

# Projet BIUM - Analyse d'OVNI



Lenaïg Tasso  
Sarah Saleh

M1 DAC

Vendredi 5 Juin 2020





# Sommaire

1 - Introduction / Problématique

2 - Données

3 - Modélisation

4 - Analyse

4.1 Vraisemblance

4.1.1 Drogue et OVNI

4.1.2 Religion et forme d'OVNI

4.1.3 Santé mentale et OVNI

4.1.4 OVNI et satellites

2. Caractéristiques

4.2.1 OVNI et mois

4.2.2 Météo et durée de l'apparition

4.2.3 Forme de l'OVNI et heure

4.2.4 Durée de l'apparition et lieu

4.2.5 Couleurs de l'OVNI et météo

5 - Conclusion

6 - Aller plus loin

# 1 - Introduction / Problématique

Afin de permettre un essor du tourisme de l'étrange aux Etats-Unis, nous avons été engagées par une start-up qui souhaite attirer les touristes amateurs d'ufologie. Dans ce projet, nous allons donc réaliser des analyses à ce sujet. Dans un premier temps, nous allons donc voir si ces reports d'apparition d'objets volants semblent réels et sont corrélés ou non avec certains phénomènes, puis nous nous intéresserons ensuite aux spécificités de ces apparitions.

En premier lieu, nous allons donc essayer de répondre aux questions concernant la vraisemblance des reports :

- Le prix de la drogue influe-t-il sur les reports d'apparition d'OVNI et plus particulièrement sur la forme de ceux-ci ?
- L'état des croyances religieuses conditionne-t-il les formes aperçues ?
- L'état de la santé mentale possède-t-il un lien avec les OVNI ?
- Les témoignages d'apparitions d'OVNI ne coïncideraient-ils pas avec les dates de lancement de satellites dans certains états ?

Puis finalement, nous allons aux caractéristiques de ces apparitions :

- Quel est le mois où les OVNI apparaissent-ils le plus ?
- Y-a-t-il un climat plus propice aux apparitions ?
- Un lien existe-t-il entre la forme d'un OVNI et l'heure à laquelle il a été aperçu ?
- Les OVNI ont-ils une préférence géographique quant à leur durée d'apparition ou de visibilité ?
- Adaptent-ils la couleur de leur soucoupe en fonction de la météo ?

## 2 - Données

Nous avons réussi à trouver les données que nous désirions, mise à part pour les données climatiques. En effet, les seules que nous avons trouvées à ce sujet, sont celles d'anomalies climatiques. Nous les avons trouvées sur les sites suivantes :

Opendatasoft : Pour les bases militaires

Ucsusa : Pour les satellites

Kaggle : Pour les données sur la santé mentale

Dataworld : Pour les données restantes (OVNI, météo, religions, et prix de la marijuana)

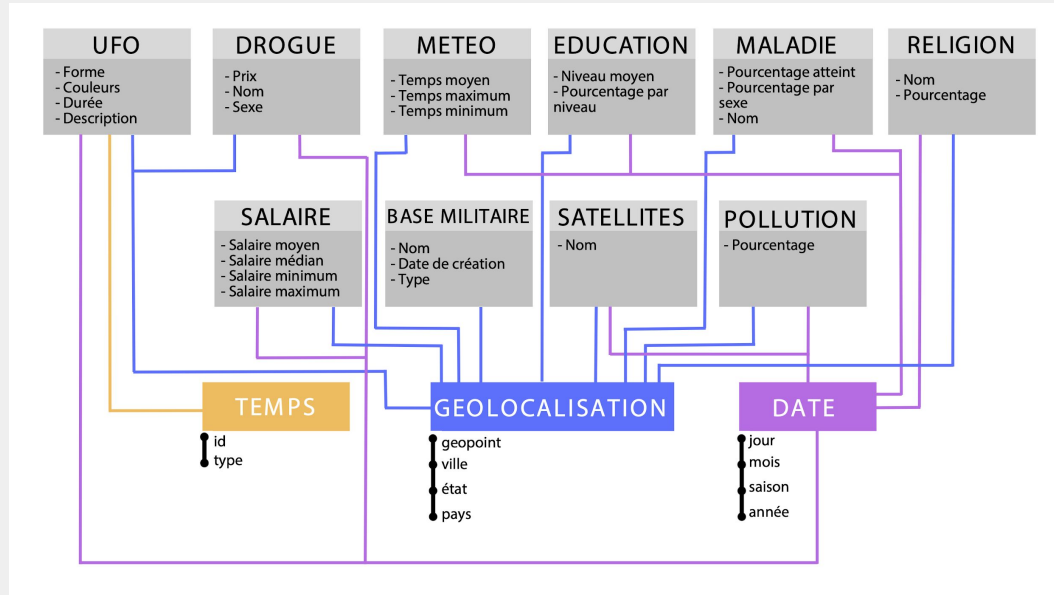
La plupart des apparitions d'OVNI se sont déroulés aux Etats-Unis, nous avons par conséquent décidé de nous intéresser seulement aux données provenant de ce pays pour nos analyses.

Cette mesure nous a semblé indispensable car dans le cas de jointures avec d'autres faits, nous aurions perdu trop d'informations. Toutefois, nous sommes conscientes qu'il aurait été plus intéressant d'avoir un maximum de données

Par ailleurs, nous avons alors remplacer le fichier fourni world.sql par un autre fichier us.cities qui possèdent à peu près la même granularité.

# 3 - Modélisation

Nous avons définis au total 6 faits et 2 dimensions.  
Voici le schéma en constellation :



# 3 - Modélisation

Le schéma logique des faits est alors :

OVNI (idOvni, idDate#, idGeo#, forme, couleurs, durée, description)

DROGUE (idDprix, idDate#, #idGeo, prixQualiSup, prixQualiMed, prixQualiInf)

METEO (idM, idDate#, idGeo#, tempMin, tempMax, ecartMoyenne, typeClimat)

SANTE\_MENTALE (idSM, )

RELIGION (idR, idDate#, chrstGen, judGen, islmGen, budGen, population, sommeReligion, chrstGenPct, judGenPct, islmGenPct, budGenPct )

SATELLITES (idS, idDate#, idGeo#, nom, propriétaire, paysOperateur, utilisateur)

Celui des dimensions est :

GEOLOCALISATION (idG, geopoint, ville, état, pays)

DATE (idDate, jour, mois, saison, année)

TEMPS (idT, type)

# 4 - Analyse

L'analyse fait partie des étapes cruciales pour parvenir à des conclusions stables. Pour chacune des questions, nous allons vous expliquer la méthodologie d'analyse que nous avons utilisé et nous allons aussi vous montrer des graphiques et des cartes que nous trouvons parlantes pour étayer nos propos.

## 4.1 Vraisemblance

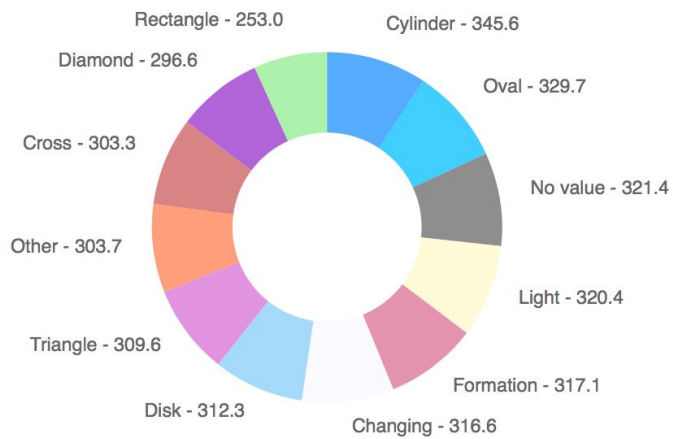
### 4.1.1 Drogue et OVNI

- Le prix de la drogue influe-il sur les reports d'apparition d'OVNI et plus particulièrement sur la forme de ceux-ci ?

Durant cette question, nous avons voulu tester la véracité d'un préjugé. En effet, nous nous sommes interrogées à propos de l'existence d'un possible lien entre le prix de la marijuana aux Etats-Unis en fonction du temps et des témoignages relatant d'apparition d'OVNI.

Nous avons dans un premier temps joint les deux faits DROGUE et OVNIS ensemble, puis réalisé un graphique simple montrant le nombre d'OVNI en fonction du prix moyen de la marijuana par mois et par forme et ce pour toutes les qualités possibles.

Avg. of HighQ\_avg by Shape

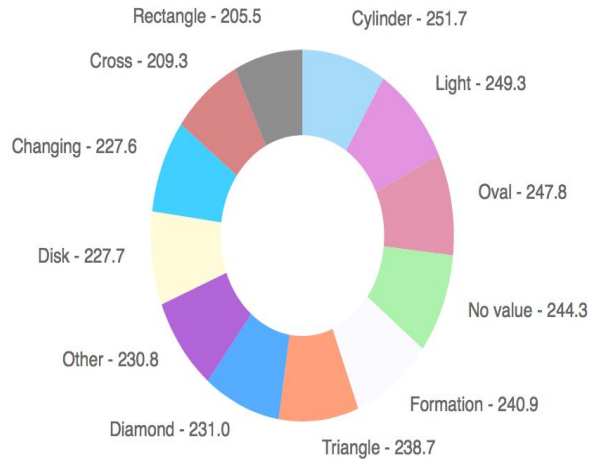


Nombre d'OVNI en fonction du prix moyen de la marijuana par mois et par forme et ce pour une qualité supérieure

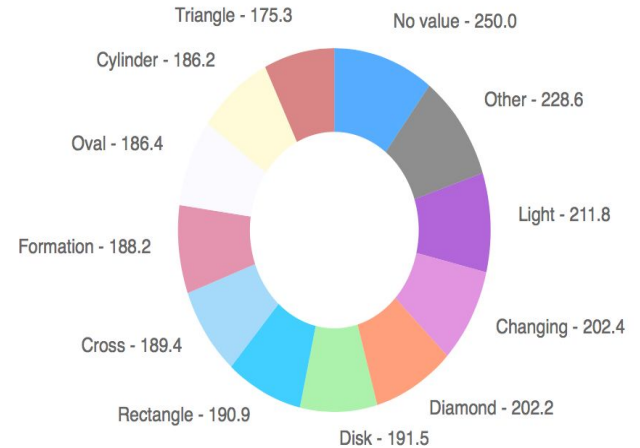


Nombre d'OVNI en fonction du prix moyen de la marijuana  
par mois et par forme et ce pour une qualité inférieure

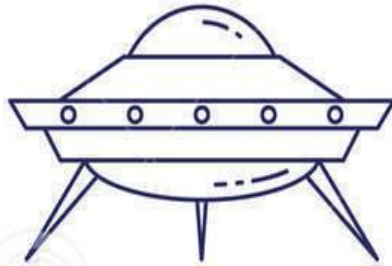
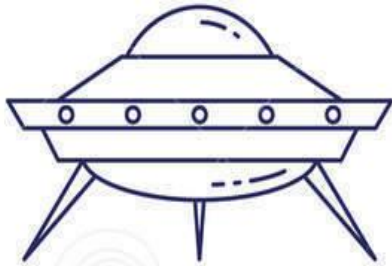
Avg. of MedQ\_avg by Shape



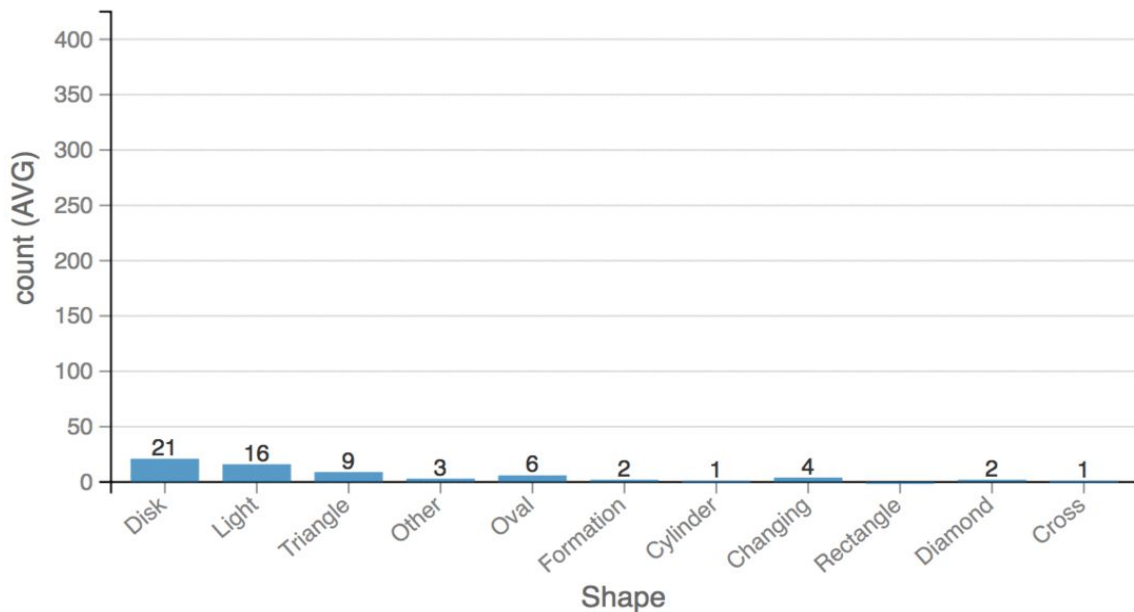
Avg. of LowQ\_avg by Shape



Nombre d'OVNI en fonction du prix moyen de la marijuana  
par mois et par forme et ce pour une qualité medium



Avg. of count by Shape

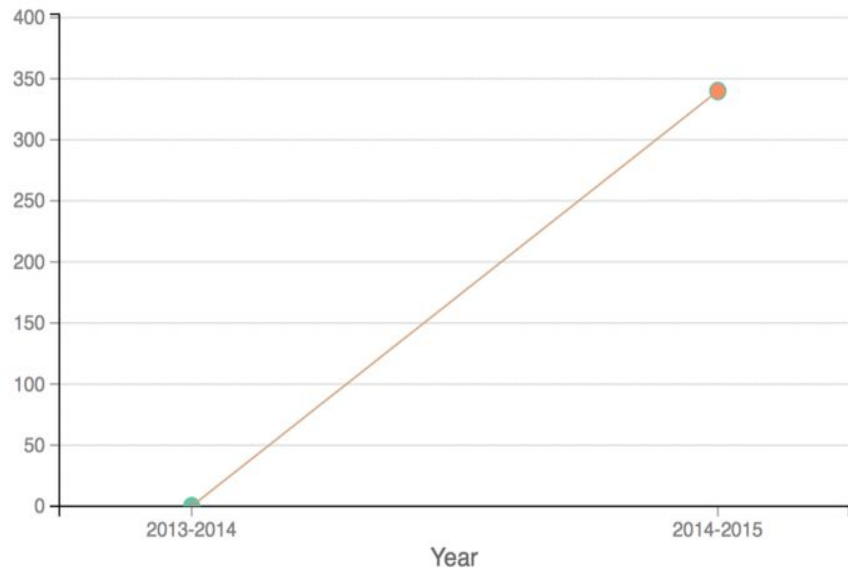


Histogramme représentant le clustering de la forme à partir du prix de la marijuana pour toutes les qualités

On remarque que bien que nous ayons effectué un clustering pour essayer de prédire la forme à partir du prix de la marijuana, les 11 clusters ne se distinguent pas réellement les uns des autres. En effet, pour la plupart la forme majoritaire est le disque. Cela est sûrement dû au fait que parmi les données que nous avons, cette forme était la forme prédominante.

On peut donc en déduire que ces deux faits ne semblent pas corrélés.

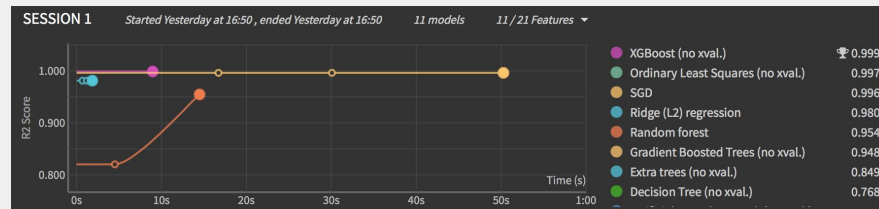
Avg. of prediction by Year Alabama



Graphique de la prédiction du prix de la marijuana de qualité supérieure (en vert) et du vrai prix de celle-ci (en orange) en fonction des différents états

Nous avons ensuite essayé de prédire le prix de la marijuana de qualité supérieure à partir des caractéristiques des témoignages d'OVNI.

Nous avons obtenu des résultats très encourageants



Cependant, cela a pu être biaisé par le fait que l'écart entre les différents prix ne sont pas vraiment significatifs.

A partir de toutes ces informations, nous pouvons conclure que ces deux faits ne semblent pas avoir de réels incidences entre eux.

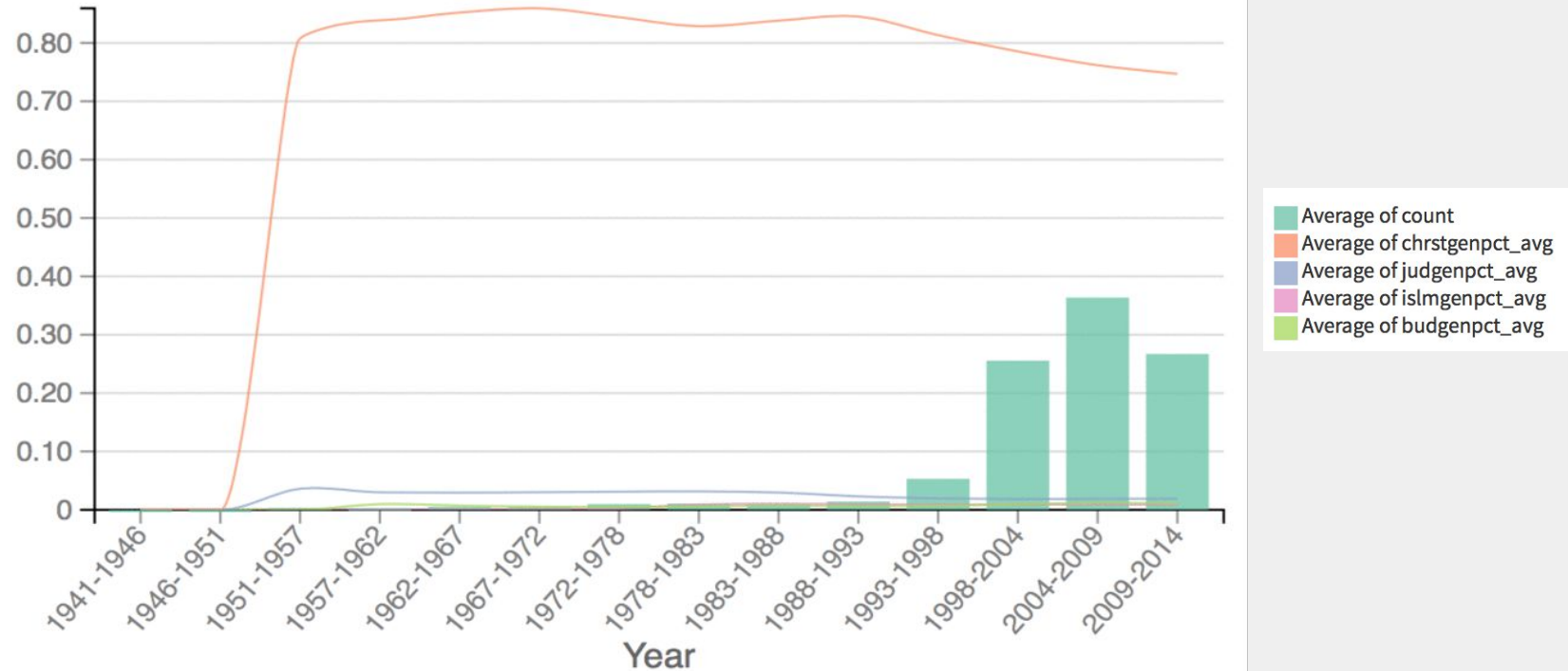
## 4.1.2 Religion et forme d'OVNI

- L'état des croyances religieuses conditionne-t-il les formes aperçues ?

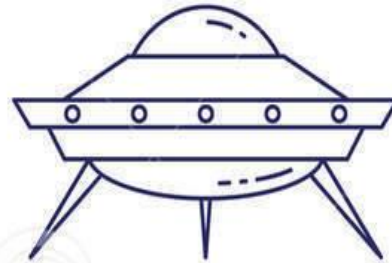
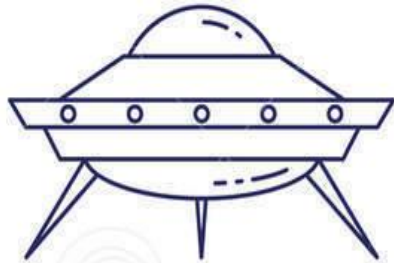
Pour cette question, nous voulions savoir si les formes des OVNIS qui ont été observés pouvait être corrélées avec l'état des différentes croyances à travers les Etats-Unis. Nous avons donc joint les deux faits RELIGION et OVNIS ensembles. Ensuite, nous les avons regroupés par année et par forme pour former ce graphique montrant le nombre d'OVNIS par forme sur plusieurs années en fonction des 4 religions les plus importantes.



Avg. of count by Year Changing

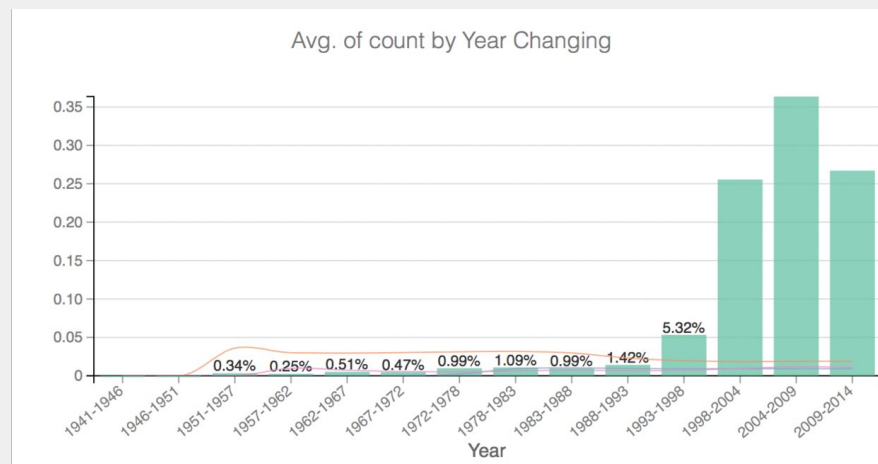


Pourcentage du nombre d'OVNIS par forme sur plusieurs années en fonction des 4 religions les plus importantes



Nous constatons dans les années 2000, une hausse de reports d'apparition d'OVNI qui se fait ressentir au même moment qu'une baisse dans le pourcentage de personnes de religion chrétienne. Nous pouvons expliquer cela avec le fait qu'avec l'avènement d'internet de plus en plus de personnes peuvent témoigner. Concernant la religion, les Etats Unis connaissent depuis toujours de fort taux d'immigration de personnes dont la foi n'est pas forcément celle qui était majoritaire auparavant, et cela est notamment vrai sur cette période.

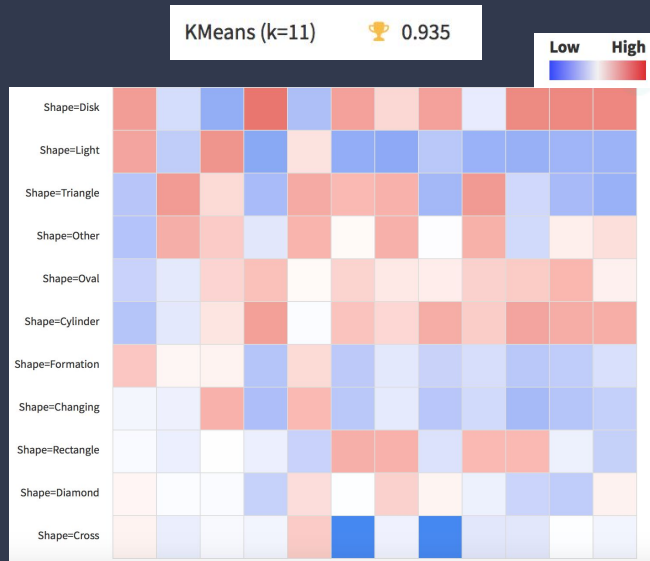
De plus, on peut dès lors remarquer que la majorité des Américains sont de confession chrétienne.



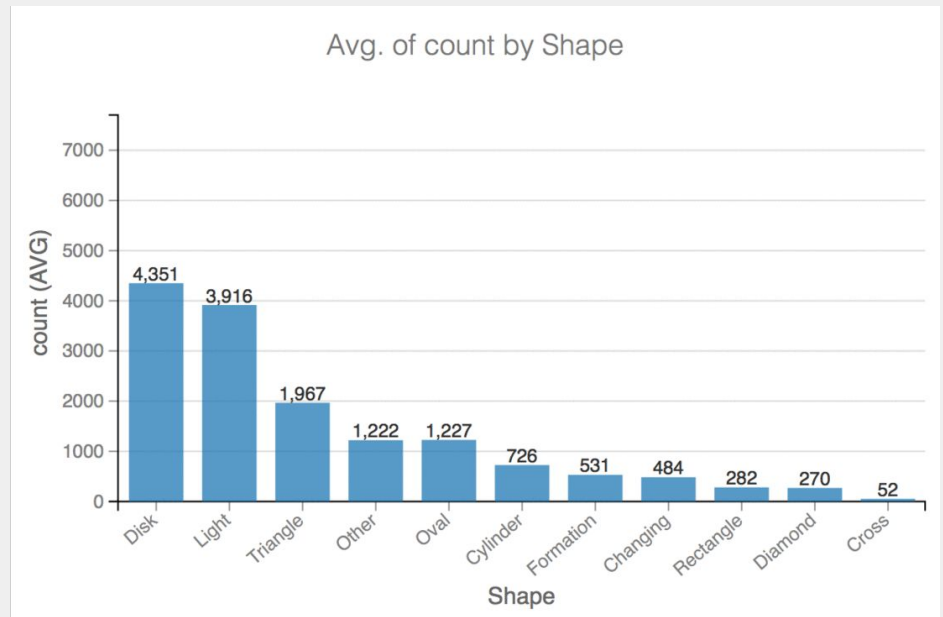
Pourcentage du nombre d'OVNIS par forme sur plusieurs années en fonction des 3 religions exceptées le christianisme les plus importantes

Nous avons réalisé ce second histogramme pour mieux apercevoir l'évolution des autres religions. Nous pouvons ainsi remarquer que la religion ne semble avoir aucune incidence sur les formes des OVNI.

Finalement, nous avons réaliser une analyse de ces deux faits. Nous avons essayé de prédire la forme d'un OVNI en fonction des croyances. Cependant bien que l'accuracy obtenue fut bonne, les résultats ne furent pas fructueux quand nous avons essayé de nous entraîner sur des données dont nous connaissions déjà la forme, comme le montrent les deux images ici présentes.



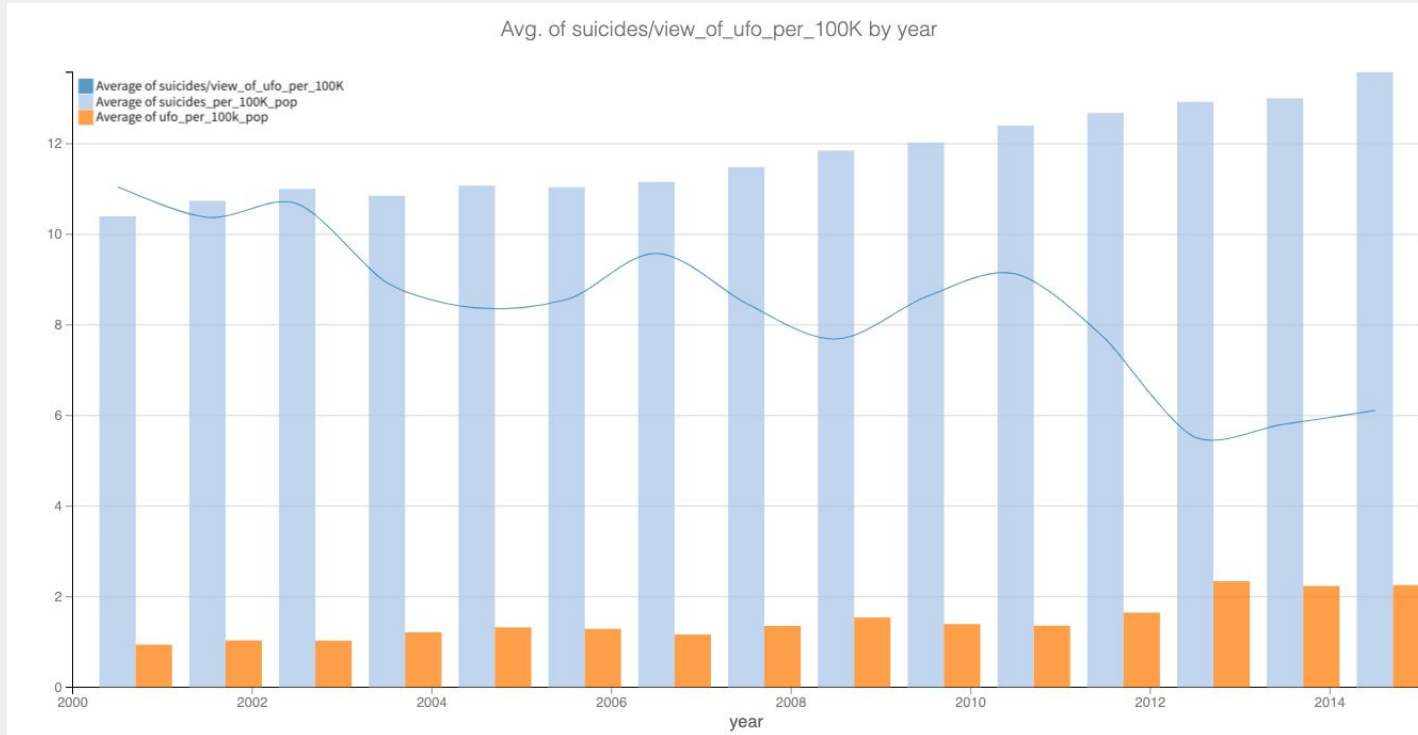
Heatmap résultant du clustering



Histogramme représentant le clustering de la forme à partir de l'état des différentes religions aux Etats-Unis en fonction de la forme

Nous pouvons donc en conclure que cette idée bien qu'intéressante en théorie n'est pas vérifiée. Cela est peut être dû, comme nous l'avons dit auparavant, au fait qu'aux Etats-Unis, l'écrasante majorité des habitants (environ 80%) possède comme religion le christianisme.

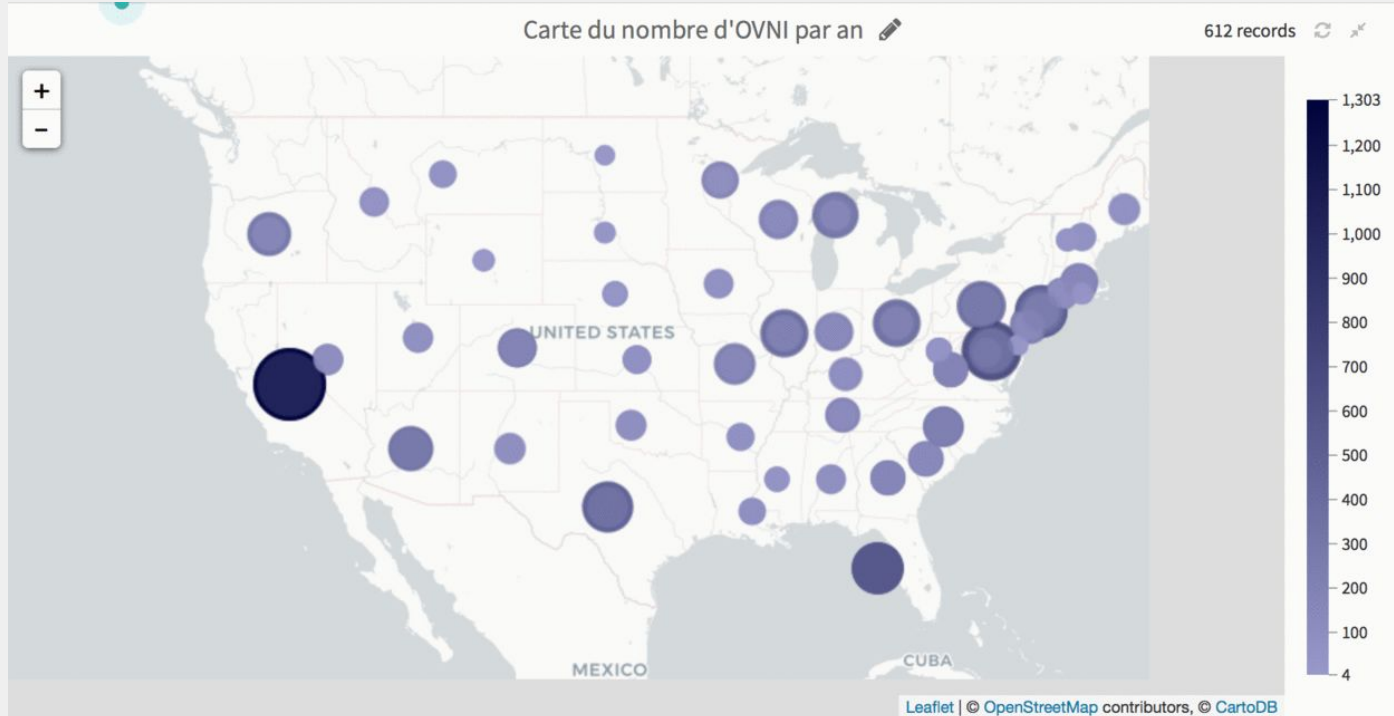
## 4.1.3 Santé mentale et OVNI





## 4.1.4 OVNI et satellites

## 4.2.1 OVNI et mois



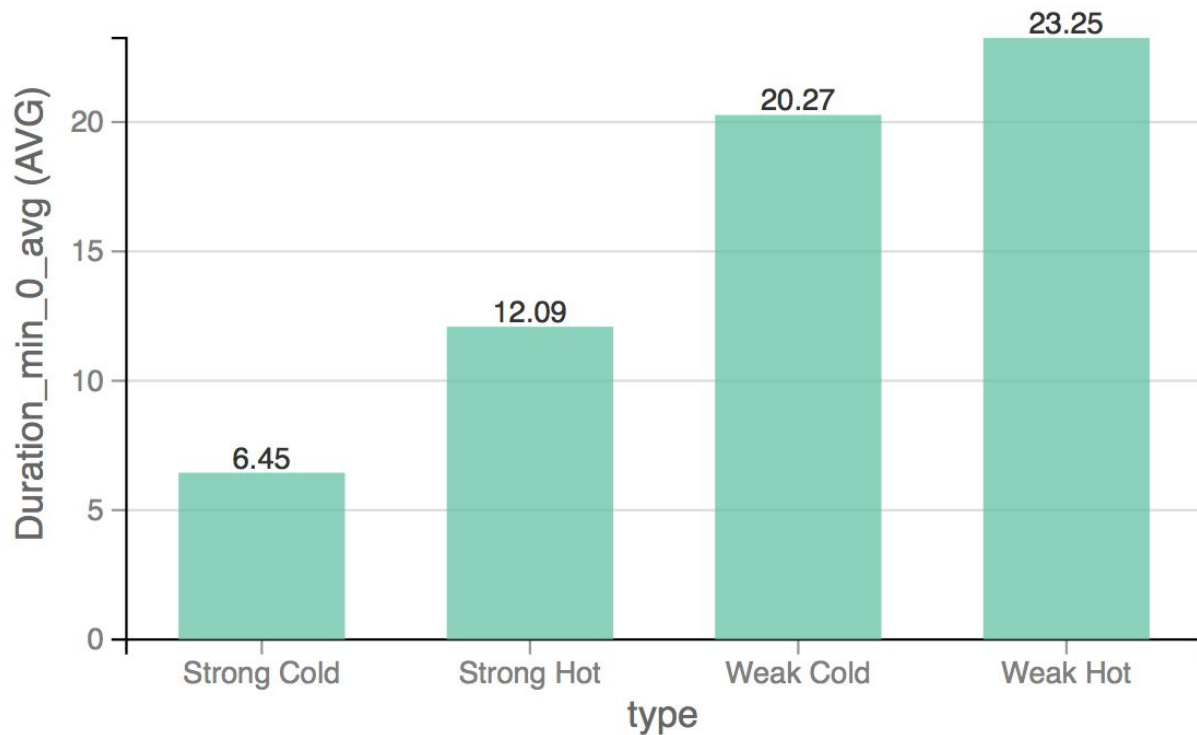
## 4.2.2 Météo et durée de l'apparition

- Y-a-t-il un climat plus propice aux apparitions ?

Au cours de cette question, nous cherchions à connaître un possible lien entre le type de climat lorsque l'OVNI a été aperçu et la durée d'apparition de celui-ci. Nous avons donc dans un premier temps converti toutes les durées en minutes pour mieux pouvoir les comparer, puis joint les deux faits METEO et OVNI. Voici un histogramme montrant la durée d'apparition en fonction du type du climat.

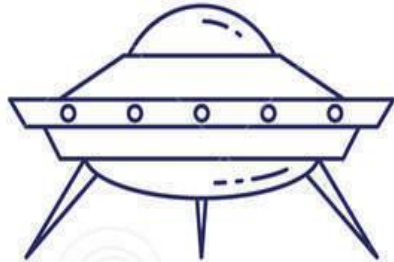


Avg. of Duration by type



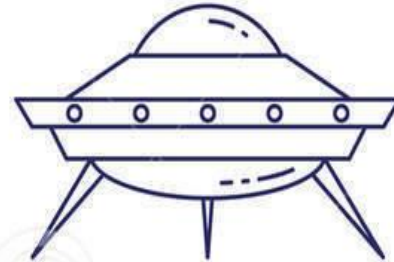
On peut remarquer que lorsque le climat est très prononcé par un extrême, soit par la chaleur ou soit par le froid, la durée de l'apparition de l'OVNI dure moins longtemps en moyenne que pour un climat moins intense.

Histogramme montrant la durée d'apparition en fonction du type du climat

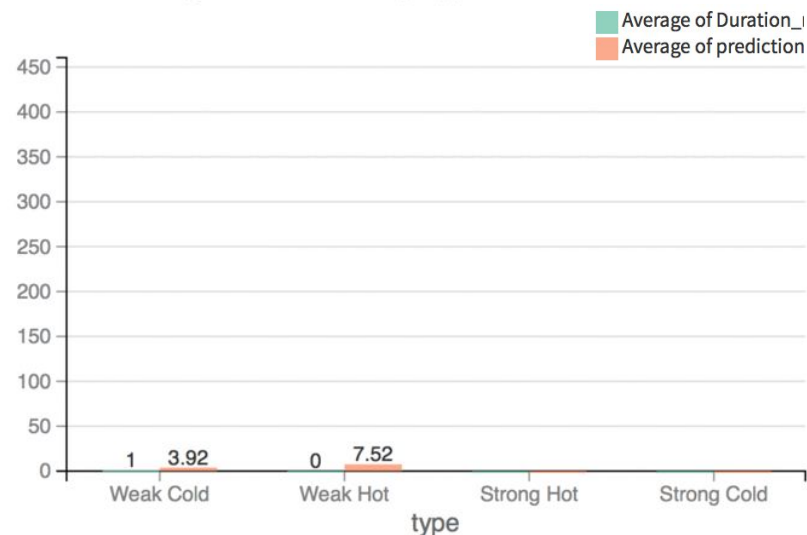


Et ensuite, nous avons réalisé une analyse pour voir si ces deux variables étaient corrélées ou non. Nous avons essayé de prédire la durée de l'apparition en fonction des différents attributs du fait METEO, mais cela n'a pas été payant car lorsque nous avons vérifié nos résultats, cela n'était pas cohérent avec les durées que nous aurions dû trouver.

En conclusion, cela prouve qu'il est difficile de prédire la durée d'apparition d'un OVNI car cela est très subjectif. En effet, peut être la durée d'apparition n'est pas semblable en fonction de qui l'observe. Cependant, on peut en déduire que si l'on veut voir un OVNI qui dure plus longtemps, il est préférable d'avoir un climat faiblement chaud tandis que si l'on désire qu'il dure moins longtemps il est préférable d'avoir un climat considérablement froid.

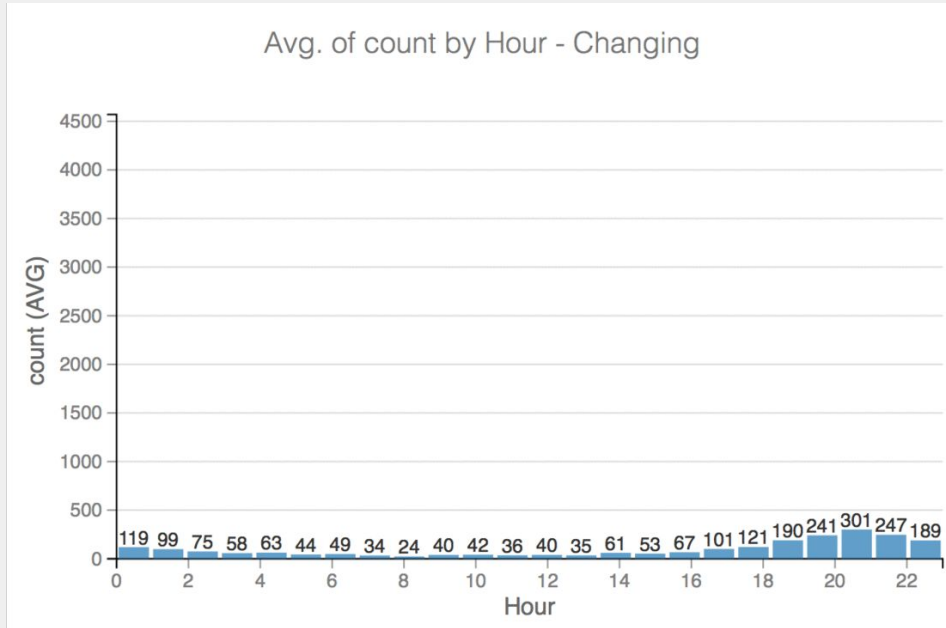


Avg. of Duration by Type Alabama

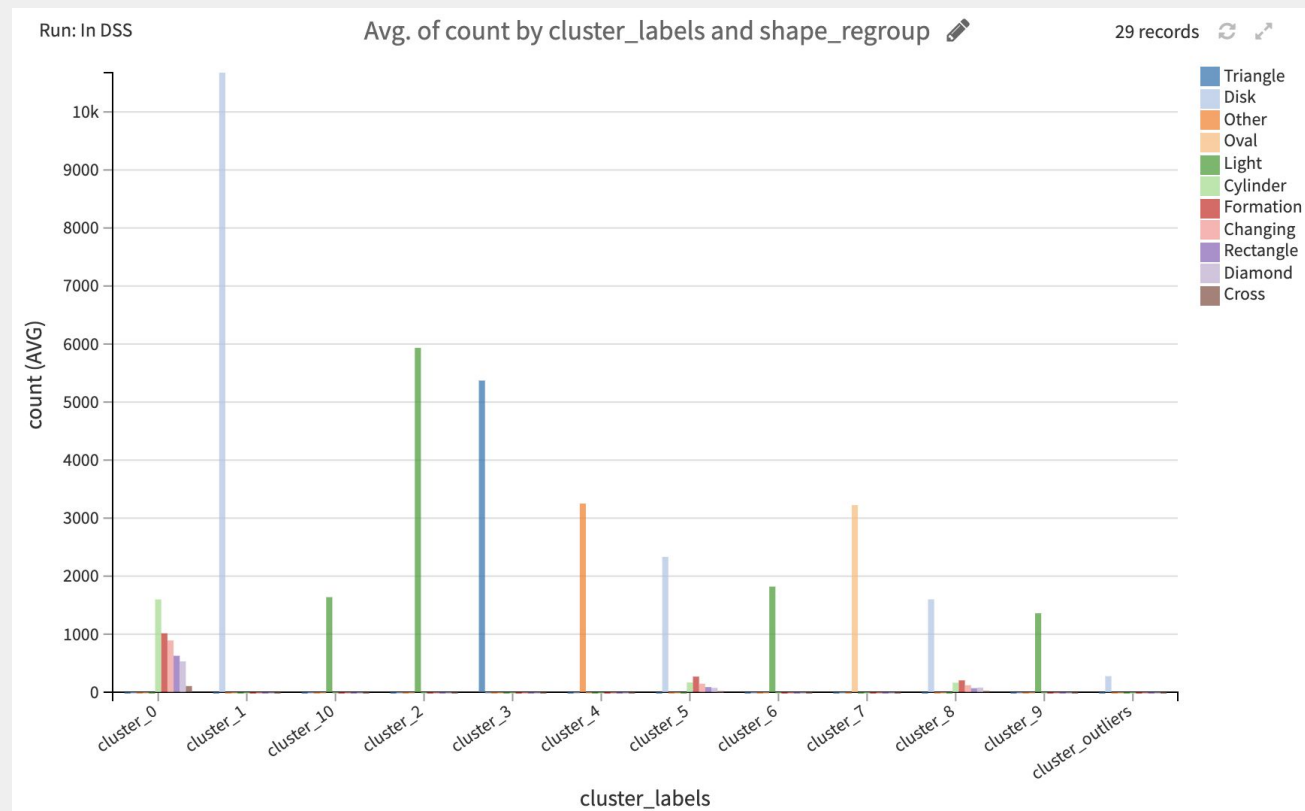
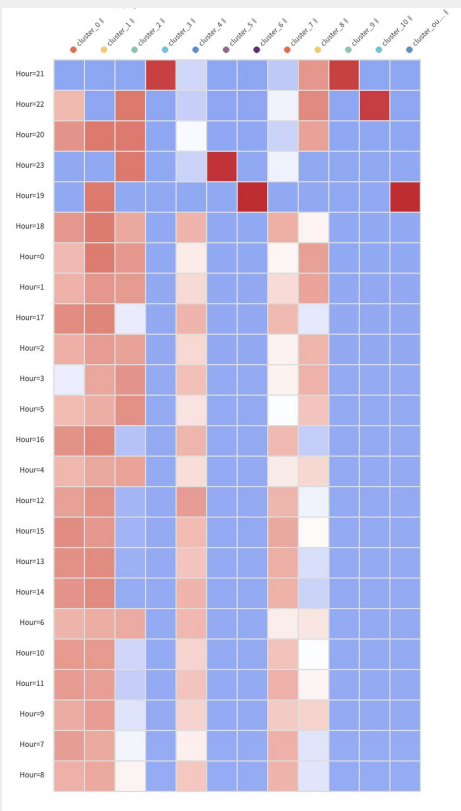


Histogramme représentant la durée d'apparition moyenne d'un OVNI en comparaison avec la prédiction obtenue en fonction du type de climat

## 4.2.3 Forme de l'OVNI et heure



Histogramme représentant le nombre d'OVNI par heure en fonction de la forme observée

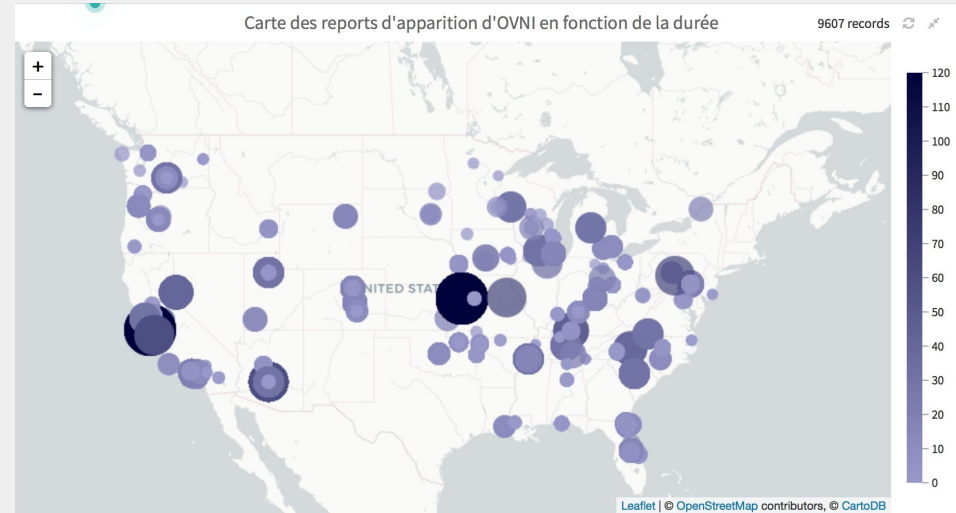


## 4.2.4 Durée de l'apparition et lieu

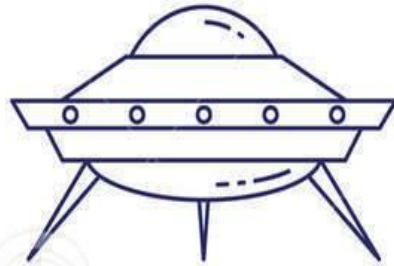
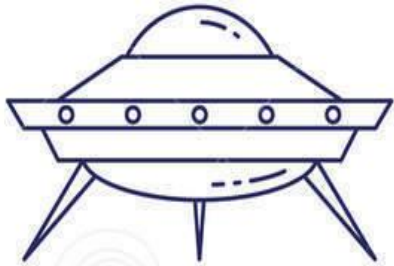
- Les OVNI ont-ils une préférence géographique quant à leur durée d'apparition ou de visibilité ?

Pour cette question, nous voulions proposer à la start-up qui nous a engagé des lieux en fonction de la durée de visibilité des OVNI. En effet, certains clients peuvent souhaiter observer des apparitions plus ou moins longues selon leurs envies.

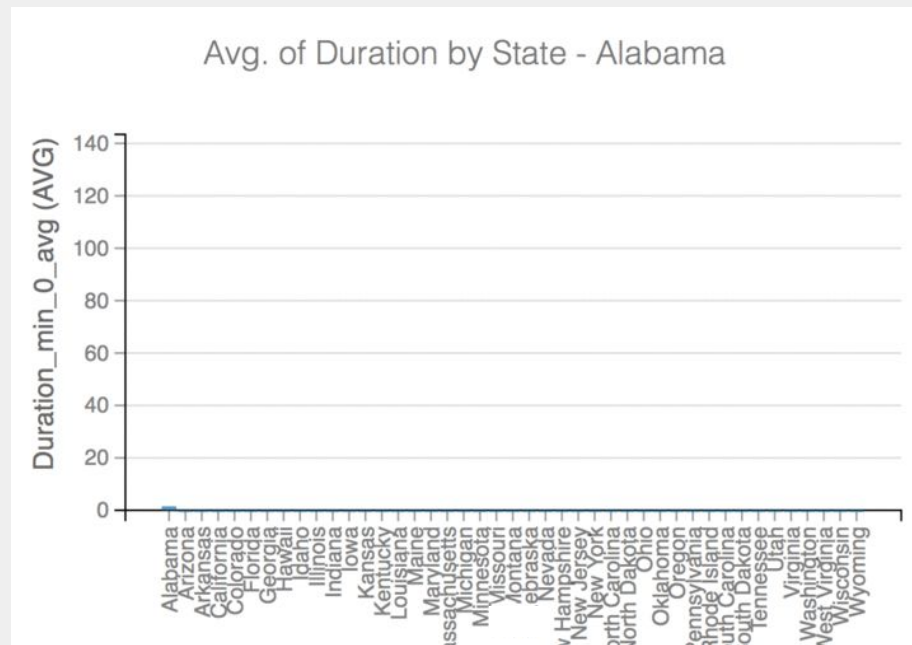
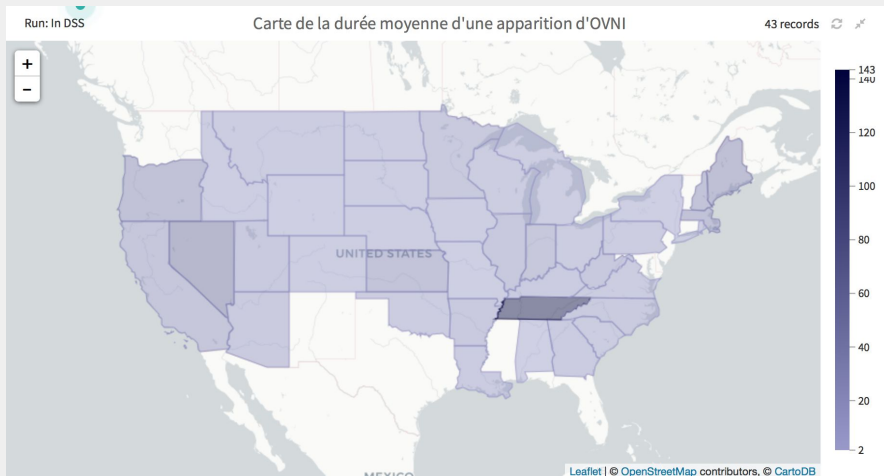
Sur la carte ci-présente, on peut apercevoir que les manifestations qui durent le plus longtemps sont celles qui se trouvent sur les côtes des Etats-Unis exceptées pour un état au centre, et donc celles qui ont une durée moins considérables sont celles à l'intérieur du pays.





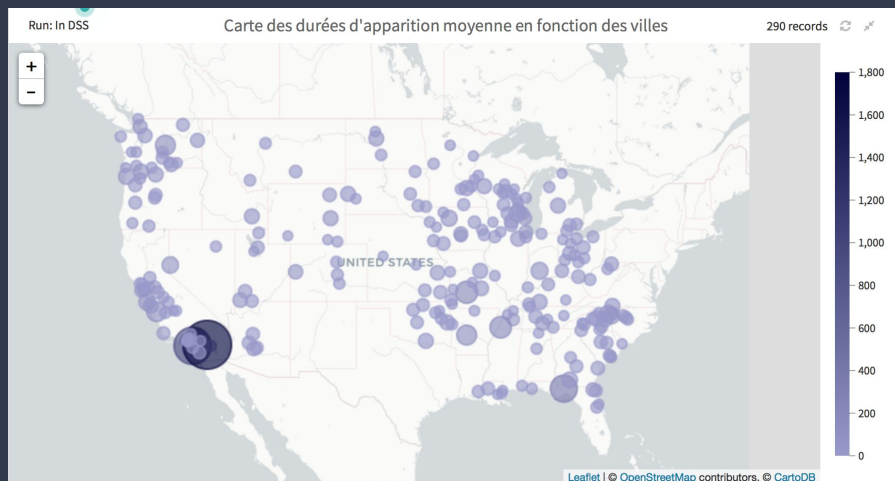


Nous avons ensuite regroupé par états les différentes déclarations et ainsi obtenues une durée moyenne.



Histogramme de la durée moyenne d'apparition d'un OVNI en fonction des états

Comme nous pouvons le relever sur les deux images précédentes, les états où les manifestations ont une durée plus importantes sont le Tennessee, le Nevada, le Maine ou encore l'Oregon avec une durée supérieure à 30 minutes. Au contraire, ceux qui ont une durée minimale sont les états de New York, le Missouri, Rhode Island et l'Alabama.



Puis nous avons réalisé la même opération, mais cette fois-ci en fonction des villes. On observe que 2 villes de Californie ont des durées très importantes : Santa Ana et Palm Springs. De plus, c'est également le cas pour Memphis dans le Tennessee, Tallahassee en Floride et Buffalo dans l'état de New York.

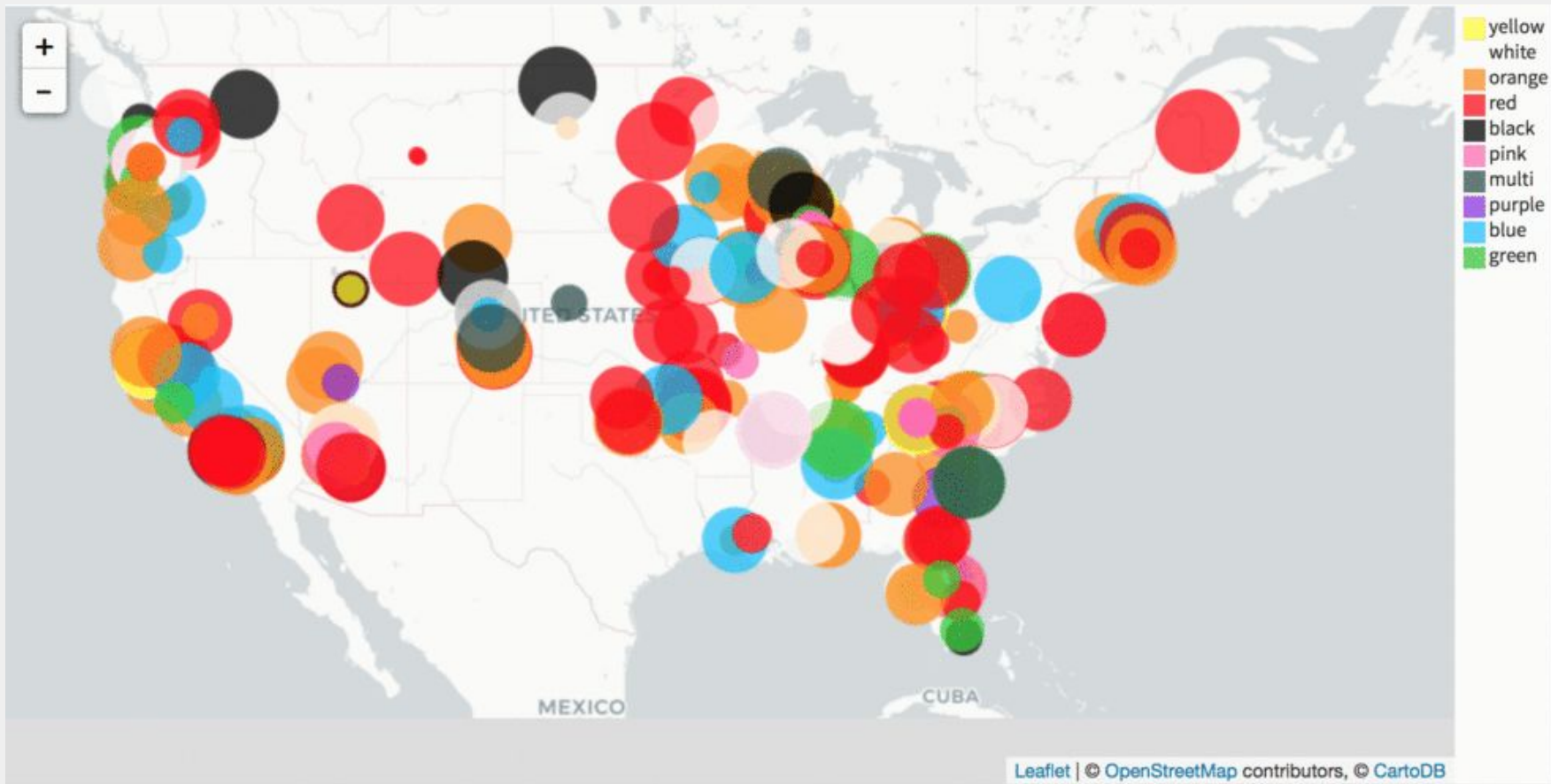
On peut donc en conclure que dans le Tennessee, les durées sont vraiment élevées et en particulier à Memphis. Par ailleurs, bien que la Californie ne fasse partie que du top 10 des états avec les durées les plus hautes, elle possède 2 villes qui ont des durées supérieures à une journée. Finalement, c'est également le cas de Floride qui a une durée moyenne d'apparition de 19 minutes, mais qui pour Tallahassee possède une durée d'environ 6 heures. La pollution, lumineuse et de particules, a cependant peut-être son rôle à jouer dans ces données car la Californie ou encore l'état de New York font partie des états les plus pollués.

## 4.2.5 Couleurs de l'OVNI et météo

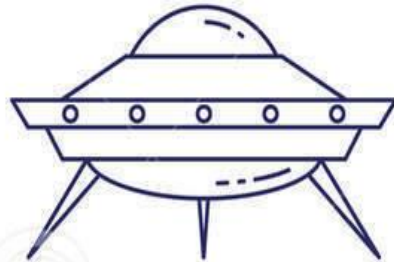
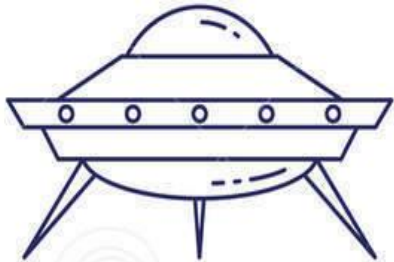
- Adaptent-ils la couleur de leur soucoupe en fonction de la météo ?

Pour cette question, le but était cette fois-ci de savoir si la couleur des OVNIS était liée ou non au climat. Nous avons donc extrait la couleur des résumés de report d'apparition des ovnis, puis joint les deux faits METEO et OVNIS. Nous avons ensuite créé une carte montrant la couleur des OVNI en fonction du temps.





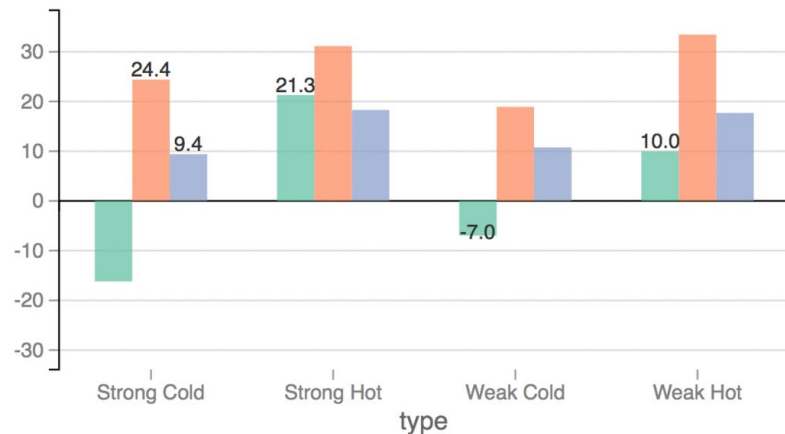
Carte montrant la couleur des OVNI en fonction des degrés d'écart à la moyenne saisonnière



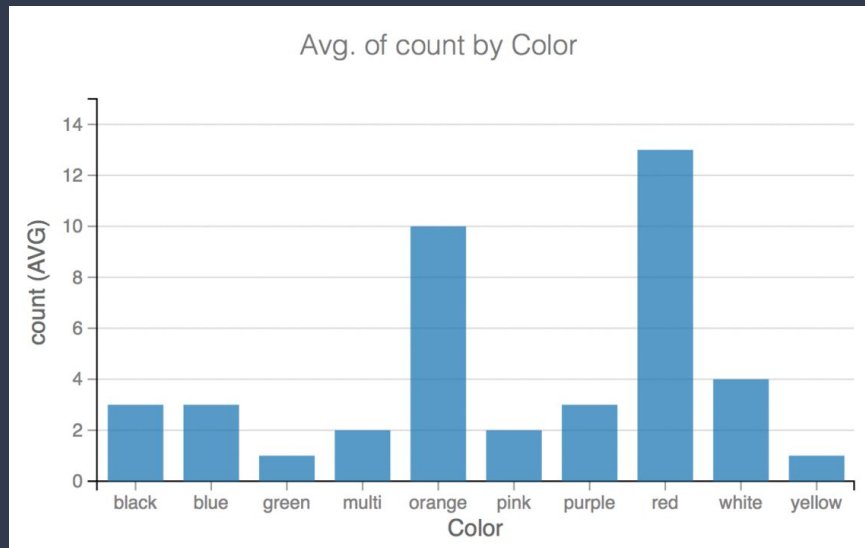
On peut dès à présent remarquer que les couleurs prédominantes observées sont des couleurs chaudes comme le rouge et l'orange. Toutefois, le jaune, couleur que nous pensions prédominante, n'est finalement pas si prépondérante.

Dans cet histogramme, on peut relever que certaines couleurs d'OVNI ne sont pas visibles dans certains climat, comme par exemple le bleu et le vert pour un climat considérablement froid, ou le violet quand le climat est extrême. On peut donc en déduire qu'il n'est pas facile d'observer un OVNI de couleur froide, il est effectivement plus aisé de voir un OVNI de couleur chaude comme nous avons pu le voir sur la carte précédente.

Avg. of degrees from mean by type - Black



Histogramme représentant le clustering de la couleur d'un OVNI à partir du type de climat et de l'écart en degrés à la moyenne saisonnière



Histogramme représentant le clustering de la couleur d'un OVNI à partir du type de climat et de l'écart en degrés à la moyenne saisonnière

Nous avons ensuite essayer de prédire la couleur d'un OVNI en fonction des caractéristiques météorologiques telles que le type de climat et l'écart de degrés à la moyenne saisonnière.

On peut constater que nous arrivons à isoler certaines couleurs comme le vert, le blanc ou le rouge, mais pas toutes malheureusement.

# 5 - Conclusion

## Faire conclusion

Avant pas beaucoup de témoignages, mais avec développement d'internet beaucoup +  
Faire des petites conclusions sur où aller pour voir quelles couleurs, quelles formes, durée plus ou moins longue, etc

Témoignages semblent corrélés avec aucun autres faits mais restent tout de même assez difficile à prédire  
comme le démontrent notre difficulté à prédire certaines caractéristiques

## 6 - Aller plus loin