

A propos du Dataset:

Ce jeu de données est extrait des données du gouvernement américains dans le domaine public .Les données suivantes expliquent le coût de l'assurance d'un petit échantillon de la population américaine grâce à leur attributs:

- âge : âge du principal bénéficiaire
- genre: sexe de l'assureur, femme, homme
- IMC : Indice de masse corporelle, permettant de comprendre le corps, les poids relativement élevés ou faibles par rapport à la taille,indice objectif de poids corporel (kg/m^2) utilisant le rapport taille/poids, idéalement 18,5 à 24,9
- enfants : Nombre d'enfants couverts par l'assurance maladie / Nombre de personnes à charge
- fumeur : Fumer
- région américaines : zone résidentielle du bénéficiaire aux États-Unis, nord-est, sud-est, sud-ouest, nord-ouest.
- charges : Frais médicaux individuels facturés par l'assurance maladie

	age	genre	imc	enfants	fumeurs	region_americaine	charges
0	19	female	27.900	0	yes	southwest	16884.92400
1	18	male	33.770	1	no	southeast	1725.55230
2	28	male	33.000	3	no	southeast	4449.46200
3	33	male	22.705	0	no	northwest	21984.47061
4	32	male	28.880	0	no	northwest	3866.85520
...
1333	50	male	30.970	3	no	northwest	10600.54830
1334	18	female	31.920	0	no	northeast	2205.98080
1335	18	female	36.850	0	no	southeast	1629.83350
1336	21	female	25.800	0	no	southwest	2007.94500
1337	61	female	29.070	0	yes	northwest	29141.36030

Si vous sélectionnez ce Dataset pour votre projet, vous serez noté sur l'implémentation d'un algorithme de régression. Votre algorithme sera donc en mesure de prédire le tarif d'assurance en fonction du profils.