

Rapport Projet Notebook/Django

PINAULT Alice, RAHALI Sofiane et PUJOL Corentin

1. Avancement :

Nous avons donc commencé par analyser la structure des données .txt récupérées sur le site du gouvernement. Nous avons pu constater que certaines colonnes n'étaient pas assez complètes afin d'être analysées et traitées pour obtenir des résultats et nous avons dû les supprimer ou les modifier (par exemple les virgules à la place des points rendaient la conversion string vers float impossible sur certaines colonnes). Il y avait aussi plus de deux millions de valeurs ce qui était extrêmement lourd à lire pour le notebook, qui prenait par moment longtemps avant de finir son exécution. Par la suite nous avons utilisé certaines fonctions mathématiques afin d'avoir des résultats plus généraux avec notamment des calculs de moyennes.

Nous avons donc fortement avancé les graphes du notebook en amont des partiels, puis nous nous sommes concentrés ensuite sur la partie du framework Django, qui a sans doute été la plus compliquée étant donné le peu de temps que nous avons eu pour la réaliser. Nous sommes également un peu frustrés de devoir rendre le projet si tôt et sans avoir pu plus nous plonger dans le développement Web, car c'est un domaine très intéressant sur lequel nous n'avions pas eu auparavant l'opportunité de nous y intéresser.

2. Problèmes rencontrés :

Lors de ce projet nous avons rencontré de nombreuses difficultés que nous allons vous décrire succinctement ci-dessous :

- Concernant les visualisations, les graphiques lineplot ont été les plus difficiles à appréhender. En effet, il faut choisir avec soin un groupBy adéquat et créer une colonne « count » dans son dataframe avec le nombre d'Index correspondant. Le changement de librairie matplotlib a aidé à cette fin : en effet la librairie seaborn est plus intuitive.
- Concernant l'utilisation des cartes choroplèthes, la principale difficulté a été de faire correspondre le fichier geojson avec le dataframe des DVF. En effet, par exemple les codes départementaux pour la France étaient dans un format différent sur le fichier geojson. Il a donc fallu réaliser de nombreuses modifications dans le dataframe original.
- Enfin au niveau de l'application django, il a tout d'abord été très compliqué de paramétrer les ordinateurs pour son utilisation. En effet, pour coder en python, nous utilisons en majorité l'IDLE Spyder d'Anaconda. Le logiciel python n'était donc pas dans les « path » de nos ordinateurs (par exemple) et

on ne pouvait pas utiliser python dans le cmd. Il a fallu un peu de temps afin que tous nos ordinateurs soient parfaitement paramétrés. Nous avons ensuite regardé de nombreux tutoriels concernant le framework django afin d'assimiler les premières bases. Nous avons commencé à créer le serveur depuis l'invite de commande de nos ordinateurs respectifs, puis nous avons choisi l'éditeur Visual Code Studio afin de développer ensuite notre application. Il a été très long de comprendre comment chaque fichier fonctionnait dans l'éditeur et dans le serveur. Il y a aussi de nombreuses dépendances entre chacun des fichiers (exemple avec les fichiers URLs, les fichiers d'aperçu et les différents templates html qui sont constamment liés). Nous sommes un peu déçus car nous avons manqué un peu de temps et qu'une semaine supplémentaire aurait sûrement permis d'obtenir une application encore plus aboutie.

3. Contributions :

Au niveau des contributions, le travail dans notre groupe a été globalement homogène et en pourcentage nous mettrions :

- Sofiane RAHALI : 35%
- Alice PINAULT : 35%
- Corentin PUJOL : 30%

Nous vous remercions de l'attention porter à notre projet.

Alice, Sofiane et Corentin.