

Iliass Ben ammar  
Nithusan Sivakanthan

# Prédiction de prix sur les voitures d'occasions Audi

---

Licence 3 informatique  
Université Paris 8  
Mai 2021

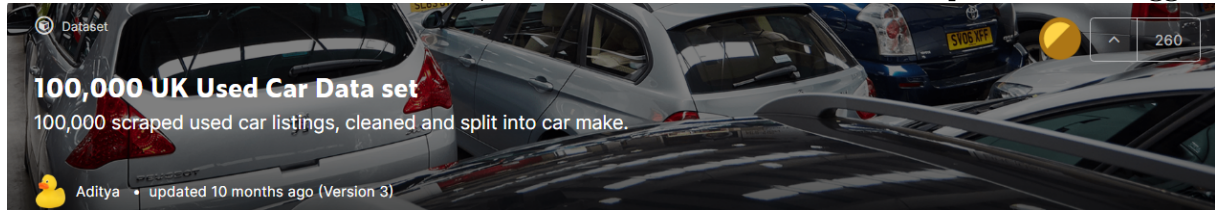
# TABLE DES MATIÈRES

## Sommaire :

1	Présentation du projet	2
2	Explication théorique	2
3	Code	3
4	Résultats	3
5	Conclusion	4
6	Source :	4

# 1 Présentation du projet

Le but de notre projet est de prédire le prix des voitures d'occasions de la marque Audi en s'aidant du dataset "100,000 UK Used Car Data set" disponible sur kaggle.



## 2 Explication théorique

Les données sont classé de cette façon :

	model	year	price	transmission	mileage	fuelType	tax	mpg	engineSize
0	A1	2017	12500	Manual	15735	Petrol	150	55.4	1.4
1	A6	2016	16500	Automatic	36203	Diesel	20	64.2	2.0
2	A1	2016	11000	Manual	29946	Petrol	30	55.4	1.4
3	A4	2017	16800	Automatic	25952	Diesel	145	67.3	2.0
4	A3	2019	17300	Manual	1998	Petrol	145	49.6	1.0
...	...	...	...	...	...	...	...	...	...
10663	A3	2020	16999	Manual	4018	Petrol	145	49.6	1.0
10664	A3	2020	16999	Manual	1978	Petrol	150	49.6	1.0
10665	A3	2020	17199	Manual	609	Petrol	150	49.6	1.0
10666	Q3	2017	19499	Automatic	8646	Petrol	150	47.9	1.4
10667	Q3	2016	15999	Manual	11855	Petrol	150	47.9	1.4

- model : fait référence au modèle du véhicule.
- year : fait référence à l'année de construction du véhicule.
- price : fait référence au prix d'achat du véhicule.
- transmission : fait référence au a la boite de vitesse (manuel ou bien automatique).
- mileage : fait référence au kilomètres que la voiture a parcouru.
- fueltype : fait référence au type de carburant.
- tax : fait référence à la taxe de circulation.
- mpg : fait référence a la mesure de la distance qu'un véhicule peut parcourir en utilisant un seul gallon de gaz.
- engineSize : fait référence à la taille du véhicule.

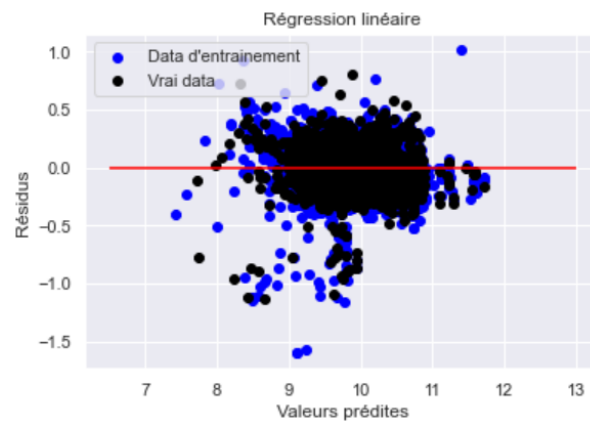
Les données vont être analyser puis séparer 70% pour la partie train et 30% pour la partie test afin de tester notre entraînement.

Pour le modèle de prédiction nous avons choisit la regrettions linéaire après de nombreuses recherches nous avons trouver que c'était le modèle le plus simple a utiliser et a mètre en pratique.

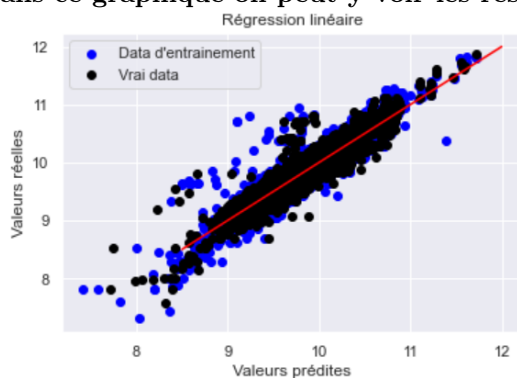
### 3 Code

- 1) On importe nos librairies .
- 2) On importe notre dataset .
- 3) On check si on a besoin de netoyer le dataset .
- 4) On check les meilleurs corrélations .
- 5) On découpe 30% de notre dataset pour le garder pour plus tard afin tester notre programme de prédiction.
- 6) On passe à l'entraînement
- 7) Calculs de précision de prédiction

### 4 Résultats



Dans ce graphique on peut y voir les résidus



Dans ce graphique on peut y voir les valeurs prédites sur les vrai valeurs

## 5 Conclusion

Pour conclure afin de réaliser la prédiction du prix des voitures d'occasion de la marque Audi nous avons utilisé le modèle de régression linéaire et les résultats obtenue paraissent satisfaisants ce qui montre que le modèle utilisé correspond bien à la tâche.

## 6 Source :

Kaggle :

<https://www.kaggle.com/adityadesai13/used-car-dataset-ford-and-mercedes?select=audi.csv>