



15/12/2020

CHEVALLIER Matteo, BOISSEAU Aurèle,
DEWAVRIN Thibault, FREMAUX Mathieu.

Rapport Technique Groupe 12



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

I - Le problème des rats

I - a) Description

La régulation des populations de rats est un problème connu depuis longtemps qui touche toutes les grandes villes. Jadis leur présence était synonyme de mort, puisqu'ils étaient porteurs du bacille de la peste, qui a tué des millions de personnes. Aujourd'hui, même si la peste a quasiment disparu de nos contrées, leur présence cause toujours des problèmes : ils sont toujours vecteurs de zoonoses comme la leptospirose, la rage, et beaucoup d'autres. En outre, s'ajoutent à ce problème les dégâts matériels qu'ils causent. En effet, ils ont besoin de ronger des choses pour utiliser leurs dents, et dans un environnement urbain, ils sont capables de tout ronger même le béton. Ils rongent les câbles électriques, ce qui peut causer des incendies, ou courts-circuits, les plafonds, sols, murs, et même les fondations d'immeubles, ce qui peut causer leurs affaissements. La présence de rats est également associée à l'insalubrité.

Pour ces raisons, le maire de New York a annoncé en 2017 un plan de 32 millions de dollars pour dératiser la ville. Cependant, réaliser une dératisation a un coût, et lorsqu'il s'agit de dératiser une ville, il faut des moyens faramineux. Il paraît alors normal de chercher à dératiser plus intelligemment afin de réduire les coûts et c'est dans cette optique que la mairie de New York nous a contacté en tant que Data Scientists.

II - b) Résultats attendus

La mairie de New York nous a contacté pour que, à partir de l'analyse de données, elle puisse mieux comprendre les raisons de présence des rats. Grâce à une approche statistique, nous pouvons trouver différents facteurs à l'origine de la présence de rats, qui n'auraient pas forcément été mis en lumière par des approches traditionnelles. Dans ce cadre là nous avons pu déterminer ci-dessous l'objectif de la mission ainsi que la problématique à laquelle il faudra répondre :

Objectif : Effectuer des recommandations à la mairie de New York sur la politique à mener vis-à-vis des restaurants, des lieux de dépôt de compost, et du ramassage des déchets.

Problématique : Peut on expliquer la présence de rats dans la ville de New York par la présence de facteurs adjutants comme les restaurants, les lieux de compost, ou encore la masse de déchets ramassés ?

Nous allons donc chercher à établir ou fabriquer des variables qui peuvent expliquer la présence de rats et les optimiser pour réduire leur population .

II - Présentation et visualisation des données

II - a) Présentation des datasets

La présence de rats est liée à de nombreux critères et nous avons eu besoin d'un nombre important de dataset (6) pour réunir les informations nécessaires. Plus exactement, nous avons utilisé :

- un dataset sur les rats: dans ce dataset était compris toutes les détections de rats dans le centre de New York depuis 10 ans, avec leur localisation (coordonnées géographiques, numéro de zipcode, quartier) et la date de la détection.
- 2 datasets sur les déchets: on y trouve la position des composteurs à ordures (quartier), la date et la quantité de déchets ramassée mensuellement en fonction du type de déchets (ménager, papier, électronique).
- un dataset sur les restaurants: on y trouve toutes les inspections sanitaires des restaurants au cours de ses 10 dernières années, avec un score de salubrité, la cause de l'inspection, la date, le lieu (zipcode, quartier et coordonnées géographiques, type de restaurants)
- un sur la population par zipcode en 2020
- un sur les espaces verts: localisation (zipcode), superficie

L'enjeu a donc été de mettre en relation et lier ces différents datasets. Il a fallu pour cela trouver une variable commune à ces datasets et nous permettant de faire le lien entre eux. Nous avons choisi une variable spatiale: le zipcode qui est l'équivalent américain du numéro d'arrondissement. Cette décision nous a confronté à un problème: un dataset sur les déchets ne localisait pas ces données en fonction du zipcode mais en fonction d'un autre indicateur spatial utilisé à New York : les districts communautaires. Le problème étant que ces districts et les zipcodes ne correspondent pas très bien, il a donc fallu faire des approximations pour faire correspondre ces districts aux zipcodes. Il nous a aussi fallu convertir la localisation en zipcode en utilisant des fichiers GeoJson sur un autre dataset. La deuxième difficulté a été que pour chaque variable explicative, il a fallu trouver un dataset pertinent et qui puisse se joindre aux autres, ce qui rajoute un travail de recherche d'informations et de datasets exploitables à l'analyse et au traitement des données.

Au niveau de la provenance de ces datasets, ils proviennent tous du portail open data officiel de la ville de New York (<https://opendata.cityofnewyork.us/>). Nous allons d'ailleurs préciser dans la partie suivante dans quel cadre juridique nous avons réalisé cette étude.

II - b) Cadre juridique

Nos datasets proviennent du portail officiel d'open data de New York, les données appartiennent donc à la ville de New York. Cependant nous sommes une équipe mandatée par la ville de New York pour réaliser cette étude. Nous avons donc un contrat qui nous spécifie une autorisation pour utiliser ces données.

En outre, les données open data de la ville de New York sont réglementées par l'open data law, une loi qui oblige les institutions publiques de la ville de New York à publier leurs données sur un portail commun, sans licence, et sans restriction quant à l'utilisation des ces données.

On peut donc librement utiliser ces données pour notre étude.

II - c) Analyse descriptive des données

Nous avons dans un premier temps décidé de représenter les données de notre variable explicative, à savoir les lieux où des rats avaient été observés. Comme ces données sont réparties dans le temps, il est utile de regarder d'effectuer une représentation évolutive en fonction du temps. (lien vers la vidéo qui montre cette évolution [ici](#))

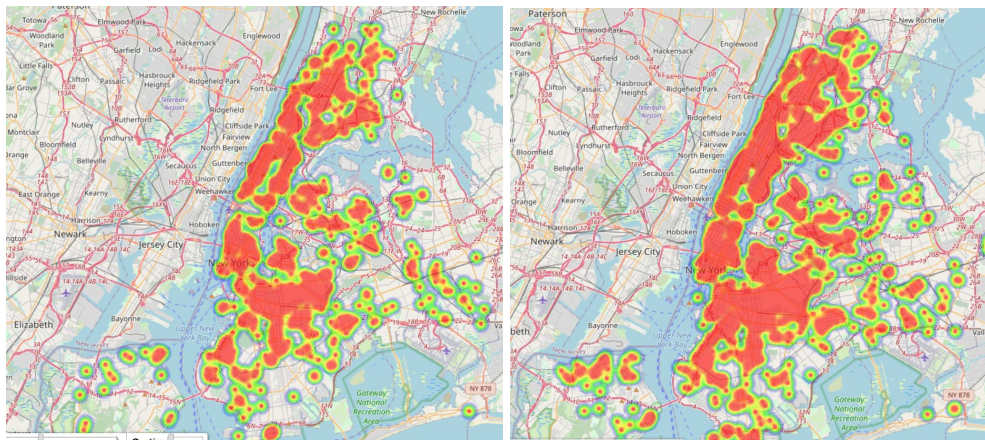


figure 1 : évolution dans le temps du nombre de rat

La visualisation sur plusieurs mois nous a permis de comprendre que sur certaines zones, la présence de rats était un problème récurrent, tandis que sur d'autres, elle est plus ponctuelle.

Nous avons donc décidé de découper la carte de la ville de New York en différentes zones. Il nous fallait une zone ni trop petite, ni trop grande, et sur laquelle nous avions des données pour suffisamment de variables explicatives. Nous avons choisi l'échelle du 'zipcode' (terminologie employée tout au long du rapport) qui correspond au code postal, au risque de devoir faire des conversions sur certains jeux de données. Par ailleurs, nous avons trouvé intéressant de travailler en utilisant la variable rats / humains plutôt que simplement les rats, en observant qu'il y avait une forte corrélation entre les deux variables. En voici donc la représentation.

On peut facilement voir qu'il y a de fortes disparités de densité de rats par zipcode. On va donc pouvoir effectuer un algorithme de partitionnement et affecter chaque zipcode à une classe de type de rats. Par ailleurs, cela nous permet d'avoir un premier outil à fournir à la ville de New York : si un zipcode passe d'une classe à la classe supérieure ou dépasse un seuil de vigilance, il faudra être particulièrement vigilant sur ce zipcode et peut-être mener une action (par exemple de dératisation) dessus.

(en noir les zones sans données)

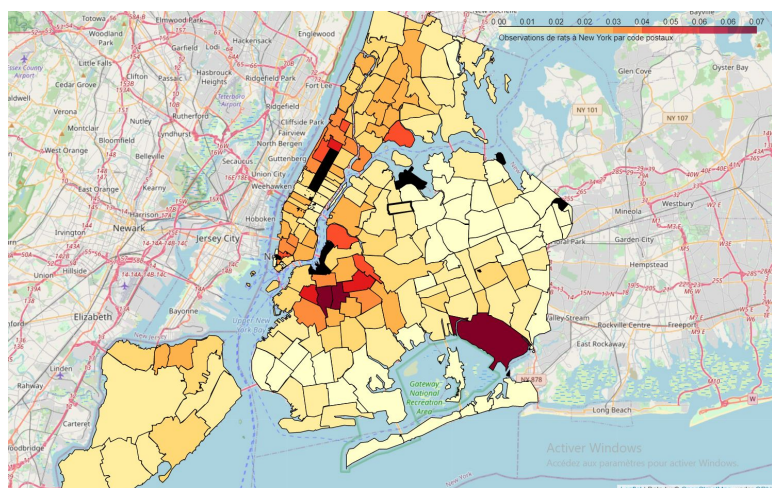


figure 2 : densité de rat par zipcode

On a aussi établi un tableau de corrélation entre les principales variables (voir ci-dessous). On remarque que le nombre d'arbres n'est pas corrélé avec les rats par habitant (corrélation négative) mais le nombre de composteurs, de parcs et de restaurants semble un peu corrélé (0.41, 0.29, 0.22).

	zipcode	count_rats	population	ratsperH	count_restaurants	Number_tree	count_bin2	count_park	groupe_rat
zipcode	1.000000	-0.141731	0.111856	-0.308848	-0.275592	0.251570	-0.391975	-0.010109	-0.233083
count_rats	-0.141731	1.000000	0.614963	0.713815	0.257488	0.119209	0.377079	0.548395	0.639373
population	0.111856	0.614963	1.000000	0.085186	0.268146	0.466058	0.160322	0.523396	0.635244
ratsperH	-0.308848	0.713815	0.085186	1.000000	0.221197	-0.172724	0.411289	0.285198	0.477812
count_restaurants	-0.275592	0.257488	0.268146	0.221197	1.000000	-0.062104	0.367809	0.199552	0.402257
Number_tree	0.251570	0.119209	0.466058	-0.172724	-0.062104	1.000000	-0.185093	0.279315	0.264897
count_bin2	-0.391975	0.377079	0.160322	0.411289	0.367809	-0.185093	1.000000	0.362374	0.348814
count_park	-0.010109	0.548395	0.523396	0.285198	0.199552	0.279315	0.362374	1.000000	0.490309
groupe_rat	-0.233083	0.639373	0.635244	0.477812	0.402257	0.264897	0.348814	0.490309	1.000000

figure 3 : corrélation entre les différentes variables

III - Utilisation des données

III - a) Analyse temporelle des données

Nous nous sommes ensuite intéressés à la temporalité et à l'évolution de la population de rats en fonction du temps, et à des corrélations éventuelles qu'il pourrait y avoir entre la présence de rats et celle de restaurants.

Nous avons commencé par représenter simplement les données :

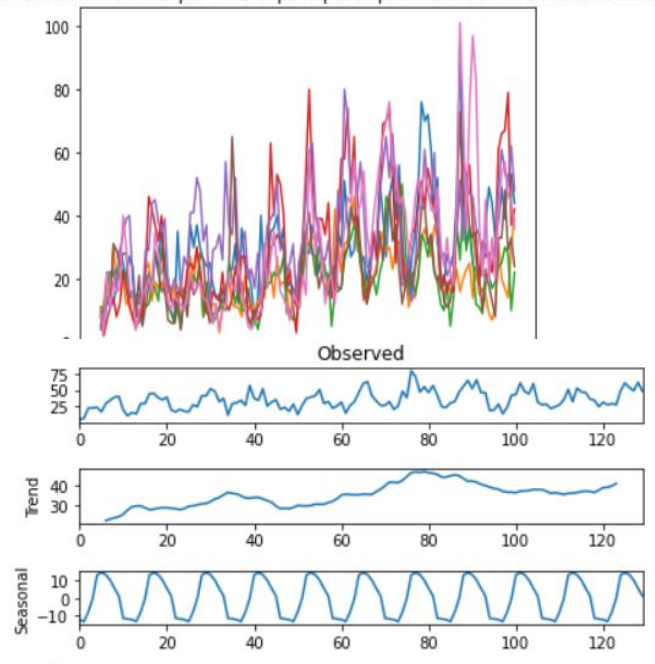
Le résultat étant illisible, il convient de visualiser le résultat séparément pour chaque zipcode.

Par ailleurs on remarque qu'il y a une potentielle saisonnalité, il peut donc être pertinent d'effectuer une décomposition saisonnière.

On peut observer sur la figure ci-jointe que la forte saisonnalité des données est confirmée. Cela nous donne une information importante dans une logique de dératisation : il peut être intéressant de choisir le bon moment pour dératiser : si c'est pour une dératisation qui s'effectue par pièges de carbone posés dans les terriers des rats pour les tuer pendant qu'ils y sont, il vaut mieux le faire pendant l'hiver où ils restent plus dans leurs terriers, et pour dératiser des lieux extérieurs, autant le faire lorsqu'il fait plus chaud et que les rats ont plus tendance à sortir. On notera que les rats peuvent se reproduire toute l'année, donc qu'il n'y a pas de saison à privilégier pour réduire le nombre de nouveau-nés.

On nuancera toutefois ces résultats en considérant la manière dont les données ont été enregistrées : à une donnée correspond une observation en extérieur. Les gens sortent aussi moins en hiver, ils ont donc moins de chance de voir des rats, ce qui renforce la saisonnalité observée dans les données.

Evolution du nombre de rats par mois et par zipcode pour les zones à fort volume de rats



On peut aussi voir que les tendances ont des creux. Cela peut s'expliquer par des dératisations effectuées à ce moment-là. Il faudrait donc corriger les tendances si l'on cherche des corrélations entre la présence de rats et un autre facteur, car cela peut fausser la corrélation.

Nous nous sommes ensuite concentrés sur la recherche de corrélation temporelle entre rats et restaurants qui ont une note $\leq C$ (C étant la pire note avec P pour suspension pour cause d'insalubrité). En effet, chercher à corrélérer les rats avec des inspections sanitaires s'avèrerait probablement inutile si l'on n'exploite pas le résultat de l'inspection !

La décomposition saisonnière n'est pas une bonne décomposition pour les données des restaurants. La tendance doit être mesurée autrement, c'est pourquoi nous avons utilisé le "moving average" (rolling) pour calculer cette tendance.

Les corrélations de tendances que nous obtenons sont faussées par les creux dans les tendances d'observations de rats, et par le caractère aléatoire des inspections sanitaires (pour savoir qu'un restaurant a une mauvaise note, il faut qu'il ait été inspecté !) et ne sont pas statistiquement significatives, (p-value trop élevée). Par ailleurs, la population de rats peut effectivement diminuer dans un endroit à cause d'un élément extérieur. Rappelons que l'analyse est faite uniquement pour les zipcodes avec un nombre de rats observés sur tout l'étalage temporel moyen à élevé.

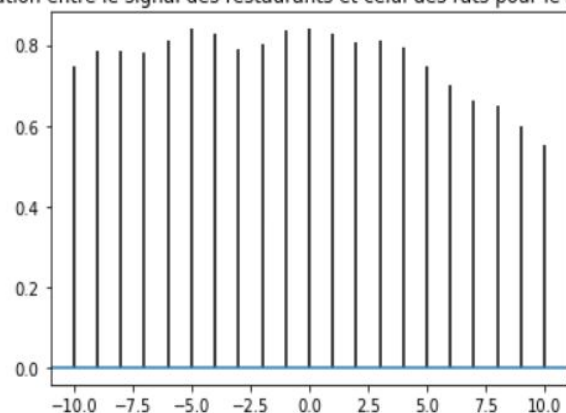
Il peut alors être intéressant de faire trois choses :

- Etudier l'intercorrélation qui permettrait de trouver une corrélation avec un retard
- définir d'autres "clusters" : population de rats globalement croissante, décroissante et neutre.
- corriger dans les courbes de tendances les creux créés par les dératisations (cela sort du cadre de cette étude).

Nous avons donc étudié l'intercorrélation, et cherché son maximum.

On peut facilement voir que le maximum de l'intercorrélation n'est jamais très éloigné de la valeur obtenue en zéro, on ne peut pas noter un réel maximum obtenu par une valeur du retard x qui signifierait que si l'on a tant de restaurants avec une mauvaise note tel mois, il est plus probable qu'on aie un tel nombre de rats obtenu dans le même zipcode x mois plus tard. L'analyse de corrélation avec un décalage ne s'avère donc pas pertinente, du moins tant que l'on a pas corrigé la tendance, en modifiant les creux créés par les dératisations.

Intercorrélation entre le signal des restaurants et celui des rats pour le zipcode 11238.0



Nous avons cependant déterminé de nouveaux groupes en fonction de si la population de rats est globalement croissante sur les deux dernières années, ou si elle est neutre ou décroissante. ce qui sera un paramètre utile lors de l'analyse spatiale et la classification en général. Nous avons pu faire de même avec les restaurants, en ajoutant la pente de la tendance régressée.

III - b) Analyse spatiale

III - b - 1) Détermination du meilleur classifieur

A partir du nouveau dataset composé ([figure](#)), on a d'abord fait un benchmark des différents classifieurs existant afin de trouver le meilleur. On a pris ici en compte pour chaque zipcode :

- Le nombre de rats par habitant
- la population
- Le nombre d'arbres recensés dans le zipcode
- le nombre de parc
- le nombre de restaurants
- le nombre de poubelle importante (non de particulier mais de restaurant par exemple)

La figure ci-dessous indique le score de chacune de ces méthodes et on note que le classifieur gradient boosting semble donner les meilleurs résultats.

```
print ("Le decision tree obtient un score de ",round(clf.score(X_test,y_test),2) )
print("Le SVM obtient un score de",round(clf1_3.score(X_test,y_test),2))
print("Le gradient Boosting Classifieur obtient un score de ",round(GBC.score(X_test,y_test),2))
```

Le decision tree obtient un score de 0.61
Le SVM obtient un score de 0.64
Le gradient Boosting Classifieur obtient un score de 0.66

Cependant ces résultats ne nous semblaient pas très élevés, on a donc cherché à les améliorer en intégrant en plus la fréquence de ramassage des déchets (par semaine) . On refait le benchmark avec la même démarche et on obtient effectivement des résultats un peu plus élevés (voir figure ci-dessous). On obtient toujours de meilleurs résultats avec le gradient Boosting, on a donc décidé de retenir ce classifieur pour la suite de l'étude.

```
print ("Le decision tree obtient un score de ",round(clf.score(X_test2,y_test2),2) )
print("Le SVM obtient un score de",round(clf1_3.score(X_test2,y_test2),2))
print("Le gradient Boosting Classifieur obtient un score de ",round(GBCg.score(X_test2,y_test2),2))
```

Le decision tree obtient un score de 0.68
Le SVM obtient un score de 0.67
Le gradient Boosting Classifieur obtient un score de 0.72

On a ensuite enlevé la population des variables du classifieur car on trouvait qu'elle n'apporterait rien. On trouvait finalement un score de 0.65.

On a ensuite cherché à établir une variable de validation.

III - 2 - b) Détermination d'une variable de validation

On avait besoin de créer une variable de validation pour valider les solutions proposées et s'assurer qu'elles réduisent bien le nombre de rats. On a fait une méthode de Coude pour déterminer le nombre optimal de catégories, on en a trouvé 3. On a donc décidé de classer les zipcodes en 3 catégories :

- le groupe 0 où le taux de rats par habitant est inférieur à 0.0095
- le groupe 1 où le taux de rats par habitant est entre 0.0095 et 0.021
- le groupe 2 où le taux de rats par habitant est supérieur à 0.021

Au départ de l'étude, on a compté 52 zipcodes dans le groupe 0, 58 zipcodes dans le groupe 1 et 55 zipcodes dans le groupe 2. On a considéré dans la suite de l'étude que nos

propositions étaient valides si elles améliorent ces chiffres, c'est-à-dire si le nombre de zipcode dans le groupe 2 baissait et celui dans le groupe 0 augmentait.

Il ne nous reste plus qu'à déterminer des variables actionnables cohérentes !

III - 2 - c) Détermination des variables activables

Parmi les corrélations de départ, celle qui est la plus corrélée avec le nombre de rats et la population. Cependant, ce ne peut pas être une variable actionnable car il est utopique de penser que l'on peut bouger de zipcode une grande quantité de personnes. On s'est donc concentré sur 2 variables qui pouvaient convenir: la fréquence de ramassage et le nombre de composteurs par zipcode.

- **Fréquence de ramassage**

On n'avait pas pu obtenir de bonne corrélation avec cette variable dans l'analyse descriptive. Il se trouve qu'en fonction des zipcodes, la ville de New York ramasse 2 ou 3 fois les déchets par semaine (avec une proportion équivalente). On a donc modifié ces données pour qu'elle ramasse 3 fois par semaine dans tous les zipcodes et on y a appliqué le classifieur défini plus haut. On s'attendait à avoir de meilleurs résultats mais ce fut finalement le contraire. Il semblerait donc qu'il y ait une corrélation inverse entre l'augmentation de 2 à 3 de la fréquence de ramassage et la baisse du nombre de rats. On peut donc en conclure que deux ramassages de déchets par semaine est suffisant et qu'il n'y a pas d'impact à augmenter cette fréquence sur le nombre de rats.

il y a 48 classes 0, contre 52 avant
il y a 52 classes 1 contre 58 avant
il y a 65 classes 2 contre 55 avant

- **Nombre de composteurs**

On avait déterminé dans l'analyse descriptive une corrélation de 0.41 entre le nombre de rats par habitant et le nombre de composteurs. On a donc voulu savoir si augmenter ou baisser ce nombre avait un impact significatif sur la population de rats. On a dans cette étude aucune information sur le degré de remplissage de ces composteurs, on a donc considéré qu'ils étaient uniformément remplis.

On a d'abord travaillé sur la répartition de ces composteurs. On a remarqué que certains zipcodes avaient beaucoup plus de composteurs que d'autres. On a donc pris la moyenne de composteur par zipcode (1.3) et on a attribué à tous les zipcodes cette moyenne avant d'y appliquer le classifieur.

On obtient de meilleurs résultats, ce qui nous incite à penser que la multiplication de lieux avec une forte concentration de composteurs a tendance à faire augmenter la population de rats. A l'inverse, la pose d'un nouveau composteur isolé semble avoir un impact moindre sur la population de rats. Il est donc intéressant de répartir le mieux possible les composteurs dans la ville et d'éviter de créer des zones avec une concentration importante de composteurs.

il y a 56 classes 0, contre 52 avant
il y a 69 classes 1 contre 58 avant
il y a 40 classes 2 contre 55 avant

Pour vérifier cette constatation, on a ensuite repris le nombre initial de composteur de chaque zipcode et on en a rajouté 3 à chaque zipcode. Les résultats sont sans appel, on a une augmentation très sensible du nombre de rats par habitant, ce qui confirme bien que la

il y a 21 classes 0, contre 52 avant
il y a 16 classes 1 contre 58 avant
il y a 128 classes 2 contre 55 avant

quantité de rats est corrélée au nombre de déchets, mais aussi qu'une concentration importante de composteurs est une cause importante du nombre élevé de rats.

Il faut bien sûr prendre ces résultats avec des pincettes. En effet, nous n'avons pas pris en compte la quantité de remplissage de ces composteurs faute d'information. Cependant il est probable que les résultats soient bien meilleurs si on ajoute 3 composteurs mais que le taux de remplissage des composteurs existants baisse drastiquement.

Autre chose, nous n'avons pas eu le temps de mettre en relation l'augmentation de conteneur et l'augmentation de la fréquence de ramassage. Il aurait été intéressant de coupler ces 2 variables pour vérifier si une augmentation du nombre de composteurs couplée à une augmentation de la fréquence de ramassage peut améliorer les résultats.

On a ensuite cherché un moyen d'étudier l'impact de la répartition "intelligente" du nombre de composteurs dans New York. Pour cela on a fait ces manipulations :

- Pour les zipcodes avec 0 composteur, on en ajoute 1
- Pour les zipcodes avec 3 composteurs, on en enlève 1
- Pour les zipcodes avec 4 composteurs, on en enlève 2 mais on en rajoute 1 à son plus proche voisin
- Pour les zipcodes avec 5 composteurs ou plus, on en enlève 3 mais on en ajoute 1 dans le zipcode le plus proche

Cela fait un peu plus de 50 composteurs de plus (augmentation de 25%). On obtient cette fois-ci des résultats équivalents malgré

l'augmentation du nombre de composteurs. `il y a 54 classes 0, contre 52 avant`

Nous n'avons pas eu le temps de le vérifier `il y a 51 classes 1 contre 58 avant`

mais il est probable qu'en prenant en compte `il y a 60 classes 2 contre 55 avant`

le taux de remplissage, comme il y a plus de

composteurs, celui-ci devrait être plus bas et cela devrait améliorer ces résultats. La ville de New York a donc tout intérêt à mieux répartir ces composteurs !

Bien sûr, cette étude a des limites, si certains zipcodes ont autant de composteurs et d'autres pas, c'est qu'il y a forcément des raisons. Malgré cela, la mairie de New York a intérêt d'y faire attention et de mieux les répartir dans la mesure du possible.

III - c) Analyse spatio-temporelle

Nous avons enfin cherché pour affiner le modèle en associant le côté spatial et le côté temporel. Pour cela nous avons entraîné un nouveau classifieur en ajoutant une variable en plus ou en remplacement de celle qui y était déjà :

- le "growingstate" : variable indiquant la rapidité de l'augmentation des restaurants ayant de mauvaises notes dans un zipcodes

On remarque ensuite que cette variable a une corrélation de 0.34 avec le nombre de rats par habitant, ce qui est significatif. Cela signifierait que plus les notes des restaurants dans un zipcode baisse avec le temps, plus le nombre de rats augmente.

Une fois le classifieur entraîné, on trouve un score de 0,68 ce qui est légèrement mieux que précédemment (0.65). Nous n'avons pas eu le temps ensuite d'aller plus loin mais on aurait certainement pu ajouter d'autres variables spatio-temporelles permettant d'affiner le modèle.

IV - Déroulement du projet

Nous allons revenir brièvement sur notre organisation et sur notre retour. Nous pensons que le plus important dans ce genre de projet est l'organisation et une attribution claire des tâches à chacun. Dans notre cas, cela c'est très bien déroulé, on se connaissait déjà avant et on a pas eu de mal à se répartir les tâches. En début de projet, nous étions 3 à nous concentrer sur l'analyse descriptive et sur la recherche et le rassemblement des datasets et l'autre était chargé de faire un début d'analyse temporel. Ensuite nous nous sommes réparti en 3 équipes, 1 personne qui se concentrait sur le cadre juridique / l'écriture des rapports / mettre en liaison les datasets, un groupe de 2 qui travaillaient sur l'analyse spatiale et le dernier sur l'analyse temporelle.

Au niveau organisation, nous avons créé un dépôt github pour le partage du code et un drive pour partager les différents documents / notebook.

Enfin, si on devait le refaire nous choisirions un sujet avec un dataset unique ou maximum 2 datasets, car on a du mal à mettre en relation les différents datasets, et notre sujet prend en compte tellement de variable que c'est difficile d'en trouver des vraiment significatives ou les prendre toutes en compte.

Conclusion

Nous avons donc, au cours de ce projet, pu :

- Trouver et regrouper les datasets sur différentes variables explicatives
- Identifier une échelle adaptée au problème
- Proposer un outil de seuil de vigilance à mettre en place pour surveiller l'évolution des rats dans le futur et proposer une manière d'améliorer l'efficacité des dératisations en utilisant la saisonnalité des données.
- Compris comment étaient réparties les données sur les observations de rats dans le temps
- Construit un modèle qui permet d'estimer, en fonction d'un jeu de paramètres donné, (population, nombre de restaurants avec des mauvaises notes, nombre de composteurs du zipcode) le nombre de rats qu'il y avait
- Trouver une variable activable (le nombre de composteurs) qui soit suffisamment corrélée avec la présence de rats.
- Déterminer qu'il n'était pas forcément nécessaire d'augmenter la fréquence de ramassage des déchets
- Proposer plusieurs solutions qui réduiraient le nombre de rats en jouant sur notre le nombre de composts et estimer leurs efficacités en utilisant le modèle de prédiction.

Les limites étant que les variables explicatives les plus corrélées ne sont pas forcément activables, et qu'il est très difficile d'obtenir des données suffisamment complètes pour proposer une solution optimale au problème, qui est par ailleurs très complexe.

Notre travail nécessite donc d'être approfondi, et d'avoir plus de moyens et des données plus complètes avant d'être complètement déployé. Une partie du problème vient en effet du fait que les données n'ont pas été saisies dans le même but que celui du projet, c'est pour cela qu'elles ne nous servent pas forcément aussi bien qu'on le voudrait.

Un travail qui pourrait être également intéressant et complémentaire au nôtre serait par exemple d'implémenter un algorithme de deep learning pour que les caméras puissent détecter les rats et prévenir les autorités compétentes et de rentrer les données sur la base de données.