Commencez ici si...

Vous avez une certaine expérience avec R ou Python et les bases de l'apprentissage automatique. Il s'agit d'un concours parfait pour les étudiants en science des données qui ont suivi un cours en ligne sur l'apprentissage automatique et qui cherchent à élargir leurs compétences avant d'essayer un concours vedette.

Description du concours



Demandez à un acheteur de maison de décrire la maison de ses rêves, et il ne commencera probablement pas par la hauteur du plafond du sous-sol ou la proximité d'une voie ferrée est-ouest. Mais l'ensemble de données de ce concours de terrain de jeu prouve que beaucoup plus influence les négociations de prix que le nombre de chambres ou une clôture blanche.

Avec 79 variables explicatives décrivant (presque) tous les aspects des maisons résidentielles à Ames, Iowa, ce concours vous met au défi de prédire le prix final de chaque maison.

Compétences pratiques

* Ingénierie des fonctionnalités créatives
* Techniques de régression avancées comme la forêt aléatoire et l'amplification de gradient

Remerciements

L' ensemble de [données Ames Housing](http://www.amstat.org/publications/jse/v19n3/decock.pdf)  a été compilé par Dean De Cock pour être utilisé dans l'enseignement de la science des données. C'est une alternative incroyable pour les scientifiques des données à la recherche d'une version modernisée et étendue de l'ensemble de données Boston Housing souvent cité.