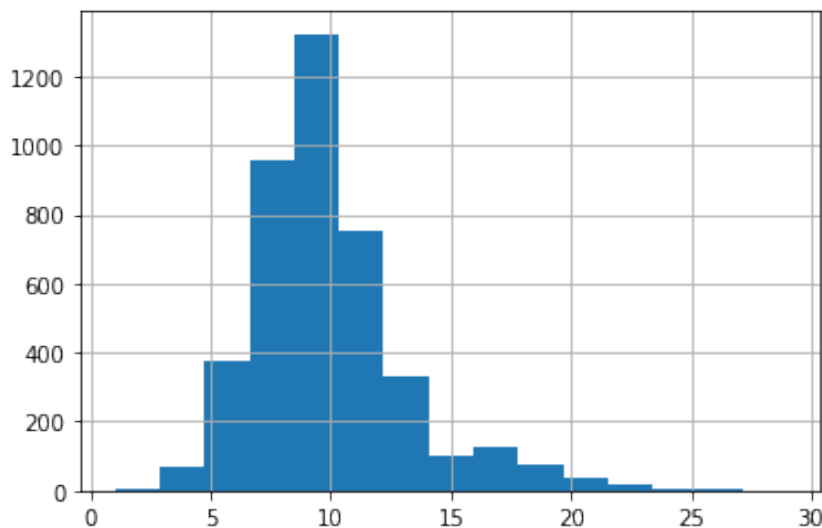


TP Abalone

- A. L'âge d'un abalone (mollusque marin) peut être déterminé en coupant sa coquille et en comptant le nombre d'anneaux sur la coquille. Notre but est de prédire l'âge (la colonne Rings) à partir des caractéristiques physiques de l'abalone.

index	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Rings
0	M	0.455	0.365	0.095	0.514	0.2245	0.101	0.15	15
1	M	0.35	0.265	0.09	0.2255	0.0995	0.0485	0.07	7
2	F	0.53	0.42	0.135	0.677	0.2565	0.1415	0.21	9
3	M	0.44	0.365	0.125	0.516	0.2155	0.114	0.155	10
4	I	0.33	0.255	0.08	0.205	0.0895	0.0395	0.055	7

Nous avons ploté l'histogramme de l'âge (rings) comme suit :



1. Ploter un histogramme de l'âge.
2. Quelles sont les variables que vous proposez dans les X du modèle avant d'appliquer un algorithme de K-NN ?
3. Nous disposons de 4177 mesures et nous avons construit un dataset de 80% d'entraînement et 20% de test. Appliquer le K-NN en variant les paramètres suivants : k et weight.
4. Calculer le RMSE (indication : autour de 2.37)
Interpréter ce résultat ?
5. Comment peut-on améliorer la régression ?
6. Peut-on appliquer la technique de bagging pour encore améliorer la performance du modèle ? comment ?