

PROJET SUR L'EDUCATION: NETTOYAGE DE DONNÉES

INTRODUCTION

En faisant ce projet j'avais deux buts. Le premier, ciblant mes compétences techniques, était d'apprendre les outils de Python pour le nettoyage de données. Le deuxième était de comprendre quelques différences entre l'éducation de lycéens dans deux grandes villes Américaines: Milwaukee, Wisconsin et Las Vegas, Nevada. Ce choix de villes était fait en raison de leurs tailles similaires et leurs démographiques assez différentes. À Milwaukee, les races sont très séparées, mais à Las Vegas, elles sont plus mêlées.

Je me suis concentrée sur les trois questions suivantes:

- (1) Existe-il une différence entre les types de lycées (publics, privés, à charte, ou alternatifs¹) fréquentés par les lycéens des deux villes?
- (2) Quelles différences existent-elles entre les compositions raciales des lycées de chaque ville et celle de la population de cette ville?
- (3) Des différences d'accès des examens *Advanced Placement (AP)*² existent-elles entre les lycéens des deux villes ou entre ces lycéens et ceux du pays entier?

Afin de répondre aux questions, j'utilisais des données tirées des sources suivantes:

- (1) The Civil Rights Data Collection³,
- (2) The National Center for Education Statistics⁴ (NCES) (et ici)
- (3) The College⁵ Board, et
- (4) The United States Census Bureau Fact Finder⁶.

QUESTION 1:

Pour traiter la première question, j'ai dû commencer par résoudre un petit problème des données. En fait, il y avait une grosse différence entre les compositions des deux secteurs scolaires. Le *Milwaukee Public School District* inclut exactement des lycées situés dans les limites de la ville de Milwaukee. Par contre, les écoles de Las Vegas appartiennent au *Clark County School District*, ce qui inclut aussi des écoles de banlieue et celles des zones rurales. Par conséquent, la comparaison directe entre ces deux secteurs ne me semblait pas très éclairante. Alors, j'avais commencé en filtrant des lycées en fonction de leurs codes de location, c'est-à-dire, la description du quartier du lycée. J'avais gardé seulement les lycées situés dans des grandes villes urbaines.

¹Aux États-Unis, les lycées alternatifs desservent plutôt des élèves avec des gros difficultés d'apprentissage ou de comportement.

²Les examens *AP* permettent des lycéens d'obtenir des crédits universitaires.

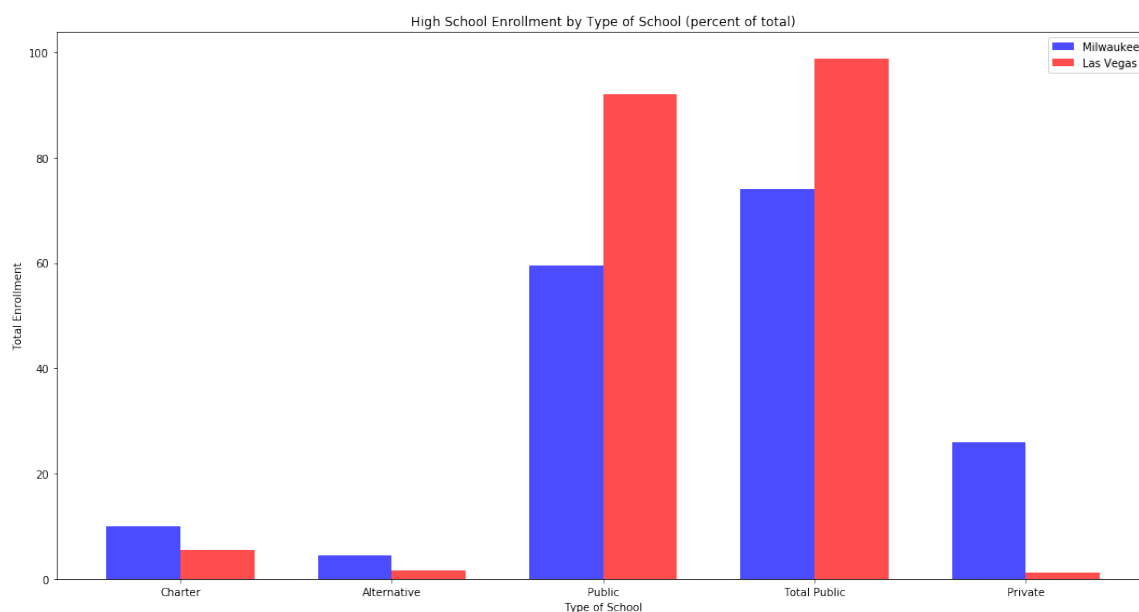
³La collecte de données sur les droits civils.

⁴Le centre américain des données statistiques en éducation.

⁵N.B. Aux États-Unis, le mot *college* fait référence à une université proposant uniquement des diplômes de premier cycle. Les collèges en France sont des *middle schools* aux États-Unis.

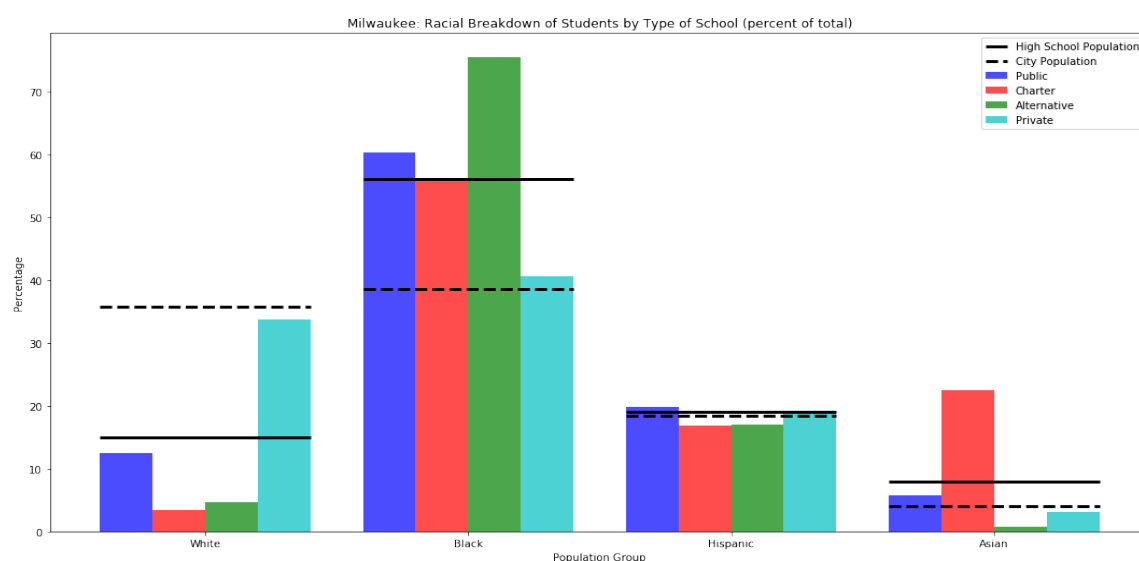
⁶Le bureau du recensement des États-Unis.

En considérant les choix de lycée dans les deux villes, on trouve les résultats suivants. En Milwaukee, il y avait une grande partie de lycéens dans les lycées privés (26%) et à charte (10%), tandis que ces valeurs en Las Vegas sont 1% et 5%, respectivement. Par contre, Milwaukee a beaucoup moins de lycéens dans les lycées publics que Las Vegas (50% contre 92%). Ici, il faudrait observer que l'état de Wisconsin possède un système robuste des chèques d'éducation⁷ ainsi que le soutien considérable pour les écoles et lycées à charte.



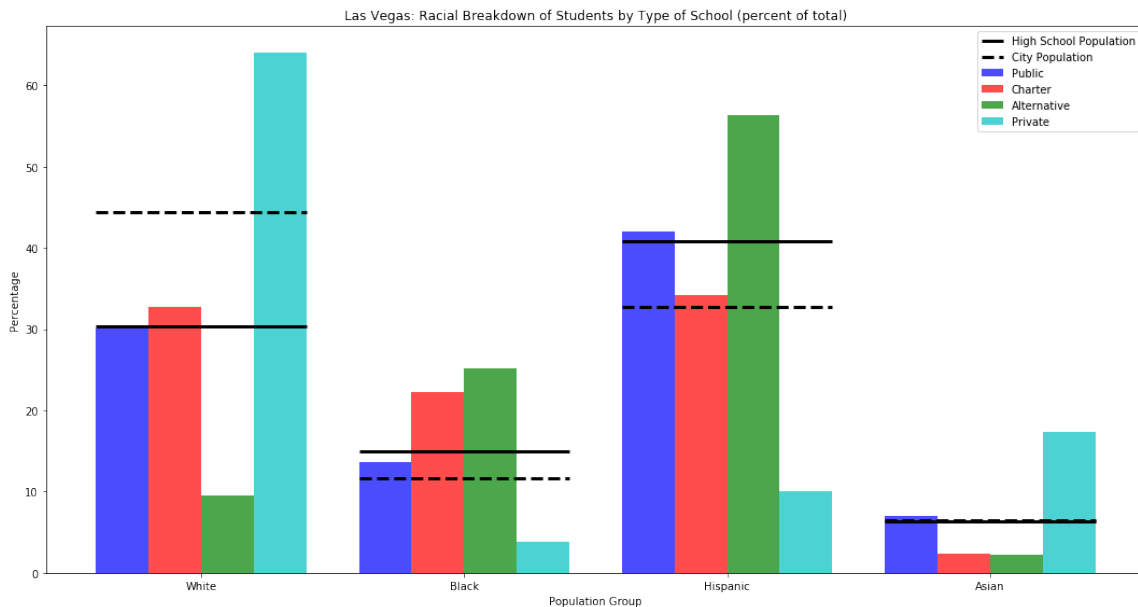
QUESTION 2:

En considérant les populations de lycéens dans les quatre types de lycée ainsi que celles des villes, on obtient les résultats suivants. À Milwaukee, la population de la ville est plus blanche que celle des lycéens. En plus, les lycées privés manquent des lycéens minoritaire, tandis que des lycéens noirs sont surreprésentés aux lycées alternatifs.



⁷Les chèques d'éducation, ou *vouchers* en anglais, permettent les élèves de payer les cours privés avec des fonds publics.

À Las Vegas, la population de la ville est aussi plus blanche que celle des lycéens. Ici c'est les lycéens blancs et asiatiques qui sont surreprésentés dans les lycées privées, alors que les lycéens noirs et latino-américains sont surreprésentés parmi la population des lycées alternatifs. De plus, les lycéens noirs sont surreprésentés parmi les lycéens à charte.



QUESTION 3:

Pour la troisième question, qui s'adresse aux examens AP, j'ai examiné les taux d'accès et de réussite aux examens. Des nombreux éléments ont compliqué ce travail. Premièrement, les sources ne décrivent pas les mêmes groupes d'élèves. Pour les examens AP de Milwaukee et de Las Vegas, j'avais les données sur les lycéens aux lycées publics, à charte, et alternatifs⁸. Je n'avais pas les données sur les lycées privés. Par contre, les données du College Board se compose de l'information sur tout lycéen du pays ayant passé un examen AP. Comme je voulais comparer les taux d'accès et de réussite des groupes de lycéens, j'ai essayé de résoudre ce problème en utilisant les données du NCES. Avec cette information, j'ai estimé le nombre de lycéens de chaque race dans chaque type de lycée au niveau national. Cela m'a permis d'estimer les taux d'accès et de réussite pour ces groupes d'élèves au niveau national.

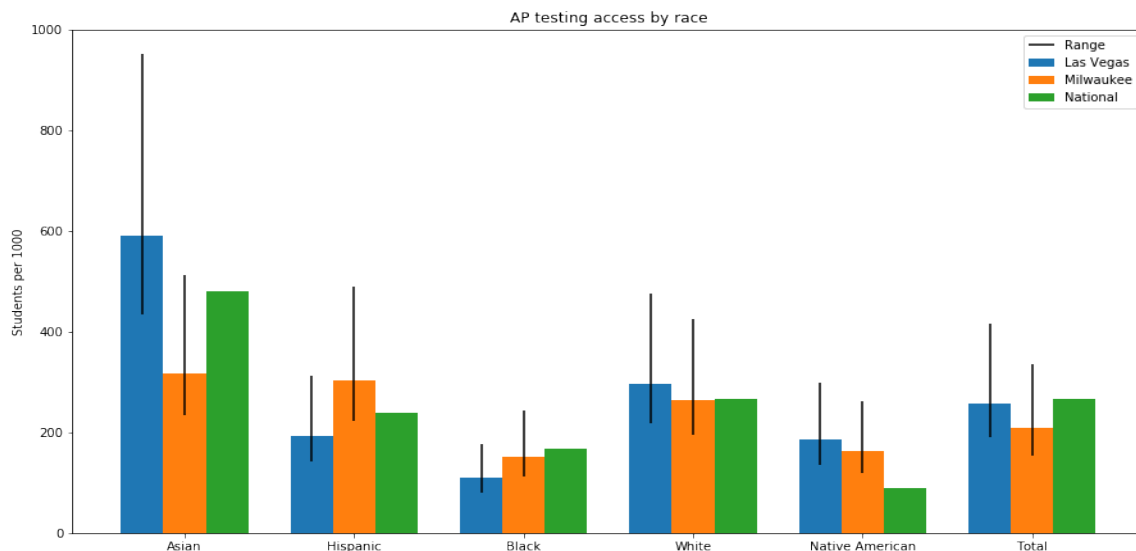
Un autre problème vient des données des villes obtenues sur le site web du *Civil Rights Data Collection*. Les sous-groupes ayant moins de trois membres ont été donnés la valeur ' ≤ 2 '. Cela ajoutait de l'incertitude autour des données. Pour démêler des données autant que possible, j'ai utilisé les valeurs connues à essayer de déterminer des valeurs manquantes⁹. En cette manière, je pourrais remplacer la plupart des valeurs ' ≤ 2 ' avec des valeurs fixes. Enfin, les données du College Board sont réparties par classe¹⁰. Par contre, dans les données de Milwaukee et de Las Vegas, les accès aux examens APs du troisième au terminale sont combinés. Ici j'aurais dû faire des suppositions. D'abord, j'ai supposé que la distribution des classes des lycéens des deux villes qui passent les examens AP est le même que celle au niveau national. Deuxièmement, j'ai

⁸Ces lycées reçoivent des fonds publics.

⁹Par exemple, si on connaît qu'il y aurait 20 lycéens qui a passé un examen, et parmi eux, il y aurait 15 lycéens noirs et 5 lycéens asiatiques, on sait déjà que aucun lycéen blancs n'a passé l'examen.

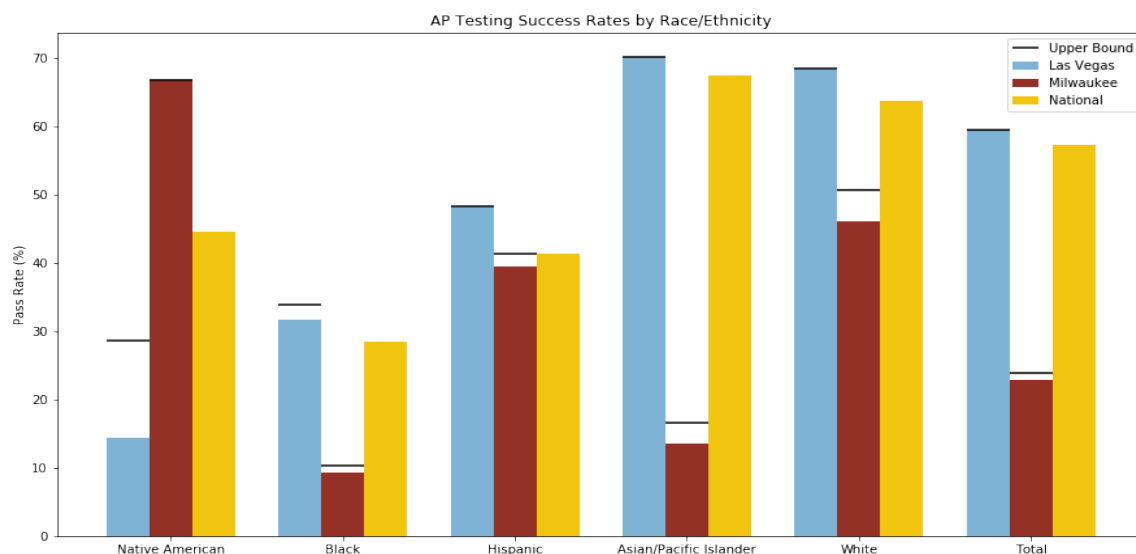
¹⁰Les lycéens peuvent passer les examens APs pendant n'importe quelle année de lycée. Ils peuvent aussi les passer pendant plusieurs années, si cela les convient.

supposé que la grande majorité d'élèves qui a passé un examen AP avant la terminale, aurait passé d'autre en terminale.



J'ai trouvé que certaines groupes de lycéens dans les villes semblent d'avoir plus d'accès aux examens AP que leurs analogues au niveau national. À Las Vegas, c'étaient les lycéens asiatiques ou amérindiens qui passent les examens avec plus de fréquence que leurs pairs au niveau nationale. À Milwaukee, c'étaient les lycéens latino-américains ou amérindiens. Par contre, il existait aussi des groupes avec un taux d'accès assez bas, comme des lycéens asiatiques à Milwaukee et des lycéens noirs à Las Vegas.

Au niveau de taux de réussite, on trouve encore des différences entre des groupes. Parmi eux, les lycéens amérindiens à Milwaukee ainsi que les lycéens blancs ou latino-américains à Las Vegas ont eu plus de réussite que leur pairs au niveau national. Par contre, les élèves amérindiens à Las Vegas ainsi que les élèves noirs, blancs, ou asiatiques à Milwaukee semblent d'avoir moins de succès que leurs pairs au niveau national. Au niveau de ville, les lycéens de Las Vegas ont connu un réussite légèrement supérieur que ceux au niveau national, tandis que les élèves de Milwaukee ont montré beaucoup moins.



CONCLUSION:

En regardant nos données, nous voyons qu'il existe des différences significatives dans les types de lycées choisis par les parents du lycéens. En particulier, à Milwaukee et à Las Vegas, nous constatons que la ségrégation entre les types d'écoles est assez forte, bien que Las Vegas ait vu moins de fuite vers les lycées privés que Milwaukee. Peut-être pour cette raison, alors que l'accès aux examens AP pour les élèves des lycées publiques des deux villes est relativement similaire à celui au niveau national, Milwaukee montre un manque de réussite frappant par rapport au niveau national.