



Bases de Données Large Echelle

Projet 2

Etudiant :

DJEGHRI Amine

MAMOU Idles

Numéro :

3801757

3803676

Décembre 2020

Description des données utilisées

The data behind the Inside Airbnb site is sourced from publicly available information from the Airbnb site.

Airbnb est un service de plateforme communautaire payant de location de logements de particuliers fondée en 2008 par des Américains (source: wikipedia)

Site d'où provient le dataset

Nous avons choisi le dataset Paris: listings

<http://insideairbnb.com/get-the-data.html>

Le dataset contient 16 attributs dont la variable à prédire qui est le prix

Lien direct du dataset (téléchargement):

<http://data.insideairbnb.com/france/ile-de-france/paris/2020-11-05/visualisations/listings.csv>

Taille du dataset: 9 mb

Lien du partage du notebook vers databricks :

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bfcf/3140631640044697/3605521943064466/6344744759096101/latest.html>