Résumé du projet fil rouge La Manu

Visualisation des risques des catastrophes naturelles comme les inondations avec un paramètre économique

Comme par exemple :

- la valeur immobilière (valeur foncière d’une maison d’un appartement)

- ou le cout des assurances habitation (MRH)

1. Description du projet

Le choix de ce projet s’est porté par rapport à l’actualité qui secoue dernièrement le monde entier à savoir le changement climatique. Il fallait ensuite trouver des données open source pour mener à bien ce projet. Les sites gouvernementaux permettent d’avoir de nombreuses données sur les risques naturelles et sur les valeurs foncières. Par contre, les données des assurances sont assez confidentielles et elles sont montrées au niveau régional ou par ville.

<https://www.lelynx.fr/assurance-habitation/comparaison/pas-cher/prix/>

<https://www.assurland.com/presse/communiques-de-presse-assurance/infographie-dans-quelle-region-l-assurance-habitation-coute-t-elle-le-plus-cher.html>

2. Cohérence par rapport au contenu de la formation (python, sql, modélisation, Power BI).

Pour concevoir ce projet, il faut utiliser un langage adapté comme python avec des IDE facilitant l’édition et l’exécution du code comme Pycharm.

Le projet est mis en valeur dans JupiterNotebook.

3. Les url des datasets ou des api utilisés. Préciser leur intérêt.

Les urls utilisés pour récupérer les données sur les risques naturels et en l’occurrence les inondations sont des api que l’on peut requêter par lot de 10 codes insee maxi:

<https://api.gouv.fr/les-api/api-georisques>

<https://api.gouv.fr/documentation/api-georisques>

Le fichier csv des codes insee est utile pour avoir le code département, le code région, les coordonnées géographiques (geo\_point\_2d et geo\_shape) :

<https://public.opendatasoft.com/explore/dataset/correspondance-code-insee-code-postal/export/>

Le fichier immo est utilisé pour avoir la valeur immobilière moyenne par code insee :

<https://www.data.gouv.fr/fr/reuses/carte-prix-immobilier-carte-dexploration-des-biens-vendus-en-france/>

<https://files.data.gouv.fr/geo-dvf/latest/csv/2022/communes/01/>

4. Les bibliothèques python utilisées. Préciser leur intérêt.

Les librairies utilisées sont :

- pandas pour manager le dataset pour le cleaning, les merges.

- csv pour lire le fichier csv insee

- requests pour faire des requêtes sur les api des zones inondables

- json pour loader et sauvegarder en format json les requêtes api

- time pour afficher le temps de traitement

- numpy pour mettre en dur le calcul du nombre de bâtiments (maisons et appartements)

- folium et webbrowser pour afficher la carte géographique de la France avec les indicateurs

5. Les difficultés rencontrées.

Les requête sur l’api des risques naturels prend beaucoup de temps si l’on fait la requête par code insee et de même si elles se font par lot de 10 qui est le maximum autorisé. La solution est de le délocaliser du programme principal pour éviter de l’exécuter à chaque fois et de le mettre dans un programme qui sauvegarde les requêtes dans un fichier json qui lui sera appeler par le programme principal.

La commande groupby ne permet pas toutes les opérations comme la proportion de maisons par rapport aux total des immeubles. La solution trouvée est de sauvegarder le nombre total de maisons et appartements et le nombre de chacun d’entre eux.

La commande merge inner s’est changé en left du fait que certains codes insee n’ont pas de zones inondables mais cela provoquait des erreurs pour le code à la suite.

La fonction lamda combiné avec les if et les colonnes du dataset ne donne pas le résultat escompté dans folium. Le conditionnement a du être sorti pour ajouter à la carte chaque cas en l’occurrence un code couleur pour chaque valeur moyenne.

6. Quelques éléments de la visualisation (copies d'écran, explications...).

Folium trouve ses limites pour afficher la carte de France : dans un 1er temps, il affiche les zones des prix immobiliers puis les marqueurs des inondations. Le zoom est trop latent pour arriver à une navigation convenable. Il est donc nécessaire de limiter la zone de visualisation dans une ou plusieurs département ou régions spécifiques.

Une image contenant carte

Description générée automatiquementUne image contenant carte

Description générée automatiquement

**Une image contenant carte

Description générée automatiquement** **Une image contenant carte

Description générée automatiquement**

7. Les éléments que vous auriez aimé traiter dans le cadre de votre projet mais qu'ils ne sont pas couverts par la formation

la visualisation avec Tableau.