



RNCP35288 CDSD

Plan your trip with Kayak

Build & Manage a Data
Infrastructure

Rédacteur : Jean-Yves Vuillequez





Contexte et enjeu



Objectif :

- Aider Kayak à mieux recommander des destinations : Combinaison de signaux externes (météo) et signaux d'offre (hébergements : quantité, notes, localisation).
- L'enjeu est d'identifier les destinations les plus attractives selon des critères objectifs (ex : météo favorable + offre d'hôtels bien notés).



Process et pipeline data

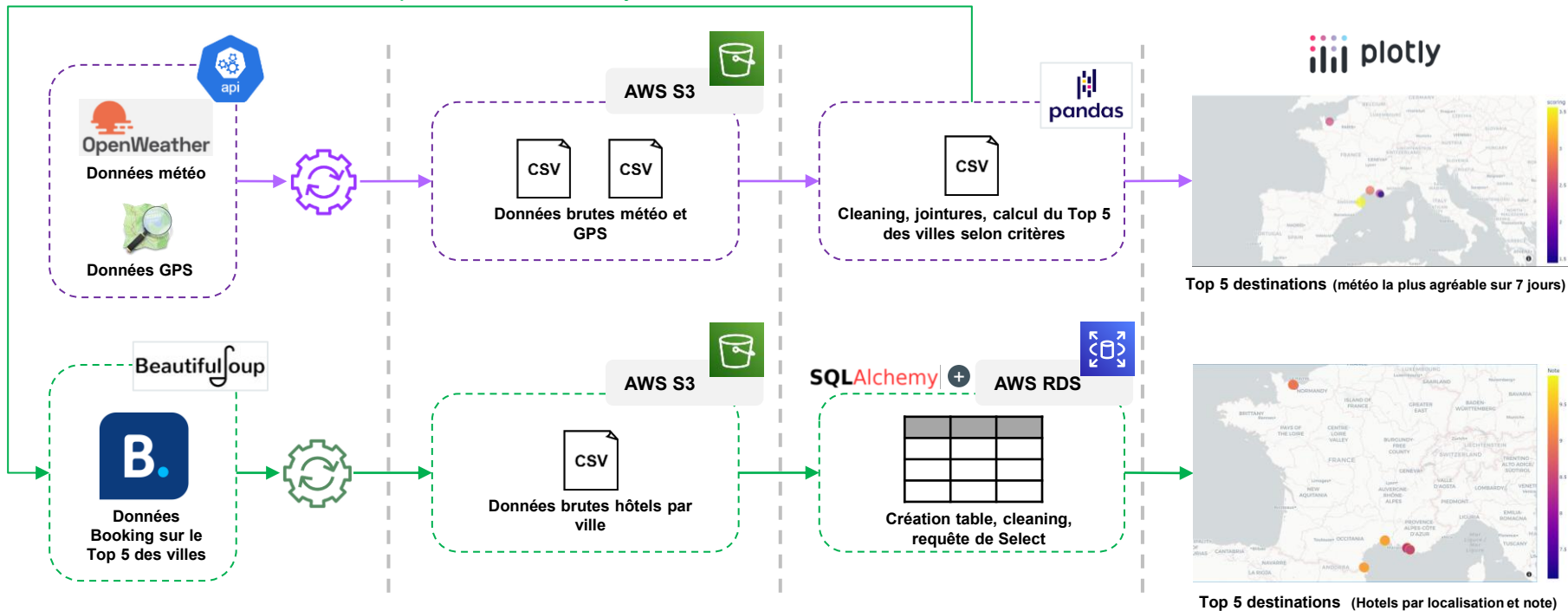
Extract

Load

Transform

Visualisation

Sélection du top 5 des villes dans un fichier json

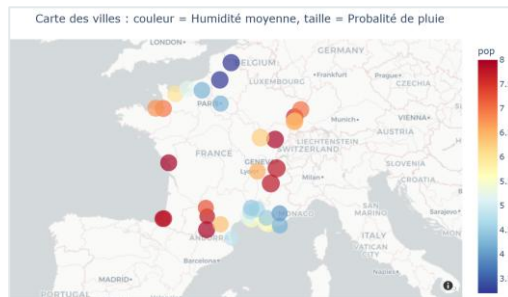
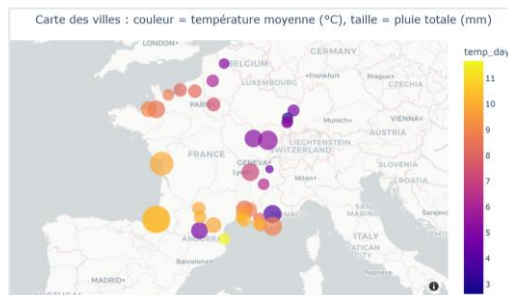




Données GPS et météo : Choix top 5 des destinations

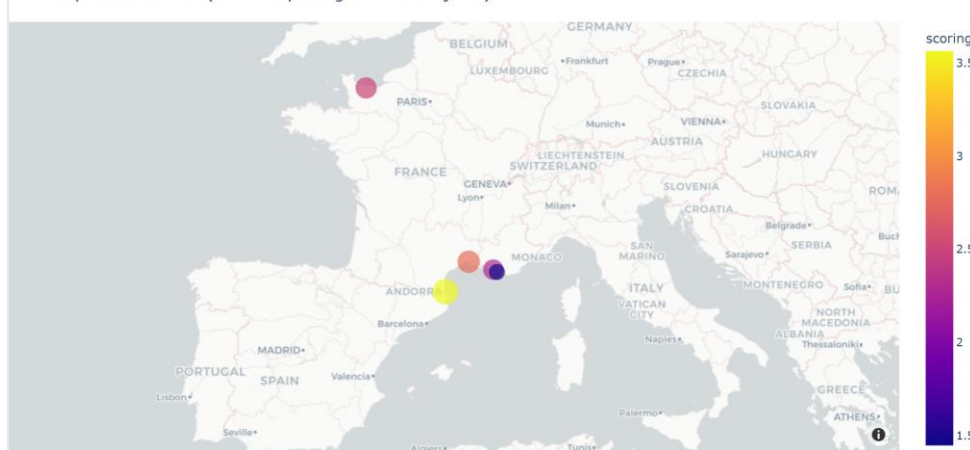
Objectif : identifier les villes offrant la météo la plus agréable sur les 7 prochains jours

Résultat : Les villes sont classées par score décroissant, et les 5 meilleures destinations sont retenues comme celles offrant la météo la plus favorable sur la période étudiée.



	city	mean_temp_day	mean_humidity	expected_rain_7d_mm	scoring
14	Collioure	11.514286	63.142857	20.1840	3.564800
0	Aigues Mortes	10.245714	70.857143	17.7900	2.783000
6	Bayeux	8.832857	75.857143	13.5300	2.498143
24	Marseille	10.024286	71.714286	18.5409	2.310587
12	Cassis	9.441429	70.857143	19.5600	1.447714

Top 5 destinations (météo la plus agréable sur 7 jours)

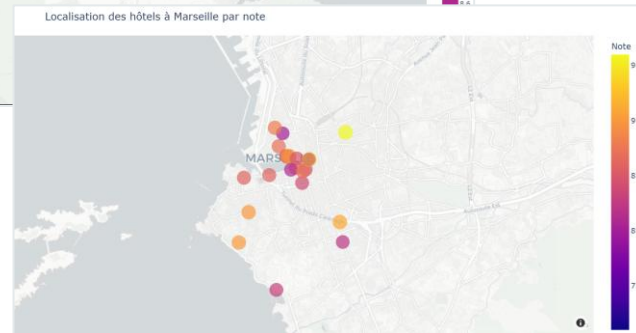
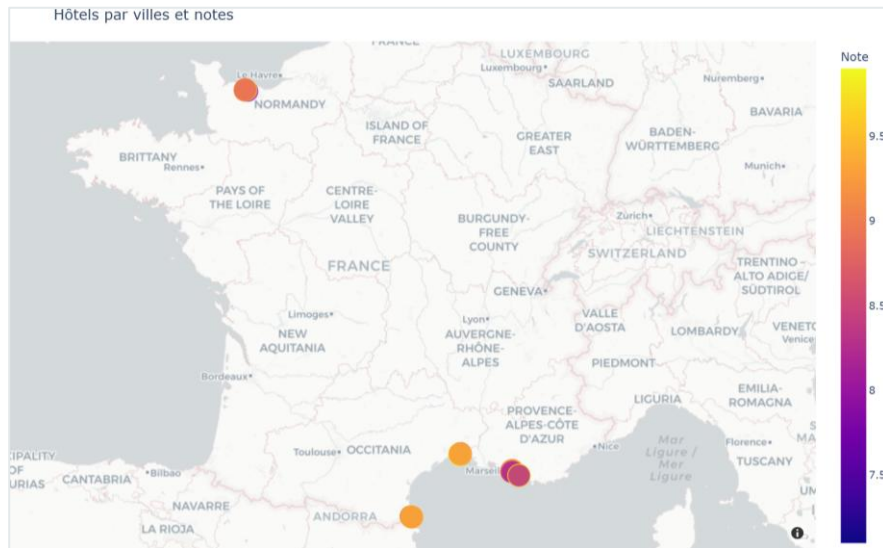




Données Booking : Visualisations obtenues à l'issu du scraping

Objectif : Visualiser la répartition géographique des hôtels des 5 meilleures destinations et comparer leur qualité (note Booking)

Résultat : Les cartes mettent en évidence où se situent les hôtels par ville, et permettent d'identifier rapidement les localisations et les établissements les mieux notés.





Jedha

Merci pour votre attention
Des questions ?





Annexes



Ressources AWS

dsfs36-bucket-01 [Info](#)

[Objects](#)[Metadata](#)[Properties](#)[Permissions](#)[Metrics](#)[Management](#)[Access Points](#)

Objects (2)

[Copy S3 URI](#)[Copy URL](#)[Download](#)[Open](#)[Delete](#)[Actions](#)[Create folder](#)[Upload](#)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	outputs/	Folder	-	-	-
<input type="checkbox"/>	raw/	Folder	-	-	-

Databases (1)

☒ Group resources[Modify](#)[Actions](#)[Create database](#)

< 1 >

<input type="checkbox"/>	DB identifier	Status	Role	Engine	Upgrade rollout order	Region ...	Size
<input type="radio"/>	kayak-db	Available	Instance	PostgreSQL	SECOND	eu-west-3c	db.t4g.micrc