

# **Données et Cartographie**

**Utilisation de fichiers CSV en situation réelle**

**Vincent BOUILLOT**

# Présentation du projet

- Aujourd'hui, les données sont souvent publiques et très souvent de tailles très importantes. C'est le phénomène d'**Open Data**.
- Nombre de données ne sont pas utilisées mais pourraient permettre de mieux comprendre certains phénomènes.
- On peut en particulier relier des données avec de la cartographie.
- Ce projet vise donc à créer une **application de cartographie** basée sur du **data-mining** qui va nous permettre, grâce à l'utilisation de **tables de grande taille**, de répondre à des questions intéressantes.

# Buts du projet

## Buts généraux

- Découvrir l'**Open Data** : bases de données publiques et libres de droits.
- Poser trois questions qui peuvent être résolues grâce à votre projet.
- Diffuser des informations via un site internet.

## Buts techniques

- Réaliser des opérations sur des tables CSV (au moins 2 tables) et des dictionnaires Python.
- Utiliser une librairie de cartographie.
- Réaliser des popups dans une librairie de cartographie.

# Buts du projet

## Buts supplémentaires (non évaluées)

- Découvrir le standard JSON.
- Explorer des outils de visualisation et de data-mining en Python.
- Comprendre l'imbrication de différentes technologies dans les applications web.

# Contraintes techniques

- Ce projet est à réaliser seul.
- La date de rendu est fixée au 10 mai.

## Librairies Python

- La lecture des fichiers CSV doit se faire grâce aux librairies développées en cours (Pandas n'est pas autorisé).
- Les fichiers d'entrée doivent être au format CSV (JSON non autorisé).

## Tables

- Les tables doivent être publiques (libres de droit).
- Les tables doivent avoir plus de 500 lignes, au format CSV et contenir des **données temporelles**.

# Rendu du projet — Écrit — 30 points

- L'écrit se fera sur votre site internet GitHub dont **l'organisation est libre** (liens, une ou plusieurs pages etc.).
- Il doit obligatoirement apparaitre les éléments suivants :
  - Présentation générale de votre projet ;
  - Présentation de la base de données publique utilisée ;
  - Présentation de votre démarche projet (buts poursuivis, problèmes rencontrés...)
  - Explication d'une partie de votre programme (processing ou data-mining)
  - Application web de cartographie dynamique.
  - Tous les programmes Python utilisés dans 1 fichier compressé ;

# Rendu du projet – Écrit – 30 points

- Grille de notation :

Compétences	Degré de maitrise					Points
	0 %	25 %	50 %	75 %	100 %	
Créer et organiser un site web						5
Manipuler plusieurs tables CSV et travailler sur des données présentant des évolutions temporelles						5
Réaliser des calculs sur des structures complexes (listes, dictionnaires) en Python						5
Utiliser une librairie de cartographie en Python						5
Créer une application permettant de répondre simultanément à 3 questions						5
Expliquer clairement la démarche d'un projet informatique						5

- Vous déposerez dans Pronote un fichier word **converti au format pdf** contenant deux lignes :  
L1 : nom prénom  
L2 : lien fonctionnel vers votre site GitHub

# Rendu du projet – Oral – 20 points

- Un oral sera organisé à la rentrée (ou à distance pas de déconfinement)
- L'oral durera 5 minutes durant lesquels :
  - vous présenterez les trois questions auxquels vous répondez ;
  - vous répondrez à ces questions en faisant une démonstration de votre application ;
  - vous expliquerez la partie traitement de données de votre code Python.
- Une grille de notation vous sera fournie à la rentrée.



# Exemple du prof

## Covid-19 et répartition territoriale en France

- 3 questions sont posées :
  - Comment se répartissent les régions touchées par l'épidémie de Covid-19 ?
  - Le pic épidémique est-il atteint en France ?
  - Où sont les patients en réanimation en France ?
  - Un confinement généralisé est-il nécessaire ?
- Lien vers l'exemple : **[lien cliquable](#)**

# Le cours d'aujourd'hui...

Aujourd'hui, nous allons mettre en place les bases de votre projet. L'idée est de savoir où vous allez à la fin de ce cours et pendant vos vacances.

1. Etude de l'exemple du prof : démarche projet.
2. Recherche documentaire sur les bases de données.
3. Recherche documentaire sur les librairies de cartographie.
4. Choix d'un sujet
5. Bonnes vacances !

# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Identifiez les points clés du projets :
  - quelles données ont été nécessaires pour réaliser ce projet ?
  - qu'est-il essentiel de savoir faire pour répondre à la problématique ?
  - qu'aurait-il été possible de faire pour améliorer ce projet ?
- Estimez le temps nécessaire à la réalisation de chaque étape.
  - dans votre projet, essayez de mettre en place un plan de bataille : combien de temps devez-vous passer sur chaque étape ?

# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Quelles données ont été nécessaires pour réaliser ce projet ?

Table contenant le nombre de malades (réa+dcd) en fonction du temps

**Table reliant le numéro du département à des coordonnées GPS.**

Table de toutes les communes de France

Données cartographiques : affichage des pays et des départements

- Qu'est-il essentiel de savoir faire pour répondre à la problématique ?

**2h-4h** Gérer des tables (fusion, sélection etc.)

**1h** Apprendre à utiliser les bibliothèques de cartographie.

**30''** Générer l'application web HTML (front-end)

**1h** Calcul sur les tables en Python

**10'** Popup pour afficher des données textuelles

**8h** Popup pour afficher des données graphiques

- Qu'aurait-il été possible de faire pour améliorer ce projet ?

**8h-10h** Rafraichissement automatique de l'application web (avec les nouvelles données)

**0,5h** Affichage d'un graphique général (dans un coin ou dans un popup dédié)

**5h** Zones des départements à améliorer

Site plus rapide / temps de chargement à réduire

Voir plus grand : Europe et/ou le monde!

**2h** Ajout d'estimations (pic épidémique est passé, à combien de jour est ce pic...)

# **Étude de l'exemple du prof**

## **Covid-19 et répartition territoriale en France**

- Estimez le temps nécessaire à la réalisation de chaque étape.
  - dans votre projet, essayez de mettre en place un plan de bataille : combien de temps devez-vous passer sur chaque étape ?

# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Recherche documentaire sur les bases de données :
  - Rendez-vous sur [data.gouv.fr](https://data.gouv.fr) et cherchez covid 19 dans la barre de recherche. Choisissez une base de données puis téléchargez une table.
  - Cette table permet-elle l'affichage sur un outil cartographique ?

# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Recherche documentaire sur les bases de données :
  - Rendez-vous sur [data.gouv.fr](https://data.gouv.fr) et cherchez covid 19 dans la barre de recherche. Choisissez une base de données puis téléchargez une table.
  - Cette table permet-elle l'affichage sur un outil cartographique ?
  - Trouvez une base de données liant numéro de département et coordonnées GPS de la préfecture.
- Cherchez d'autres bases de données en **Open Data**.

# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Recherche documentaire sur les outils de cartographie en Python. 5'

**Pandas / Geopandas**  
**GeoPy**

**MapBox**  
**BaseMap**

**Cartopy**  
**Leaflet.js**

**Matplotlib**  
**Plotly**



**ShapeFiles**  
**PyGal**

**Folium !**  
**Leaflet.js**



# Étude de l'exemple du prof

## Covid-19 et répartition territoriale en France

- Recherche documentaire sur les outils de cartographie en Python :
  - Matplotlib : peu adapté mais, pourquoi pas?
  - Geopandas : très adapté pour traiter les données, pas pour faire des cartes
  - Cartopy : un peu daté
  - Folium :   [Recherche Google : Folium Python](#)

# Étude de VOTRE projet !

- Deux manières de concevoir votre projet :
  - Réfléchissez à une question qui vous interpelle et qui demande d'étudier des données dans le temps, puis naviguez dans l'Open Data pour voir si les données existent.
  - Naviguez dans l'Open Data pour trouver des données temporelles intéressantes et essayez de trouver une question qui vous interpelle !

# Bonnes vacances

- Étrangement, je ne prévois pas de partir très loin de chez moi pendant ces vacances.
- Je reste en repli technique pour vous aider sur votre projet :
  - vous pouvez me contacter sur Discord pendant les vacances ;
  - très bonnes vacances à tous.