

# Divvy Bike Project

*Rémy Garnier - Camille Palmier*

*28 février 2018*

## I- Introduction

But et intérêts de l'étude Remerciements

## II - Présentation des données

### Le jeu de données

Divvy est le système de partage de vélos de la ville de Chicago, avec 6000 vélos disponibles à plus de 570 stations à travers Chicago et Evanston. Divvy, comme les autres systèmes de vélos de libre-service, se compose d'une flotte de vélos robustes et durants qui sont enfermés dans un réseau de stations d'accueil dans toute la région. Les vélos peuvent être déverrouillés à partir d'une station et renvoyés à n'importe quelle autre station du système. Divvy offre aux résidents et aux visiteurs une option de transport pratique pour se déplacer et explorer Chicago.

Divvy est un programme du Département des Transports de Chicago (CDOT), qui possède les vélos, les stations et les véhicules de la ville. Le financement initial du programme provenait de subventions fédérales pour des projets qui favorisent le redressement économique, réduisent la congestion du trafic et améliorent la qualité de l'air, ainsi que des fonds supplémentaires provenant du programme de financement des augmentations d'impôt de la Ville. En 2016, Divvy s'est étendu à la banlieue voisine d'Evanston grâce à une subvention de l'État de l'Illinois.

Sur le site de Divvy, les données sur les trajets et les stations sont disponibles au public. Tous les ans, deux jeux de données sont publiés : un premier allant de janvier à juillet et un second allant de août à décembre.

Chaque trajet à vélo est anonymisé et comprend :

- Jour et heure du début du voyage
- Fin du voyage jour et heure
- Station de départ
- Station d'arrivée
- Type de coureur (membre, tour unique et passe d'exploration)
- Si un membre voyage, il inclura également le sexe et l'année de naissance du membre

Exemple de 5 trajets réalisés en 2014 :

##	trip_id	starttime	stoptime	bikeid	tripduration
## 1	2355134	6/30/2014 23:57	7/1/2014 0:07	2006	604
## 2	2355133	6/30/2014 23:56	7/1/2014 0:00	2217	263
## 3	2355130	6/30/2014 23:33	6/30/2014 23:35	2798	126
## 4	2355129	6/30/2014 23:26	7/1/2014 0:24	173	3481
## 5	2355128	6/30/2014 23:16	6/30/2014 23:26	173	638
##	from_station_id	from_station_name	to_station_id		
## 1	131	Lincoln Ave & Belmont Ave	303		
## 2	282	Halsted St & Maxwell St	22		

```
## 3          327 Sheffield Ave & Webster Ave          225
## 4          134   Peoria St & Jackson Blvd          194
## 5          320   Loomis St & Lexington St          134
##          to_station_name  usertype gender birthyear
## 1 Broadway & Cornelia Ave Subscriber   Male    1988
## 2      May St & Taylor St Subscriber   Male    1992
## 3 Halsted St & Dickens Ave Subscriber   Male    1993
## 4      State St & Wacker Dr Subscriber Female    1988
## 5 Peoria St & Jackson Blvd Subscriber Female    1988
```

Les 5 premières lignes de la table des stations de 2014 :

```
##      id                      name latitude longitude dpcapacity
## 1  43 Michigan Ave & Washington St 41.88389 -87.62465         43
## 2  44      State St & Randolph St 41.88473 -87.62773         27
## 3  33      State St & Van Buren St 41.87718 -87.62784         27
## 4 199      Wabash Ave & Grand Ave 41.89174 -87.62694         15
## 5  51      Clark St & Randolph St 41.88458 -87.63189         31
##      online date
## 1  2013-06-16
## 2  2013-06-16
## 3  2013-06-25
## 4  2013-08-10
## 5  2013-06-17
```

## Analyses descriptives des données

Table 1: Nombre de trajets et stations par année.

Année	Stations	Trajets
2013	300	759 788
2014	300	2 454 634
2015	474	3 183 439
2016	581	3 595 383
2017	585	3 829 014

Les vélos ont été mis en service en juin 2013 d'où un nombre de trajets plus faible. Ces chiffres nous montre une augmentation du nombre de trajets sur le réseau et une augmentation du réseau lui-même au cours du temps. C'est entre 2014 et 2015 que ces augmentations sont les plus importantes avec environ 30% de trajets supplémentaires et 174 nouvelles stations.

## Visualisation géographique

Pour visualiser l'emplacement des stations à Chicago, nous les avons tracés sur un fond de carte de la ville.

Extension du réseau de stations entre 2014 et 2016

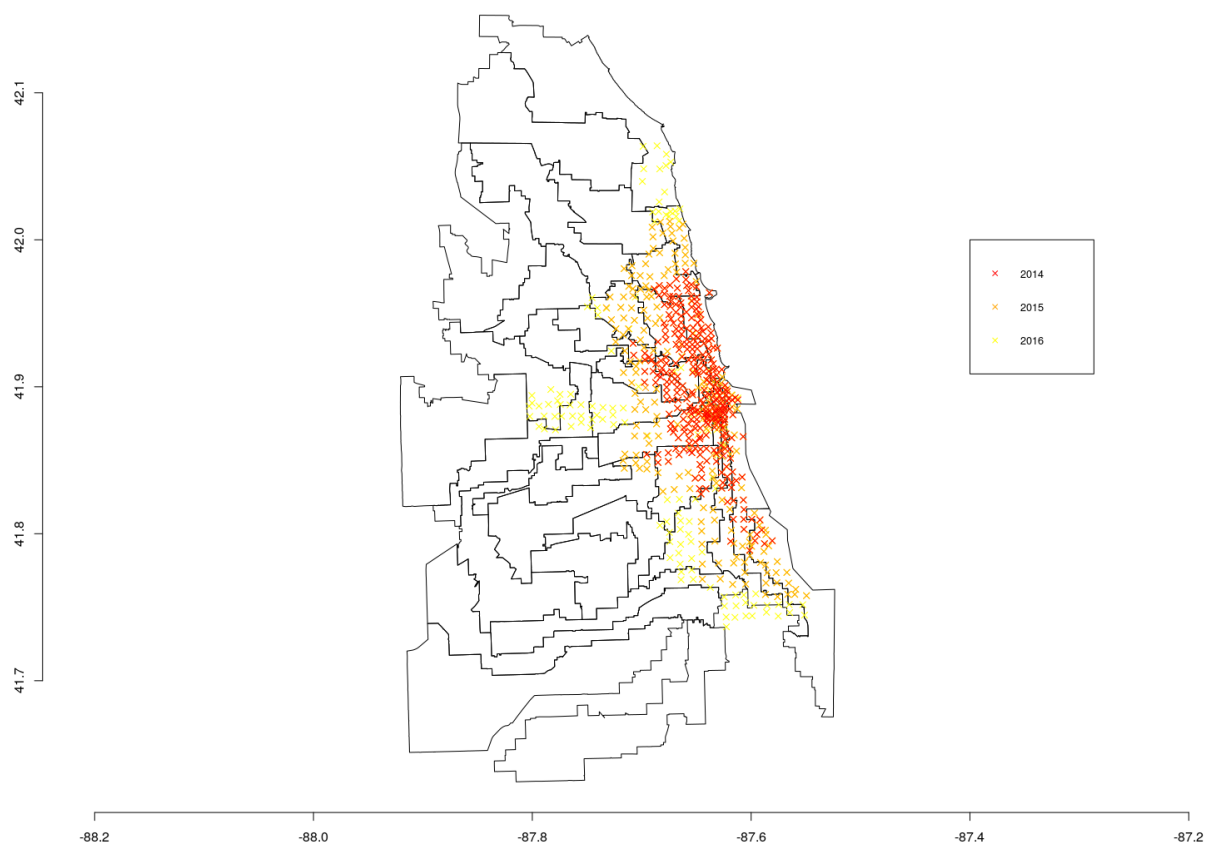


Figure 1:

Classement par fréquence des stations... : graphiques (stations qui bougent)

### **III - Transformation des données**

Agrégation géographie/temps Définition de notre série d'intérêt Justification séparation en district ? Cartes  
Ajouts nouvelles variables (dow, pluvio...) Quelques visualisations (série sur 2014/2015)

### **IV - Les modèles**

Tendance / Saisonalité Modèle GAM (GAMLSS) - significativité des variables Random Forests (Adaboost)  
Modèles auto-régressifs Agrégation

### **V - Conclusion**