

Rapport Business Intelligence



Accident de voitures aux Etats-Unis

1. Introduction

Nombreux sont les accidents routiers qui arrivent chaque jour dans le monde. Les Etats-Unis n'échappent pas à la règle avec environ 38000 morts et plus de 4 millions de blessé par ces mêmes accidents. Aujourd'hui, l'objectif est de baisser le nombre d'accident de la route. Pour atteindre cet objectif, nous allons nous poser 3 questions :

- Quelles sont les régions les plus touchées par les accidents de voitures ?
- Les accidents graves sont majoritaires ou minoritaires ?
- Quelles sont les indicateurs environnementaux causant le plus d'accidents ?

Répondre à ces questions nous permettra de pouvoir prendre les mesures nécessaires dans les zones adéquats selon la gravité des accidents.

Pour répondre à ces questions, nous allons étudier :

1. Le nombre d'accidents aux Etats-Unis
2. Le nombre d'accidents dans le temps
3. Les facteurs environnementaux de l'accidents

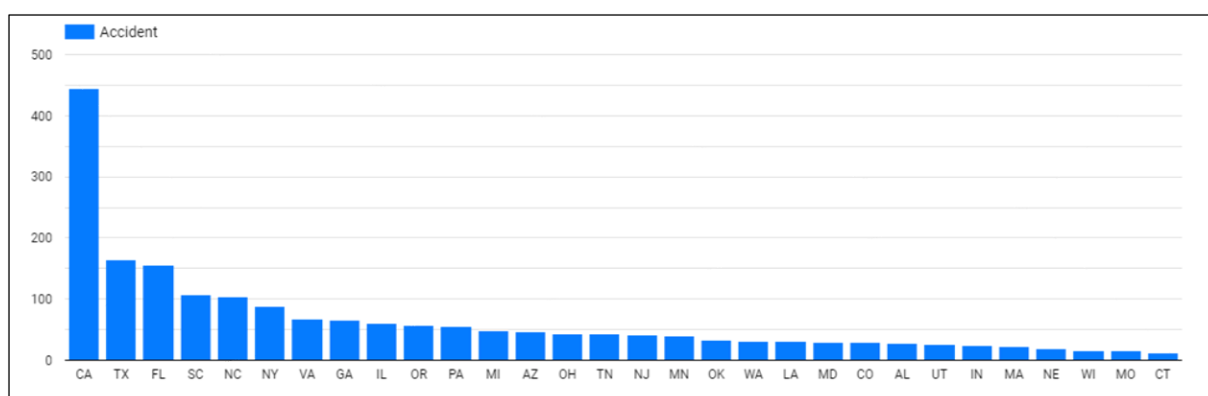
2. Adaptation du Jeu de données

Dans un premier temps, on a dû adapter notre jeu de données car il contenait beaucoup trop de données (quelques millions de lignes). Le problème était que, faute de la quantité de données, le temps d'exécution était augmenté et donc, les performances en été impacté. Il a donc été décidé de créer une version mini de note jeu de données avec une diminution des données à 2000 lignes toutes choisis aléatoirement permettant tout de mêmes d'avoir des résultats satisfaisant malgré la quantité de données réduite.

Dans un second temps, il a été décidé de ne pas enlever les valeurs nulles. La raison étant que même si elles sont nulles, de pouvoir tout de mêmes être au courant qu'il y a eu un accident et de pouvoir traiter l'accident avec les colonnes qui ne seront pas nulle.

3. Etude du nombre d'accidents en fonction de l'état.

Nous allons tout d'abord essayer de voir quelles sont les états les plus touchées par les accidents de voitures :



Annexe 1 : Histogramme du nombre d'accidents par Etat

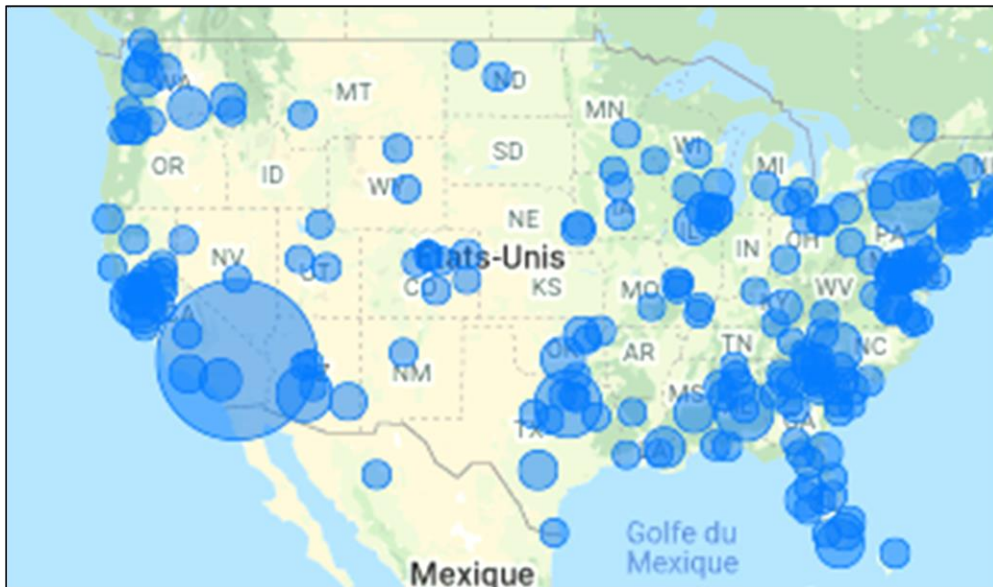
Nous pouvons voir dans le diagramme ci-dessus que certains états ont beaucoup plus d'accidents que d'autres. En l'occurrence :

- La Californie (CA) avec environs 490 accidents qui est bien au-dessus des autres états
- Le Texas (TX) avec 220 accidents
- La Floride (FL) avec 210 accidents
- La Caroline du Sud (SC) avec 110 accidents

ont beaucoup plus d'accidents que la plupart des autres pays. Après, avoir fait des recherches, je me suis aperçu qu'il y avait une corrélation entre le nombre d'accidents d'un état et sa population. En effet :

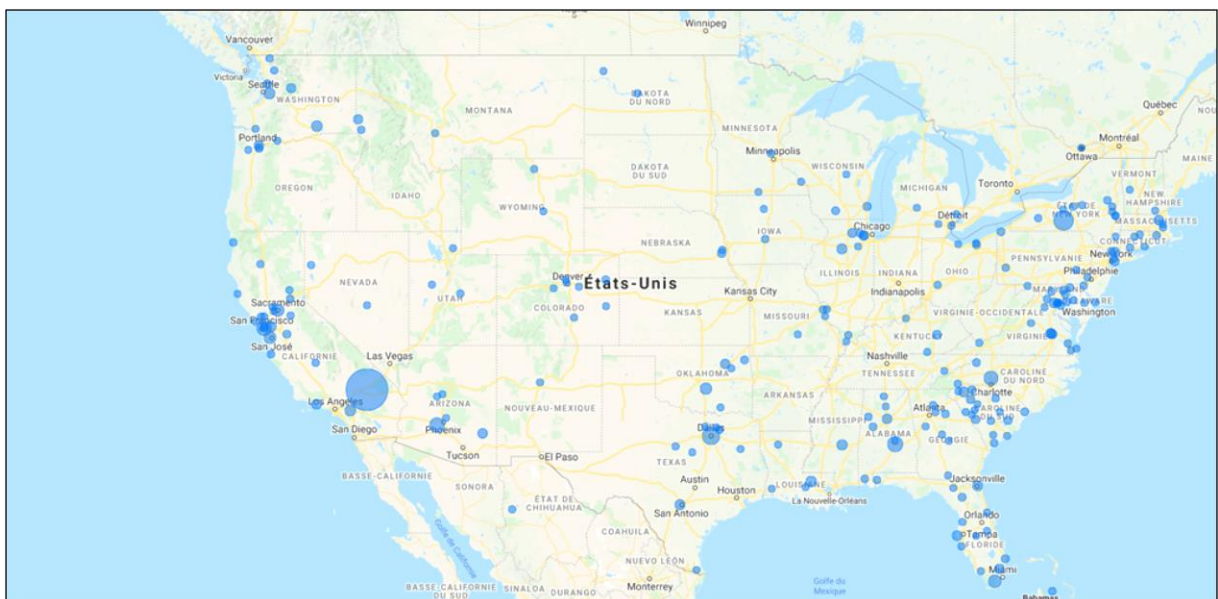
Etats	Habitants
Californie	39 557 045
Texas	28 701 845
Floride	21 299 325
New York	19 542 209
Pennsylvanie	12 807 060

Annexe 2 : Tableau des 5 Etats les plus peuplés des Etats-Unis



Annexe 3 : Carte des Etats-Unis

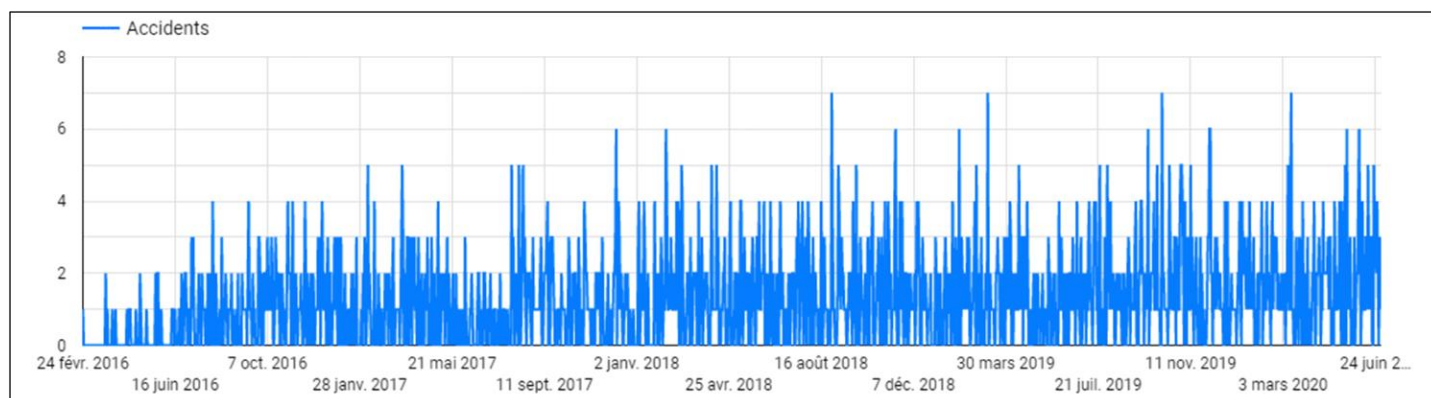
Cette carte nous donne une vision globale des accidents aux Etats-Unis. On peut voir que les accidents sont beaucoup plus importants dans les zones proches de la mer et les grands états. Cette carte vient nous conforter dans notre hypothèse disant que ce sont les états les plus peuplés qui sont le plus touchées par le nombre d'accidents.



Annexe 4 : Carte des Etats-Unis plus global.

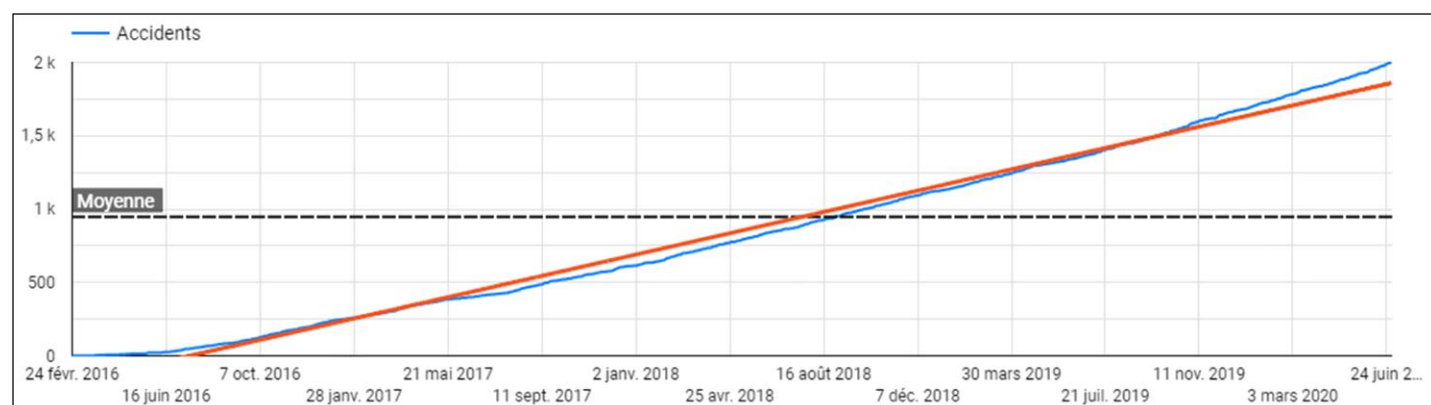
4. Etude du nombre de d'accidents dans le temps

Nous allons désormais étudier l'évolution du nombre d'accidents dans le temps. Comme nous pouvons le voir dans le diagramme ci-dessous, le nombre d'accidents est variable selon le jour :



Annexe 5 : Représentation de l'évolution des accidents dans le temps.

Cependant, le diagramme ci-dessous va nous apporter plus de précision :



Annexe 6 : Représentation de l'évolution des accidents dans le temps.

La droite bleue fait un cumule des données permettant d'avoir une vision plus précise de l'évolution du nombre d'accident. Ainsi on s'aperçoit que plus l'on avance dans le temps, plus le nombre d'accidents augmente. La droite rouge représentant la ligne de tendance vient nous conforter dans cela en nous montrant une droite ascendante.

Cela pourrait s'expliquer par le fait que la population mondiale ne fait qu'augmenter comme nous pouvons le voir dans un article de « LE MONDE » (voir Annexe). Et comme nous avons pu le voir précédemment, plus la population est forte, plus le nombre d'accidents est élevé.

5. Facteurs environnementaux de l'accident

Nous nous sommes intéressés aux différents facteurs environnementaux présent lors de l'accidents. Plus précisément, les facteurs suivants :

- Présence d'un croisement
- Accident au lever/couché du soleil
- Panneaux de signalisation
- Température (F)
- Humidité de l'air

En plus de nos de KPI :

- Météo lors de l'accident
- Gravité de l'accident (entre 1 et 4)

Voici les résultats obtenus :

Facteurs	Commentaires
Croisements	Le nombre d'accident est plus élevé lorsqu'il n'y a pas de croisements.
Lever/couché du soleil	Le nombre d'accidents est plus élevé lors du lever que du coucher de soleil.
Panneaux de signalisation	Lorsqu'il n'y a pas de panneau de signalisation, le nombre d'accident est plus grand.
Température	Au niveau de la température c'est assez hétérogène.
Humidité	Au niveau de l'humidité de l'air, il y a plus d'accident lorsque l'air est humide.
Gravité	Il y a plus d'accident grave soit avec un fort impact sur la circulation que l'inverse.
Conditions météorologique	Il y a plus d'accident lorsque le temps est clair et dégagé.

Ces résultats pourraient s'expliquer par le fait que les personnes sont moins vigilantes lorsque l'environnement a moins d'impact sur eux, par exemple :

- Lorsqu'il ne pleut pas.
- Lorsque la voie semble dégagée.
- Lorsqu'il n'y a pas de feu tricolore.

6. Conclusion

En conclusion, nous avons plus cibler les zones les plus touchées aux Etats-Unis par les accidents de voitures. Nous savons également que les accidents les plus graves, soit avec une gravité de 4 sont majoritaires. De plus, nous savons maintenant que :

- Les régions les plus touchées par le nombre d'accidents sont celles qui sont le plus peuplées
- Que dans l'état, aux vues du faite que la population est en perpétuelle augmentation, le nombre d'accidents est amenée à augmenter s'il n'y a aucun changement
- Hypothétiquement, il y a plus d'accident lorsque les usagers de la route sont moins vigilants de part des facteurs environnementaux moins contraignant.

Il faudra alors être plus vigilant sur :

- Dans les zones peuplées
- Dans les zones moins restreintes par la signalisation

7. Glossaire

Mot	Définition
Ligne de tendance	Une ligne de tendance est une droite ascendante (ou descendante) qui fait support (ou résistance). Les lignes de tendances sont à la base de l'analyse technique. Elles vous permettent d'identifier la tendance de fond d'un actif et peuvent être utilisées sur l'ensemble des unités de temps.
KPI	Un indicateur clé de performance (ICP) – KPI en anglais – est un indicateur utilisé pour l'aide à la décision dans les organisations. Les ICP sont utilisés particulièrement en gestion de la performance organisationnelle.

8. Annexe

Le Monde :

https://www.lemonde.fr/planete/article/2017/06/22/la-population-mondiale-atteindra-9-8-milliards-d-habitants-en-2050-selon-l-onu_5149088_3244.html#:~:text=La%20population%20continue%20d'augmenter,aux%207%2C55%20milliards%20actuels.

KPI :

https://fr.wikipedia.org/wiki/Indicateur_cl%C3%A9_de_performance

<https://www.baumann-avocats.com/definition/kpi.php>

Ligne de tendance :

<https://www.centralcharts.com/fr/gm/1-apprendre/7-analyse-technique/25-debutant/390-les-lignes-de-tendance#:~:text=Une%20ligne%20de%20tendance%20est,ensemble%20des%20unit%C3%A9s%20de%20temps.>