**Livrable 1**

Groupe :

* Thomas Favrel
* Martin Gille
* Mohamed Haijoubi

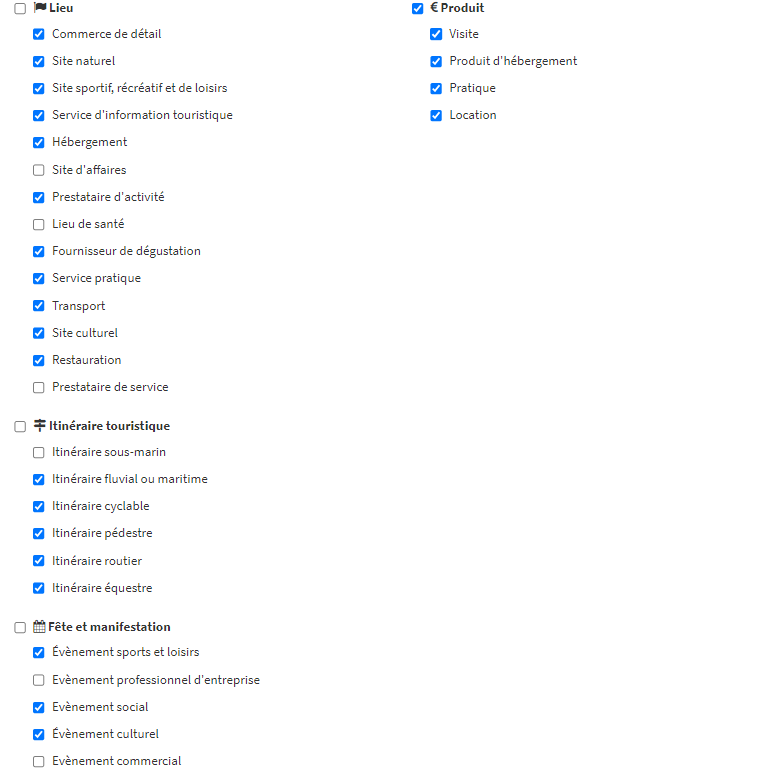
Tutrice : Mage

Sujet : Itinéraire de vacances

**Étape 1 / Découverte de la source de données disponible**

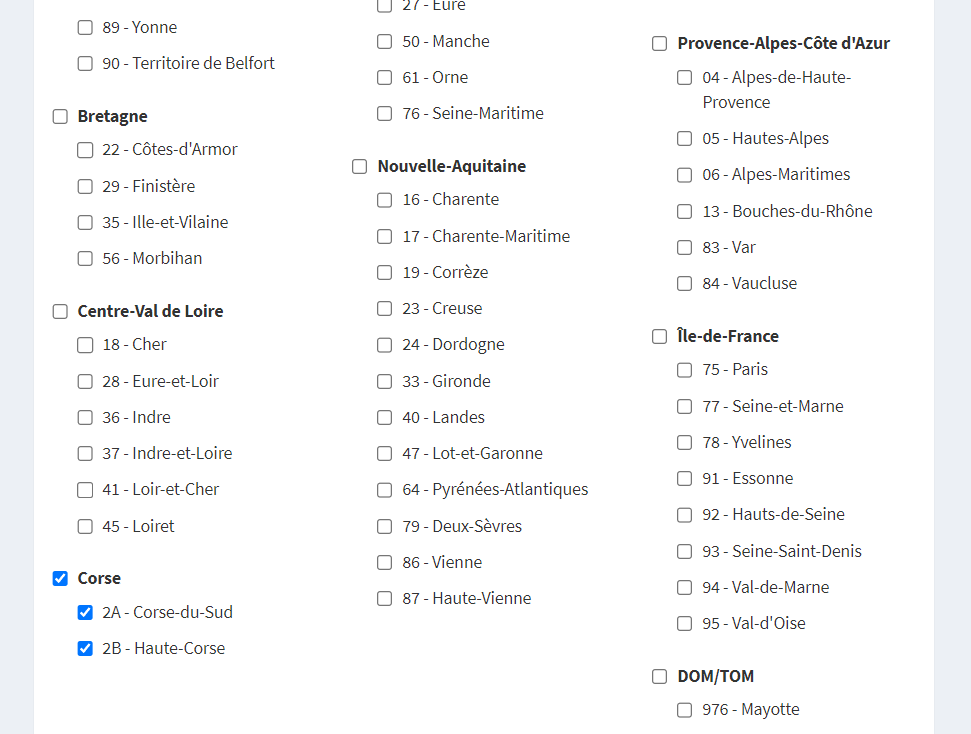
La récolte des données se fait depuis le site :

* [DATAtourisme](https://www.datatourisme.gouv.fr/) : facile, mis régulièrement à jour et rapide.

Voici les catégories de points d’intérêt sélectionnés :

Les points d’intérêt ont été choisi pour leurs pertinences et le respect des objectifs ciblés par notre groupe : des vacances agréables. Certains des critères ont pu être modifiés par la suite.

Les régions sélectionnées sont :



La région Corse a été choisi pour sa beauté, son attractivité, et sa pertinence touristique, … De plus son nombre de points d’intérêt restreint facilite le développement et le stockage de notre solution. Une ou plusieurs régions supplémentaires seront ajoutées pendant le déroulement du projet pour remplir les conditions de données massives.

Voici le descriptif des variables du fichier CSV récupéré depuis la source Datatourisme :

* Nom\_du\_POI : nom du point d'intérêt ;
* Categories\_de\_POI : lien vers une page d'aide à la spécification de catégorie d'un POI ;
* Latitude : distance angulaire de la terre à l'équateur ;
* Longitude : distance angulaire à un méridien d'origine
* Adresse\_postale : adresse postale ;
* Code\_postal\_et\_commune : code postal suivi du nom de la commune (00000#Ville) ;
* Periodes\_regroupees : dates (2023-09-22<->2023-09-24) ;
* Covid19\_mesures\_specifiques : mesures anti-covid déployées au POI (le plus souvent vide) ;
* Covid19\_periodes\_d\_ouvertures\_confirmees : périodes d'ouvertures en considérant la Covid19 ;
* Createur\_de\_la\_donnee : nom du créateur du POI dans la base de données Datatourisme ;
* SIT\_diffuseur : plateforme permettant à chaque territoire de saisir, enrichir et diffuser ses données touristiques ;
* Date\_de\_mise\_a\_jour : date de la dernière mise à jour des données ;
* Contacts\_du\_POI : numéro de téléphone et adresse mail à joindre suivi de l'adresse du site internet (#+33 0 00 00 00 [00#adress@gmail.com#https://](mailto:00#adress@gmail.com#https://)...) ;
* Classements\_du\_POI : classement par catégorie ;
* Description : description du POI ;
* URI\_ID\_du\_POI : lien vers la page du site Datatourisme conteant les informations correspondant au POI.

Extrait du fichier CSV :

Voici le descriptif des variables du fichier JSON récupéré depuis la source Datatourisme :

* @id: Lien vers le POI.
* dc:identifier: Identifiant unique
* @type: La classification de POI. Par exemple : festival, musée, chambre d’hôtes...
* rdfs:comment: Commentaire du POI
* rdfs:label: Nom du Point Of Interrest
* hasBeenCreatedBy: L'agent qui a créé ce POI dans le système d'information :
  + @id: Lien vers l'agent sur DATAtourisme
  + dc:identifier: Identifiant unique
  + schema:email: E-mail de l'agent
  + schema:legalName: Nom officiel de l'agent
  + @type: Type de l'agent
  + foaf:homepage: Lien vers le site de l'agent
* hasBeenPublishedBy: L'agent qui a publié ce POI dans le système d'information:
  + @id: Lien vers l'agent sur DATAtourisme
  + schema:legalName: Nom officiel de l'agent
  + @type: Type de l'agent
  + foaf:homepage: Lien vers le site de l'agent
* hasContact: L'agent à contacter pour affaires générales relatives à ce POI:
  + @id: Lien vers l'agent sur DATAtourisme
  + schema:email: E-mail de l'agent
  + schema:telephone: Telephone de l'agent
  + @type: Type de l'agent
  + foaf:homepage: Lien vers le site de l'agent
* hasDescription: Description textuelle courte ou longue du POI pouvant être associée à une audience. Par exemple, un POI peut avoir une description dédiée aux écoles et une autre dédiée au grand public. Si une description n'a aucune audience renseignée, on suppose qu'elle est dédiée à tous les publics:
  + @id: Lien vers la description
  + @type: Type(s) de la descripton
  + hasTranslatedProperty: Informations de traductions:
    - @id: Lien vers les informations de la traduction de la description
    - dc:contributor: Traducteur (api.deepl.com/v2/translate)
    - rdf:language: Langue de la traduction
    - rdf:predicate: Type de la description
  + shortDescription: Dictionnaire des descriptions traduit en plusieurs langues:
    - fr: En fr
    - de: ...
* hasTheme: Liste des thèmes:
  + @id: Identifiant du thème au format kb:[theme en anglais sans espace]
  + @type: Type du thème
  + rdfs:label: Thème en différente langue:
    - fr: fr
    - de: ...
* hasTranslatedProperty: liste d'informations de traductions sur le commentaire:
  + @id: Lien vers les informations de la traduction du commentaire
  + dc:contributor: Traducteur (api.deepl.com/v2/translate)
  + rdf:language: Langue de la traduction
  + rdf:predicate: Table de jointure entre @id et comment
* isLocatedAt: Informations sur l'emplacement géographique du POI:
  + @id: Lien vers les informations sur l'emplacement
  + schema:address: Informations sur l'adresse postale
    - @id: Lien vers les informations de l'adresse postale
    - schema:addressLocality: Ville
    - schema:postalCode: Code postale
    - schema:streetAddress: Adresse postale (certaines données ne sont pas propres)
    - @type: Type de l'adresse
    - hasAddressCity: Informations complémentaires sur l'adresse postale
      * @id: Identifiant
      * @type: Type de localisation (ex: City)
      * rdfs:label: Liste de la ville traduite ['fr']
      * insee: Code Insee
      * isPartOfDepartment: Informations du département
        + @id: Identifiant
        + @type: Type de localisation (ex: Department)
        + rdfs:label: Liste de la ville traduite ['fr']
        + insee: Code Insee
        + isPartOfRegion: Informations du département:

@id: Identifiant

@type: Type de localisation (ex: Region)

rdfs:label: Liste de la ville traduite ['fr']

insee: Code Insee

isPartOfCountry: Informations du Pays: \* @id: Identifiant \* @type: Type de localisation (ex: Country) \* rdfs:label: Liste de la ville traduite ['fr']

* + schema:geo:
    - @id: Lien vers les informations coordonnées gépgraphiques
    - schema:latitude: Latitude
    - schema:longitude: Longitude
    - @type: Type
  + schema:openingHoursSpecification: Informations des horraires d'ouverture et fermeture
    - @id: Lien vers les informations des horraires
    - schema:closes: Heure de fermeture
    - schema:opens: Heure d'ouverture
    - schema:validFrom: Date de début de période d'ouverture
    - schema:validThrough: Date de fin de période d'ouverture
    - @type: Type
    - additionalInformation: Liste d'inforamtions supplémentaires traduit en plusieurs langues
    - hasTranslatedProperty: Liste des informations de traduction des infos supplémentaires
      * @id: Lien vers les infos de trad
      * dc:contributor: Traducteur (api.deepl.com/v2/translate)
      * rdf:language: Langue de la traduction
      * rdf:predicate: Table de jointure entre @id et additionalInformation
  + @type: Type de place
* isOwnedBy: Informations de l'agent propriétaire
  + @id: Lien vers l'agent sur DATAtourisme
  + schema:email: E-mail de l'agent
  + schema:legalName: Nom de l'agent
  + @type: Type de l'agent
  + foaf:homepage: Lien vers le site de l'agent
* lastUpdate: Dernière mise à jour du POI
* lastUpdateDatatourisme: Dernière mise à jour du POI sur DATAtourisme

Extrait du fichier JSON :

