

UNIVERSITÉ DE MONTRÉAL

Projet 1

Par Pierre-Briac METAYER—MARIOTTI
Matricule : 2017 9280

Baccalauréat en neurosciences cognitives

Faculté des arts et sciences

Travail présenté à
Alexandru Hanganu

Dans le cadre du cours PSY3019

Cueillette et traitement de données en neurosciences cognitives

5 février 2023

Le projet consiste à analyser l'évolution des libertés civiles dans le monde de 1950 à 2021. L'idée serait d'afficher une carte interactive du monde montrant l'évolution de ces libertés de manière visuelle.

Afin d'entreprendre ce projet j'ai utilisé une banque de donnée disponible sur World Data.

Ma banque de donnée contient 5 colonnes à l'origine :

- Entity : nom du pays
- Code : Abréviation du pays
- Year : année concernée
- civ_libs_vdem_owid : score de liberté civile dans un pays x pour une année x
- region : Continent du pays

Les banque de données contient des informations sur les scores de libertés civiles pour divers pays et régions du monde entre 1850 et 2021.

J'ai commencé par nettoyer les données en supprimant toutes les données antérieures à 1950 et en remplaçant le nom d'un pays pour des raisons de simplification (Yemen People's Republic par Yemen). J'ai nettoyé car les données géographiques ont trop changé de 1850 à 1950 (guerres etc...).

Ensuite, j'ai utilisé une visualisation des données manquantes à l'aide de la bibliothèque Seaborn. Les données nettoyées sont ensuite regroupées par région et par année, puis enregistrées dans un nouveau fichier CSV.

Les hypothèses qu'on cherche à explorer sont les suivantes :

- Existe-t-il des différences des libertés civiles dans le monde en fonction des continents et des années ?
- Y a-t-il globalement à l'échelle mondiale une évolution des libertés civiles dans le monde ?
- Comment la localisation géographique pourrait-elle prédire les libertés civiles ?

Afin de répondre à ces hypothèses nous allons utiliser comme variable prédictrice seront la localisation des pays et régions (donnée catégorielle) et l'année (considéré ici comme variable continue). La variable prédite sera le score de liberté civile (variable continue situer entre 0.009 et 0.978). On utilisera aussi un score présentant la consommation de compotes par pays qui sera entre 0 et 100, et elle est une variable quantitative discrète (c'est une consommation fictive). Avec ce score on regardera si la consommation de compotes peut prédire le score de liberté civiles.

Il y a ensuite deux visualisations de données : un graphique de ligne montrant les scores des libertés civile moyennes pour chaque région entre 1950 et 2021, et un histogramme montrant les scores moyens pour chaque région.

Le projet continue avec une analyse de la tendance temporelle des scores de liberté civile moyens pour l'ensemble du monde. Il y a un graphique de points qui montre l'évolution du score de liberté civile moyen pour chaque année entre 1950 et 2021.

Ensuite on va créer les fausses données de consommations de compotes. Après leur création on va les rajouter au fichier de donnée initial pour en créer un nouveau. On va réaliser quelques statistiques dessus et afficher un graphique.

Le projet se termine par une visualisation interactive du score de liberté civile pour chaque pays dans le monde. Une carte interactive est créée à l'aide de la bibliothèque Plotly Express. Les scores de liberté civile pour chaque pays sont projetés sur une carte du monde en format HTML, avec la possibilité de changer les années pour observer les changements dans les scores de liberté civile au fil du temps.

Pseudo code :

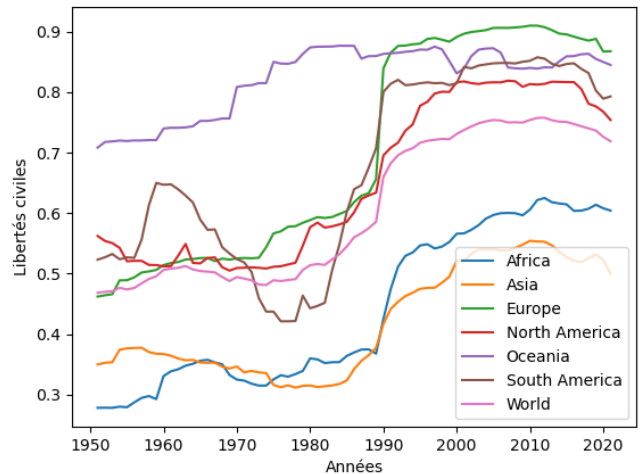
1. Importer les bibliothèques nécessaires.
2. Charger le fichier CSV contenant les données à analyser à l'aide de pandas.
3. Supprimer toutes les données inférieures à 1950 et remplacer "Yemen People's Republic" par "Yemen".
4. Vérifier et afficher les données manquantes et les enregistrer dans un fichier PNG et CSV.
5. Regrouper et enregistrer les données par région et par année en utilisant la méthode `mean()`.
6. Générer un graphique pour le niveau de liberté dans chaque continent entre 1950 et 2021.
7. Créer un histogramme avec la moyenne totale dans le temps pour chaque région.
8. Générer un nuage de points avec une droite de régression pour illustrer l'évolution du score de libertés civiles au fil des ans.
9. Générer des graphiques interactifs pour illustrer l'évolution de la liberté individuelle dans le monde.
10. Rajouter de nouvelles données et les assembler.
11. Faire des statistiques dessus.
12. Charger les données `gapminder` à partir de la bibliothèque `plotly express`. Puis créer un dataframe contenant des pays et leur code `iso_alpha` pour pouvoir les localiser.
13. Attribuer à chaque pays une localisation lorsque celui-ci est dans le dataframe.

14. Projeter les données sur une carte en fonction des années et des pays.

Les hypothèses :

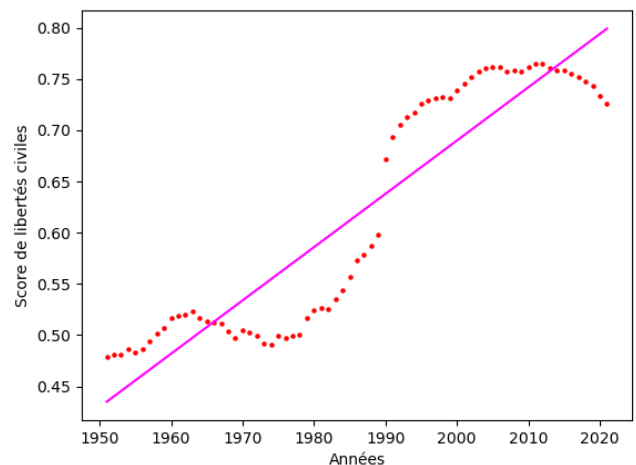
- Existe-t-il des différences au niveau des libertés civiles dans le monde en fonction des continents et des années?

On voit des évolutions marquées et positives vers le haut des libertés civiles dans le monde et qui s'avère différente en fonction des continents. Globalement les libertés civiles dans le monde ont explosé à la chute de l'URSS en 1991.



- Y a-t-il globalement à l'échelle mondiale une évolution des libertés civiles dans le monde ?

On voit une augmentation des libertés civiles dans le monde au fur et à mesure avec toujours cette augmentation brutale au début des années 90. On remarque aussi un déclin depuis la crise de 2008.



- Comment la localisation géographique pourrait-elle prédire les libertés civiles ?

→ Voir graphique [Graphique_libertés_civiles.html](#)

On constate des zones où les libertés civiles sont plus grandes.

Calendrier / échéancier :

Pour réaliser ce projet j'ai d'abord réfléchi à ce que je voulais faire : une carte interactive.

Ensuite j'ai essayé de faire du code régulièrement avec des séances plus intensives. Je l'ai commencé il y a trois semaines, et je codais le lundi, mardi, mercredi et dimanche sur ce projet. Globalement le projet m'a pris deux heures mais ce qui fut plus c'est la réalisation de la carte avec les essais de plusieurs librairies etc... En tout j'ai dû prendre 6 heures.