



#### LE DATASET

- https://archive.ics.uci.edu/ml/datasets/Seoul+Bike+Sharing+Demand
- Ce dataset contient des données concernant la location de vélo à Seoul du 01/12/2017 jusqu'au 30/11/2018
- Il fait également part des conditions météorologiques ainsi que les informations sur la date.



#### LE DATASET

Le dataset se présente comme suis :

Date: year-month-day

Rented Bike count - Count of bikes rented at each hour

Hour - Hour of the day

Temperature-Temperature in Celsius

Humidity - %

Windspeed - m/s

Visibility - 10m

Dew point temperature - Celsius

Solar radiation - MJ/m2

Rainfall - mm

Snowfall - cm

Seasons - Winter, Spring, Summer, Autumn

Holiday - Holiday/No holiday

Functional Day - NoFunc(Non Functional Hours), Fun(Functional hours)



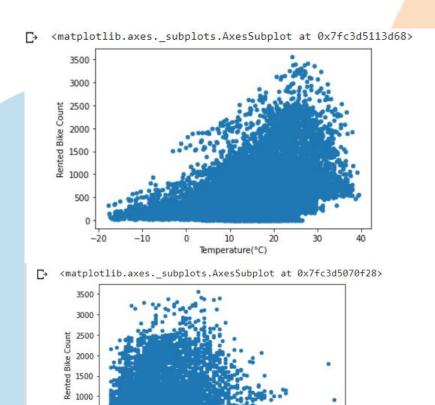


## NOS RÉFLEXIONS

- Dans un premier temps, nous avons observer et visualiser les différentes variables du dataset.
- Après diverses observations, il nous est apparu assez évident le but de ce projet.
  Nous devions prévoir les vélos loués en fonction des différents paramètres fournis.



### NOS OBSERVATIONS

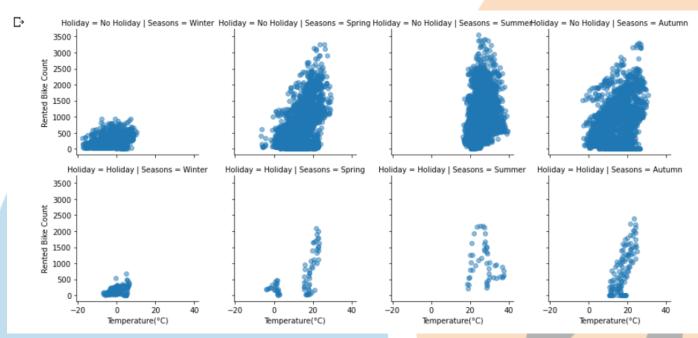


Nous avons dans un premier temps essayer de trouver visuellement des corrélations entre les différentes variables. Par exemple, nous avions pu observer une certaine linéarité entre la température et le nombre de vélo loués.

De plus nous avons chercher une corrélation entre la vitesse du vent le nombre de vélos mais sans succès. Le graphique tel qu'il est, est difficilement exploitable.



#### NOS OBSERVATIONS

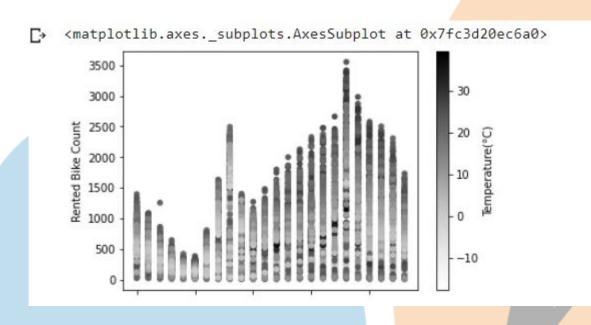


cikit

- Ici nous avons l'évolution globale du nombre de vélos loués en fonction des saisons, de la température, et la différence entre les vacances et les jours de travail.
- Globalement, nous pouvons voir que le nombre de vélo loués est bine plus important durant les saisons printanière, hivernal et automnale.



#### NOS OBSERVATIONS



scikit

- Enfin, nous avons observer les corrélations entre la température, l'heure ainsi que le nombre de vélo loués.
- Dans la globalité, nous pouvons observer un pic de location entre 8h et 9h ainsi qu'aux alentours de 19h.
- De plus le nombre de vélo loués tend à s'accroître avec la température comme observer au premier graphique.



#### LE DATASET

Afin que notre dataset soit parfaitement exploitable par la suite, nous avons dû apporter des modifications à ce dernier :

- Nous avons décidé de retirer la colonne « Date », celle-ci n'ayant que peu d'impact sur notre cible, sachant que cette dernière est intrinsèquement lié à la variable « Seasons », « Holiday », « Functionning Day ». De plus ce format est difficile à
- Nous avons remplacer les valeurs des 4 saisons par [1,2,3,4].
- Nous avons également remplacer les valeurs de la variable « Holiday » comme un boolean (entre 0 et 1).
- Nous avons fait de même pour la variable « Functionning Day ».
- Enfin lors la phase de train split, nous avons standardiser nos valeurs pour éliminer les valeurs négatives de température, nous permettant d'utiliser certaine fonction d'erreur.



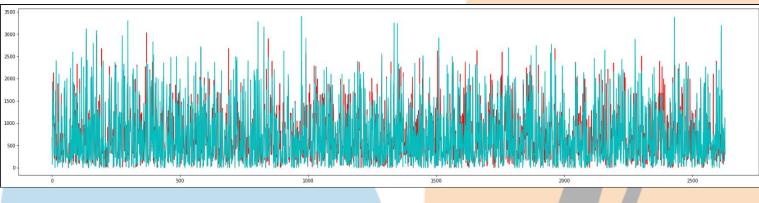
#### NOS PREDICTIONS

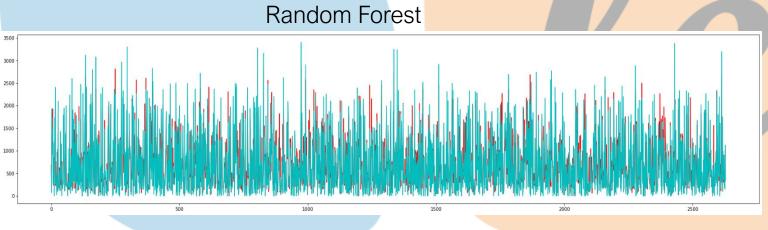
- Après différents tests de modèle, nous avons pris la décision d'utiliser deux modèles qui correspondaient au mieux à notre problématique : Decision Tree Regressor ainsi que Random Forest.
- En effet, à l'aide de GridSearchCV, nous avons pu établir les meilleurs hyperparamètres pour nos deux modèles respectifs.
- Sur nos deux modèles, nous avons eu un score respectifs de : 0,84 et 0,87 sur nos hyperparamètres trouvés par GridSearchCV.



### LES PERFORMANCES DE NOS PREDICTIONS

#### Decision Tree Regressor





En vert nous avons y\_test qui représente les résultats réels de x\_test.

# cikit

En rouge nous avons les résultats prédits par nos deux modèles respectifs



#### NOTRE API DJANGO

#### Welcome to our prediction API

Fill in the values and hit submit to find out how many bikes would be rented in Seoul in your specific meteorological conditions.

2	-
Temperature(°C):	
-7.8	-
Humidity(%):	
1051	-
Wind speed(m/s):	
1.9	-
Visibility(m):	
1627	-
Dew point temperature(°	C):
-16.1	-
Solar Radiation(MJ/m2):	
0	-
Rainfall(mm):	
0	-
Snowfall(cm):	
0	-
Seasons:	
Winter ~	
□ Holiday	
☐ Function day	

- Ceci est la page d'accueil de notre API.
- Sur cette page nous pouvons renseigner les différentes à valeur à fournir à nos predicteurs



#### NOTRE API DJANGO

The projected number of bikes that will be rented on your day is: 103.84615384 (DecisionTree Algorithm)

The projected number of bikes that will be rented on your day is: 115.71563174603172 (Random Forest Algorithm)

 Nous obtenons donc les deux prédictions liées à nos deux modèles.