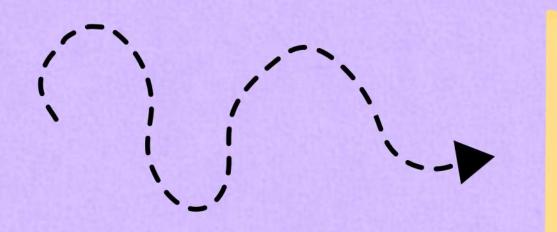


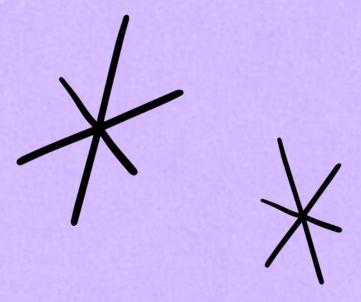
Kesako?

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:

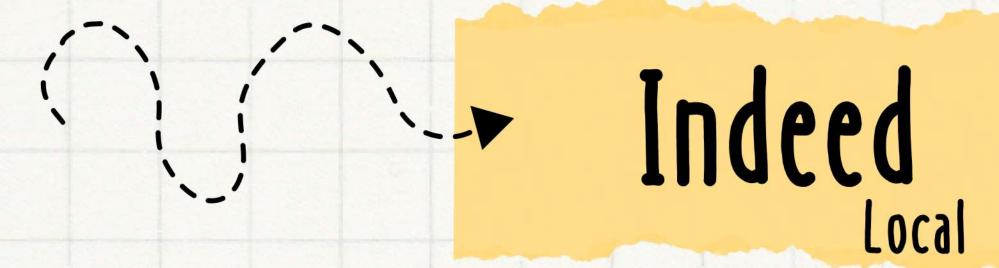
- 1- Collecte de données : Développer des scripts en python pour récupérer le contenu HTML des pages web.
- 2- Nettoyage et préparation des données : nettoyer les valeurs, convertir les types de données si nécessaire, gérer les valeurs manquantes, etc.
- 3-Stockage des données : Proposer une méthode 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.



```
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')

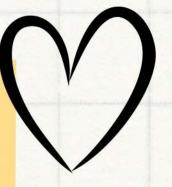
    job_listings = soup.find_all('div', class_='truncated-text-fix')
    jobsdata = []
    for job in job_listings:
        title = job.find('h2').text.strip() if job.find('h2') else ''
        company = job.find(class_="d-inline-block mt-1 mb-0 ng-star-inserted").text.strip() if job.find(class_="mr-1 mb-0 ng-star-inserted").text.strip() if job.find(class_="m
```


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

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

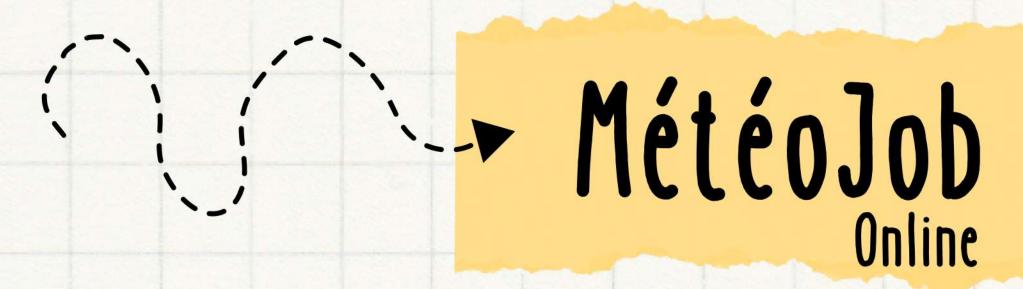
df jobs = extract job data(url)

if df_jobs is not None:
 display(df_jobs)



Récupération de toutes les informations pertinentes

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





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 20
                types = job.find(class ="mr-1 mb-0 ng-star-inserted").text.strip() if job.find(class ="mr-1 mb-0 ng-star-inserted").
                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
   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

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

	Titre	Company	Type_de_contrat	date_de_publication	salaire
0	Architecte Data (H/F)	Gif	CDI	II y a 8 jours	
1	Architecte Data (H/F)	Gif		II 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		II y a 3 jours Nouveau	
316	Coordonnateur de Données - F/H	RTE - Réseau de transport d'électricité	CDI	II y a 5 jours	
317	Coordonnateur de Données - F/H	RTE - Réseau de transport d'électricité		II y a 5 jours	
318	Responsable gouvernance de données (H/F)	Prosol	CDI	II y a 3 jours Nouveau	
319	Responsable gouvernance de données (H/F)	Prosol		II y a 3 jours Nouveau	

320 rows × 5 columns

Meteo Job Anne

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

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

```
# Importation des librairies
import pandas as pd
import requests
from bs4 import BeautifulSoup
def scrape job data(url, num pages):
   # initialistion 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 requète GET à URL
       response = requests.get(scroll url)
       # férification du succés de la requète
       if response.status code == 200:
           # Extraction du contenue HTML
           htmlData = response.content
           # Parser le contenu HTML en utiloisant 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").text.strip() if job.find(class = "d-inline-block mt-1 mb-0 ng-star-inserted").text.strip() if job.find(c
                                types = job.find(class_="mr-1 mb-0 ng-star-inserted").text.strip() if job.find(class_="mr-1 mb-0 ng-star-inserted") &
                                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 aves la cellule salaire remplie
                                          jobs_data.append({"Titre": title, "Company": company, "Type_de_contrat": types, "date_de_publication": date, "sal
                     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 Ingé	nieur Talend H/	F	Amilton
		2	Alter	nance - Product	Owner data F/	Н	OpenClassroon
		3		Chef	f de projet IT H/	F	AÉSIO Mutuel
		4		Ingé	nieur réseau H/	F	AÉSIO Mutuel
		5		(H/F) Releas	se Manager SA	P	ADSEARC
		6	CDII	CONSULTANT T SER	RANSACTION VICES LYON		ENDRIX LY
		7		Développe	ur Back End H/	F	Activu
		8		S	crum Master H/	F	Activu
	Lyon 07 (69)	0		Arch	itecte Data (H/F	-)	G
		1	(Coordonnateur d	e Données - F/	H RTE	- Réseau de transpo d'électrici
	Lyon 09 (69)	0		Data Controlle	r Comptable H/	F	Talents AE
	Saint-Genis- Laval (69)	0	GEST	ONNAIRE MAS	TER DATA (F/H	()	Randsta
	Saint-Priest (69)	0		(H/F) Déve	eloppeur Data E	31	ADSEARC
	Solaize (69)	0	C	ata scientist en	alternance (H/F	-)	OpenClassroon
F400	Villeurbanne (69)	0		Da	ata Scientist (h/	f) LHH	Recruitment Solution
3.						AND DESCRIPTION	
CNVFILN	Prosol Chaponn	ay (6	9)	CI	11	a 2 jours Nouveau	
CNVFILN	Prosol Chaponn ERGIE Chassi	\$181		CDD - 6 moi	n ∏ya		30 000 € - 35 000 € par an
CNVFILN	ERGIE Chassi	\$181	9)		s IIya	Nouveau 7 heures	A THE RESIDENCE OF THE PARTY OF
CNVFILM	ERGIE Chassi Emploi Ly	eu (6	9)	CDD - 6 moi	on IIya is IIya	Nouveau 7 heures Nouveau a 2 jours	A THE RESIDENCE OF THE PARTY OF
CNVFILM	ERGIE Chassi Emploi Ly niltone Ly	eu (6 on (6	9) 9) 9)	CDD - 6 moi	is II y a DI II y a	Nouveau 7 heures Nouveau a 2 jours Nouveau	€ par an 28 000 € - 38 000
CNVFILN	ERGIE Chassi Emploi Ly miltone Ly srooms Ly	eu (6 on (6 on (6	9) 9) 9) 9)	CDD - 6 moi	II y a II y	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 €
CNVFILI	ERGIE Chassi Emploi Ly miltone Ly rooms Ly utuelle Ly	eu (6 on (6 on (6 on (6	9) 9) 9) 9) 9)	CDD - 6 moi	ol IIya ol IIya ol IIya ol IIya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours a 4 jours	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 €
CNYFILM	ERGIE Chassi Emploi Ly miltone Ly rooms Ly utuelle Ly utuelle Ly	eu (6 on (6 on (6 on (6	9) 9) 9) 9) 9)	CDD - 6 moi	II y a II	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours a 4 jours Nouveau	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 €
CNYFILM	ERGIE Chassi Emploi Ly miltone Ly rooms Ly utuelle Ly utuelle Ly ARCH Ly	eu (6 on (6 on (6 on (6 on (6	9) 9) 9) 9) 9) 9)	CDD - 6 moi CE Alternance Apprentissag CE	Ilya Ilya Ilya Ilya Ilya Ilya Ilya Ilya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours A 1 jours nouveau a 7 jours a 2 jours	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 € par mois
	ERGIE Chassi Emploi Ly miltone Ly srooms Ly utuelle Ly utuelle Ly ARCH Ly X LYO Ly	eu (6 on (6 on (6 on (6 on (6	9) 9) 9) 9) 9) 9) 9)	CDD - 6 moi	Ilya Ilya Ilya Ilya Ilya Ilya Ilya Ilya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours A 1 jours Nouveau a 7 jours a 2 jours Nouveau a 4 jours	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 € par mois
	ERGIE Chassi Emploi Ly miltone Ly srooms Ly utuelle Ly utuelle Ly IARCH Ly X LYO Ly Activus Ly	eu (6 on (6 on (6 on (6 on (6 on (6	9) 9) 9) 9) 9) 9) 9)	CDD - 6 moi	Ilya Ilya Ilya Ilya Ilya Ilya Ilya Ilya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours A 1 jours Nouveau a 7 jours Nouveau a 2 jours Nouveau a 4 jours Nouveau o heures	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 € par mois
	ERGIE Chassi Emploi Ly miltone Ly srooms Ly utuelle Ly utuelle Ly IARCH Ly X LYO Ly Activus Ly	eu (6 on (6 on (6 on (6 on (6 on (6 on (6	9) 9) 9) 9) 9) 9) 9) 9)	CDD - 6 moi	Ilya Ilya Ilya Ilya Ilya Ilya Ilya Ilya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours A 1 jours Nouveau a 7 jours Nouveau a 7 jours Nouveau a 7 jours Nouveau a 10 jours Nouveau	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 € par mois
	ERGIE Chassi Emploi Ly miltone Ly srooms Ly utuelle Ly utuelle Ly ARCH Ly Activus Ly Activus Ly	eu (6 on (6	9) 9) 9) 9) 9) 9) 9) 9) 9)	CDD - 6 moi	Ilya Ilya Ilya Ilya Ilya Ilya Ilya Ilya	Nouveau 7 heures Nouveau a 2 jours Nouveau a 17 jours a 7 jours A 1 jours Nouveau a 7 jours Nouveau a 7 jours Nouveau a 10 jours Nouveau Nouveau	€ par an 28 000 € - 38 000 € par an 905 € - 1 600 € par mois



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
                           types = job.find(class_="mr-1 mb-0 ng-star-inserted").text.strip() if job.find(class_="mr-1 mb-0 ng-star-inserted") (
                           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
                           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'é
                                                                 "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
           types = job.find(class_="mr-1 mb-0 ng-star-inserted").text.strip() if job.find(class_="mr-1 mb-0 ng-star-inserted") a
           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
           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'
                            "Salaires": salaire})
        df_jobs = pd.DataFrame(jobsdata)
        df jobs = df jobs.iloc[::1]
        return of 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:
                                                                                         Triages des
    # 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)
                                                                                         annonces par
    # Afficher les DataFrames par lieu
    for location, df in dfs_by_location.items():
        print(f"Emploi à {location}:")
       display(df)
    print("Une erreur s'est produite lors de la récupération des données.")
```

lieux dans des dataframe dédiés

					f de projet
Emploi à Chassieu	(69):	Lieu	Type d'offre	Date d'a	eur résea
Titre de l'offre O Data Engineer F/H	SYNERGIE	Chassieu (69)	CDD-01116		Manager
0 Data Engineel F/11				So	ck End
Emploi à Lyon (6	9):	Titr	e de l'offre	Le Merca	SILY
		DATA Eng	gineer Cloud	And of	
0			ur Talend H/F	Am	
1		ance - Product C	owner data F/H	OpenClass	
2	Alterna	ance - Floor	de projet IT H/F	AÉSIO	
3			nieur réseau H/	F AESIO	
				Δ	DSEA
4		(H/F) Rele	ase Manager SA	AF.	Ad
5		Dévelop	peur Back End	H/F	
6			-		

Titre de l'offre	Société	Lieu		
Engineer Cloud	Le Mercato de l'Emploi	Lyon	Type d'offr	- Jule U
nieur Talend H/F	Amiltone	(69) Lyon	CD	llya2jo
Owner data F/H Ope	enClassrooms	(69) Lyon	Alternance /	11
le projet IT H/F AÉ	SIO Mutuelle	(69) Lyon	Apprentissage	
eur réseau H/F AÉS	BIO Mutuelle	(69) Lyon	CDI	II y a 4 jou
Nananar oan	ADSEARCH	(69) Lyon	CDI	
ck End H/F	Activus	(69) Lyon	CDI	II y a 2 jou
SILYON 7 EN	DRIX LYO	(69) Lyon	CDI	II y
ster H/F	Activus	(69)	CDI	II y a 4 jou
	- todayus	(69)	CDI	н
Société	Lieu	Type d'offre	Date d'actualisation	
Gif Isport d'électricité	Lyon 07 (69)	CDI	Il y a 7 jours	Salaires
EA	Lyon 07 (69)	CDI	II y a 4 jours Nouveau	
Ad Type d'offre	Date d'actuali	sation	A2 1	
CDI	II y a 3 jours No		Salaires 00 € - 35 000 € par an	

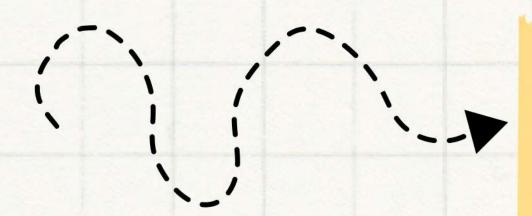
Resultat

Lieu	Type dies		
Lyon	Type d'offre	Date d'actualisation	Salai
(69)	CDI	II y a 2 jours Nouveau	Salaire
Lyon (69)	CD		
Lyon		Il y a 17 jours	28 000 € - 38 000 € pai
(69)	Alternance / Apprentissage	ll v a 7 iour	an
Lyon (69)			905 € - 1 600 € par mois
21	CDI	II y a 4 jours Nouveau	
(69)	CDI	II v o 7 :-	
Lyon	Market 1	" y a / Jours	
	CDI	Il y a 2 jours Nouveau	50 000 € - 60 000 €
(69)	CDI		-2 000 € - 00 000 €
Lyon		Nouveau	
	CDI	ll y a 4 jours Nouveau	
(69)	CDI	PCL C	
	Lyon (69)	Lyon (69) CD	Lyon (69) CDI II y a 2 jours Nouveau Lyon (69) CDI II y a 17 jours Lyon (69) Alternance / Apprentissage II y a 7 jours Lyon (69) CDI II y a 4 jours Nouveau Lyon (69) CDI II y a 2 jours Nouveau Lyon (69) CDI II y a 2 jours Nouveau Lyon (69) CDI II y a 2 jours Nouveau Lyon (69) CDI II y a 2 jours Nouveau Lyon (69) CDI II y a 10 heures Nouveau Lyon (69) CDI II y a 4 jours Nouveau Lyon (69) CDI II y a 4 jours Nouveau Lyon (69) CDI II y a 4 jours Nouveau

Salaires

000 € par an

Meteo Job



Stockage



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").text.strip() if
                    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
                    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, "Type d'actualisation, "Type d'actualisation, "Type d'actualisation, "Type d'actualisation de l'actualisation de l'actualisation
                                                               "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)
```

