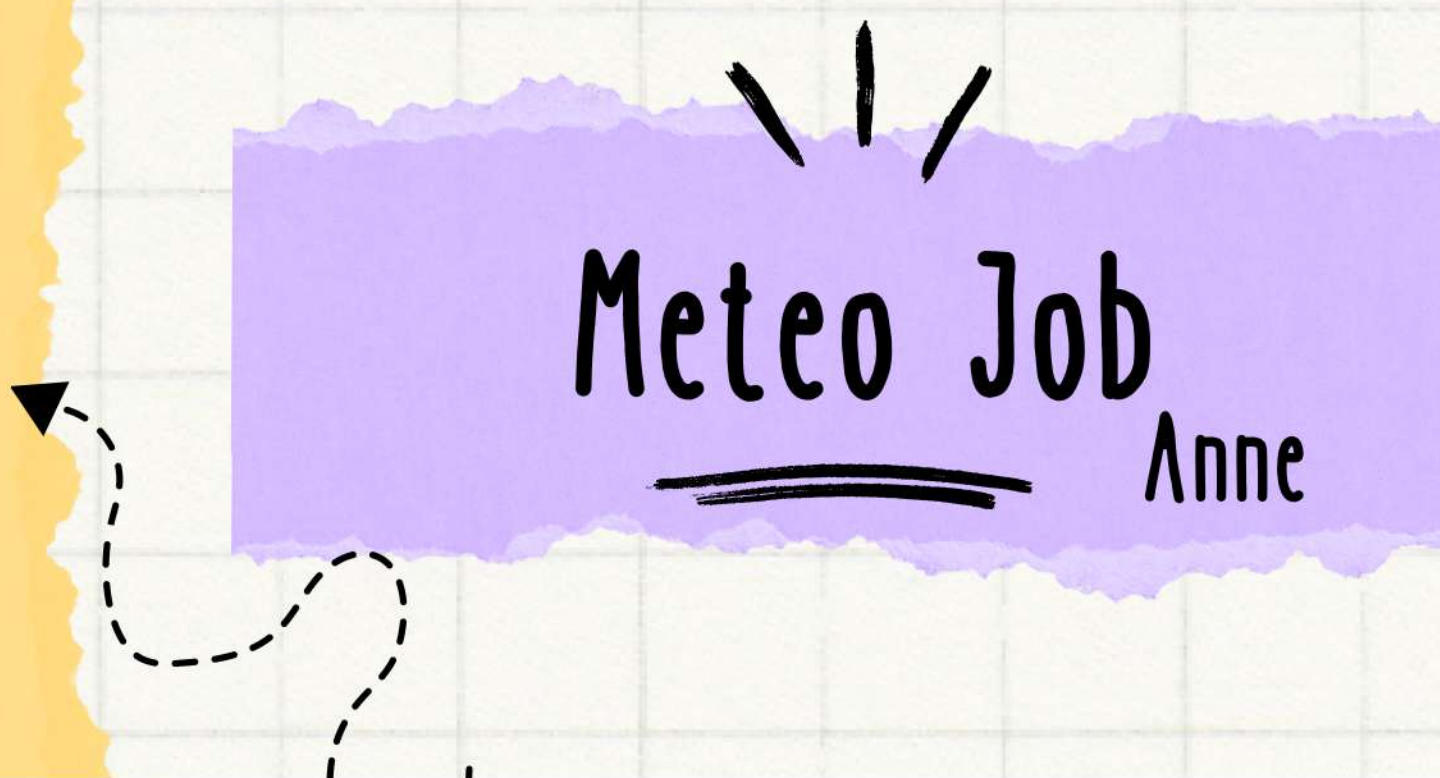
    df_jobs = pd.DataFrame(jobs_data)
    display(df_jobs)

# Utilisation de la fonction
url = "https://www.meteojob.com/jobs?what=Data&where=Lyon"
scrape_job_listings(url, pages=5) # Scrapes job listings de la pages 2 à 5
```

	Titre	Company	Type_de_contrat	date_de_publication	salaire
0	Architecte Data (H/F)	Gif	CDI	Il y a 7 jours	
2	Data Scientist (h/f)	LHH Recruitment Solutions	CDI	Il y a 12 jours	40 000 € - 46 000 €
4	DATA Engineer Cloud	Le Mercato de l'Emploi	CDI	Il y a 2 jours Nouveau	
6	Data Controller Comptable H/F	Talents AEC	CDI	Il y a 3 jours Nouveau	25 000 € - 35 000 € par an
8	Data Ingénieur Talend H/F	Amiltone	CDI	Il y a 17 jours	28 000 € - 38 000 € par an
10	Data Engineer F/H	SYNERGIE	CDD - 6 mois	Il y a 7 heures Nouveau	30 000 € - 35 000 € par an
12	Data scientist en alternance (H/F)	OpenClassrooms	Alternance / Apprentissage	Il y a 7 jours	905 € - 1 600 € par mois
14	(H/F) Développeur Data BI	ADSEARCH	CDI	Il y a 8 jours	40 000 € - 50 000 €
16	GESTIONNAIRE MASTER DATA (F/H)	Randstad	Intérim - 6 mois	Il y a 6 jours	À partir de 2 200 € par mois
18	PMO DATA STAGE (H-F)	CESI	Stage	Il y a 12 jours	
20	Alternance - Product Owner data F/H	OpenClassrooms	Alternance / Apprentissage	Il y a 7 jours	905 € - 1 600 € par mois
22	Stage - Data Marketing et Analytics BtoB H/F	The Adecco Group	Stage	Il y a 13 jours	6 000 € - 8 000 €
24	Chef de projet IT H/F	AÉSIO Mutuelle	CDI	Il y a 4 jours Nouveau	
26	Ingénieur réseau H/F	AÉSIO Mutuelle	CDI	Il y a 7 jours	
28	(H/F) Release Manager SAP	ADSEARCH	CDI	Il y a 2 jours Nouveau	50 000 € - 60 000 €
30	Coordonnateur de Données - F/H	RTE - Réseau de transport d'électricité	CDI	Il y a 4 jours Nouveau	
32	Responsable gouvernance de données (H/F)	Prosol	CDI	Il y a 2 jours Nouveau	
34	CDI CONSULTANT TRANSACTIONS SERVICES LYON 7	ENDRIX LYO	CDI	Il y a 4 jours Nouveau	

	Titre	Company	Type_de_contrat	date_de_publication	salaire
0	Architecte Data (H/F)	Gif	CDI	Il y a 8 jours	
1	Architecte Data (H/F)	Gif		Il y a 8 jours	
2	DATA Engineer Cloud	MERCATO DE L'EMPLOI	CDI	Hier Nouveau	
3	DATA Engineer Cloud	MERCATO DE L'EMPLOI		Hier Nouveau	
4	Data Ingénieur Talend H/F	Amiltone	CDI	Hier Nouveau	
...
315	(H/F) Release Manager SAP	ADSEARCH		Il y a 3 jours Nouveau	
316	Coordonnateur de Données - F/H	RTE - Réseau de transport d'électricité	CDI	Il y a 5 jours	
317	Coordonnateur de Données - F/H	RTE - Réseau de transport d'électricité		Il y a 5 jours	
318	Responsable gouvernance de données (H/F)	Prosol	CDI	Il y a 3 jours Nouveau	
319	Responsable gouvernance de données (H/F)	Prosol		Il y a 3 jours Nouveau	

320 rows × 5 columns



Scaping avec scroll

```
import pandas as pd
import requests
from bs4 import BeautifulSoup

def scrape_job_listings(url, pages):
    jobs_data = []

    for i in range(2, pages + 5):
        scroll_url = f"{url}&page={i}"
        response = requests.get(scroll_url)

        if response.status_code == 200:
            htmlData = response.content
            soup = BeautifulSoup(htmlData, 'html.parser')
            job_listings = soup.find_all('div', class_='truncated-text-fix')

            for job in job_listings:
                title = job.find('h2').text.strip() if job.find('h2') else ''
                company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
                types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
                date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
                salary = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''
                jobs_data.append({"Titre": title, "Company": company, "Type_de_contrat": types, "date_de_publication": date, "salaire": salary})

    df_jobs = pd.DataFrame(jobs_data)
    display(df_jobs)

# Utilisation de la fonction
url = "https://www.meteojob.com/jobs?what=Data&where=Lyon"
scrape_job_listings(url, pages=5) # Scrapes job listings de la pages 2 à 5
```


	Titre	Company	Type_de_contrat	date_de_publication	salaire
0	Data Scientist (h/f)	LHH Recruitment Solutions	CDI	Il y a 12 jours	40 000 € - 46 000 €
1	Data Controller Comptable H/F	Talents AEC	CDI	Il y a 3 jours Nouveau	25 000 € - 35 000 € par an
2	Data Ingénieur Talend H/F	Amilstone	CDI	Il y a 17 jours	28 000 € - 38 000 € par an
3	Data Engineer F/H	SYNERGIE	CDD - 6 mois	Il y a 11 heures Nouveau	30 000 € - 35 000 € par an
4	Data scientist en alternance (H/F)	OpenClassrooms	Alternance / Apprentissage	Il y a 3 heures Nouveau	905 € - 1 600 € par mois
5	(H/F) Développeur Data BI	ADSEARCH	CDI	Il y a 8 jours	40 000 € - 50 000 €
6	GESTIONNAIRE MASTER DATA (F/H)	Randstad	Intérim - 6 mois	Il y a 6 jours	À partir de 2 200 € par mois
7	Alternance - Product Owner data F/H	OpenClassrooms	Alternance / Apprentissage	Il y a 3 heures Nouveau	905 € - 1 600 € par mois
8	Stage - Data Marketing et Analytics BtoB H/F	The Adecco Group	Stage	Il y a 13 jours	6 000 € - 8 000 €
9	(H/F) Release Manager SAP	ADSEARCH	CDI	Il y a 2 jours Nouveau	50 000 € - 60 000 €
10	Data Scientist (h/f)	LHH Recruitment Solutions	CDI	Il y a 12 jours	40 000 € - 46 000 €
11	Data Controller Comptable H/F	Talents AEC	CDI	Il y a 3 jours Nouveau	25 000 € - 35 000 € par an
12	Data Ingénieur Talend H/F	Amilstone	CDI	Il y a 17 jours	28 000 € - 38 000 € par an
13	Data Engineer F/H	SYNERGIE	CDD - 6 mois	Il y a 11 heures Nouveau	30 000 € - 35 000 € par an
14	Data scientist en alternance (H/F)	OpenClassrooms	Alternance / Apprentissage	Il y a 3 heures Nouveau	905 € - 1 600 € par mois
15	(H/F) Développeur Data BI	ADSEARCH	CDI	Il y a 8 jours	40 000 € - 50 000 €
16	GESTIONNAIRE MASTER DATA (F/H)	Randstad	Intérim - 6 mois	Il y a 6 jours	À partir de 2 200 € par mois
17	Alternance - Product Owner data F/H	OpenClassrooms	Alternance / Apprentissage	Il y a 3 heures Nouveau	905 € - 1 600 € par mois
18	Stage - Data Marketing et Analytics BtoB H/F	The Adecco Group	Stage	Il y a 13 jours	6 000 € - 8 000 €
19	(H/F) Release Manager SAP	ADSEARCH	CDI	Il y a 2 jours Nouveau	50 000 € - 60 000 €
20	Data Scientist (h/f)	LHH Recruitment Solutions	CDI	Il y a 12 jours	40 000 € - 46 000 €

Importation des librairies

import pandas **as** pd

import requests

from bs4 **import** BeautifulSoup

def scrape_job_data(url, num_pages):

initialisation de la table vide

jobs_data = []

boucle for qui vas faire chaque page

for i **in** range(1, num_pages + 1):

Construction de l'url pour le scrolling

scroll_url = f"{url}&page={i}"

envoyer une requête GET à URL

response = requests.get(scroll_url)

vérification du succès de la requête

if response.status_code == 200:

Extraction du contenu HTML

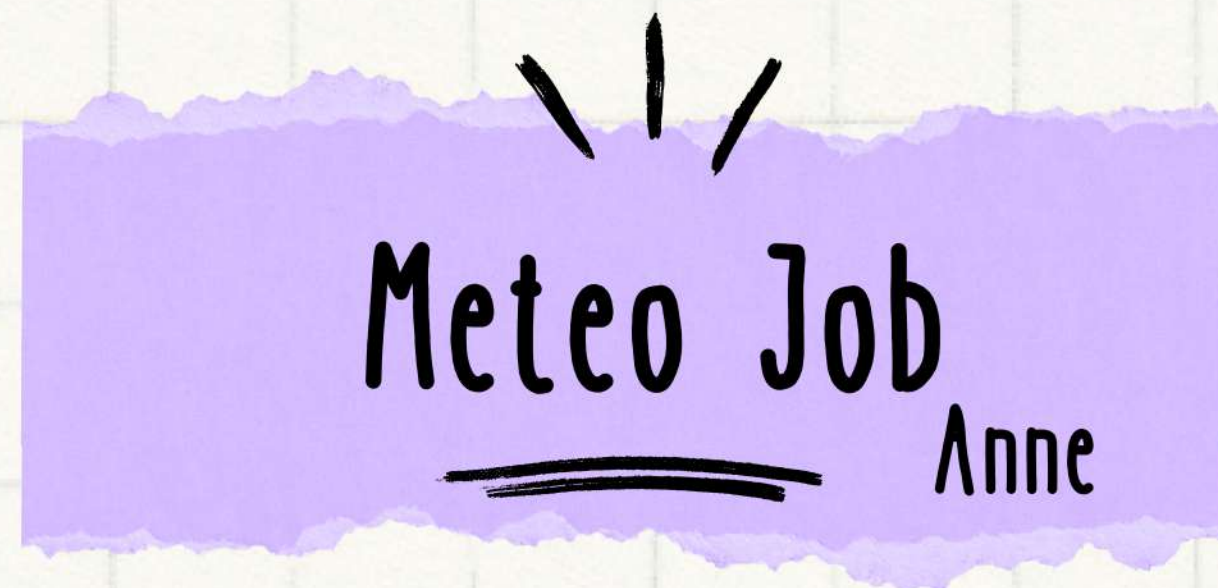
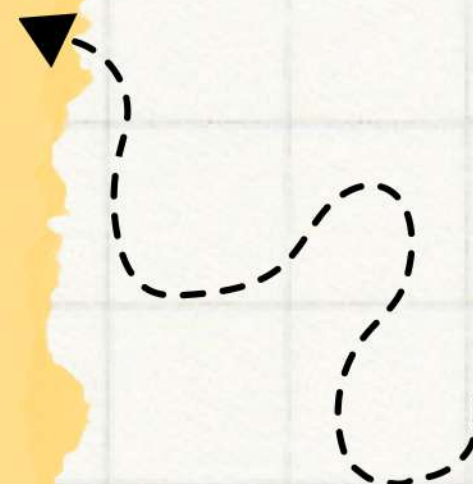
htmlData = response.content

Parser le contenu HTML en utilisant BeautifulSoup

soup = BeautifulSoup(htmlData, 'html.parser')

Trouver la liste des offres

job_listings = soup.find_all('div', class_='truncated-text-fix')



Scraping avec scroll trie des offres avec salaires

Trouver la liste des offres

job_listings = soup.find_all('div', class_='truncated-text-fix')

boucle pour récupérer chaque information sur chaque page

for job **in** job_listings:

Extraction des détail de l'offre

title = job.find('h2').text.strip() **if** job.find('h2') **else** ''

company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() **if** job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') **else** ''

types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() **if** job.find(class_='mr-1 mb-0 ng-star-inserted') **else** ''

date = job.find(class_='cc-font-size-small').text.strip() **if** job.find(class_='cc-font-size-small') **else** ''

salary = job.find(class_='cc-tag-primary-light').text.strip() **if** job.find(class_='cc-tag-primary-light') **else** ''

Vérification de la cellule Salary

if salary:

Ajouter les offres avec la cellule salaire remplie

jobs_data.append({"Titre": title, "Company": company, "Type_de_contrat": types, "date_de_publication": date, "salaire": salary})

else:

print(f"Failed to retrieve data from page {i}.")

Create a DataFrame from the job data list

df_jobs = pd.DataFrame(jobs_data)

Display the DataFrame

display(df_jobs)

1	Data Ingénieur Talend H/F	Amilton
2	Alternance - Product Owner data F/H	OpenClassroom
3	Chef de projet IT H/F	AÉSIO Mutuel
4	Ingénieur réseau H/F	AÉSIO Mutuel
5	(H/F) Release Manager SAP	ADSEARCO
6	CDI CONSULTANT TRANSACTIONS SERVICES LYON 7	ENDRIX LY
7	Développeur Back End H/F	Activu
8	Scrum Master H/F	Activu
Lyon 07 (69)	0	Architecte Data (H/F)
	1	Coordonnateur de Données - F/H RTE - Réseau de transport d'électricité
Lyon 09 (69)	0	Data Controller Comptable H/F Talents AE
Saint-Genis-Laval (69)	0	GESTIONNAIRE MASTER DATA (F/H) Randsta
Saint-Priest (69)	0	(H/F) Développeur Data BI ADSEARCO
Solaize (69)	0	Data scientist en alternance (H/F) OpenClassroom
Villeurbanne (69)	0	Data Scientist (h/f) LHH Recruitment Solution

Prosol	Chaponnay (69)	CDI	Il y a 2 jours Nouveau	
ERGIE	Chassieu (69)	CDD - 6 mois	Il y a 7 heures Nouveau	30 000 € - 35 000 € par an
Emploi	Lyon (69)	CDI	Il y a 2 jours Nouveau	
Amilton	Lyon (69)	CDI	Il y a 17 jours	28 000 € - 38 000 € par an
rooms	Lyon (69)	Alternance / Apprentissage	Il y a 7 jours	905 € - 1 600 € par mois
utuelle	Lyon (69)	CDI	Il y a 4 jours Nouveau	
utuelle	Lyon (69)	CDI	Il y a 7 jours	
ARCH	Lyon (69)	CDI	Il y a 2 jours Nouveau	50 000 € - 60 000 €
X LYO	Lyon (69)	CDI	Il y a 4 jours Nouveau	
Activus	Lyon (69)	CDI	Il y a 10 heures Nouveau	
Activus	Lyon (69)	CDI	Hier Nouveau	
Gif	Lyon 07 (69)	CDI	Il y a 7 jours	
nsport tricité	Lyon 07 (69)	CDI	Il y a 4 jours Nouveau	
is AEC	Lyon 09 (69)	CDI	Il y a 3 jours Nouveau	25 000 € - 35 000 € par an

Meteo Job Lou

Triages des annonces par lieux

```
import pandas as pd
import requests
from bs4 import BeautifulSoup

def extract_job_data(url):
    response = requests.get(url)
    if response.status_code == 200:
        htmlData = response.content
        soup = BeautifulSoup(htmlData, 'html.parser')

        job_listings = soup.find_all('div', class_='truncated-text-fix')
        jobsdata = []
        for job in job_listings:
            title = job.find('h2').text.strip() if job.find('h2') else ''
            company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
            types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
            localisation = job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0').text.strip() if job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0') else ''
            date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
            salaire = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''

            jobsdata.append({"Titre de l'offre": title, "Société": company, "Lieu": localisation, "Type d'offre": types, "Date d'ajout": date, "Salaires": salaire})

        df_jobs = pd.DataFrame(jobsdata)
        df_jobs = df_jobs.iloc[:,1:]
        return df_jobs
    else:
        return None

# Exemple d'utilisation de la fonction
url = "https://www.meteojob.com/jobs?what=data&where=Lyon"
df_jobs = extract_job_data(url)
if df_jobs is not None:
    # Créer un DataFrame par lieu
    df_jobs_by_location = df_jobs.groupby('Lieu').apply(lambda x: x.reset_index(drop=True))
    display(df_jobs_by_location)
else:
    print("Une erreur s'est produite lors de la récupération des données.")
```



```

import pandas as pd
import requests
from bs4 import BeautifulSoup

def extract_job_data(url):
    response = requests.get(url)
    if response.status_code == 200:
        htmlData = response.content
        soup = BeautifulSoup(htmlData, 'html.parser')

        job_listings = soup.find_all('div', class_='truncated-text-fix')
        jobsdata = []
        for job in job_listings:
            title = job.find('h2').text.strip() if job.find('h2') else ''
            company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
            types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
            localisation = job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0').text.strip() if job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0') else ''
            date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
            salaire = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''

            jobsdata.append({"Titre de l'offre": title, "Société": company, "Lieu": localisation, "Type d'offre": types, "Date d'actualisation": date, "Salaires": salaire})

        df_jobs = pd.DataFrame(jobsdata)
        df_jobs = df_jobs.iloc[:,1:]
        return df_jobs
    else:
        return None

# Exemple d'utilisation de la fonction
url = "https://www.meteojob.com/jobs?what=data&where=Lyon"
df_jobs = extract_job_data(url)
if df_jobs is not None:
    # Créer un DataFrame par Lieu
    dfs_by_location = {}
    for location, group in df_jobs.groupby('Lieu'):
        dfs_by_location[location] = group.reset_index(drop=True)

    # Afficher les DataFrames par Lieu
    for location, df in dfs_by_location.items():
        print(f"Emploi à {location}:")
        display(df)
else:
    print("Une erreur s'est produite lors de la récupération des données.")

```

Triages des
annonces par
lieux dans des
dataframe dédiés

Meteo Job
Lou

Emploi à Chassieu (69):

	Titre de l'offre	Société	Lieu	Type d'offre	Date d'actualisation
0	Data Engineer F/H	SYNERGIE	Chassieu (69)	CDD - 6 mois	Il y a 7 heures

Emploi à Lyon (69):

	Titre de l'offre	Société	Lieu	Type d'offre	Date d'actualisation
0	DATA Engineer Cloud	Le Mercato de l'Emploi	Lyon (69)	CDI	Il y a 2 jours
1	Data Ingénieur Talend H/F	Amilstone	Lyon (69)	CDI	Il y a 4 jours
2	Alternance - Product Owner data F/H	OpenClassrooms	Lyon (69)	Alternance / Apprentissage	Il y a 2 jours
3	Chef de projet IT H/F	AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 7 jours
4	Ingénieur réseau H/F	AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 4 jours
5	(H/F) Release Manager SAP	ADSEARCH	Lyon (69)	CDI	Il y a 2 jours
6	Développeur Back End H/F	Activus	Lyon (69)	CDI	Il y a 10 heures

Resultat

Titre de l'offre	Société	Lieu	Type d'offre	Date d'actualisation	Salaires
DATA Engineer Cloud	Le Mercato de l'Emploi	Lyon (69)	CDI	Il y a 2 jours	Nouveau
Ingénieur Talend H/F	Amilstone	Lyon (69)	CDI	Il y a 4 jours	Nouveau
Product Owner data F/H	OpenClassrooms	Lyon (69)	Alternance / Apprentissage	Il y a 2 jours	Nouveau
Chef de projet IT H/F	AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 7 jours	Nouveau
Ingénieur réseau H/F	AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 4 jours	Nouveau
Release Manager SAP	ADSEARCH	Lyon (69)	CDI	Il y a 2 jours	Nouveau
Développeur Back End H/F	Activus	Lyon (69)	CDI	Il y a 10 heures	Nouveau

Société	Lieu	Type d'offre	Date d'actualisation	Salaires
Gif	Lyon 07 (69)	CDI	Il y a 7 jours	Nouveau
Transport d'électricité	Lyon 07 (69)	CDI	Il y a 4 jours	Nouveau

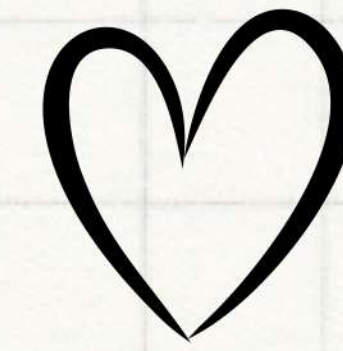
Type d'offre	Date d'actualisation	Salaires
CDI	Il y a 3 jours	Nouveau 25 000 € - 35 000 € par an

Société	Lieu	Type d'offre	Date d'actualisation	Salaires
Le Mercato de l'Emploi	Lyon (69)	CDI	Il y a 2 jours	Nouveau
Amilstone	Lyon (69)	CDI	Il y a 17 jours	28 000 € - 38 000 € par an
OpenClassrooms	Lyon (69)	Alternance / Apprentissage	Il y a 7 jours	905 € - 1 600 € par mois
AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 4 jours	Nouveau
AÉSIO Mutuelle	Lyon (69)	CDI	Il y a 7 jours	Nouveau
ADSEARCH	Lyon (69)	CDI	Il y a 2 jours	Nouveau 50 000 € - 60 000 €
Activus	Lyon (69)	CDI	Il y a 10 heures	Nouveau
ENDRIX LYO	Lyon (69)	CDI	Il y a 4 jours	Nouveau
Activus	Lyon (69)	CDI	Hier	Nouveau



Stockage

Local



Stockage local pour la
récupération de ces données

```
import pandas as pd
import re
import requests
from bs4 import BeautifulSoup

url = "https://www.meteojob.com/jobs?what=data&where=Lyon"
response = requests.get(url)
if response.status_code == 200:
    htmlData = response.content
    soup = BeautifulSoup(htmlData, 'html.parser')

    job_listings = soup.find_all('div', class_='truncated-text-fix')
    jobsdata = []
    for job in job_listings:
        title = job.find('h2').text.strip() if job.find('h2') else ''
        company = job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='d-inline-block mt-1 mb-0 ng-star-inserted') else ''
        types = job.find(class_='mr-1 mb-0 ng-star-inserted').text.strip() if job.find(class_='mr-1 mb-0 ng-star-inserted') else ''
        localisation = job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0').text.strip() if job.find(class_='mr-3 mb-1 cc-font-size-small mb-lg-3 mr-lg-0') else ''
        date = job.find(class_='cc-font-size-small').text.strip() if job.find(class_='cc-font-size-small') else ''
        salaire = job.find(class_='cc-tag-primary-light').text.strip() if job.find(class_='cc-tag-primary-light') else ''

        jobsdata.append({"Titre de l'offre": title, "Société": company, "Lieu": localisation, "Type d'offre": types, "Date d'actualisation": date, "Salaires": salaire})

    df_jobs = pd.DataFrame(jobsdata)
    df_jobs.to_csv('jobs.csv', index=False) # Cette ligne ajoute l'exportation au format CSV

display(df_jobs)
```


Github

Github > network





Thank you!

for your attention