

Dashboard BIUM

Haoran LI / Marcelin SEBLE

Contents

1	Introduction et Axe d'analyse:	2
2	Schéma en étoile	3
3	Carte géographique	4
3.1	Crime	4
3.2	Cost	5
3.3	Healthcare	6
3.4	Education	7
3.5	Economy	8
3.6	Age(Demography)	9
3.7	Total Score	10
4	Histogramme	11
5	Résultats	12

1 Introduction et Axe d'analyse:

Problématique : Quel est l'État américain le plus approprié pour élever ses enfants?

Telle est la question à laquelle on a décidé d'apporter une réponse en faisant appel aux données récoltées au cours des années précédentes. Ces données concernent les aspects suivants :

- Crime :

Taux de criminalité par État \ Mort par arme à feu

- Coût de la vie :

Prix moyen d'une maison par État \ Coût des soins à plein temps \ Coût moyen de clôture

- Soins de santé:

Qualité de l'accès aux soins \ Santé publique \ Dépenses de santé du gouvernement par enfant

- Education:

Dépenses du gouv d'éducation \ nombre des établissements d'enseignement supérieur.

- Economie:

PIB(Produit Intérieur Brut) par habitant \ Revenu personnel

- Structure d'âge (Démographie):

Age moyen \ Pourcentage des habitants moins de 18 ans

2 Schéma en étoile

Dans notre projet, on analyse chaque dimension et enfin on a un fait qui montre le score total en utilisant la moyenne pondérée.

Donc on n'utilise pas SCHÉMA EN CONSTELLATION, on vous montre que le Schéma en étoile ci-dessous:

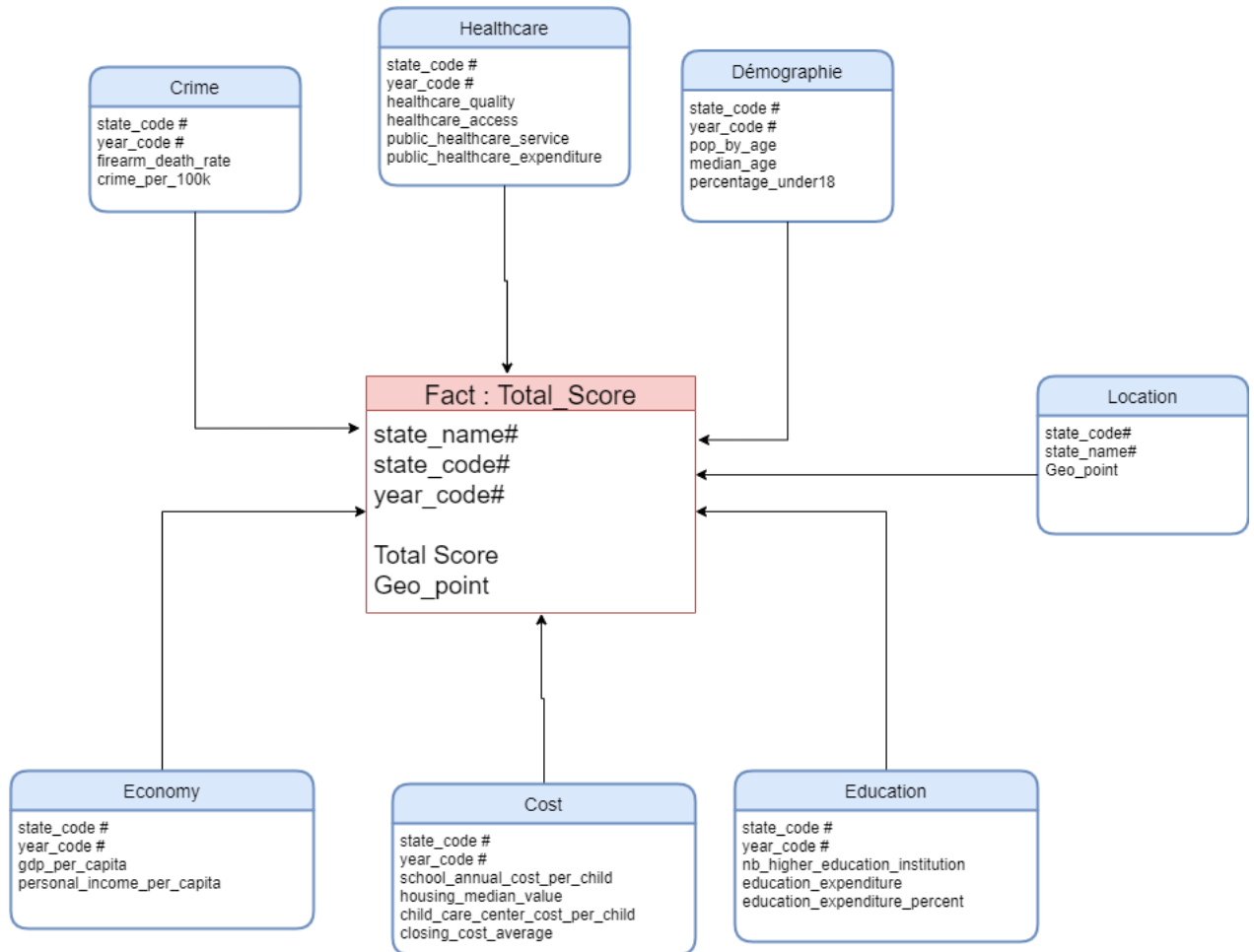


Figure 1: Schéma en étoile

3 Carte géographique

La principe de Nomalisation des scores a déjà expliqué dans [Technique Rapport](#).

3.1 Crime

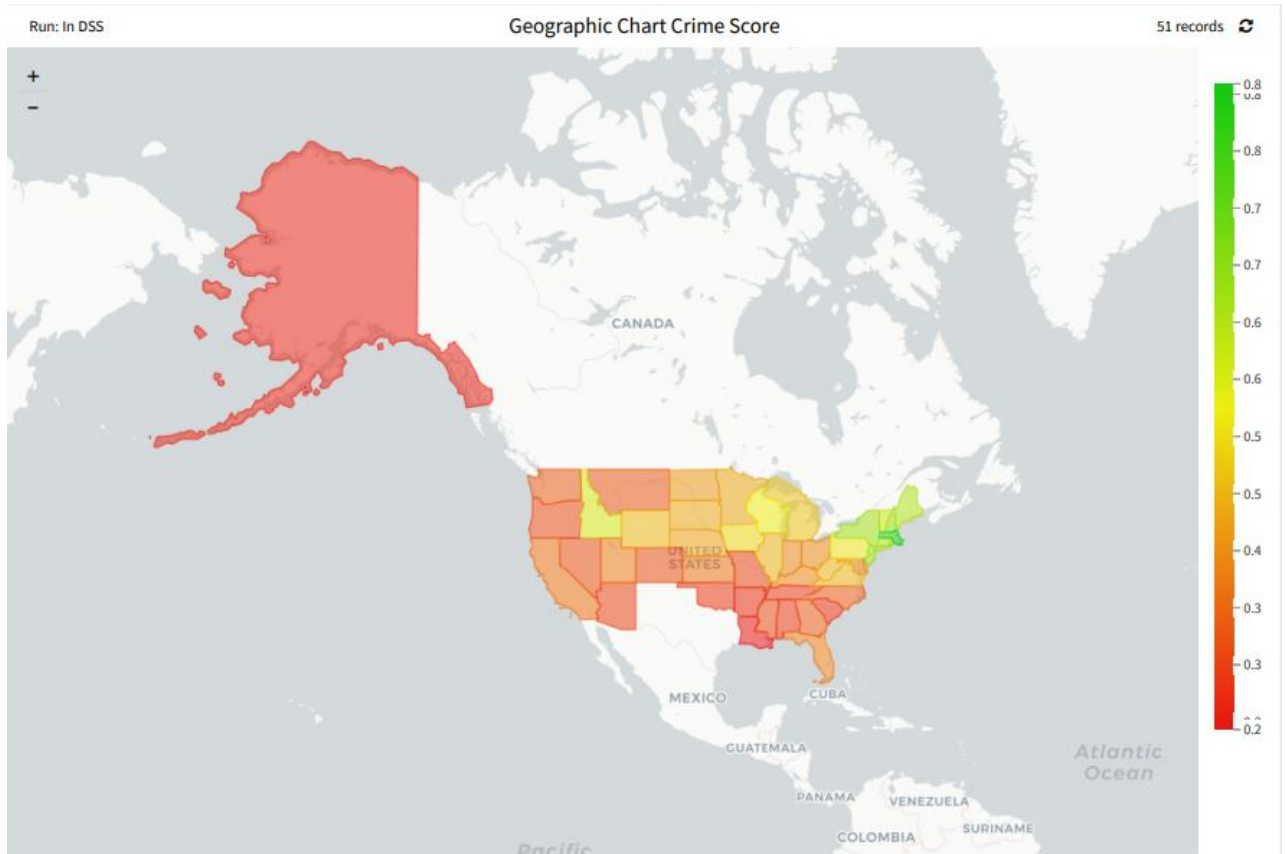


Figure 2: Crime Score

On utilise la formule :

$$\text{Crime score} = 0.5 * \frac{0.185 - \text{firearmDeathRate}}{0.185 - 0.021000} + 0.5 * \frac{4367.1 - \text{crime per 100k}}{4367.1 - 1179.8}$$

- Remarque:

Pour le poids, on pense que **taux de criminalité** et **blessures par balle** sont même importants. Donc les poids est [0.5, 0.5].

Sur cette carte, on voit que normalement, l'Est est plus sécurisé que l'Ouest. Par contre, en Alaska on remarque que le taux de sécurité est bas.

3.2 Cost

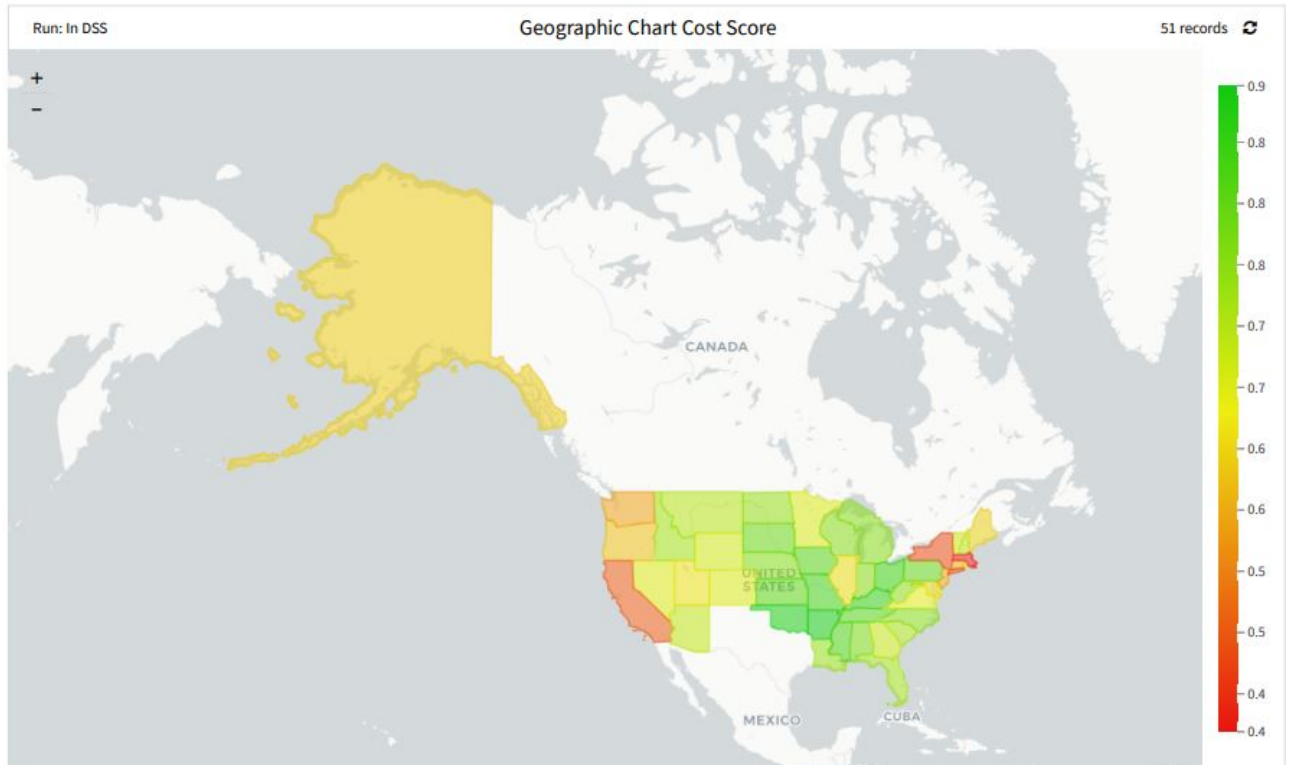


Figure 3: Cost Score

On utilise la formule :

$$\text{Cost score} = 0.25 * \frac{615300 - \text{MedianValue}}{615300 - 119000} + 0.25 * \frac{7280 - \text{cost for school age child}}{7280 - 2423} + 0.25 * \frac{2648 - \text{Average closing costs}}{2648 - 1734} + 0.25 * \frac{24.08 - \text{Full time care cost}}{24.08 - 5.7600}$$

- Remarque:

Ici, on trouve les coûts relatifs de la vie, on fait l'hypothèse qu'ils ont la même importance [0.25, 0.25, 0.25, 0.25]. En effet, Score de coût ne représentait que peu pour le score total.

3.3 Healthcare

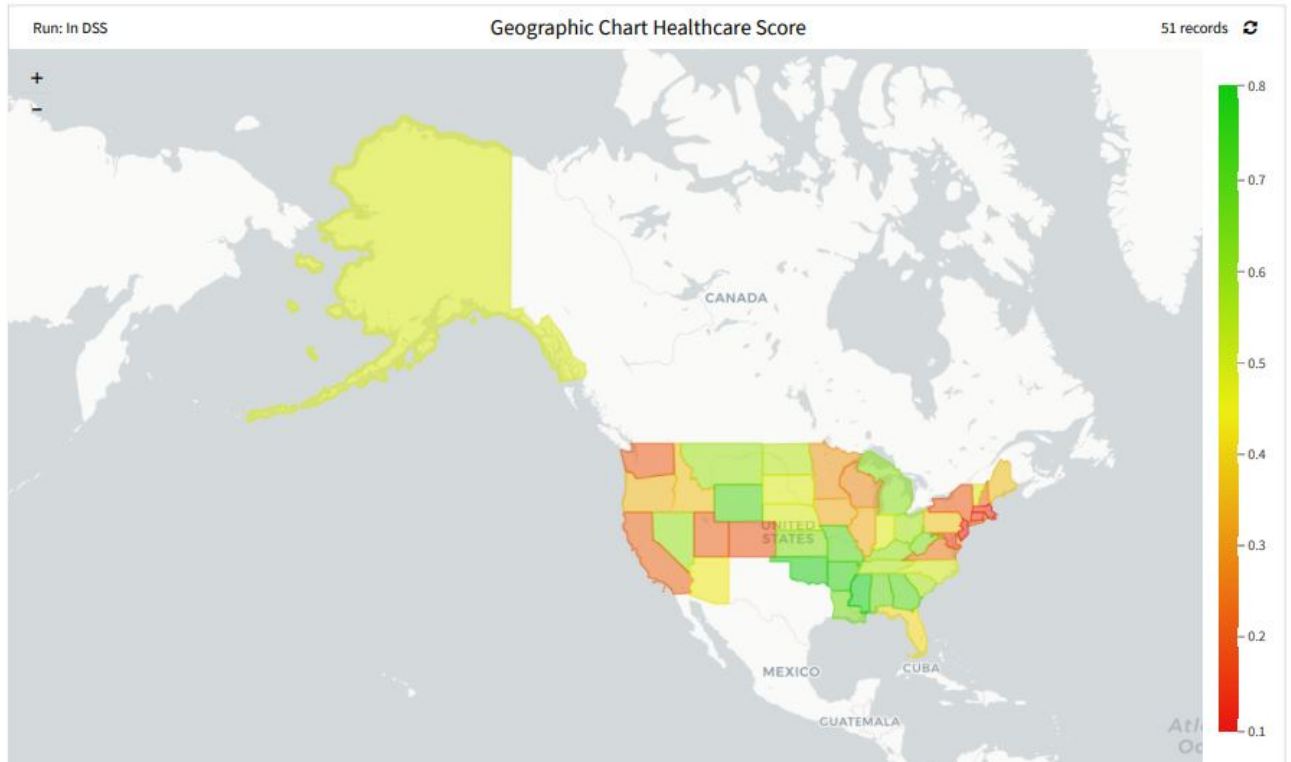


Figure 4: Healthcare Score

On utilise la formule :

$$\text{Healthcare score} = 0.2 * \frac{\text{health expenditure per child} - 465.39}{3436.9 - 465.39} + 0.8 * \frac{\text{health care rank} - 8}{147 - 8}$$

- Remarque:

Parce que *healthcare rank* vient de USnews(Sourcé donné en Technique Rapport). On fait l'hypothèse qu'il est plus fiable et important, donc le poids est [0.2, 0.8].

Le niveau médical et les ressources dans le sud-est sont les meilleurs.

3.4 Education

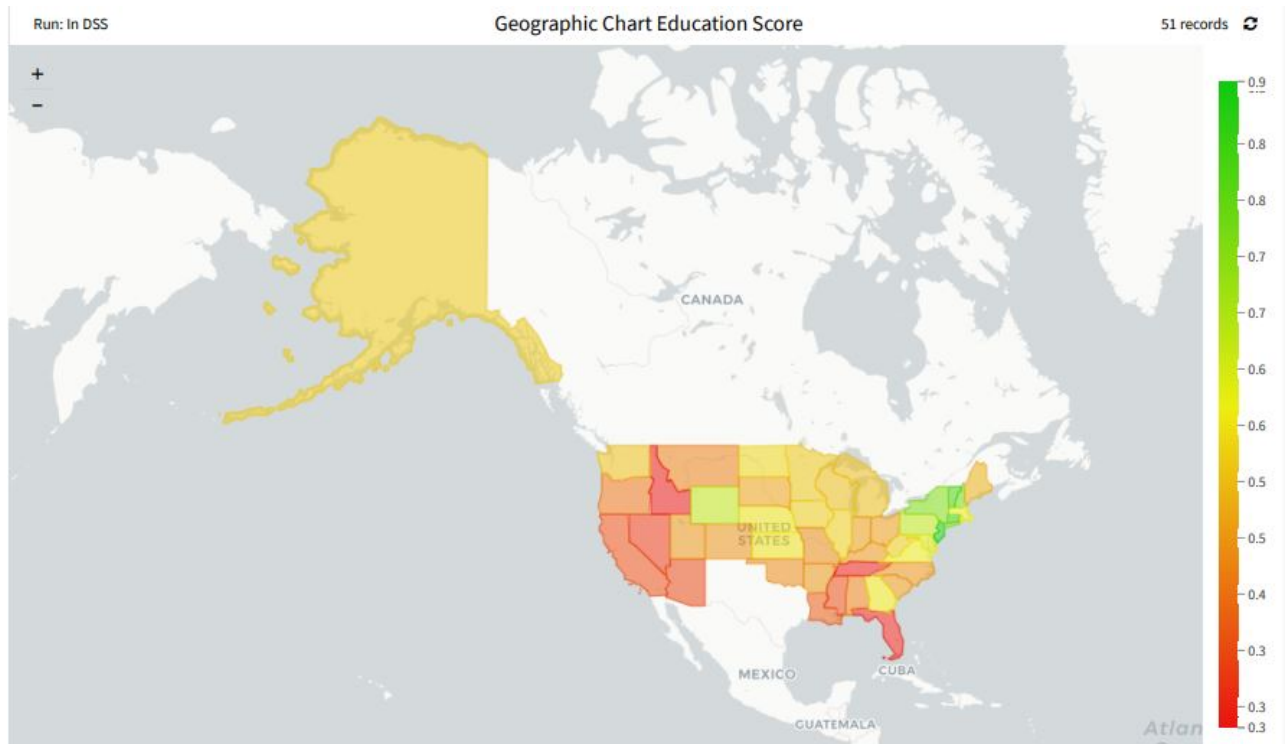


Figure 5: Education Score

On utilise la formule :

$$\text{Education score} = 0.495 * \frac{\text{elementary and secondary school expenditure} - 6747}{24749 - 6747} + 0.495 * \frac{\text{percent} - 16.56}{35 - 16.56} + 0.01 * \frac{\text{number total} - 8}{416 - 8}$$

- Remarque:

Rappelons que notre problématique porte sur l'Etat idéal dans lequel élevé son enfant, donc le nombre des d'écoles d'enseignements supérieurs n'est pas trop important. Par contre, écoles primaires et secondaires sont plus importantes. En se basant sur cette logique, le poids est [0.495, 0.495, 0.01].

3.5 Economy

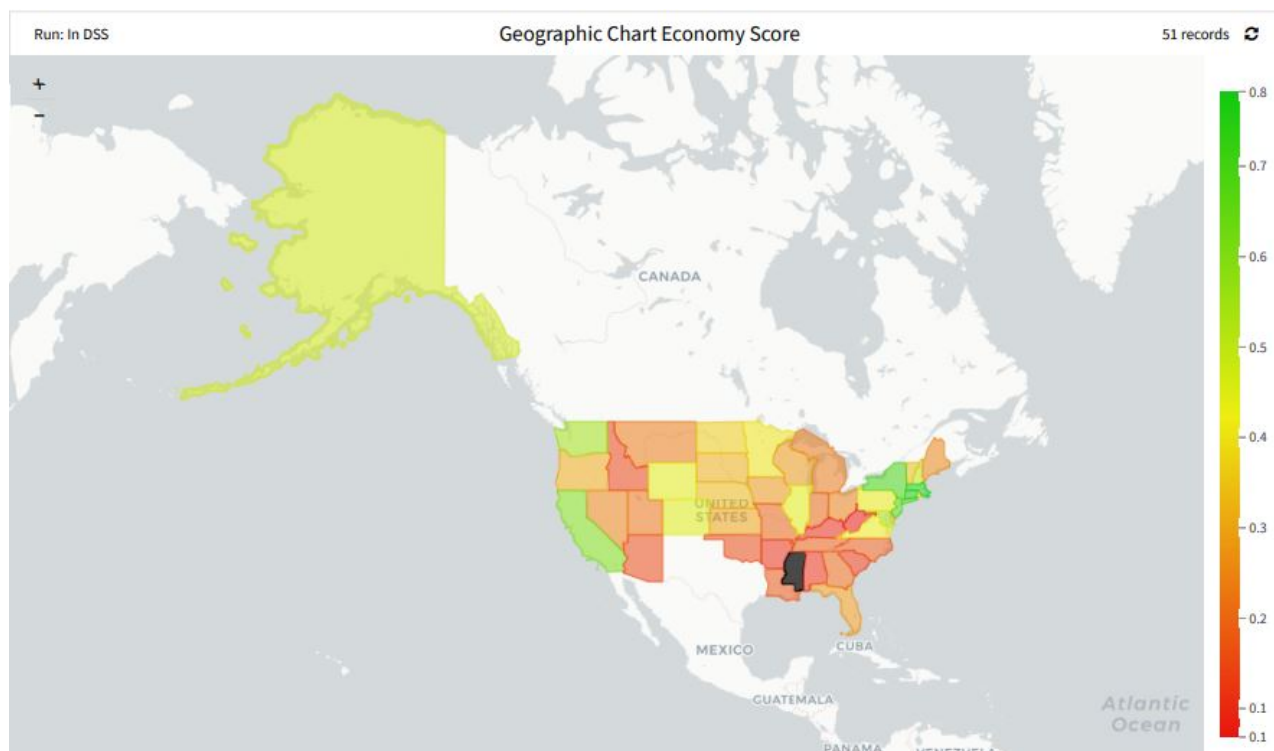


Figure 6: Economy Score

On utilise la formule :

$$\text{Economy score} = 0.1 * \frac{GDP \text{ per capita} - 38493}{201360 - 38493} + 0.9 * \frac{Income - 41745}{87064 - 41745}$$

- Remarque:

Le PIB est généralement pour un État dans son ensemble. En revanche, le revenu personnel est plus pertinent ici. Donc le poids est [0.1, 0.9]. Avec cette hypothèse et à partir des données qu'on a à notre disposition on remarque que l'Etat du Mississippi économie a le pire.

3.6 Age(Demography)

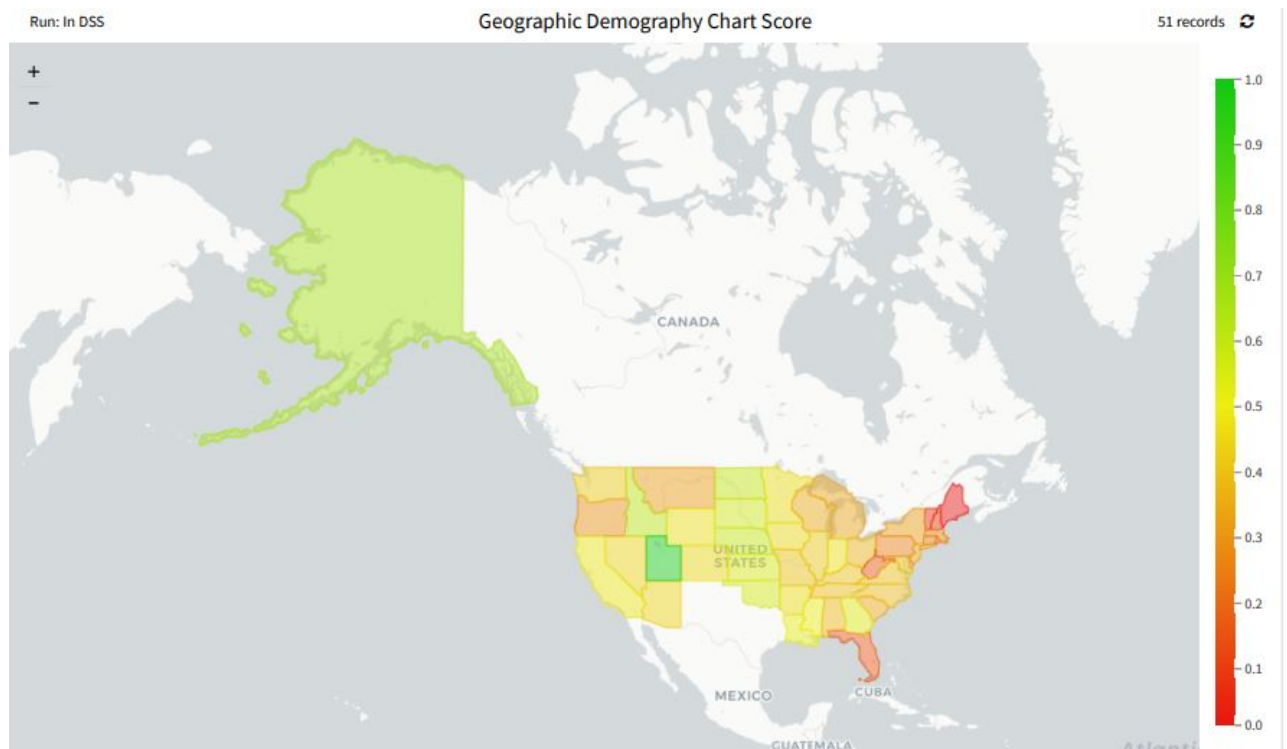


Figure 7: Demography Score

On utilise la formule :

$$\text{Age(Demography) score} = 0.5 * \frac{44.9 - \text{Median}_{age}}{44.9 - 31} + 0.5 * \frac{\text{PCNT}_{under18} - 18.2}{29 - 18.2}$$

- Remarque:

Nous avons trouvé deux facteurs très importants, l'âge moyen et la proportion de mineurs. Donc le poids est [0.5, 0.5].

3.7 Total Score

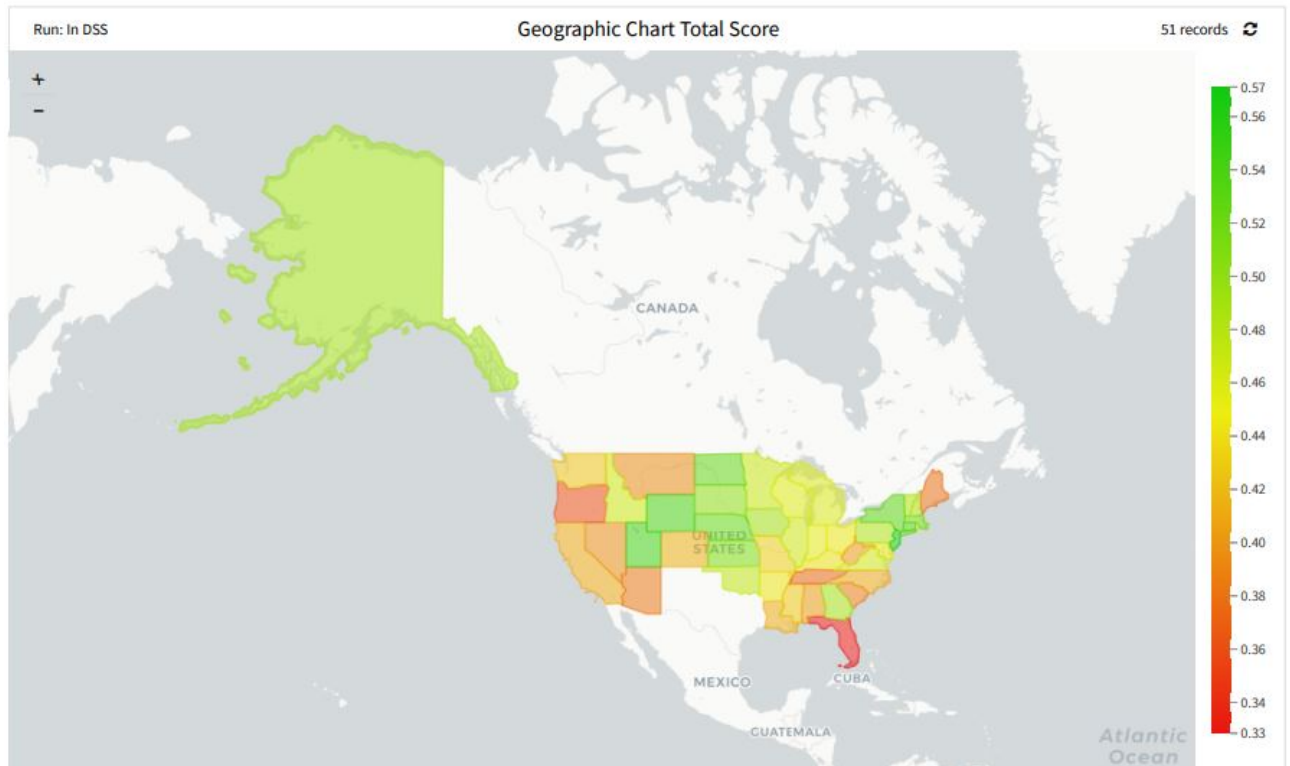


Figure 8: Age(Demography) Score

On utilise la formule :

$$\text{Total score} = (0.1 * \text{Cost Score} + 0.3 * \text{Age Score} + 0.1 * \text{Economy Score} + 0.1 * \text{Healthcare Score} + 0.3 * \text{Education Score} + 0.3 * \text{Crime Score}) / 1.2$$

- Remarque:

Pour le score total, c'est encore moyenne pondérée finalement.

Après réflexion, on a fait l'hypothèse que les axes "Éducation", "Age(Demography)" et "Crime", sont les 3 facteurs les plus importants. Parce qu'on recherche un environnement propice pour l'éducation des enfants.

En conclusion, le poids de score total est $[0.1, 0.1, 0.1, 0.3, 0.3] / 1.2$.

4 Histogramme

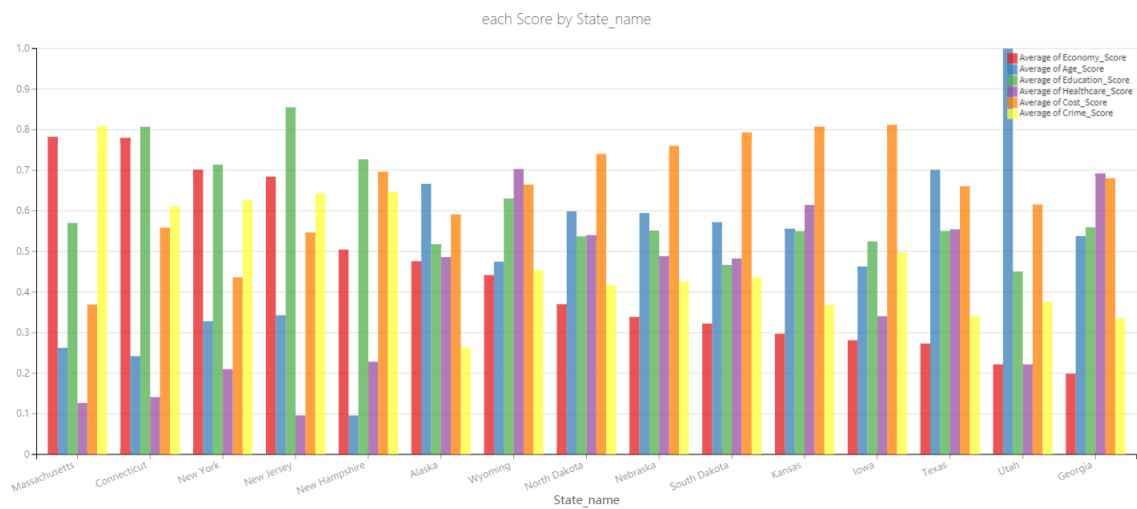


Figure 9: Chaque Score pour 15 premiers

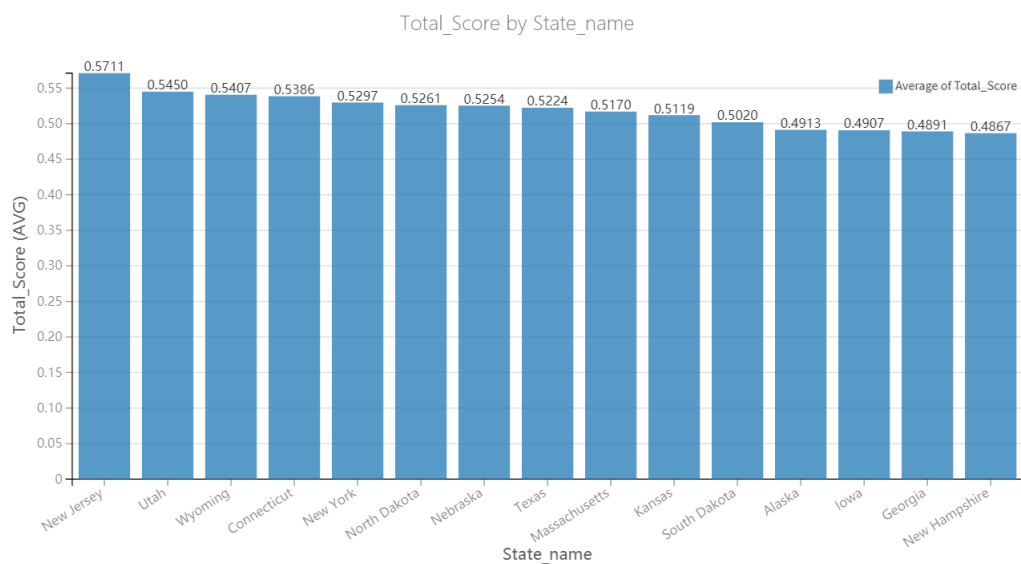


Figure 10: Score Total pour 15 premiers

5 Résultats

● Notre palmarès

Les 15 premiers États américains les plus appropriés pour élever ses enfants sont :

- | | | | | | | |
|---------------|-------------------|------------------|----------------|------------------|-----------------|----------|
| 1. New Jersey | 2. Utah | 3. Wyoming | 4. Connecticut | 5. New York | 6. North Dakota | |
| 7. Nebraska | 8. Texas | 9. Massachusetts | 10. Kansas | 11. South Dakota | 12. Alaska | 13. Iowa |
| 14. Georgia | 15. New Hampshire | | | | | |

● Palmarès de moneyrates

- | | | | | | |
|------------------|------------------------|---------------|----------------|------------------|------------|
| 1. New Jersey | 2. (tie) Massachusetts | 2. (tie) Utah | 4. Connecticut | 5. New Hampshire | 6. Wyoming |
| 7. (tie) Florida | 7. (tie) Wisconsin | 9. Ohio | 10. New York | | |

● Palmarès de USNews (according to WalletHub)

- | | | | | | |
|------------------|---------------|-----------------|-------------|------------|------------------|
| 1. Massachusetts | 2. Minnesota | 3. North Dakota | 4. New York | 5. Vermont | 6. New Hampshire |
| 7. New Jersey | 8. Washington | 9. Connecticut | 10. Utah | | |

Les 2 sources sont données en [Technique rapport](#).

On les compare, et on trouve ce qu'on fait est bien, ils sont en bon accord. Sûrement, il y a des différences ce qui est le cas entre les deux autres classement, leurs palmarès sont différents. On rappelle aussi que nous ne disposons pas des mêmes données et surtout que les données peuvent changer avec le temps. Nous ne nous sommes sûrement pas basées sur les mêmes axes d'analyses pour réaliser ce classement, c'est l'autre raison qui pourrait également expliquer les différences.

En conclusion, ce projet nous a permis d'explorer les outils permettant de faire le business intelligence. À la fin, on peut dire que notre problématique (**Quel est l'État américain le plus approprié pour élever ses enfants ?**) a été résolue. Comme amélioration, on a pensé à faire usage des techniques du machine learning afin d'éviter de faire des hypothèses fortes comme on a eu à faire avec le choix des poids pour le calcul des scores.

● Remarque:

Si tu veux consulter toutes les cartes géographiques, je vous invite d'aller dans notre Dashboard de Dataiku. En fait, vous pouvez consulter les détails (le nom et le score) de tous les états sur la carte.

Username :lih / Password : haoran