

# Ads\_clustering\_Joel\_KabanguKabobo

December 14, 2023

```
[1]: from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

## 0.1 Machine Learning Project:

### 0.1.1 Regroupement des annonces immobilières par l'apprentissage non supervisé(Unsupervised Learning)

Par Joel Kabangu Kabobo

Au sein des plateformes des annonces immobilières, il n'est pas rare de constater la présence répétitive d'annonces concernant un même bien, formulées de manière distincte sur différentes plateformes et assorties de descriptions variées. En effet, de nombreux acteurs professionnels, voire des particuliers, diffusent des annonces identiques relatives à un bien immobilier sur divers portails spécialisés. Afin d'optimiser le processus de recherche et d'obtenir des données statistiques plus précises, cette étude s'est focalisée sur la détection d'annonces similaires provenant des plateformes diverses. Ces annonces ont été regroupées de manière cohérente, tandis que les annonces uniques ont été traitées individuellement. Cette approche permet de distinguer les annonces uniques, considérées comme des exclusivités propres à la plateforme de publication. Notre analyse s'est concentrée sur les annonces immobilières en cours de publication pour la ville de Levallois-Perret au mois d'août 2019.

**Note :** si vous pouvez, veuillez excuter ce notebook avec le GPU pour une rapidité d'exécution sinon il fonctionne aussi avec le CPU mais l'exécution sera lente.

### 0.1.2 L'installation des bibliothèques et packages utilisées

```
[2]: !pip install hdbscan cuba sentence_transformers umap-learn transformers
↳ torchvision >/dev/null 2>/dev/nul
```

### Importation des bibliothèques utilisées

```
[ ]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import os
```

```

from sklearn.preprocessing import LabelEncoder
# using plotly to visualize clusters
import plotly.express as px
import plotly.graph_objects as go
import plotly.offline as pyo
from sklearn.decomposition import PCA
from sklearn.metrics import pairwise_distances
import copy
from sklearn.feature_extraction.text import CountVectorizer
from nltk import sent_tokenize, pos_tag
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
from transformers import AutoModel, AutoTokenizer
import torch
import torchvision
import random

pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)
pd.set_option('display.max_colwidth', None)

```

```

[ ]: import nltk
nltk.download('stopwords')
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')

```

```

[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data] Unzipping corpora/stopwords.zip.
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data] Unzipping tokenizers/punkt.zip.
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data] /root/nltk_data...
[nltk_data] Unzipping taggers/averaged_perceptron_tagger.zip.

```

```

[ ]: True

```

```

[ ]: def seed_everything(seed: int):
    random.seed(seed)
    os.environ['PYTHONHASHSEED'] = str(seed)
    np.random.seed(seed)
    torch.manual_seed(seed)
    torch.cuda.manual_seed(seed)
    torch.backends.cudnn.deterministic = True
    torch.backends.cudnn.benchmark = True

```

### 0.1.3 Chargement des données

```
[ ]: df = pd.read_csv("/content/drive/MyDrive/dataset/Dataset - Ads _L  
↳Levallois-Perret - 2019-08 - export-ads-levallois-perret-2019-08-27.csv")  
df.head(10)
```

```
[ ]:                                     ID \  
0  22c05930-0eb5-11e7-b53d-bbead8ba43fe  
1  8d092fa0-bb99-11e8-a7c9-852783b5a69d  
2  44b6a5c0-3466-11e9-8213-25cc7d9bf5fc  
3  e9e07ed0-812f-11e8-82aa-61eacebe4584  
4  872302b0-5a21-11e9-950c-510fefc1ed35  
5  de04afa0-2e5f-11e8-bda5-3334b193df7f  
6  530c7ac0-36a5-11e7-a435-b3bf21527190  
7  5135e0a0-62dd-11e8-91b3-8d7b603ef3b8  
8  bc076370-45ec-11e9-a00a-897941d6ca9c  
9  08491e50-8dfe-11e9-8a42-a16638394d93  
  
                                     URL \  
0  http://www.avendrealouer.fr/location/levallois-  
perret-92/b-appartement/3-pieces/loc-101-42762/fd-15565447.html  
1  
https://www.bienici.com/annonce/ag440414-165474561  
2  https://www.bellesdemeures.com/annonces/vente/appartement-luxe/levallois-  
perret-92/ile-de-la-jatte-square-baudin/144215779/  
3  
https://www.seloger.com/annonces/locations/bureaux/levallois-perret-92/villiers-  
danton/135859129.htm  
4  https://www.bellesdemeures.com/annonces/vente/maison-luxe/levallois-  
perret-92/hotel-de-ville-planchette/145918109/  
5  
https://immobilier.lefigaro.fr/annonces/annonce-15788745.html  
6  
https://www.seloger.com/annonces/locations/bureaux/levallois-  
perret-92/116903887.htm  
7  
https://immobilier.lefigaro.fr/annonces/annonce-16671439.html  
8  
https://www.leboncoin.fr/locations/1583612210.htm  
9  
https://www.leboncoin.fr/ventes_immobilieres/1631696940.htm  
  
CRAWL_SOURCE PROPERTY_TYPE NEW_BUILD \  
0  A_VENDRE_A_LOUER      APARTMENT      False  
1  BIEN_ICI              APARTMENT      False  
2  BELLES_DEMEURES       APARTMENT      False  
3  SE_LOGGER             PREMISES       False
```

4	BELLES_DEMEURES	HOUSE	False
5	EXPLORIMMO	APARTMENT	False
6	SE_LOGGER	PREMISES	False
7	EXPLORIMMO	PARKING	NaN
8	LE_BON_COIN	PARKING	NaN
9	LE_BON_COIN	APARTMENT	False

## DESCRIPTION \

0

Au rez de chaussée d'un bel immeuble récent, appartement de standing se composant d'une entrée, un grand séjour avec grande baie, cuisine entièrement équipée et aménagée (Frigidaire, four...), 2 chambres avec rangements, salle de bain en marbre avec double vasque, WC séparés avec lave-mains. Eau chaude et chauffage compris dans les charges. Parquet. Cave, parking. Très lumineux et calme. Mise ? jour le 16/04/2017 R?f?rence annonce AVENDREALOUER : L-95-09 9-7621

1

Je vous propose un appartement dans la rue Collange dans un immeuble de standing, moderne et sécurisé. 49m2 loi carrez situé au premier étage avec ascenseur comprenant une entrée, un séjour-salon, une cuisine équipée et aménagée, une chambre, une salle de bain et un toilette séparé. Appartement ayant une exposition sud donnant sur le jardin baigné de lumière et bénéficiant d'une belle hauteur sous plafond, parquet, balcon-terrasse ainsi que une cave. A proximité des écoles, des commerces et transports. Les charges de copropriété sont de 1440 euros/an. Appartement sans perte de place et vue verdoyante. Prix 429.000 euros honoraires charges vendeur. Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que conseillère indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389 Mandat 176430 Copropriété de 140 lots (). Charges annuelles : 1440 euros.

2

Dans un cadre arboré, calme et fleuri, un penthouse de 267 m<sup>2</sup> en très bon état se situant au 7ème et dernier étage. Cet appartement d'exception bénéficie d'une vue panoramique située sur la pointe de l'Île de la Jatte. Il dispose d'une terrasse de 300 m<sup>2</sup> de plain-pied face au salon. Il se compose d'un séjour de 75 m<sup>2</sup>, d'une grande cuisine dînatoire de 30 m<sup>2</sup> avec buanderie, d'une suite de 30 m<sup>2</sup> avec salle de bains et dressing, de 2 autres chambres avec salle de bains, d'un bureau ou 4e chambre. Annexes: une cave, un double box et une place de parking. Cris POCHON.

3 "Le meilleur coworking flexible de la ville, 50 m<sup>2</sup> de salons d'affaires et d'espaces de coworking pour vous. Meilleur rapport qualité-prix. Espaces flexibles de qualité disponibles dès aujourd'hui à Levallois à partir de 89 € par mois. Cliquez ici ! Venez nous rendre visite et commencez à travailler aujourd'hui ! Vous bénéficierez d'un programme d'adhésion qui vous donne accès à plusieurs espaces de coworking dans les lieux les plus prestigieux de la ville et vous permet de travailler selon des horaires de bureau flexibles. Il inclut un accès à un salon d'affaires entièrement équipé avec des bureaux partagés ou des box individuels, une connexion Internet haut débit, des rafraîchissements, du café, du thé, etc. En plus de nos programmes d'adhésion, nous proposons des

bureaux privatifs, réservés et flexibles à partir de 60 m<sup>2</sup>. Nous pouvons personnaliser votre espace de travail avec 1 à 100 bureaux ou plus. Nous fournissons tous les avantages du bureau pendant que votre équipe se concentre sur ses activités. La société IWG est à la tête de la révolution de l'espace de travail. Dans le cadre de nos activités, nous aidons plus de 2,5 millions de professionnels et entreprises à être plus productifs. En effet, nous leur offrons une sélection d'espaces de travail, communautés et services professionnels, inspirants et collaboratifs. Nous possédons plus de 3 000 sites et plus de 100 000 bureaux à la location allant de 5 m<sup>2</sup> à 1 000 m<sup>2</sup>, voire plus. Renseignez-vous dès maintenant ! ".

4

Levallois - Parc de la Planchette A toute proximité du parc de la Planchette et du quartier commerçant, dans une rue calme, maison d'architecte construite en 2011 d'une surface totale (maison + patio) de 330 m<sup>2</sup>, s'articulant sur trois niveaux principaux et en plus un sous-sol. Le premier niveau présente des réceptions avec une grande hauteur sous plafond et une spacieuse cuisine dînatoire qui donnent, l'un et l'autre, sur un agréable patio. Au premier étage s'organise l'espace des enfants autour de trois belles chambres, une salle de bains ainsi qu'une salle de douche. Le deuxième niveau accueille deux chambres supplémentaires dont la chambre de maître qui ouvre sur une petite terrasse. Au sous-sol, une grande salle de jeux, une buanderie et une chambre d'amis. La maison, entièrement climatisée, est présentée en parfait état. Baignée de lumière, elle offre des volumes ouvrant partiellement les uns sur les autres, le tout procurant à ses occupants une circulation très fluide. Location d'un emplacement de parking possible dans un immeuble mitoyen.

5

LEVALLOIS - FRONT DE SEINE - 5-6 PIECES Appartement Familial traversant de 142.76m<sup>2</sup> au 3ème étage d'une résidence récente de standing avec gardien et ascenseurs. Ce bien vous offre : entrée, double séjour de 41.69m<sup>2</sup> donnant sur balcon, cuisine dînatoire donnant sur balcon terrasse, 3 chambres (12.29m<sup>2</sup>, 14.36m<sup>2</sup> et 15.46m<sup>2</sup>) - possibilité 4ème chambre, 2 salles d'eau, 1 salle de bains et 3 wc. Nombreux rangements. Lumineux et ensoleillé. Une place de parking complètent ce bien. Possibilité d'acquérir une deuxième place de parking en sus. Soumis au statut de la copropriété PRIX DE VENTE : 1.420.000€ FAI dont 3.65 % d'honoraires à la charge de l'acquéreur. Building Partners, fort d'une expertise de plus de 20 ans, est spécialisé dans la transaction de biens 'Familiaux de Standing' et bénéficie d'implantations de choix dans l'ouest parisien. C'est avec plus de 450 appartements, maisons et hôtels particuliers que Building Partners offre le plus large choix de biens à la vente. 65, rue Président Wilson 92300 Levallois (33) 1.47.31.01. 74.

6

LEVALLOIS, Paul Vaillant Couturier, près gare Levallois Clichy, Local commercial, avec belle vitrine sur rue superficie environ 76 au sol, 63 Carrez, se décomposant en grand espace de 37 m<sup>2</sup>, partie surélevée 11 m<sup>2</sup>, bureau, WC,, petite pièce en sous-sol ouvert 13,75 m<sup>2</sup> Loyer 1620€ dont 120 € de charges - honoraires 6000 E. Honoraires TTC en sus : 6000 €.

7

Louez au mois un parking privé au 2 rue Maurice Ravel à Levallois-Perret pour 60€. Aucun engagement ni caution requis pour un abonnement au mois dans ce parking souterrain. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Présence d'un ascenseur : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Pour essayer ce parking retrouvez toutes les informations utiles sur le site internet Yespark avec la référence : 92300-1495-1Attention, ce parking ne peut accueillir que des deux-roues : scooter ou moto

8

LOUE place de parking dans résidence de standing sécurisée.\nAccès facile avec badge.\nLoyer mensuel : 100 € toutes charges comprises.\nDisponibilité immédiate.

9

A VENDRE APPARTEMENT DE 2 PIÈCES MAIRIE DE LEVALLOIS\n\nA vendre à Levallois-Perret, à proximité de l'hôtel de Ville, du parc de la planchette, des commerces et du métro Anatole France, dans un immeuble bien tenu et dans un environnement calme et lumineux, appartement de 2 pièces situé au 4ème étage avec ascenseur, il se compose d'une entrée, d'un séjour donnant sur un balcon, d'une cuisine aménagée, d'une chambre spacieuse, d'une salle de bains et d'un WC séparé. Une cave au sous-sol. DPE NC.\nPour toute information complémentaire ou pour prendre rendez-vous contactez nous au 01.47.48.90.10.\nRéférence annonce : V19775\nLes honoraires sont à la charge du vendeur

IMAGES \

0

["https://cf-medias.avendrealouer.fr/image/\_873908158\_d.jpg","https://cf-medias.avendrealouer.fr/image/\_873908159\_d.jpg","https://cf-medias.avendrealouer.fr/image/\_873908160\_d.jpg","https://cf-medias.avendrealouer.fr/image/\_873908161\_d.jpg","https://cf-medias.avendrealouer.fr/image/\_873908162\_d.jpg"]

1

["http://photos.ubiflow.net/440414/165474561/photos/1.jpg?20190401022226","http://photos.ubiflow.net/440414/165474561/photos/2.jpg?20190401022226","http://photos.ubiflow.net/440414/165474561/photos/3.jpg?20190401022226","http://photos.ubiflow.net/440414/165474561/photos/4.jpg?20190401022226","http://photos.ubiflow.net/440414/165474561/photos/5.jpg?20190401022226","http://photos.ubiflow.net/440414/165474561/photos/6.jpg?20190401022226"]

2

["https://v.seloger.com/s/width/965/visuels/0/9/3/7/0937mk01pff7dkq2n4dhqcwehdoa4okzpmzc1zgw.jpg","https://v.seloger.com/s/width/965/visuels/0/a/j/o/0ajo0vk9eb r4wr14ywp0zkm3bxwdit1dzkkdozlkw.jpg","https://v.seloger.com/s/width/965/visuels/0/w/l/2/0wl2vlttoepcvz24qoays30lyhph3xwn9s0mh1pgcg.jpg","https://v.seloger.com/s/width/965/visuels/1/s/o/a/1soaklv0bt3eun1wziqooyh6hxjp4igc7imo1mo00.jpg","https://v.seloger.com/s/width/965/visuels/2/6/d/b/26dbyhdhwukb7n30t9jub1jtjueprd3is4v1mytj4.jpg","https://v.seloger.com/s/width/965/visuels/0/z/e/o/0zeoexpb067nj17jmr ybbeuzvssav7qvl5mgls000.jpg","https://v.seloger.com/s/width/965/visuels/1/2/2/s/

122sdljffjcr83o312bn9b080jkhpeg160mrorn4zk.jpg", "https://v.seloger.com/s/width/965/visuels/1/r/m/p/1rmpeucxc94cfboiunpicj3tfu5s7jtqi660lxm9s.jpg", "https://v.seloger.com/s/width/965/visuels/1/c/o/0/1co082mqmawxfgvqvc94qxb7ey9l4xys8eu20xkw.jpg", "https://v.seloger.com/s/width/965/visuels/0/j/g/e/0jgex95s10o2b64jw032v9dlmdh3ak9a41b7ns9vk.jpg", "https://v.seloger.com/s/width/965/visuels/1/0/p/8/10p82gy6ecftk5navepv662ycffnq3vzy0m4fx7nk.jpg", "https://v.seloger.com/s/width/965/visuels/1/y/f/1/1yfl55uj9znq1j5q3go2pd2dcfwt2ixhbd73g0hds.jpg", "https://v.seloger.com/s/width/965/visuels/0/f/9/4/0f94q9liybfc2d17wnp926f4rdvt7k9pldgjr40.jpg", "https://v.seloger.com/s/width/965/visuels/1/p/a/8/1pa882qnalcr98q4dck5n34pc59mchmp23jv54bgg.jpg", "https://v.seloger.com/s/width/965/visuels/1/m/y/s/1mysnxtw0s79l4rdpbzcyv00l3zffw0vokqft1j4.jpg", "https://v.seloger.com/s/width/965/visuels/1/7/m/2/17m2rc4n8b4rawqpwxrvnsfst6gqnqbludobpuyo.jpg"]

3

["https://pix.yanport.com/ads/e9e07ed0-812f-11e8-82aa-61eacebe4584/image\_fc7d9181f6042f4a630c7fdb9f40ba3.jpg", "https://v.seloger.com/s/width/800/visuels/1/j/4/5/1j4582b04i58fwekzdtgiu20uq3a640e7w3papd4.jpg", "https://v.seloger.com/s/width/800/visuels/1/y/c/d/1ycdcnohlrp0qttl6edgfs4cjajfz8nmo02ox7fs.jpg", "https://v.seloger.com/s/width/800/visuels/1/p/0/1/1p0lztg2uxdey4rb322p2lk8iq17l1c8w1ukk6bjc.jpg", "https://v.seloger.com/s/width/800/visuels/0/g/d/o/0gdod5pizzrepy7g3gi4h4aq4b34ycr2lq6do10bs.jpg", "https://v.seloger.com/s/width/800/visuels/2/0/o/g/20ogztaombstv4wfenyhqgj3r906mur2vbbvp70jw.jpg"]

4 ["https://v.seloger.com/s/width/966/visuels/0/1/i/d/01idozdpkiqhx4m3geazwntv1tlizbmrgbsowuxqr.jpg", "https://v.seloger.com/s/width/966/visuels/1/1/m/p/11mpaj3fh0fkm5dibbwp0kvjb14m0nvft66zm9oib.jpg", "https://v.seloger.com/s/width/966/visuels/1/3/6/n/136njd5eerd2fanb5ziopgm81ryica9nvyh4y95jn.jpg", "https://v.seloger.com/s/width/966/visuels/0/x/e/p/0xepdzjiv3ke7dyfvzawmbq0ovblihi695uue95nbn.jpg", "https://v.seloger.com/s/width/966/visuels/2/a/u/a/2auaewifd1qo5wo93iq6lvfyql5n9j5d7b37zanv7.jpg", "https://v.seloger.com/s/width/966/visuels/1/9/4/j/194jfb4daqvqv2ansxsfni8bnash3shkn8cb6bzxar.jpg", "https://v.seloger.com/s/width/967/visuels/2/3/2/f/232fmf6i6n031m2pe9x4zll7ogosljsfxli3ciij8.jpg", "https://v.seloger.com/s/width/967/visuels/0/9/x/1/09x1ynwrcv4pagxpwxw3fdikqoi9a9q2c9skm7jv8.jpg", "https://v.seloger.com/s/width/966/visuels/0/4/c/z/04czv4qf9hfdzz3ft4jwztk8q6lmhlj2hup3gnu7.jpg", "https://v.seloger.com/s/width/966/visuels/1/6/b/w/16bwx1fvqeewn59mg8pxd1dmy034kehggf0dj9tub.jpg", "https://v.seloger.com/s/width/966/visuels/1/3/8/b/138b9v6clewev7xyaeiume4kempngkomsgvdbt5f.jpg", "https://v.seloger.com/s/width/967/visuels/1/9/0/i/190iqbmcrmqbz9qq6aacdbf8xn2pppkxirik5ij4k.jpg", "https://v.seloger.com/s/width/966/visuels/1/p/d/m/1pdmf81qcmf3w5rem5g3weok32wzsj7athuz9p103.jpg", "https://v.seloger.com/s/width/966/visuels/2/9/v/o/29vo678c6nenmx26g5dgeie2z5r6p5w5agu864ker.jpg", "https://v.seloger.com/s/width/967/visuels/0/1/4/8/01488yy8bb6shyy0t5l8oev0k0xkaqb90zf9rga50.jpg", "https://v.seloger.com/s/width/966/visuels/1/e/7/7/1e7715joruuqf9ui72k32o3m3d9p85h0f94xk7yr.jpg", "https://v.seloger.com/s/width/966/visuels/2/7/h/w/27hw4yaluzbvabxopyk604bf7n5pomb9qi95ryt03.jpg", "https://v.seloger.com/s/width/966/visuels/1/9/x/o/19xoysqdj9x7ffhesk7rgi1quy5k4xz1xclbf8r4j.jpg", "https://v.seloger.com/s/width/966/visuels/1/r/t/v/1rtvfc3376l76w6wx86rb4c7awvewaujupos08wgj.jpg", "https://v.seloger.com/s/width/966/visuels/1/g/d/w/1gdwgnwsbsxtksuwvrrf936grkogrn43ni1k447n.jpg", "https://v.seloger.com/s/width/966/visuels/0/1/c/2/0lc29posdw7obwi801c5a0nyr9clkiugqf5wj24ib.jpg", "https://v.seloger

.com/s/width/966/visuels/1/9/k/6/19k6d63jynnbkmo8xy7natw1h7tlz41stusyub83.jpg",  
"https://v.seloger.com/s/width/966/visuels/2/7/7/p/277p5gupde1ptrnr1zsvw7uzw2u55  
0lgki0lmm03.jpg", "https://v.seloger.com/s/width/967/visuels/0/m/b/b/0mbbj39bm49  
w7un6591jzrwq8rs5xdw8v60bb6el0.jpg", "https://v.seloger.com/s/width/967/visuels/1  
/7/q/1/17q1baqe67nrs627d557b7oirw61n1vul5mf6l3dg.jpg", "https://v.seloger.com/s/w  
idth/966/visuels/2/2/u/c/22ucaq0k1xx4st1gilasf8664w5qk2keg03erdw0j.jpg"]

5

["http://photo3.pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867599&lettre=a&for  
mat=1&d=20170124143804&ext=.jpg", "http://photo2.pericles.fr/photo\_get.php?S1=382  
9&S2=01&IO=14867599&lettre=b&format=1&d=20170317125339&ext=.jpg", "http://photo1.  
pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867599&lettre=c&format=1&d=20170317  
125339&ext=.jpg", "http://photo3.pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867  
599&lettre=d&format=1&d=20170317125339&ext=.jpg", "http://photo2.pericles.fr/phot  
o\_get.php?S1=3829&S2=01&IO=14867599&lettre=e&format=1&d=20170317125339&ext=.jpg"  
, "http://photo1.pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867599&lettre=f&for  
mat=1&d=20170317125339&ext=.jpg", "http://photo3.pericles.fr/photo\_get.php?S1=382  
9&S2=01&IO=14867599&lettre=g&format=1&d=20170317125339&ext=.jpg", "http://photo2.  
pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867599&lettre=h&format=1&d=20170317  
125339&ext=.jpg", "http://photo1.pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867  
599&lettre=i&format=1&d=20170317125339&ext=.jpg", "http://photo3.pericles.fr/phot  
o\_get.php?S1=3829&S2=01&IO=14867599&lettre=j&format=1&d=20170317125339&ext=.jpg"  
, "http://photo2.pericles.fr/photo\_get.php?S1=3829&S2=01&IO=14867599&lettre=k&for  
mat=1&d=20170317125339&ext=.jpg"]

6

["https://v.seloger.com/s/cdn/x/visuels/2/2/r/j/22rj7oi6hkj2k8munr2nbos8liscngds  
i5pxadpqt.jpg", "https://v.seloger.com/s/cdn/x/visuels/0/k/0/e/0kOes2zsugmf0rubi3  
w1m5fe5dbtedcy9mrpksej.jpg", "https://v.seloger.com/s/cdn/x/visuels/2/2/n/6/22n6  
ae1hjhnhksor4toavr4fy5g5qzp8mg8qorfhls.jpg", "https://v.seloger.com/s/cdn/x/visuel  
s/0/5/a/x/05axd8jx27taagy1ci57neddjf3y57081rhxx1tr4.jpg", "https://v.seloger.com/  
s/cdn/x/visuels/0/v/t/p/0vtph6okje67f50g19k4m3f5ozcyqhij4f2iahalc.jpg"]

7

["http://photos.ubiflow.net/751982/151820943/photos/1.jpg?20190307173853", "http:  
//photos.ubiflow.net/751982/151820943/photos/2.jpg?20190307173853", "http://photo  
s.ubiflow.net/751982/151820943/photos/3.jpg?20190319194136"]

8

["https://pix.yanport.com/ads/bc076370-45ec-11e9-a00a-  
897941d6ca9c/image\_780fb54c298609615d78800f53c94a26.jpg", "https://img1.leboncoin  
.fr/ad-image/406bafcf6c9c42034c33f5b1eb0414ac4fe699fc.jpg", "https://img1.lebonco  
in.fr/ad-image/fba3a393cc064d3c925ee72e706ffd41072bc2fd.jpg"]

9

["https://pix.yanport.com/ads/08491e50-8dfe-11e9-8a42-  
a16638394d93/image\_5c33df7018a70d9dc968a9cc0e56ae0c.jpg", "https://img5.leboncoin  
.fr/ad-image/4b689d66f16e5cb23adc27c0f9537167d706e2d8.jpg", "https://img3.lebonco  
in.fr/ad-image/82f37c98fbbe0c3d8b58aa1a0dafc93e70d57163.jpg", "https://img0.lebon  
coin.fr/ad-image/6c201c105084f122b00530454540e45e1ef3ecc6.jpg", "https://img5.leb  
oncoin.fr/ad-image/82bcce611206c5532de7782d57e72782b4c27cd2.jpg"]



	SURFACE	LAND_SURFACE	BALCONY_SURFACE	TERRACE_SURFACE	ROOM_COUNT	\
0	72.00	NaN	NaN	NaN	3.0	
1	48.00	NaN	NaN	NaN	2.0	
2	267.00	NaN	NaN	NaN	6.0	
3	50.00	NaN	NaN	NaN	0.0	
4	330.00	NaN	NaN	NaN	8.0	
5	142.76	NaN	NaN	NaN	5.0	
6	63.00	NaN	NaN	NaN	0.0	
7	16.00	NaN	NaN	NaN	NaN	
8	11.00	NaN	NaN	NaN	NaN	
9	57.00	NaN	NaN	NaN	2.0	

	BEDROOM_COUNT	BATHROOM_COUNT	LUNCHROOM_COUNT	TOILET_COUNT	FURNISHED	\
0	NaN	NaN	NaN	NaN	False	
1	1.0	NaN	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	
5	NaN	NaN	NaN	NaN	NaN	
6	NaN	NaN	NaN	NaN	NaN	
7	NaN	NaN	NaN	NaN	NaN	
8	NaN	NaN	NaN	NaN	NaN	
9	NaN	NaN	NaN	NaN	NaN	

	FIREPLACE	AIR_CONDITIONING	GARDEN	SWIMMING_POOL	BALCONY	TERRACE	\
0	NaN	NaN	NaN	NaN	NaN	NaN	
1	NaN	NaN	1.0	NaN	NaN	NaN	
2	NaN	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	NaN	
5	NaN	NaN	NaN	NaN	NaN	NaN	
6	NaN	NaN	NaN	NaN	NaN	NaN	
7	NaN	NaN	NaN	NaN	NaN	NaN	
8	NaN	NaN	NaN	NaN	NaN	NaN	
9	NaN	NaN	NaN	NaN	NaN	NaN	

	CELLAR	PARKING	PARKING_COUNT	HEATING_TYPES	HEATING_MODE	FLOOR	\
0	NaN	True	1.0	[]	COLLECTIVE	NaN	
1	NaN	False	NaN	[]	INDIVIDUAL	1.0	
2	NaN	False	NaN	[]	NaN	NaN	
3	NaN	False	NaN	[]	NaN	NaN	
4	NaN	False	NaN	[]	NaN	NaN	
5	NaN	False	NaN	[]	NaN	NaN	
6	NaN	False	NaN	[]	NaN	NaN	
7	NaN	False	NaN	[]	NaN	NaN	
8	NaN	False	NaN	[]	NaN	NaN	
9	NaN	False	NaN	[]	NaN	NaN	

	FLOOR_COUNT	CONSTRUCTION_YEAR	ELEVATOR	CARETAKER	ENERGY_CONSUMPTION	\
0	NaN	NaN	True	NaN	NaN	
1	6.0	NaN	True	False	NaN	
2	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	
5	NaN	NaN	NaN	NaN	NaN	
6	NaN	NaN	NaN	NaN	NaN	
7	NaN	NaN	NaN	NaN	NaN	
8	NaN	NaN	NaN	NaN	NaN	
9	NaN	NaN	NaN	NaN	NaN	

	GREENHOUSE_GAS_CONSUMPTION	MARKETING_TYPE	PRICE	PRICE_M2	\
0	NaN	RENT	1979.0	27.49	
1	NaN	SALE	429000.0	8937.50	
2	NaN	SALE	4970000.0	18614.23	
3	NaN	RENT	89.0	1.78	
4	NaN	SALE	2790000.0	8454.55	
5	NaN	SALE	1420000.0	9946.76	
6	NaN	RENT	1740.0	27.62	
7	NaN	RENT	60.0	3.75	
8	NaN	RENT	100.0	9.09	
9	NaN	SALE	570000.0	10000.00	

	PRICE_EVENTS	\
0	[{"price":1979,"date":"2017-03-22"}]	
1	[{"price":452000.0,"date":"2018-09-18"}, {"price":429000.0,"date":"2018-09-25"}]	
2	[{"price":4970000.0,"date":"2019-02-19"}]	
3	[{"price":89.0,"date":"2018-07-06"}]	
4	[{"price":3150000.0,"date":"2019-04-08"}, {"price":2980000.0,"date":"2019-05-09"}, {"price":2790000.0,"date":"2019-06-14"}]	
5	[{"price":1420000.0,"date":"2018-03-23"}]	
6	[{"price":1925,"date":"2017-04-27"}, {"price":1845,"date":"2017-05-19"}, {"price":1925,"date":"2017-05-22"}, {"price":1967,"date":"2017-06-09"}, {"price":1965.0,"date":"2018-10-25"}, {"price":1845.0,"date":"2019-03-29"}, {"price":1800.0,"date":"2019-05-10"}, {"price":1740.0,"date":"2019-07-10"}]	
7	[{"price":55.0,"date":"2018-05-29"}, {"price":60.0,"date":"2019-07-10"}]	
8	[{"price":90.0,"date":"2019-03-14"}, {"price":100.0,"date":"2019-05-14"}]	

9

```
[{"price":570000.0,"date":"2019-06-13"}, {"price":559000.0,"date":"2019-07-13"}]
```

	RENTAL_EXPENSES	RENTAL_EXPENSES_INCLUDED	DEPOSIT	FEES	FEES_INCLUDED	\
0	276.0	True	NaN	NaN	NaN	
1	NaN	NaN	NaN	NaN	False	
2	NaN	NaN	NaN	NaN	NaN	
3	NaN	NaN	NaN	NaN	NaN	
4	NaN	NaN	NaN	NaN	NaN	
5	NaN	NaN	NaN	NaN	False	
6	120.0	NaN	NaN	NaN	NaN	
7	NaN	True	NaN	NaN	NaN	
8	NaN	True	NaN	NaN	NaN	
9	NaN	NaN	NaN	NaN	NaN	

	EXCLUSIVE_MANDATE	AGENCIES_UNWANTED	OCCUPIED	DEALER_NAME	\
0	False	NaN	NaN	Lamirand Et Associes	
1	False	NaN	False	Proprietes Privees	
2	False	NaN	False	Propriétés Parisiennes	
3	False	NaN	NaN	Iwg	
4	False	NaN	False	Daniel Feau Neuilly	
5	False	NaN	False	Building Partners	
6	False	NaN	NaN	Cabinet De L'ouest Parisien	
7	False	NaN	NaN	Yespark	
8	False	False	NaN	baudoin	
9	False	NaN	False	Impact Immo	

	DEALER_TYPE	CITY_ID	CITY	ZIP_CODE	DEPT_CODE	\
0	AGENCY	54178039	Levallois-Perret	92300.0	92	
1	MANDATARY	54178039	Levallois-Perret	92300.0	92	
2	AGENCY	54178039	Levallois-Perret	92300.0	92	
3	AGENCY	54178039	Levallois-Perret	92300.0	92	
4	AGENCY	54178039	Levallois-Perret	92300.0	92	
5	AGENCY	54178039	Levallois-Perret	92300.0	92	
6	AGENCY	54178039	Levallois-Perret	92300.0	92	
7	AGENCY	54178039	Levallois-Perret	92300.0	92	
8	PRIVATE	54178039	Levallois-Perret	92300.0	92	
9	AGENCY	54178039	Levallois-Perret	92300.0	92	

	PUBLICATION_START_DATE	PUBLICATION_END_DATE	LAST_CRAWL_DATE	\
0	2017-03-22T04:07:56.095	NaN	2017-04-21T18:52:35.733	
1	2018-09-18T11:04:44.461	NaN	2019-06-06T10:08:10.89	
2	2019-02-19T16:49:03.547	NaN	2019-06-13T08:22:14.314	
3	2018-07-06T15:18:59.805	NaN	2019-06-18T10:40:07.405	
4	2019-04-08T17:12:20.123	NaN	2019-06-14T15:02:18.155	
5	2018-03-23T06:03:10.963	NaN	2019-07-10T15:06:39.615	
6	2017-04-27T01:35:00.674	NaN	2019-07-10T15:06:17.12	

```

7  2018-05-29T01:12:11.485      NaN  2019-07-10T02:46:51.652
8  2019-03-14T00:04:24.651      NaN  2019-07-13T11:03:51.699
9  2019-06-13T17:09:20.737      NaN  2019-07-13T18:17:01.782

```

```

    LAST_PRICE_DECREASE_DATE
0                          NaN
1          2018-09-25
2                          NaN
3                          NaN
4          2019-06-14
5                          NaN
6          2019-07-10
7                          NaN
8                          NaN
9          2019-07-13

```

```
[ ]:
```

#### 0.1.4 Description et traitement des données

```
[ ]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2164 entries, 0 to 2163
Data columns (total 57 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ID                    2164 non-null  object
1   URL                   2164 non-null  object
2   CRAWL_SOURCE          2164 non-null  object
3   PROPERTY_TYPE         2164 non-null  object
4   NEW_BUILD             1973 non-null  object
5   DESCRIPTION           2160 non-null  object
6   IMAGES                2164 non-null  object
7   SURFACE               2050 non-null  float64
8   LAND_SURFACE          3 non-null     float64
9   BALCONY_SURFACE       0 non-null     float64
10  TERRACE_SURFACE       25 non-null    float64
11  ROOM_COUNT            1835 non-null  float64
12  BEDROOM_COUNT         696 non-null   float64
13  BATHROOM_COUNT        0 non-null     float64
14  LUNCHROOM_COUNT       0 non-null     float64
15  TOILET_COUNT          0 non-null     float64
16  FURNISHED             467 non-null   object
17  FIREPLACE             0 non-null     float64
18  AIR_CONDITIONING      0 non-null     float64
19  GARDEN                14 non-null    float64

```

20	SWIMMING_POOL	6 non-null	float64
21	BALCONY	0 non-null	float64
22	TERRACE	86 non-null	float64
23	CELLAR	0 non-null	float64
24	PARKING	2164 non-null	bool
25	PARKING_COUNT	157 non-null	float64
26	HEATING_TYPES	2164 non-null	object
27	HEATING_MODE	653 non-null	object
28	FLOOR	660 non-null	float64
29	FLOOR_COUNT	465 non-null	float64
30	CONSTRUCTION_YEAR	503 non-null	float64
31	ELEVATOR	552 non-null	object
32	CARETAKER	247 non-null	object
33	ENERGY_CONSUMPTION	0 non-null	float64
34	GREENHOUSE_GAS_CONSUMPTION	0 non-null	float64
35	MARKETING_TYPE	2164 non-null	object
36	PRICE	2104 non-null	float64
37	PRICE_M2	1991 non-null	float64
38	PRICE_EVENTS	2164 non-null	object
39	RENTAL_EXPENSES	441 non-null	float64
40	RENTAL_EXPENSES_INCLUDED	600 non-null	object
41	DEPOSIT	55 non-null	float64
42	FEES	94 non-null	float64
43	FEES_INCLUDED	511 non-null	object
44	EXCLUSIVE_MANDATE	2164 non-null	bool
45	AGENCIES_UNWANTED	201 non-null	object
46	OCCUPIED	1209 non-null	object
47	DEALER_NAME	2077 non-null	object
48	DEALER_TYPE	2164 non-null	object
49	CITY_ID	2164 non-null	int64
50	CITY	2164 non-null	object
51	ZIP_CODE	2163 non-null	float64
52	DEPT_CODE	2164 non-null	int64
53	PUBLICATION_START_DATE	2164 non-null	object
54	PUBLICATION_END_DATE	0 non-null	float64
55	LAST_CRAWL_DATE	2164 non-null	object
56	LAST_PRICE_DECREASE_DATE	340 non-null	object

dtypes: bool(2), float64(29), int64(2), object(24)

memory usage: 934.2+ KB

```
[ ]: #voir le nombre des lignes et des colonnes de nos données
print('Number of rows: ', df.shape[0])
print('Number of columns: ', df.shape[1])
```

Number of rows: 2164  
Number of columns: 57

```
[ ]: #voir les nombre de données manquantes pour chaque variable
df.isnull().sum()
```

```
[ ]: ID                                0
URL                                    0
CRAWL_SOURCE                          0
PROPERTY_TYPE                         0
NEW_BUILD                             191
DESCRIPTION                           4
IMAGES                                0
SURFACE                               114
LAND_SURFACE                         2161
BALCONY_SURFACE                      2164
TERRACE_SURFACE                     2139
ROOM_COUNT                           329
BEDROOM_COUNT                       1468
BATHROOM_COUNT                      2164
LUNCHROOM_COUNT                     2164
TOILET_COUNT                        2164
FURNISHED                           1697
FIREPLACE                           2164
AIR_CONDITIONING                    2164
GARDEN                              2150
SWIMMING_POOL                       2158
BALCONY                              2164
TERRACE                             2078
CELLAR                              2164
PARKING                              0
PARKING_COUNT                       2007
HEATING_TYPES                        0
HEATING_MODE                         1511
FLOOR                                1504
FLOOR_COUNT                         1699
CONSTRUCTION_YEAR                   1661
ELEVATOR                            1612
CARETAKER                           1917
ENERGY_CONSUMPTION                  2164
GREENHOUSE_GAS_CONSUMPTION          2164
MARKETING_TYPE                      0
PRICE                                60
PRICE_M2                             173
PRICE_EVENTS                         0
RENTAL_EXPENSES                     1723
RENTAL_EXPENSES_INCLUDED            1564
DEPOSIT                             2109
FEES                                 2070
FEES_INCLUDED                       1653
```

```

EXCLUSIVE_MANDATE          0
AGENCIES_UNWANTED          1963
OCCUPIED                    955
DEALER_NAME                 87
DEALER_TYPE                 0
CITY_ID                     0
CITY                        0
ZIP_CODE                     1
DEPT_CODE                   0
PUBLICATION_START_DATE      0
PUBLICATION_END_DATE        2164
LAST_CRAWL_DATE              0
LAST_PRICE_DECREASE_DATE    1824
dtype: int64

```

```

[ ]: #voir les information statistique de nos données
df.describe()

```

```

[ ]:
count    SURFACE  LAND_SURFACE  BALCONY_SURFACE  TERRACE_SURFACE  \
count    2050.000000      3.000000           0.0      25.000000
mean     128.136068      30.000000           NaN      31.363600
std      423.339898      8.660254           NaN      57.467242
min        6.000000     25.000000           NaN       5.590000
25%       36.200000     25.000000           NaN      13.000000
50%       55.000000     25.000000           NaN      18.000000
75%       93.000000     32.500000           NaN      26.000000
max     10287.000000     40.000000           NaN     300.000000

```

```

count    ROOM_COUNT  BEDROOM_COUNT  BATHROOM_COUNT  LUNCHROOM_COUNT  \
count    1835.000000     696.000000           0.0           0.0
mean        2.243597      1.847701           NaN           NaN
std         1.715621      1.104508           NaN           NaN
min          0.000000      1.000000           NaN           NaN
25%          1.000000      1.000000           NaN           NaN
50%          2.000000      1.000000           NaN           NaN
75%          3.000000      2.000000           NaN           NaN
max         10.000000      8.000000           NaN           NaN

```

```

count    TOILET_COUNT  FIREPLACE  AIR_CONDITIONING  GARDEN  SWIMMING_POOL  \
count          0.0          0.0           0.0     14.0          6.0
mean          NaN          NaN           NaN      1.0          1.0
std           NaN          NaN           NaN      0.0          0.0
min           NaN          NaN           NaN      1.0          1.0
25%           NaN          NaN           NaN      1.0          1.0
50%           NaN          NaN           NaN      1.0          1.0
75%           NaN          NaN           NaN      1.0          1.0
max           NaN          NaN           NaN      1.0          1.0

```

	BALCONY	TERRACE	CELLAR	PARKING_COUNT	FLOOR	FLOOR_COUNT	\
count	0.0	86.0	0.0	157.000000	660.000000	465.000000	
mean	NaN	1.0	NaN	1.605096	2.916667	5.827957	
std	NaN	0.0	NaN	1.299694	2.123793	2.432820	
min	NaN	1.0	NaN	1.000000	0.000000	0.000000	
25%	NaN	1.0	NaN	1.000000	1.000000	5.000000	
50%	NaN	1.0	NaN	1.000000	2.000000	6.000000	
75%	NaN	1.0	NaN	2.000000	4.000000	7.000000	
max	NaN	1.0	NaN	9.000000	14.000000	19.000000	

  

	CONSTRUCTION_YEAR	ENERGY_CONSUMPTION	GREENHOUSE_GAS_CONSUMPTION	\
count	503.000000	0.0	0.0	
mean	1965.081511	NaN	NaN	
std	37.403519	NaN	NaN	
min	1870.000000	NaN	NaN	
25%	1930.000000	NaN	NaN	
50%	1976.000000	NaN	NaN	
75%	1997.000000	NaN	NaN	
max	2021.000000	NaN	NaN	

  

	PRICE	PRICE_M2	RENTAL_EXPENSES	DEPOSIT	\
count	2.104000e+03	1991.000000	441.000000	55.000000	
mean	4.261142e+05	5483.773420	406.687347	2988.506000	
std	6.486603e+05	5384.760624	1329.451157	8071.172419	
min	3.300000e+01	1.780000	2.000000	70.000000	
25%	1.600000e+03	32.375000	60.000000	860.500000	
50%	2.200000e+05	7586.210000	100.000000	1350.000000	
75%	6.097500e+05	10056.750000	151.000000	2450.000000	
max	6.000000e+06	89000.000000	14287.330000	60000.000000	

  

	FEEs	CITY_ID	ZIP_CODE	DEPT_CODE	PUBLICATION_END_DATE
count	94.000000	2164.0	2163.0	2164.0	0.0
mean	1531.935957	54178039.0	92300.0	92.0	NaN
std	6269.065130	0.0	0.0	0.0	NaN
min	29.500000	54178039.0	92300.0	92.0	NaN
25%	378.000000	54178039.0	92300.0	92.0	NaN
50%	639.670000	54178039.0	92300.0	92.0	NaN
75%	806.700000	54178039.0	92300.0	92.0	NaN
max	60000.000000	54178039.0	92300.0	92.0	NaN

Nous voyons que nos données contiennent 2164 lignes et 57 colonnes. Nous pouvons observer aussi d'autres informations comme la valeur moyenne, minimum, maximum, etc des variables de nos données.

Nous remarquons aussi que nos données contiennent des variables qui sont en nombre(float, int) ou en lettre(object) ou en booléenne(bool).



Nous remarquons aussi que nos données contiennent plus de valeurs manquantes et même des variables qui ne contiennent aucune valeur que des manquantes alors nous allons supprimer ces variables là.

```
[ ]: df['BALCONY'].head()
```

```
[ ]: 0    NaN
      1    NaN
      2    NaN
      3    NaN
      4    NaN
      Name: BALCONY, dtype: float64
```

```
[ ]: #Suppression des lignes où toutes les valeurs sont null
df = df.dropna(how='all')
print('Number of rows: ', df.shape[0])

# Suppression des colonnes où toutes les valeurs sont null
nan_cols = df.columns[df.isna().all()].tolist()
df = df.drop(nan_cols, axis=1)
print('Number of columns: ', df.shape[1])
```

```
Number of rows:  2164
Number of columns:  46
```

```
[ ]:
```

```
[ ]: # Affichage des nombres de valeurs uniques pour chaque variable
for col in df.columns:
    print(col, df[col].nunique())
```

```
ID 2164
URL 2164
CRAWL_SOURCE 13
PROPERTY_TYPE 5
NEW_BUILD 2
DESCRIPTION 1914
IMAGES 1947
SURFACE 454
LAND_SURFACE 2
TERRACE_SURFACE 18
ROOM_COUNT 11
BEDROOM_COUNT 7
FURNISHED 2
GARDEN 1
SWIMMING_POOL 1
TERRACE 1
PARKING 2
```

```

PARKING_COUNT 8
HEATING_TYPES 5
HEATING_MODE 3
FLOOR 12
FLOOR_COUNT 16
CONSTRUCTION_YEAR 48
ELEVATOR 2
CARETAKER 2
MARKETING_TYPE 2
PRICE 764
PRICE_M2 997
PRICE_EVENTS 1733
RENTAL_EXPENSES 129
RENTAL_EXPENSES_INCLUDED 2
DEPOSIT 44
FEES 71
FEES_INCLUDED 2
EXCLUSIVE_MANDATE 2
AGENCIES_UNWANTED 2
OCCUPIED 2
DEALER_NAME 553
DEALER_TYPE 3
CITY_ID 1
CITY 1
ZIP_CODE 1
DEPT_CODE 1
PUBLICATION_START_DATE 2103
LAST_CRAWL_DATE 2164
LAST_PRICE_DECREASE_DATE 145

```

Nous pouvons bien voir nos variables et le nombre des valeurs uniques qu'elles contiennent.

Explorons en details certaines de ces variables pour pouvoir resoudre le problème des valeurs manquantes

```
[ ]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2164 entries, 0 to 2163
Data columns (total 46 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ID                    2164 non-null  object
1   URL                   2164 non-null  object
2   CRAWL_SOURCE          2164 non-null  object
3   PROPERTY_TYPE         2164 non-null  object
4   NEW_BUILD             1973 non-null  object
5   DESCRIPTION           2160 non-null  object
6   IMAGES                2164 non-null  object

```

7	SURFACE	2050 non-null	float64
8	LAND_SURFACE	3 non-null	float64
9	TERRACE_SURFACE	25 non-null	float64
10	ROOM_COUNT	1835 non-null	float64
11	BEDROOM_COUNT	696 non-null	float64
12	FURNISHED	467 non-null	object
13	GARDEN	14 non-null	float64
14	SWIMMING_POOL	6 non-null	float64
15	TERRACE	86 non-null	float64
16	PARKING	2164 non-null	bool
17	PARKING_COUNT	157 non-null	float64
18	HEATING_TYPES	2164 non-null	object
19	HEATING_MODE	653 non-null	object
20	FLOOR	660 non-null	float64
21	FLOOR_COUNT	465 non-null	float64
22	CONSTRUCTION_YEAR	503 non-null	float64
23	ELEVATOR	552 non-null	object
24	CARETAKER	247 non-null	object
25	MARKETING_TYPE	2164 non-null	object
26	PRICE	2104 non-null	float64
27	PRICE_M2	1991 non-null	float64
28	PRICE_EVENTS	2164 non-null	object
29	RENTAL_EXPENSES	441 non-null	float64
30	RENTAL_EXPENSES_INCLUDED	600 non-null	object
31	DEPOSIT	55 non-null	float64
32	FEES	94 non-null	float64
33	FEES_INCLUDED	511 non-null	object
34	EXCLUSIVE_MANDATE	2164 non-null	bool
35	AGENCIES_UNWANTED	201 non-null	object
36	OCCUPIED	1209 non-null	object
37	DEALER_NAME	2077 non-null	object
38	DEALER_TYPE	2164 non-null	object
39	CITY_ID	2164 non-null	int64
40	CITY	2164 non-null	object
41	ZIP_CODE	2163 non-null	float64
42	DEPT_CODE	2164 non-null	int64
43	PUBLICATION_START_DATE	2164 non-null	object
44	LAST_CRAWL_DATE	2164 non-null	object
45	LAST_PRICE_DECREASE_DATE	340 non-null	object

dtypes: bool(2), float64(18), int64(2), object(24)

memory usage: 748.2+ KB

Nous remarquons beaucoup des variables ne contiennent preque pas des valeurs. Nous sommes tenter à le supprimer mais ces variables peuvent être important pour la detection de la simiralité des annonces. Par exemple la variable “SWIMING\_POOL” qui contient que 6 valeurs, cela peut être une information importante à tenir en compte car ce n’est pas toutes les maisons qui ont des piscines alors si on en a 6 valeurs dans nos données c’est fort probable qu’il s’agisse des annonces doublons car uniquement ses annonces ont des piscines. Explorons la variable

```
[ ]: ## Affichage des données des annonces avec piscine
df_swim = df[~df['SWIMMING_POOL'].isna()]
df_swim.head(6)
```

```
[ ]:                                     ID \
503      b8b93640-5f7c-11e9-adc1-59e67c527828
1010     ef5df090-30ac-11e9-8213-25cc7d9bf5fc
1261     2929ba90-2c05-11e9-8330-99e77f092ba1
1431     e3cdb250-c086-11e9-9fab-c3006e339e11
2115     e25b4110-49aa-11e9-ba7a-0ff7ef4fe44b
2128     f6dab810-c159-11e9-81e9-5f24299f2ef3
```

```
                                     URL \
503      https://www.seloger.com/annonces/achat-de-
prestige/maison/levallois-perret-92/henri-barbusse/146174921.htm
1010      https://www.seloger.com/annonces/achat-de-
prestige/maison/levallois-perret-92/villiers-l-michel/144077349.htm
1261      https://www.seloger.com/annonces/achat-de-
prestige/appartement/levallois-perret-92/trezel/143880175.htm
1431
https://www.seloger.com/annonces/locations/appartement/levallois-
perret-92/maurice-ravel/150342287.htm
2115      https://www.seloger.com/annonces/achat-de-prestige/hotel-
particulier/levallois-perret-92/henri-barbusse/145175875.htm
2128
https://www.seloger.com/annonces/locations/appartement/levallois-
perret-92/wilson-rivay/150358697.htm
```

```
      CRAWL_SOURCE PROPERTY_TYPE NEW_BUILD \
503      SE_LOGGER      HOUSE      False
1010     SE_LOGGER      HOUSE      False
1261     SE_LOGGER     APARTMENT     False
1431     SE_LOGGER     APARTMENT     False
2115     SE_LOGGER      HOUSE      False
2128     SE_LOGGER     APARTMENT     False
```

```
      DESCRIPTION \
503
En plein coeur de Levallois, quartier Villiers-Louise Michel, au calme, hôtel
particulier d'exception au style contemporain et design. Il est composé de 6
pièces sur 3 niveaux avec ascenseur et offre de belles prestations piscine,
Hamam, salle de cinéma, 3 terrasses, Patio et jardin privatif. 423 m² au sol
(loi Carrez 289 m² + annexes 134 m²). Il comprend un grand séjour avec cuisine
équipée, 1 suite parentale avec salle de bain, salle d'eau et dressing
attenants, 4 chambres avec salles d'eau attenantes, WC séparés, buanderie. Dont
3.78 % honoraires TTC à la charge de l'acquéreur.
1010
```

Hôtel Particulier contemporain sur 3 niveaux avec Ascenseur. 6 pièces avec piscine, hammam, salle de cinéma, 3 terrasses, balcon, Patio et jardin privatif. 423 m<sup>2</sup> au sol (Loi Carrez 289 m<sup>2</sup> + Annexe 134 m<sup>2</sup>) Grand séjour avec cuisine aménagée et équipée, 1 suite parentale avec salle de bain, sa salle d'eau et son Dressing, 4 chambres avec leur salle d'eau, WC séparés, buanderie. Très rare. Bien d'exception.

1261

EXCLUSIVITÉ 2 PIÈCES LEVALLOIS PLACE JEAN ZAY Exclusivité Levallois Place Jean Zay. Appartement de 2 pièces exposé Ouest en bon état composé d'une pièce à vivre avec cuisine américaine, d'une chambre, d'une salle de douche avec toilettes refaite à neuf. Une cave complète ce bien. Faibles charges. Idéal Pied à terre ou Investisseur.

1431 Charmant appartement de 37 m<sup>2</sup> refait à neuf, il est meublé et entièrement équipé, lumineux et agréable, dans une résidence calme et sécurisée. Disponible en début de mois prochain. Il dispose de: - une cuisine entièrement équipée (plaque à induction, four à pyrolyse, micro-ondes, lave-vaisselle, réfrigérateur/congélateur) ouverte sur le séjour. - un séjour avec un grand canapé d'angle, un fauteuil, un grand rangement avec penderie et étagères à chaussures et une table qui permet de prendre des repas à 4 personnes. - une chambre séparée comprend un lit double en 160x200 ainsi qu'un grand dressing - une salle de bain avec baignoire, toilettes et machine à laver séchante. Une place de parking privée se trouve au 2ème sous-sol accessible par ascenseur depuis l'intérieur de l'immeuble et par badge sécurisé depuis l'extérieur. Loyer: 1320 euros + 130 euros de charges (appartement et parking sécurisé) Dépôt de garantie: 2 mois Chauffage électrique individuel. Visites possibles en fin de mois. Le quartier calme dispose de: - un accès à 4 minutes à pieds à la gare Clichy-levallois pour se rendre en moins de 10 minutes à la gare St Lazare - un accès à différents commerces (moyenne surface de proximité, boulangerie, épicerie, pharmacie, bureau de presse et de tabac..) - un marché tous les mercredis et samedi matin sur la place Jean Zay (primeurs, poissonnier, boucher, charcutier, fleuriste..) - le centre nautique de Levallois à 3 minutes à pied (piscines, sauna, hammam..) - les quais de Seine à 5 minutes à pieds offrant pelouses, aires de jeux et de repos et un accès à l'île de la Jatte. - l'accès aux restaurants et boutiques de Paris et Levallois en moins de 10 minutes. Me contacter par mail en indiquant: projet immobilier + garanties financières apportées + coordonnées + questions éventuelles. Garantie : 2640 €. Disponible le 01/09/2019.

2115

En plein coeur de Levallois, quartier Villiers-Louise Michel, au calme, hôtel particulier d'exception au style contemporain et design. Il se compose de 6 pièces sur 3 niveaux avec ascenseur et offre de belles prestations telles que piscine, Hammam, salle de cinéma, 3 terrasses, patio et jardin privatif. 423 m<sup>2</sup> au sol (loi Carrez 289 m<sup>2</sup> + annexes 134 m<sup>2</sup>). Il comprend un grand séjour avec cuisine équipée, une suite parentale avec salle de bain, salle d'eau et dressing attenants, 4 chambres avec leur salle d'eau attenante, des WMC séparés, une buanderie. Contact: Nathalie Millereau + 33 (0) 6 75 04 25 02 dont 3.78 % honoraires TTC à la charge de l'acquéreur.

2128

Bonjour, Location d'un studio 26 m<sup>2</sup>, très clair. Dans le centre ville de Levallois Perret (rue Louis Rouquier), à proximité des commerces de salle de sport/piscine municipale/centre commercial So Ouest.. Et situé entre métro 3 Louise Michel et Anatole France et à moins de 10mn pour la Ligne L Clichy Levallois > St Lazare. Situé au 3ème étage d'un immeuble de 6 étages avec ascenseur Grand salon: contenant un lit avec sommier plus un canapé convertible et plusieurs rangements (placard buffet et armoire) - cuisine avec table murale pour manger, micro-onde, plaque de cuisson, réfrigérateur, deux chaises hautes, vaisselles..) Cuisine et salle de bains indépendantes au salon. Loyer: 950 Euro (s) charges comprises: 800e loyer + 150e charges (chauffage collectif inclus dans les charges). Disponible: immédiatement Je suis plus réactif au messages reçu à travers le site. Garantie : 1600 €.

IMAGES \

503 ["https://v.seloger.com/s/width/800/visuels/1/a/r/g/1arg135bgu63se8dhysct6m4tg7n3dox9a7427ec4.jpg", "https://v.seloger.com/s/width/800/visuels/1/3/e/9/13e9inwuvh25c61bl08m4tlw1ldfvzxbmyupwls9g.jpg", "https://v.seloger.com/s/width/800/visuels/0/2/f/a/02fa4q5v3gg4th786htxi9bzwmg8of0oxixqxxlg.jpg", "https://v.seloger.com/s/width/800/visuels/1/e/h/m/1ehmr08n5me4wr0otz3bmfff5f54dda2fqre0qi50.jpg", "https://v.seloger.com/s/width/800/visuels/0/z/s/r/0zsr4oez95o0q194eadkoyiirufuz772i101nn84.jpg", "https://v.seloger.com/s/width/800/visuels/1/q/3/q/1q3qj0al4kmrpeqy8ivynrctub7teudfy3cpf03tg.jpg", "https://v.seloger.com/s/width/800/visuels/1/z/9/l/1z9l7pmwokld4df71oy3r78n43556nyumtd1uiwd0.jpg", "https://v.seloger.com/s/width/800/visuels/1/y/7/y/1y7yo233yguin11p74s5rdk2eaamn6knlwylbeixg.jpg", "https://v.seloger.com/s/width/800/visuels/1/4/h/7/14h71z1mwpp5gx0xgfcfpnr4ysfo4d9y28gub2nh0.jpg", "https://v.seloger.com/s/width/800/visuels/1/f/f/e/1ffekez2uup08wek8m3mcjwi3b7uut5iq68hhbo4.jpg", "https://v.seloger.com/s/width/800/visuels/0/n/m/z/0nmzdah0d9xkqg2uucj8cqyxr2actni6cu9zpgllg.jpg", "https://v.seloger.com/s/width/800/visuels/1/q/p/t/1qptu7cj4o98059a73xp0swbgth678edlos9u3plg.jpg", "https://v.seloger.com/s/width/800/visuels/1/a/i/6/1ai6adehnnzv809i2swcqhn272aiqo6cm9j0fqwk4.jpg", "https://v.seloger.com/s/width/800/visuels/0/q/q/v/0qqvm3tbxedntlfct8w2gnr93phd8bisebzkkyous.jpg", "https://v.seloger.com/s/width/800/visuels/0/9/z/t/09ztf7gf22qwdy8ku4qralfbcfm7669zy5s1hkmz8.jpg", "https://v.seloger.com/s/width/800/visuels/0/u/7/h/0u7hj3nfx4hb01n0oqp0otwxwij876oxtyrfyta38.jpg", "https://v.seloger.com/s/width/800/visuels/1/g/e/9/1ge91zh8jnnhup7f7ka6k3xxtd0l974nwri1eaoz.jpg", "https://v.seloger.com/s/width/800/visuels/1/h/y/h/1hyhviu26jiigyq9qq15nb93i8xsk0fneowze6qpi.jpg"]

1010

["https://v.seloger.com/s/width/800/visuels/1/b/y/f/1byfv5nwjl7x7nsgajdlwt5x9o0iijzz430ome4wk.jpg", "https://v.seloger.com/s/width/800/visuels/1/4/s/q/14squkw8fw4e4e975tweoe01bs7xhq33dj0kcb49g.jpg", "https://v.seloger.com/s/width/800/visuels/1/v/s/6/1vs6de172zrns4bd3h3nxbodvnsnfdzm9y65nu6ac.jpg", "https://v.seloger.com/s/width/800/visuels/0/s/1/i/0s1iwul4h9llq27e4e30hvibf6mgtuedhc3jpme6s.jpg", "https://v.seloger.com/s/width/800/visuels/1/y/k/p/1ykpq0gwtujgwmhsfdaipp8oqv7i8nu53djm xoxlw.jpg", "https://v.seloger.com/s/width/800/visuels/0/x/a/g/0xagcz26vunq8k0nop tk0t2o70q92lpe2ph4l6xxw.jpg", "https://v.seloger.com/s/width/800/visuels/2/4/j/0/

24j0xr2yu451iz91mul5yx5a4aj0qhohz0r0kdphw.jpg", "https://v.seloger.com/s/width/800/visuels/2/4/m/v/24mvatphm5rabn9s6uoleyoadkkl7icow0h2zothw.jpg", "https://v.seloger.com/s/width/800/visuels/0/5/m/e/05me9ru0s948itkqfrux3trcsk7c9o7opr4cdtq38.jpg", "https://v.seloger.com/s/width/800/visuels/1/p/l/3/1pl3mgkc3eac5lghsawb8gjrc7lgbnvy3f9pvfqdw.jpg", "https://v.seloger.com/s/width/800/visuels/0/1/d/7/0ld7w8dxwl3c47ns7fyd8z2yxqzjewk626pbv0l78.jpg"]

1261

["https://pix.yanport.com/ads/2929ba90-2c05-11e9-8330-99e77f092ba1/image\_109baac1effb81d2871d0abad4f29b32.jpg", "https://v.seloger.com/s/width/800/visuels/1/w/r/4/1wr4fuj7pnvohocmcezrxvach8grd5216slx97z3.jpg"]

1431

["https://pix.yanport.com/ads/e3cdb250-c086-11e9-9fab-c3006e339e11/image\_b89ccbd948e384dfedd254b149f65ec2.jpg", "https://v.seloger.com/s/width/800/visuels/2/7/g/m/27gm67s1zzwx2u3fypcr9xh98oosuqw1a1gt73x5s.jpg", "https://v.seloger.com/s/width/800/visuels/0/k/s/q/0ksqq0zb0gm3ma34pkbd2irbf6qgou454570seqo.jpg", "https://v.seloger.com/s/width/800/visuels/1/5/z/o/15zoct7mivn3glpc78f32n8fk9vdbjs331ncpmd9c.jpg", "https://v.seloger.com/s/height/800/visuels/0/b/i/3/0bi30zsq8mtsqc7jjwt9o5guf7xoteexke9b05hko.jpg", "https://v.seloger.com/s/width/800/visuels/0/v/t/c/0vtc6yc3j520v43uribgyuoansyahlf3b1xb6qj6o.jpg", "https://v.seloger.com/s/width/800/visuels/2/6/o/7/26o7ay7euzmiesl29nnlzd52xot0wigqnz08a17ds.jpg", "https://v.seloger.com/s/width/800/visuels/0/e/q/y/0eqyehmr9eq4aop1vwawz4p5xv2z7aryc13ock9s.jpg", "https://v.seloger.com/s/width/800/visuels/0/z/d/7/0zd7rhxvn7l4dwbngqtknmtqhrym9chopm969wy3k.jpg"]

2115 ["https://v.seloger.com/s/width/800/visuels/1/3/e/9/13e9inwuvh25c6l1b108m4tlw1ldfvzxbmyupwls9g.jpg", "https://v.seloger.com/s/width/800/visuels/0/2/f/a/02fa4q5v3gg4th786htxi9bzwmg8of0oxxixqxxlg.jpg", "https://v.seloger.com/s/width/800/visuels/1/a/r/g/larg135bgu63se8dhysct6m4tg7n3d0x9a7427ec4.jpg", "https://v.seloger.com/s/width/800/visuels/1/e/h/m/1ehmr08n5me4wr0otz3bmffff5f54dda2fqre0qi50.jpg", "https://v.seloger.com/s/width/800/visuels/0/z/s/r/0zsr4oez95o0q194eadkoyiirufuz772i101nn84.jpg", "https://v.seloger.com/s/width/800/visuels/1/q/3/q/1q3qj0al4kmrpeqy8ivynrctub7teudfy3cpf03tg.jpg", "https://v.seloger.com/s/width/800/visuels/1/z/9/l/1z9l7pmwokld4df71oy3r78n43556nyumtdl1uiwd0.jpg", "https://v.seloger.com/s/width/800/visuels/1/y/7/y/1y7yo233yguin11p74s5rdk2eaamn6kn1wy1beixg.jpg", "https://v.seloger.com/s/width/800/visuels/1/4/h/7/14h71z1mwpp5gx0xgfcfpnr4ysfo4d9y28gub2nh0.jpg", "https://v.seloger.com/s/width/800/visuels/1/f/f/e/1ffekez2uup08wek8m3mcjwi3b7uut5iq68hhbo4.jpg", "https://v.seloger.com/s/width/800/visuels/0/n/m/z/0nmzdah0d9xkqg2uucj8cqyxr2actni6cu9zpgllg.jpg", "https://v.seloger.com/s/width/800/visuels/1/q/p/t/1qptu7cj4o98059a73xp0swbgth678edlos9u3plg.jpg", "https://v.seloger.com/s/width/800/visuels/1/a/i/6/1ai6adehnnzv809i2swcqhn272aiqo6cm9j0fqwk4.jpg", "https://v.seloger.com/s/width/800/visuels/0/q/q/v/0qqvm3tbxedntlftct8w2gnr93phd8bisebzkkyous.jpg", "https://v.seloger.com/s/width/800/visuels/0/9/z/t/09ztf7gf22qwdy8ku4qralfbcfm7669zy5s1hkmz8.jpg", "https://v.seloger.com/s/width/800/visuels/0/u/7/h/0u7hj3nfx4hb01n0oqp0otwxwij876oxtyrftyta38.jpg", "https://v.seloger.com/s/width/800/visuels/1/g/e/9/1ge91zh8jnnhup7f7ka6k3xxt0l974nwri1eaoz.jpg", "https://v.seloger.com/s/width/800/visuels/1/h/y/h/1hyhviu26jiigyq9qq15nb93i8xsk0fneowze6qpi.jpg"]

2128

["https://pix.yanport.com/ads/f6dab810-c159-11e9-81e9-5f24299f2ef3/image\_4a148f6115bf0ec2e135a14d11e3bd46.jpg", "https://v.seloger.com/s/width/800/visuels/0/6/a/8/06a8x4khrmbttrtdw03zoon70b1c7xj1054ku3zeo.jpg", "https://v.seloger.com/s/width/800/visuels/1/c/x/z/1cxzcnubhs99egqg1n7ddc5eft4li4xhtsqi5yps0.jpg", "https://v.seloger.com/s/height/800/visuels/1/e/e/6/1ee6utius5js13uimaps96aykrt05vnjftfep06ps.jpg"]

	SURFACE	LAND_SURFACE	TERRACE_SURFACE	ROOM_COUNT	BEDROOM_COUNT	\
503	290.00	NaN	NaN	6.0	5.0	
1010	289.00	40.0	NaN	6.0	5.0	
1261	29.68	NaN	NaN	2.0	1.0	
1431	37.00	NaN	NaN	2.0	1.0	
2115	290.00	NaN	NaN	6.0	5.0	
2128	26.00	NaN	NaN	1.0	NaN	

	FURNISHED	GARDEN	SWIMMING_POOL	TERRACE	PARKING	PARKING_COUNT	\
503	NaN	NaN	1.0	1.0	False	NaN	
1010	NaN	NaN	1.0	1.0	False	NaN	
1261	NaN	NaN	1.0	NaN	False	NaN	
1431	True	NaN	1.0	NaN	False	NaN	
2115	NaN	NaN	1.0	1.0	False	NaN	
2128	True	NaN	1.0	NaN	False	NaN	

	HEATING_TYPES	HEATING_MODE	FLOOR	FLOOR_COUNT	CONSTRUCTION_YEAR	\
503	["ELECTRIC"]	INDIVIDUAL	0.0	NaN	NaN	
1010	["ELECTRIC"]	NaN	0.0	NaN	NaN	
1261	[]	INDIVIDUAL	1.0	4.0	1930.0	
1431	[]	NaN	NaN	NaN	NaN	
2115	[]	NaN	0.0	NaN	NaN	
2128	[]	NaN	3.0	6.0	NaN	

	ELEVATOR	CARETAKER	MARKETING_TYPE	PRICE	PRICE_M2	\
503	NaN	NaN	SALE	3269000.0	11272.41	
1010	NaN	NaN	SALE	3269000.0	11311.42	
1261	NaN	NaN	SALE	279000.0	9400.27	
1431	True	NaN	RENT	1450.0	39.19	
2115	NaN	NaN	SALE	3269000.0	11272.41	
2128	True	NaN	RENT	950.0	36.54	

PRICE\_EVENTS \

503  
 [{"price":3429000.0,"date":"2019-04-15"}, {"price":3273000.0,"date":"2019-05-10"}, {"price":3269000.0,"date":"2019-05-11"}]  
 1010 [{"price":3500000.0,"date":"2019-02-14"}, {"price":3450000.0,"date":"2019-03-13"}, {"price":3429000.0,"date":"2019-03-29"}, {"price":3419000.0,"date":"2019-05-05"}, {"price":3269000.0,"date":"2019-05-09"}]  
 1261



```
[{"price":279000.0,"date":"2019-02-09"}]
1431
[{"price":1450.0,"date":"2019-08-17"}]
2115
[{"price":3429000.0,"date":"2019-03-18"}, {"price":3273000.0,"date":"2019-05-10"}, {"price":3269000.0,"date":"2019-05-11"}]
2128
[{"price":950.0,"date":"2019-08-18"}]
```

	RENTAL_EXPENSES	RENTAL_EXPENSES_INCLUDED	DEPOSIT	FEES	FEES_INCLUDED	\
503	NaN	NaN	NaN	NaN	NaN	
1010	NaN	NaN	NaN	NaN	NaN	
1261	NaN	NaN	NaN	NaN	NaN	
1431	130.0	True	NaN	NaN	NaN	
2115	NaN	NaN	NaN	NaN	NaN	
2128	150.0	True	NaN	NaN	NaN	

	EXCLUSIVE_MANDATE	AGENCIES_UNWANTED	OCCUPIED	\
503	False	NaN	False	
1010	False	NaN	False	
1261	True	NaN	False	
1431	False	False	NaN	
2115	False	False	False	
2128	False	False	NaN	

	DEALER_NAME	DEALER_TYPE	CITY_ID	CITY	\
503	Marc Foujols Groupe Immobilier	AGENCY	54178039	Levallois-Perret	
1010	Loulimmo	AGENCY	54178039	Levallois-Perret	
1261	Consultants Immobilier	AGENCY	54178039	Levallois-Perret	
1431	NaN	PRIVATE	54178039	Levallois-Perret	
2115	MARC FOUJOLS GROUPE IMMOBILIER	PRIVATE	54178039	Levallois-Perret	
2128	NaN	PRIVATE	54178039	Levallois-Perret	

	ZIP_CODE	DEPT_CODE	PUBLICATION_START_DATE	LAST_CRAWL_DATE	\
503	92300.0	92	2019-04-15T12:48:07.996	2019-08-08T06:51:44.022	
1010	92300.0	92	2019-02-14T23:04:51.287	2019-08-23T17:59:32.752	
1261	92300.0	92	2019-02-09T00:53:48.028	2019-08-24T18:15:52.573	
1431	92300.0	92	2019-08-16T19:53:00	2019-08-26T00:39:04.642	
2115	92300.0	92	2019-03-18T18:23:08.584	2019-08-24T13:24:25.094	
2128	92300.0	92	2019-08-18T00:08:00	2019-08-25T03:56:42.436	

	LAST_PRICE_DECREASE_DATE
503	2019-05-11
1010	2019-05-09
1261	NaN
1431	NaN
2115	2019-05-11

```
[ ]: #Affichage des descriptions des annonces avec piscine
```

```
df_swim['DESCRIPTION'].apply(lambda x: print("\n new =",x))
```

new = En plein coeur de Levallois, quartier Villiers-Louise Michel, au calme, hôtel particulier d'exception au style contemporain et design. Il est composé de 6 pièces sur 3 niveaux avec ascenseur et offre de belles prestations piscine, Hammam, salle de cinéma, 3 terrasses, Patio et jardin privatif. 423 m<sup>2</sup> au sol (loi Carrez 289 m<sup>2</sup> + annexes 134 m<sup>2</sup>). Il comprend un grand séjour avec cuisine équipée, 1 suite parentale avec salle de bain, salle d'eau et dressing attenants, 4 chambres avec salles d'eau attenantes, WC séparés, buanderie. Dont 3.78 % honoraires TTC à la charge de l'acquéreur.

new = Hôtel Particulier contemporain sur 3 niveaux avec Ascenseur. 6 pièces avec piscine, hammam, salle de cinéma, 3 terrasses, balcon, Patio et jardin privatif. 423 m<sup>2</sup> au sol (Loi Carrez 289 m<sup>2</sup> + Annexe 134 m<sup>2</sup>) Grand séjour avec cuisine aménagée et équipée, 1 suite parentale avec salle de bain, sa salle d'eau et son Dressing, 4 chambres avec leur salle d'eau, WC séparés, buanderie. Très rare. Bien d'exception.

new = EXCLUSIVITÉ 2 PIÈCES LEVALLOIS PLACE JEAN ZAY Exclusivité Levallois Place Jean Zay. Appartement de 2 pièces exposé Ouest en bon état composé d'une pièce à vivre avec cuisine américaine, d'une chambre, d'une salle de douche avec toilettes refaite à neuf. Une cave complète ce bien. Faibles charges. Idéal Pied à terre ou Investisseur.

new = Charmant appartement de 37 m<sup>2</sup> refait à neuf, il est meublé et entièrement équipé, lumineux et agréable, dans une résidence calme et sécurisée. Disponible en début de mois prochain. Il dispose de: - une cuisine entièrement équipée (plaque à induction, four à pyrolyse, micro-ondes, lave-vaisselle, réfrigérateur/congélateur) ouverte sur le séjour. - un séjour avec un grand canapé d'angle, un fauteuil, un grand rangement avec penderie et étagères à chaussures et une table qui permet de prendre des repas à 4 personnes. - une chambre séparée comprend un lit double en 160x200 ainsi qu'un grand dressing - une salle de bain avec baignoire, toilettes et machine à laver séchante. Une place de parking privée se trouve au 2ème sous-sol accessible par ascenseur depuis l'intérieur de l'immeuble et par badge sécurisé depuis l'extérieur. Loyer: 1320 euros + 130 euros de charges (appartement et parking sécurisé) Dépôt de garantie: 2 mois Chauffage électrique individuel. Visites possibles en fin de mois. Le quartier calme dispose de: - un accès à 4 minutes à pieds à la gare Clichy-levallois pour se rendre en moins de 10 minutes à la gare St Lazare - un accès à différents commerces (moyenne surface de proximité, boulangerie, épicerie, pharmacie, bureau de presse et de tabac..) - un marché tous les mercredis et samedi matin sur la place Jean Zay (primeurs, poissonnier, boucher,

charcutier, fleuriste..) - le centre nautique de Levallois à 3 minutes à pied (piscines, sauna, hammam..) - les quais de Seine à 5 minutes à pieds offrant pelouses, aires de jeux et de repos et un accès à l'île de la Jatte. - l'accès aux restaurants et boutiques de Paris et Levallois en moins de 10 minutes. Me contacter par mail en indiquant: projet immobilier + garanties financières apportées + coordonnées + questions éventuelles. Garantie : 2640 €. Disponible le 01/09/2019.

new = En plein coeur de Levallois, quartier Villiers-Louise Michel, au calme, hôtel particulier d'exception au style contemporain et design. Il se compose de 6 pièces sur 3 niveaux avec ascenseur et offre de belles prestations telles que piscine, Hammam, salle de cinéma, 3 terrasses, patio et jardin privatif. 423 m<sup>2</sup> au sol (loi Carrez 289 m<sup>2</sup> + annexes 134 m<sup>2</sup>). Il comprend un grand séjour avec cuisine équipée, une suite parentale avec salle de bain, salle d'eau et dressing attenants, 4 chambres avec leur salle d'eau attenante, des WMC séparés, une buanderie. Contact: Nathalie Millereau + 33 (0) 6 75 04 25 02 dont 3.78 % honoraires TTC à la charge de l'acquéreur.

new = Bonjour, Location d'un studio 26 m<sup>2</sup>, très clair. Dans le centre ville de Levallois Perret (rue Louis Rouquier), à proximité des commerces de salle de sport/piscine municipale/centre commercial So Ouest.. Et situé entre métro 3 Louise Michel et Anatole France et à moins de 10mn pour la Ligne L Clichy Levallois > St Lazare. Situé au 3ème étage d'un immeuble de 6 étages avec ascenseur Grand salon: contenant un lit avec sommier plus un canapé convertible et plusieurs rangements (placard buffet et armoire) - cuisine avec table murale pour manger, micro-onde, plaque de cuisson, réfrigérateur, deux chaises hautes, vaisselles..) Cuisine et salle de bains indépendantes au salon. Loyer: 950 Euro (s) charges comprises: 800e loyer + 150e charges (chauffage collectif inclus dans les charges). Disponible: immédiatement Je suis plus réactif au messages reçu à travers le site. Garantie : 1600 €.

```
[ ]: 503      None
      1010     None
      1261     None
      1431     None
      2115     None
      2128     None
      Name: DESCRIPTION, dtype: object
```

Par les annonces avec piscine nous remarquons clairement que l'annonce de l'index 503 et 2115 sont les mêmes car les valeurs de leur variables sont similiaire d'ou ils sont doublons et doivent être dans un même cluster dans notre regroupement.

### 0.1.5 Traitement de données

Dans ce partie, nous allons soigner nos données dans un format propre pour être accepté par notre model et pour avoir des resultats souhaitable. Mais nous allons traiter la colonne la plus importante "DESCRIPTION" separement car elle contient des textes.

```
[ ]: data = df.copy()
```

```
[ ]: # colonnes à supprimer
col_del = ['DEPOSIT', 'FEES', 'CITY_ID', 'CITY', 'ZIP_CODE', 'DEPT_CODE', 'ID',
           ↪ 'DESCRIPTION', 'URL',
           ↪ 'PRICE_EVENTS', 'LAST_PRICE_DECREASE_DATE', 'IMAGES',
           ↪ 'LAND_SURFACE', 'TERRACE_SURFACE', 'PARKING_COUNT',
           ↪ 'RENTAL_EXPENSES', 'FLOOR_COUNT', 'FLOOR']

data = df.drop(col_del, axis=1)
```

Nous avons supprimés les variables dont nous avons juger non importantes pour notre algorithme, et dont nous avons remarquer ayant extremement beaucoup des valeurs manquantes

```
[ ]: # Les variables categorielles
categorical_cols = ['CRAWL_SOURCE', 'PROPERTY_TYPE', 'NEW_BUILD', 'FURNISHED',
                   ↪ 'PARKING', 'HEATING_MODE', 'ELEVATOR', 'CARETAKER', 'MARKETING_TYPE',
                   ↪ 'RENTAL_EXPENSES_INCLUDED',
                   ↪ 'FEES_INCLUDED', 'EXCLUSIVE_MANDATE', 'AGENCIES_UNWANTED',
                   ↪ 'OCCUPIED', 'DEALER_TYPE', 'DEALER_NAME', 'CONSTRUCTION_YEAR',
                   ↪ 'HEATING_TYPES']
categorical_label_encoders = {}

# La conversion des valeurs en lettre en nombre
for col in categorical_cols:
    categorical_label_encoders[col] = LabelEncoder()
    data[col] = categorical_label_encoders[col].fit_transform(data[col])
```

Nous avons convertir les variables qui étaient en caractère en nombre pour qu'elles soient accepter par l'algorithme. Nous avons utiliser le Label encoder car ces variables avaient tous des variables nominales.

Comme Nous avons des colonnes en date que nous trouvons importantes, nous l'avons converti en temps unix. Le temps unix, est une mesure du temps basée sur le nombre de secondes écoulées depuis le 1er janvier 1970 00:00:00 UTC, hors secondes intercalaires.

```
[ ]: date_cols = ['PUBLICATION_START_DATE', 'LAST_CRAWL_DATE']
def convert_to_timestamp(date_string):
    if isinstance(date_string, str):
        dt = pd.to_datetime(date_string)
        timestamp = dt.value // 10**9
        return timestamp
    return date_string

data[date_cols] = data[date_cols].applymap(convert_to_timestamp)
```

```
[ ]: data.head()
```

```

[ ]:  CRAWL_SOURCE  PROPERTY_TYPE  NEW_BUILD  SURFACE  ROOM_COUNT  BEDROOM_COUNT  \
0      0      0      0      72.0      3.0      NaN
1      2      0      0      48.0      2.0      1.0
2      1      0      0     267.0      6.0      NaN
3     11      4      0      50.0      0.0      NaN
4      1      2      0     330.0      8.0      NaN

      FURNISHED  GARDEN  SWIMMING_POOL  TERRACE  PARKING  HEATING_TYPES  \
0      0      NaN      NaN      NaN      1      4
1      2      1.0      NaN      NaN      0      4
2      2      NaN      NaN      NaN      0      4
3      2      NaN      NaN      NaN      0      4
4      2      NaN      NaN      NaN      0      4

      HEATING_MODE  CONSTRUCTION_YEAR  ELEVATOR  CARETAKER  MARKETING_TYPE  \
0      0      48      1      2      0
1      1      48      1      0      1
2      3      48      2      2      1
3      3      48      2      2      0
4      3      48      2      2      1

      PRICE  PRICE_M2  RENTAL_EXPENSES_INCLUDED  FEES_INCLUDED  \
0     1979.0     27.49      1      2
1    429000.0    8937.50      2      0
2   4970000.0   18614.23      2      2
3      89.0      1.78      2      2
4   2790000.0    8454.55      2      2

      EXCLUSIVE_MANDATE  AGENCIES_UNWANTED  OCCUPIED  DEALER_NAME  DEALER_TYPE  \
0      0      2      2      297      0
1      0      2      0      422      1
2      0      2      0      423      0
3      0      2      2      269      0
4      0      2      0      164      0

      PUBLICATION_START_DATE  LAST_CRAWL_DATE
0      1490155676      1492800755
1      1537268684      1559815690
2      1550594943      1560414134
3      1530890339      1560854407
4      1554743540      1560524538

```

**Traitement des valeurs manquantes** Nous avons resolu le problèmes des valeurs manquantes des deux facons: \* Pour les variables “piscine”, “jardin”, et “terrace”, nous avons remplacer les valeurs manquantes par 0 car nous pensons que ses informations ne soit pas ecrites à l’annonce ce que le bien immobilier n’a pas une piscine ou un jardin ou une terrasse d’où pour ce bien là on mettra ces variables à 0

- Pour certaines variables qui sont déjà en nombre et que nous trouvons importantes mais qui manquent des valeurs, nous avons remplacé les valeurs manquantes par la median de la variable. Car la mediane represente bien la valeur du milieu des observations.

```
[ ]: col_fill = ['SWIMMING_POOL', 'GARDEN', 'TERRACE']
col_fill_med = ['SURFACE', 'ROOM_COUNT', 'BEDROOM_COUNT', 'PRICE_M2', 'PRICE']

for col in col_fill:
    data[col].fillna(0, inplace=True)

for col in col_fill_med:
    data[col].fillna(data[col].median(), inplace=True)

data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2164 entries, 0 to 2163
Data columns (total 28 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   CRAWL_SOURCE                          2164 non-null   int64
1   PROPERTY_TYPE                        2164 non-null   int64
2   NEW_BUILD                            2164 non-null   int64
3   SURFACE                              2164 non-null   float64
4   ROOM_COUNT                           2164 non-null   float64
5   BEDROOM_COUNT                        2164 non-null   float64
6   FURNISHED                            2164 non-null   int64
7   GARDEN                               2164 non-null   float64
8   SWIMMING_POOL                        2164 non-null   float64
9   TERRACE                             2164 non-null   float64
10  PARKING                              2164 non-null   int64
11  HEATING_TYPES                        2164 non-null   int64
12  HEATING_MODE                         2164 non-null   int64
13  CONSTRUCTION_YEAR                   2164 non-null   int64
14  ELEVATOR                             2164 non-null   int64
15  CARETAKER                           2164 non-null   int64
16  MARKETING_TYPE                      2164 non-null   int64
17  PRICE                               2164 non-null   float64
18  PRICE_M2                            2164 non-null   float64
19  RENTAL_EXPENSES_INCLUDED            2164 non-null   int64
20  FEES_INCLUDED                       2164 non-null   int64
21  EXCLUSIVE_MANDATE                   2164 non-null   int64
22  AGENCIES_UNWANTED                   2164 non-null   int64
23  OCCUPIED                            2164 non-null   int64
24  DEALER_NAME                         2164 non-null   int64
25  DEALER_TYPE                         2164 non-null   int64
26  PUBLICATION_START_DATE              2164 non-null   int64
27  LAST_CRAWL_DATE                     2164 non-null   int64
```

```
dtypes: float64(8), int64(20)
memory usage: 473.5 KB
```

```
[ ]:
```

**Traitement de la colonne “DESCRIPTION” avec NLTK** En utilisant la bibliothèque NLTK, nous avons commencé nous tirer quelques informations de la colonne “Description” telles que le nombre de token des mots pour chaque ligne de cette variable, le nombre des mots vides(stop words) et le nombre de mots clé regroupant les adjectifs, noms, verbes, etc.

Nous avons pas voulu utiliser d’autre informations comme la longueur des mots ou des phrases et d’autre operation de Nltk car nous pensons que ça sera pas utile car par exemple deux annonces peuvent avoir une même longueur ou les mêmes nombre d’adjectifs et verbes mais sont totalement deux differentes annonces.

```
[ ]: tokenizer = AutoTokenizer.from_pretrained("microsoft/deberta-v3-base")

df_ = pd.DataFrame()

STOP_WORDS = set(stopwords.words('french'))

def extract_keywords(text):
    tokens = word_tokenize(str(text))
    pos_tags = pos_tag(tokens)

    # Define relevant parts of speech for keywords (e.g., nouns, adjectives)
    relevant_pos = ['NN', 'NNS', 'NNP', 'NNPS', 'JJ', 'JJR', 'JJS']

    keywords = [word for word, pos in pos_tags if pos in relevant_pos]

    return keywords

def count_stopwords(text: str) -> int:
    stopword_list = set(stopwords.words('french'))
    words = str(text).split()
    stopwords_count = sum(1 for word in words if word.lower() in stopword_list)
    return stopwords_count

# This function applies all the above preprocessing functions on a text feature

def run(df: pd.DataFrame) -> pd.DataFrame:

    df_["desc_word_len"] = df["DESCRIPTION"].apply(lambda x: len(word_tokenize(str(x))))

    df_["desc_word_cnt"] = df["DESCRIPTION"].apply(lambda x: count_stopwords(x))
```

```

# Automatically extract keywords from the prompt
df_['desc_keywords'] = df['DESCRIPTION'].apply(lambda x:
↳len(extract_keywords(x)))

return df_

```

```

tokenizer_config.json: 0%|          | 0.00/52.0 [00:00<?, ?B/s]
config.json: 0%|          | 0.00/579 [00:00<?, ?B/s]
spm.model: 0%|          | 0.00/2.46M [00:00<?, ?B/s]
/usr/local/lib/python3.10/dist-
packages/transformers/convert_slow_tokenizer.py:473: UserWarning:

```

The sentencepiece tokenizer that you are converting to a fast tokenizer uses the byte fallback option which is not implemented in the fast tokenizers. In practice this means that the fast version of the tokenizer can produce unknown tokens whereas the sentencepiece version would have converted these unknown tokens into a sequence of byte tokens matching the original piece of text.

```

[ ]: result = run(df)
result.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2164 entries, 0 to 2163
Data columns (total 3 columns):
#   Column          Non-Null Count  Dtype
---  -
0   desc_word_len    2164 non-null   int64
1   desc_word_cnt    2164 non-null   int64
2   desc_keywords    2164 non-null   int64
dtypes: int64(3)
memory usage: 50.8 KB

```

```

[ ]: # Melange de nos données numeriques et les variables des premieres traitement
↳de la colonne DESCRIPTION
data =data.join(result)
data.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2164 entries, 0 to 2163
Data columns (total 31 columns):
#   Column          Non-Null Count  Dtype
---  -
0   CRAWL_SOURCE     2164 non-null   int64

```



1	PROPERTY_TYPE	2164	non-null	int64
2	NEW_BUILD	2164	non-null	int64
3	SURFACE	2164	non-null	float64
4	ROOM_COUNT	2164	non-null	float64
5	BEDROOM_COUNT	2164	non-null	float64
6	FURNISHED	2164	non-null	int64
7	GARDEN	2164	non-null	float64
8	SWIMMING_POOL	2164	non-null	float64
9	TERRACE	2164	non-null	float64
10	PARKING	2164	non-null	int64
11	HEATING_TYPES	2164	non-null	int64
12	HEATING_MODE	2164	non-null	int64
13	CONSTRUCTION_YEAR	2164	non-null	int64
14	ELEVATOR	2164	non-null	int64
15	CARETAKER	2164	non-null	int64
16	MARKETING_TYPE	2164	non-null	int64
17	PRICE	2164	non-null	float64
18	PRICE_M2	2164	non-null	float64
19	RENTAL_EXPENSES_INCLUDED	2164	non-null	int64
20	FEES_INCLUDED	2164	non-null	int64
21	EXCLUSIVE_MANDATE	2164	non-null	int64
22	AGENCIES_UNWANTED	2164	non-null	int64
23	OCCUPIED	2164	non-null	int64
24	DEALER_NAME	2164	non-null	int64
25	DEALER_TYPE	2164	non-null	int64
26	PUBLICATION_START_DATE	2164	non-null	int64
27	LAST_CRAWL_DATE	2164	non-null	int64
28	desc_word_len	2164	non-null	int64
29	desc_word_cnt	2164	non-null	int64
30	desc_keywords	2164	non-null	int64

dtypes: float64(8), int64(23)  
memory usage: 524.2 KB

### 0.1.6 Mise à l'échelle des données

Nous avons mis à l'échelle de nos données numériques pour une comparaison équitable des caractéristiques car si les caractéristiques de nos données ont des échelles très différentes, certaines caractéristiques pourraient dominer l'apprentissage du modèle par rapport aux autres.

```
[ ]: # scale numeric columns
numeric_cols = [col for col in data.columns]

from sklearn.compose import ColumnTransformer

# scale numeric columns and categorical columns
from sklearn.preprocessing import StandardScaler
scaler_transformer = ColumnTransformer(
    transformers=[
```

```

        ('Scaler', StandardScaler(), numeric_cols)
    ],
    remainder='passthrough'
)
new_cols = numeric_cols + [col for col in data.columns if col not in numeric_cols]
df_scaled = pd.DataFrame(scaler_transformer.fit_transform(data),
    columns=new_cols)

```

```
[ ]: df.shape, df_scaled.shape
```

```
[ ]: ((2164, 46), (2164, 31))
```

```
[ ]: #Affichage de nos données mises en echelle
df_scaled.head(10)
```

```
[ ]:
CRAWL_SOURCE  PROPERTY_TYPE  NEW_BUILD  SURFACE  ROOM_COUNT  \
0      -1.662557      -0.693198  -0.378610  -0.126821    0.501598
1      -1.168934      -0.693198  -0.378610  -0.185036   -0.130585
2      -1.415745      -0.693198  -0.378610   0.346180    2.398147
3       1.052371       1.620849  -0.378610  -0.180185   -1.394951
4      -1.415745       0.463825  -0.378610   0.498996    3.662513
5      -0.922122      -0.693198  -0.378610   0.044818    1.765964
6       1.052371       1.620849  -0.378610  -0.148651   -1.394951
7      -0.922122       1.042337   2.999998  -0.262657   -0.130585
8      -0.675310       1.042337   2.999998  -0.274785   -0.130585
9      -0.675310      -0.693198  -0.378610  -0.163205   -0.130585

BEDROOM_COUNT  FURNISHED  GARDEN  SWIMMING_POOL  TERRACE  PARKING  \
0      -0.368106  -2.465859  -0.080695   -0.052729  -0.203435  3.575394
1      -0.368106   0.492898  12.392394   -0.052729  -0.203435 -0.279689
2      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
3      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
4      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
5      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
6      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
7      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
8      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689
9      -0.368106   0.492898  -0.080695   -0.052729  -0.203435 -0.279689

HEATING_TYPES  HEATING_MODE  CONSTRUCTION_YEAR  ELEVATOR  CARETAKER  \
0       0.381856    -1.998750         0.485786  -1.473941   0.336513
1       0.381856    -1.120161         0.485786  -1.473941  -3.860686
2       0.381856     0.637018         0.485786   0.561546   0.336513
3       0.381856     0.637018         0.485786   0.561546   0.336513
4       0.381856     0.637018         0.485786   0.561546   0.336513
5       0.381856     0.637018         0.485786   0.561546   0.336513

```

6	0.381856	0.637018	0.485786	0.561546	0.336513
7	0.381856	0.637018	0.485786	0.561546	0.336513
8	0.381856	0.637018	0.485786	0.561546	0.336513
9	0.381856	0.637018	0.485786	0.561546	0.336513

	MARKETING_TYPE	PRICE	PRICE_M2	RENTAL_EXPENSES_INCLUDED	\
0	-1.135758	-0.653427	-1.082623	-1.501738	
1	0.880469	0.013431	0.632448	0.605759	
2	0.880469	7.104895	2.495104	0.605759	
3	-1.135758	-0.656379	-1.087572	0.605759	
4	0.880469	3.700492	0.539486	0.605759	
5	0.880469	1.561029	0.826719	0.605759	
6	-1.135758	-0.653800	-1.082598	0.605759	
7	-1.135758	-0.656424	-1.087193	-1.501738	
8	-1.135758	-0.656362	-1.086165	-1.501738	
9	0.880469	0.233624	0.836967	0.605759	

	FEES_INCLUDED	EXCLUSIVE_MANDATE	AGENCIES_UNWANTED	OCCUPIED	DEALER_NAME	\
0	0.532470	-0.289181	0.317393	1.119918	0.292692	
1	-2.095272	-0.289181	0.317393	-0.901352	1.102483	
2	0.532470	-0.289181	0.317393	-0.901352	1.108962	
3	0.532470	-0.289181	0.317393	1.119918	0.111299	
4	0.532470	-0.289181	0.317393	-0.901352	-0.568925	
5	-2.095272	-0.289181	0.317393	-0.901352	-1.067757	
6	0.532470	-0.289181	0.317393	1.119918	-0.983538	
7	0.532470	-0.289181	0.317393	1.119918	1.640184	
8	0.532470	-0.289181	-3.204856	1.119918	1.685533	
9	0.532470	-0.289181	0.317393	-0.901352	0.065951	

	DEALER_TYPE	PUBLICATION_START_DATE	LAST_CRAWL_DATE	desc_word_len	\
0	-0.390246	-4.227868	-7.235731	-0.417299	
1	1.262380	-1.196145	-0.422071	0.697967	
2	-0.390246	-0.338600	-0.361224	0.116772	
3	-0.390246	-1.606592	-0.316460	2.520091	
4	-0.390246	-0.071638	-0.349999	1.216330	
5	-0.390246	-2.192521	-0.121573	1.263453	
6	-0.390246	-4.028304	-0.121575	-0.605795	
7	-0.390246	-1.821135	-0.126086	0.116772	
8	2.915005	-0.214603	-0.096700	-1.328361	
9	-0.390246	0.295300	-0.094058	0.163896	

	desc_word_cnt	desc_keywords
0	-0.621698	-0.389521
1	0.835014	0.682765
2	0.457348	0.173429
3	3.640532	2.478844
4	1.698250	1.486979

5	0.996871	0.897222
6	-0.999364	-0.603979
7	0.025730	0.334272
8	-1.269125	-1.327772
9	0.619205	0.414693

Nous avons maintenant nos variables importantes en proper format et sans valeurs manquantes.

Nous procedons maintenant au traitement approfondie de la variable “Description”

### 0.1.7 Traitement de la colonne “DESCRIPTION” avec l’embedding

La variable “DESCRIPTION” est la variable qui aura la grande impacte dans l’entrainement de notre algorithme.

Pour mieux traiter la variable qui contiennent les description, nous devons pas ce filier au nombre des mots ou des mots vides(stopword) ou d’autres methodes des bases mais nous devons étudier et trouver les valeurs de similarité entre les descriptions des annonces. Pour cela nous avons utilisées deux methodes et faire de prediction avec chacune des méthodes pour evaluer notre regroupement des annonces similaires:

1. Méthode avec l’encodage du text par le modèle d’integration(embedding):
2. Méthode avec la similarité cosinus(cosine similarity) du text

Nous avons utilisés pour ce deux méthodes le model pre-entrainé “**Sentence-CamemBERT-Large**” Sentence-CamemBERT-Large est le modèle d’intégration pour le français développé par La Javaness. Le but de ce modèle d’intégration est de représenter le contenu et la sémantique d’une phrase française dans un vecteur mathématique qui lui permet de comprendre la signification du texte au-delà des mots individuels dans les requêtes et les documents, offrant ainsi une recherche sémantique puissante. Vous pouvez retrouver ce modèle au site d’hugging face

**1)Traitement des “descriptions” avec l’integration(embedding)** L’idée derrière l’embedding est de projeter des données textuelles dans un espace vectoriel où des similarités sémantiques sont préservées. En d’autres termes, des éléments similaires auront des représentations vectorielles similaires. Cette technique est largement utilisée pour résoudre des tâches NLP telles que la classification de texte, la recherche d’informations, la traduction automatique, la recommandation de contenu, etc.

Les embeddings sont souvent créés en utilisant des modèles d’apprentissage profond, tels que des réseaux de neurones, qui apprennent à représenter les données textuelles de manière à capturer les relations sémantiques entre les mots. Certains des embeddings les plus populaires incluent Word2Vec, GloVe (Global Vectors for Word Representation) et les embeddings issus de modèles de langage pré-entraînés comme BERT, GPT-3, etc.

Nous utilisons un modèle d’intégration(embedding) au lieu de tfidf, word2vec pour gérer efficacement d’autres annonces ne figurant pas dans les données d’entrainement de notre algorithme, pour avoir des meilleures performances, et la gestion des mots hors vocabulaire.

Le modèle d’intégration prend un maximum de 512 jetons et le code dans un vecteur de 768. Nous utilisons un grand modèle pour de meilleures performances, mais il existe des alternatives plus rapides pour des résultats légèrement moins bons.

Nous

```
[ ]: from sentence_transformers import SentenceTransformer, util
import umap

# utilisation du GPU si disponible pour l'exécution rapide du modèle mais si ce
# n'est pas le cas le modèle s'exécutera au CPU comme d'habitude

device = 'cuda' if torch.cuda.is_available() else 'cpu'
model = SentenceTransformer('dangvantuan/sentence-camembert-large')
model.to(device)
if device == 'cuda':
    model = model.half()

# Ajouter la colonne description en spécifiant que l'ajout de 500 caractères
# chaque fois car le modèle attribut que 514 jetons à la fois
max_desc_len = 500
description_embeddings = df['DESCRIPTION'].apply(lambda x: f"query: {str(x)[:
# max_desc_len]}").tolist()
description_embeddings = df['DESCRIPTION'].apply(lambda x: str(x)[:
# max_desc_len]).tolist()
description_embeddings = model.encode(description_embeddings,
# show_progress_bar=True, convert_to_tensor=False)

# faire une réduction de dimensionnalité en utilisant umap
desc_reducer = umap.UMAP(n_neighbors=2, n_components=100, metric='euclidean',
# verbose=True)

description_embeddings = desc_reducer.fit_transform(description_embeddings)

# faire un dataframe avec l'encodage des descriptions
df_embeddings = pd.DataFrame(description_embeddings)
```

```
.gitattributes: 0%|          | 0.00/1.23k [00:00<?, ?B/s]
README.md: 0%|          | 0.00/5.25k [00:00<?, ?B/s]
config.json: 0%|          | 0.00/683 [00:00<?, ?B/s]
model.safetensors: 0%|          | 0.00/1.35G [00:00<?, ?B/s]
pytorch_model.bin: 0%|          | 0.00/1.35G [00:00<?, ?B/s]
sentence_bert_config.json: 0%|          | 0.00/27.0 [00:00<?, ?B/s]
sentencepiece.bpe.model: 0%|          | 0.00/809k [00:00<?, ?B/s]
special_tokens_map.json: 0%|          | 0.00/298 [00:00<?, ?B/s]
```

```

tokenizer_config.json: 0%|          | 0.00/400 [00:00<?, ?B/s]

WARNING:sentence_transformers.SentenceTransformer:No sentence-transformers model
found with name /root/.cache/torch/sentence_transformers/dangvantuan_sentence-
camembert-large. Creating a new one with MEAN pooling.

Batches: 0%|          | 0/68 [00:00<?, ?it/s]

UMAP(n_components=100, n_neighbors=2, verbose=True)
Wed Dec 13 19:05:10 2023 Construct fuzzy simplicial set
Wed Dec 13 19:05:15 2023 Finding Nearest Neighbors
Wed Dec 13 19:05:18 2023 Finished Nearest Neighbor Search
Wed Dec 13 19:05:22 2023 Construct embedding

Epochs completed: 0%|          | 0/500 [00:00]
    completed 0 / 500 epochs
    completed 50 / 500 epochs
    completed 100 / 500 epochs
    completed 150 / 500 epochs
    completed 200 / 500 epochs
    completed 250 / 500 epochs
    completed 300 / 500 epochs
    completed 350 / 500 epochs
    completed 400 / 500 epochs
    completed 450 / 500 epochs
Wed Dec 13 19:05:30 2023 Finished embedding

```

```
[ ]: # voir les nombres des lignes et variables qu'a produit notre modèle
df_embeddings.shape
```

```
[ ]: (2164, 100)
```

## 0.2 Apprentissage du modèle

Nous avons remarqué que dans nos données, nous avons pas une colonne cible par laquelle nous pouvons dire si nos données sont similaire ou pas. Mais nous devons chercher les éléments de similarité entre les annonces regrouper les annonces similaire, il s'agit ici d'un problème de l'apprentissage non supervisé(Unserpervised Learning) et de l'approche "clustering" de l'apprentissage non supervisé.

La méthode du regroupement(clustering) est une approche consistant à trouver les éléments de similarité entre les données afin de le regrouper ensemble. Utilisant cette technique, nous serons à mesure de regrouper les annonces similaires ensemble.

### 0.2.1 Train et test split

Nous melangeons nos données numeriques traités et nos données resultantes du traitement de la variable "DESCRIPTION"

Nous divisons l'ensemble de données en deux parties, à savoir l'ensemble d'entraînement et de test avec un rapport de 80 % pour l'ensemble d'entraînement et de 20 % pour l'ensemble de test.

Nous faisons cette division pour éviter le sur-ajustement(overfitting), c'est pourquoi nous divisons notre ensemble de données en divisions d'entraînement et de test, ce qui nous donne une meilleure idée de la façon dont notre algorithme va fonctionner pendant la phase d'entraînement et de test.

Nous utilisons la méthode `train_test_split()` de `sklearn` pour faire cette division et nous initialisons le `random_seed` à 42.

le `random_seed` est utilisé pour initialiser le générateur de nombres aléatoires interne, qui décidera de la division des données en indices de train et de test dans votre cas.

```
[ ]: from sklearn.model_selection import train_test_split

test_size = 0.2

df_with_all_features = pd.concat([df_scaled, df_embeddings], axis=1)
df_with_all_features['CRAWL_SOURCE_original'] = df['CRAWL_SOURCE']
df_with_all_features = pd.concat([df_with_all_features, df['DESCRIPTION']],
    ↪axis=1)
X_train_full, X_test_full = train_test_split(df_with_all_features,
    ↪test_size=test_size, random_state=42)
X_train, X_test = X_train_full.drop(col_del+['CRAWL_SOURCE_original'], axis=1,
    ↪errors='ignore'), X_test_full.drop(col_del+['CRAWL_SOURCE_original'],
    ↪axis=1, errors='ignore')
```

```
[ ]: # voir la forme de nos données apres divisions
X_train.shape, X_test.shape, X_train_full.shape, X_test_full.shape
```

```
[ ]: ((1731, 131), (433, 131), (1731, 133), (433, 133))
```

```
[ ]:
```

## 0.2.2 L'aglorithme de regroupement(clustering)

Il existe plusieurs algorithme de regroupement dans l'apprentissage automatique mais pour notre problème, nous avons choisi d'utiliser l'algorithme HDBSCAN.

HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) est un algorithme de regroupement (clustering) de données basé sur la densité. Il a été développé pour identifier des clusters dans des ensembles de données complexes et bruités, en particulier lorsque la forme, la taille et la densité des clusters peuvent varier.

HDBSCAN est une extension de l'algorithme DBSCAN (Density-Based Spatial Clustering of Applications with Noise), qui est également basé sur la densité. Cependant, HDBSCAN introduit une approche hiérarchique pour le regroupement, ce qui signifie qu'il peut identifier des clusters de différentes densités et hiérarchies dans les données. Il utilise un arbre de regroupement (clustering tree) pour représenter la hiérarchie des clusters.

Les avantages de HDBSCAN incluent sa capacité à détecter automatiquement le nombre de clusters dans les données, sa résistance au bruit et sa capacité à gérer des données de formes et de densités variées. Il est largement utilisé dans l'exploration de données, la visualisation de données et

l'analyse de clusters dans divers domaines, tels que l'apprentissage automatique, la bioinformatique, la géolocalisation, la vision par ordinateur, etc.

```
[ ]: import hdbscan
seed_everything(42)
#Nous avons initialiser l'algorithme en spécifiant certaines paramètres et en
↳laissant les autres par défaut.
# min_cluster_size: Le nombre minimum d'échantillons dans un groupe pour que ce
↳groupe soit considéré comme une grappe ; les groupements plus petits que
↳cette taille seront laissés comme bruit.
# alpha : Un paramètre d'échelle de distance utilisé dans une liaison simple
↳robuste.
# min_samples : Le nombre d'échantillons dans un quartier pour qu'un point soit
↳considéré comme un point central. Cela inclut le point lui-même

clusterer = hdbscan.HDBSCAN(min_cluster_size=2,
                             min_samples=1, # min_cluster_size by default
                             alpha=5.0, # 1.0 by default
                             prediction_data=True
                             )

# entraînement du modèle
clusterer.fit(X_train)
```

```
[ ]: HDBSCAN(alpha=5.0, min_cluster_size=2, min_samples=1, prediction_data=True)
```

```
[ ]:
```

### 0.2.3 Visualisation de nos résultats de regroupement

Nous avons utilisé la méthode de réduction de dimensionnalité “umap” pour nous permettre de visualiser nos regroupements sur une dimension 2D que sur la dimensionnalité de nos variables.

```
[ ]: # Utilisation d'Umap
seed_everything(42)
reducer = umap.UMAP(n_neighbors=6, n_components=2)
X_train_pca = reducer.fit_transform(X_train)

[ ]: # plot clusters

# tronquer la description à 100 caractères
X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION'].str[:600]
X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION_wrapped'].str.
↳wrap(100)
X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION_wrapped'].
↳apply(lambda x: str(x).replace('\n', '<br>'))

fig = go.Figure()
```



```

fig.add_trace(go.Scatter(
    x=X_train_pca[:, 0],
    y=X_train_pca[:, 1],
    mode='markers',
    marker=dict(
        size=5,
        color=clusterer.labels_,
    ),
    text=X_train.index,
    customdata=X_train_full['DESCRIPTION_wrapped'],
    hovertemplate='Description: %{customdata}<br>Index: %{text}<br>Cluster:␣
↳%{marker.color}<extra></extra>'
))

# increase size of the plot
fig.update_layout(
    title='Clusters',
    xaxis_title='PCA 1',
    yaxis_title='PCA 2',
    legend_title='Cluster',
    font=dict(
        family='Courier New, monospace',
        size=18,
        color='#7f7f7f'
    ),
    width=1200,
    height=800
)
X_train_full.drop('DESCRIPTION_wrapped', axis=1, inplace=True)

fig.show()

```

[ ]:

Nous pouvons aller au dessus de chaque point de notre visualisation et nous verrons la description et le cluster qu'appartient l'annonce.

Comme la visualisation n'est pas assez visible, nous allons l'afficher sous forme tabulaire

## 0.2.4 Ajout de la colonne de regroupement

Nous allons ajouter la colonne "cluster" de regroupement des annonces dans nos données d'entraînement pour nous permettre de bien voir les annonces en fonction de regroupement et pour pouvoir évaluer notre algorithme

```
[ ]: def add_cluster_column(clusterer, X_train_):
    X_train = copy.deepcopy(X_train_).reset_index(drop=True)
    labels = clusterer.labels_
    n_clusters_ = len(set(labels)) - (1 if -1 in labels else 0)
    print('Estimated number of clusters: %d' % n_clusters_)
    print('Number of outliers/unique ads: %d' % list(labels).count(-1))

    # Affichage du nombre maximum du cluster ayant le plus des annonces
    print('Max cluster: %d' % np.sort(np.unique(labels,
↪return_counts=True)[1])[:-1][1]) # ignore -1, car -1 represente les valeurs
↪aberrantes d'ou des annonces considerés uniques

    # Affichage de la moyenne des annonces dans des clusters
    print('Mean cluster: %d' % np.unique(labels, return_counts=True)[1].mean())

    # Affichage du nombre minimum
    print('Min cluster: %d' % np.unique(labels, return_counts=True)[1].min())

    # Revoir les réels valeurs de nos données avant la mise en echelle
    _X_train = pd.DataFrame(scaler_transformer.named_transformers_['Scaler'].
↪inverse_transform(X_train[numeric_cols]), columns=numeric_cols)
    X_train[numeric_cols] = _X_train[numeric_cols]

    # ajout de la colonne regroupement
    X_train['CLUSTER'] = labels

    # créez une colonne contenant la distance par rapport au centre de gravité
↪du cluster. -1 distance signifie que le point est une valeur aberrante
    X_train['CLUSTER_DISTANCE'] = clusterer.outlier_scores_
    return X_train

X_train_full_clustered = add_cluster_column(clusterer, X_train_full)
```

```
Estimated number of clusters: 226
Number of outliers/unique ads: 1131
Max cluster: 15
Mean cluster: 7
Min cluster: 2
```

Nous voyons que notre modèle a réussi de faire 226 groupements des annonces similaires et à classer 1131 annonces comme étant uniques.

Nous pouvons aussi remarquer que le plus gros groupe a 15 annonces similaires, le plus petit groupe en a 2 et la moyenne des annonces dans les groupes est de 7

```
[ ]: #cheque les plateformes source des annonces
df['CRAWL_SOURCE'].value_counts()
```

```
[ ]: SE_LOGGER      877
LE_BON_COIN      246
A_VENDRE_A_LOUER  208
EXPLORIMMO      196
MEILLEURS_AGENTS  193
BIEN_ICI        178
LOGIC_IMMO      137
PAP             49
PARU_VENDU       31
PROPRIETES_DE_FRANCE  21
BELLES_DEMEURES  19
LUX_RESIDENCE    7
SURFACE_PRIVEE   2
Name: CRAWL_SOURCE, dtype: int64
```

Nous avons checker les valeurs uniques de sources des annonces c'est 12, ce qui pourrait dire normalement qu'on devait pas avoir un cluster avec plus des 12 annonces similaire mais cela devrait être le cas si dans chaque platform par exemple le bon coin chaque utilisateur peut creer q'un seul compte et publier l'annonce du bien qu'une fois mais defois la personne ou plusieurs personnes peut publier l'annonce du même bien immobilier 2 fois voir meme 3 fois si les membres de sa famille le publie aussi ou bien si lui meme fait une deuxième annonce du meme bien en changeant ou en rajoutant quelque mots à la description. Notre modèle est capable de detecter ses genres d'annonces, c'est pourquoi nous avons de cluster avec plus de 12 annonces indiquant que c'est le même bien immobilier mais plusieurs fois publiés.

### 0.2.5 Afficher nos resultats en fonction des clusters

```
[ ]: # voir nos resultas en groupant par les regroupements des annoces
X_train_full_clustered.sort_values(by='CLUSTER', ascending=False)[['CLUSTER',
↪ 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'SURFACE',
↪ 'CRAWL_SOURCE_original']].head(50)
```

```
[ ]:      CLUSTER  CLUSTER_DISTANCE  \
658      225      0.000000
911      225      0.390353
1475     225      0.394520
551      225      0.407556
654      225      0.000000
27       225      0.000000
1595     225      0.000000
753      224      0.000000
492      224      0.000000
780      223      0.000000
950      223      0.000000
```

1305	222	0.000000
1580	222	0.000000
684	221	0.266986
1530	221	0.000000
1255	221	0.266986
584	221	0.000000
1542	220	0.000000
1636	220	0.196617
1008	220	0.000000
850	220	0.000000
78	220	0.000000
161	220	0.135229
749	219	0.325555
655	219	0.000000
937	219	0.325555
1158	219	0.000000
83	218	0.000000
1027	218	0.000000
1004	217	0.000000
637	217	0.000000
1014	216	0.000000
1502	216	0.406694
528	216	0.204884
951	216	0.000000
619	216	0.351104
967	215	0.000000
1155	215	0.235487
980	215	0.110958
1267	215	0.235487
766	215	0.000000
960	214	0.000000
614	214	0.000000
386	213	0.658768
859	213	0.000000
1466	213	0.000000
854	212	0.300081
287	212	0.000000
151	212	0.000000
319	211	0.000000

#### DESCRIPTION \

658

URGENT A LOUER PARKING 96 RUE ARISTIDE BRIAN\npremier sous sol avec bip pour  
petite voiture

911

Nous mettons en location une place de parking en parfait état situé au 2ème  
sous-sol d'un immeuble moderne à proximité du métro Pont de Levallois (rue

Marius Aufan, proche rue Paul Vaillant-Couturier).\n\nLarge rampe d'accès.\nTaille standard, pour tout type de véhicule.\nRésidence sécurisée, porte automatisée (bip).\n\nDisponible 1er septembre.

1475

LOUE place de parking dans résidence de standing sécurisée.\nAccès facile avec badge.\nLoyer mensuel : 100 € toutes charges comprises.\nDisponibilité immédiate.

551

Bonjour,\n\nParticulier loue place de stationnement dans le parking souterrain d'une résidence proche de la station de métro Pont de Levallois.\n\nDisponible immédiatement.

654

place de parking au 77 rue P.V. Couturier \nen sous-sol avec accès par télécommande \nconvient pour une citadine ou une compacte

27

DISPONIBLE FIN AOÛT\nLoue grande place de parking rue camille pelletan.\nAccessible via une rampe très très large.\nEtage -2.

1595

Location d'une place de parking spacieuse dans un immeuble standing.\n(1er sous-sol, facile d'accès avec bip, vidéo-surveillance.)\nSitué à l'angle rue Chaptal et rue Kléber.\nDisponible de suite.

753

BOXE FERME EN 1ER SOUS SOL\n\nDans une résidence sécurisée, BOX SIMPLE, situé au 11 Greffulhe, à proximité du métro ligne 3: Pont de Levallois-Bécon\nRéférence annonce : LEVALBSL\nNon soumis au DPE

492

PLACE DE STATIONNEMENT EN 1ER SOUS SOL\n\nDans une résidence sécurisée, PLACE DE STATIONNEMENT située au 1er sous-sol, située au 11 rue Greffulhe, à proximité du métro ligne 3 : Pont de Levallois-Bécon.\nRéférence annonce : LEVALPSL\nNon soumis au DPE

780

2 BIS, RUE BAUDIN - METRO PONT DE LEVALLOIS - EMPLACEMENT DE PARKING AU 2° SOUS SOL - ACCES PAR RAMPE. LOYER 150 EUROS. HONORAIRES A LA CHARGE DU LOCATAIRE 250 EUROS TTC. DEPOT DE GARANTIE 300 EUROS .

950

2 BIS, RUE BAUDIN - METRO PONT DE LEVALLOIS - EMPLACEMENT DE PARKING AU 2° SOUS SOL - ACCES PAR RAMPE. LOYER 150 EUROS. HONORAIRES A LA CHARGE DU LOCATAIRE 250 EUROS TTC. DEPOT DE GARANTIE 300 EUROS .

1305

Cession de bail 92300 Levallois-Perret Le groupe POINT DE VENTE vous propose à proximité du métro Anatole France : un commerce idéal pour Restauration sans nuisance ou alimentaire d'une surface totale de 73 m<sup>2</sup> - Au rez de chaussée, une surface de 53 m<sup>2</sup>, avec point d'eau, WC. - Au sous-sol, une surface de 20 m<sup>2</sup>.  
Prix : 65 000 € Loyer : 1 100 € /HC /HT /MOIS Honoraires : 7 920 € H.T Ref  
annonce : P1-30948 Agence : Point de vente 01 77 39 27 91 Conseil en immobilier commercial, Point de vente vous accompagne dans vos recherches de fonds de commerce, droit au bail, locations pures, murs de boutiques. Mise à jour le

27/08/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE :  
P1-30948

1580

Cession de bail 92300 Levallois-Perret Le groupe POINT DE VENTE vous propose un commerce d'une surface totale de 55 m<sup>2</sup> : Au rez-de-chaussée, une surface de 38 m<sup>2</sup>. Au sous-sol, une surface de 17 m<sup>2</sup>, accessible depuis escalier intérieur. L'emplacement est idéal pour tous types d'activités dont restauration froide sans nuisances. Ce local a un point d'eau, WC, Prix : 132 000 € Loyer : 2 004, 20 € /HC /HT /MOIS Honoraires : 13 200 € H.T Ref annonce : P1-31945 Agence : Point de vente 01 77 39 27 91 Conseil en immobilier commercial, Point de vente vous accompagne dans vos recherches de fonds de commerce, droit au bail, locations pures, murs de boutiques. Mise à jour le 16/08/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : P1-31945

684

Le groupe POINT DE VENTE vous propose dans l'un des meilleurs emplacements de Levallois-Perret un commerce d'une surface totale de 47 m<sup>2</sup>: Au rez-de-chaussée, une surface de 47 m<sup>2</sup> avec point d'eau et WC. Local idéal pour du prêt-à-porter, une agence immobilière, un cabinet d'assurances ou de la restauration sans nuisance.

1530

Le groupe POINT DE VENTE vous propose à proximité du métro Louise Michel: un commerce d'une surface totale de 69 m<sup>2</sup> - Au rez-de-chaussée, une surface de 32 m<sup>2</sup>, avec point d'eau, WC. - Au premier étage, une surface de 25 m<sup>2</sup>, accessible depuis escalier intérieur. Honoraires TTC en sus : 16200.00 €.

1255

Le groupe POINT DE VENTE vous propose un commerce d'une surface totale de 55 m<sup>2</sup>: Au rez-de-chaussée, une surface de 38 m<sup>2</sup>. Au sous-sol, une surface de 17 m<sup>2</sup>, accessible depuis escalier intérieur. L'emplacement est idéal pour tous types d'activités dont restauration froide sans nuisances. Ce local a un point d'eau, WC.

584

Le groupe POINT DE VENTE vous propose à proximité du métro Anatole France: un commerce idéal pour Restauration sans nuisance ou alimentaire d'une surface totale de 73 m<sup>2</sup> - Au rez-de-chaussée, une surface de 53 m<sup>2</sup>, avec point d'eau, WC. - Au sous-sol, une surface de 20 m<sup>2</sup>.

1542 Louez au mois un parking privé au 2 rue Maurice Ravel à Levallois-Perret pour 146€. Aucun engagement ni caution requis pour un abonnement au mois dans ce parking souterrain. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Présence d'un ascenseur : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-1407-1 Trouver un garage à louer n'aura jamais été aussi simple. Quelle que soit la taille de votre voiture neuve ou d'occasion, les emplacements sont suffisamment grands pour vous permettre de manoeuvrer sans difficulté. Tous nos parkings sont surveillés par un gardien et propres, il est donc très rassurant d'y laisser sa

voiture.

1636

Parking à louer à levallois-perret 2 rue maurice ravel, levallois perret, Location de parking privé à 60 par mois au 2 rue Maurice Ravel à Levallois-Perret. Parking souterrain sans engagement et sans caution. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Présence d'un ascenseur : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Pour louer une place : rendez-vous sur le site internet avec la référence 92300-1495-1Attention, ce parking ne peut accueillir que des deux-roues : scooter ou moto Ascenseur 1008

Parking à louer à Levallois-Perret\n\nPlace de parking à louer au 2 rue Maurice Ravel à Levallois-Perret à 60€ par mois.\nParking souterrain sans engagement et sans caution. 2 journées d'essai gratuit pour tester le parking avant l'abonnement.\nParking sécurisé : conforme à la charte qualité Yespark.\nPrésence de caméras de surveillance : Oui.\nPrésence d'un gardien : Oui.\nPrésence d'un ascenseur : Oui.\nService client : 24h/24.\nOuverture sans badge : Oui.\nHauteur maximum : 1,90m\nPour louer une place : rendez-vous sur le site internet avec la référence 92300-1495-1Attention, ce parking ne peut accueillir que des deux-roues : scooter ou moto\nRéférence annonce : 92300-1495-1\nNon soumis au DPE

850

Louez au mois un parking privé au 22 rue Thierry Le Luron à Levallois-Perret pour 201€. Aucun engagement ni caution requis pour un abonnement au mois dans ce parking souterrain. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Présence d'un ascenseur : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Pour essayer ce parking retrouvez toutes les informations utiles sur le site internet Yespark avec la référence : 92300-85-1 La location d'une place de parking sécurisée par la présence d'un gardien n'aura jamais été aussi facile. Vous tenez à protéger votre véhicule neuf ou d'occasion ? Louez une place dans un de nos parkings où les places sont larges et où il est aisé de se garer même avec un grand véhicule.

78

Place de parking à louer au 2 rue Maurice Ravel à Levallois-Perret à 60€ par mois. Parking souterrain sans engagement et sans caution. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Présence d'un ascenseur : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-1495-1Attention, ce parking ne peut accueillir que des deux-roues : scooter ou moto. Mise à jour le 03/08/2019 Demander des infos sur ce bien Référence annonce AVENDREALOVER : 92300-1495-1

161

Location de parking privé à 131€ par mois au 2 rue Pablo Neruda à Levallois-Perret. Parking souterrain sans engagement et sans caution. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Parking sécurisé : conforme à la charte qualité Yespark. Présence de caméras de surveillance : Oui. Présence d'un gardien : Oui. Service client : 24h/24. Ouverture sans badge : Oui. Hauteur maximum : 1,90m Pour essayer ce parking retrouvez toutes les informations utiles sur le site internet Yespark avec la référence : 92300-514-1 Louez une place dans un parking sécurisé pour bénéficier d'une place large et spacieuse. La location vous permettra de garer votre voiture ou moto neuve ou d'occasion. Nos parkings sont lumineux et il est simple d'y manoeuvrer.

749

Le groupe POINT DE VENTE vous propose à proximité des enseignes Franprix, Point Soleil, Carrefour, Crédit Mutuel... : un commerce d'une surface totale de 35 m<sup>2</sup> disposant d'une façade de 5 m : - Au rez-de-chaussée, une surface de 25 m<sup>2</sup>, point d'eau, WC, ainsi qu'un sous-sol d'une superficie de 10 m<sup>2</sup>, accessible par parties communes. Le local est Restauration sans nuisance ou alimentaire..

655

Le groupe POINT DE VENTE vous propose à Levallois-Perret, à quelques pas du métro Anatole France, un local commercial d'une surface de 60 m<sup>2</sup>, un sous-sol de 30 m<sup>2</sup> directement accessible par une trappe intérieure. 13 m<sup>2</sup> de vitrine. Sanitaire et point d'eau à l'intérieur du local. Le produit se situe dans un environnement commerçant au pied des enseignes connues telles que: Le Crédit Agricole, La Poste, Leader Price..

937

Le groupe POINT DE VENTE vous propose dans la ville de Levallois à proximité du métro Anatole France: un Local commercial d'une surface totale de 34 m<sup>2</sup> disposant d'une façade de 5 m - Au rez-de-chaussée, une surface de 25 m<sup>2</sup>, avec point d'eau, WC. - Au sous-sol, une surface de 9 m<sup>2</sup>. Cette boutique dispose d'un environnement commerçant composé de: Mairie de Levallois - Agences Immobilières - cafés - restaurants - traiteur - alimentation générale. Quartier du centre ville de Levallois très dynamique et commerçant, type Bureau et Habitation.

1158

Le groupe POINT DE VENTE vous propose à proximité du métro Anatole France: un commerce idéal pour Restauration sans nuisance ou alimentaire, Terrasse d'une surface totale de 125 m<sup>2</sup> disposant d'une façade de 15 m - Au rez-de-chaussée, une surface de 100 m<sup>2</sup>, avec point d'eau, WC. - Au sous-sol, une surface de 25 m<sup>2</sup>, accessible depuis escalier intérieur. Cette boutique dispose d'un environnement commerçant composé de: Café, bar, restaurants, BP, banque, boulangerie, supermarché, coiffeur, fleuriste.

83

Le groupe POINT DE VENTE vous propose à proximité du métro Anatole France: un commerce d'une surface totale de 56m<sup>2</sup> disposant d'une façade de 5 m - Au rez-de-chaussée, une surface de 56m<sup>2</sup>, avec point d'eau, WC.

1027

L'agence POINT DE VENTE vous propose à proximité du métro Louise Michel: un commerce d'une surface totale de 40 m<sup>2</sup> disposant d'une façade de 5 m - Au rez-de-chaussée, une surface de 30 m<sup>2</sup>, avec point d'eau, WC. - Au sous-sol, une



surface de 10 m<sup>2</sup>, accessible depuis un escalier intérieur. Cette boutique dispose d'un environnement commerçant composé de: Carrefour city, Franprix, Crédit mutuel, Point soleil, BNP..

1004

Bail commercial à céder pour cause de départ en retraite. Boutique "dépot/ vente vêtements enfants" située au centre de Levallois, proche métro Louise Michel et rue piétonne, superficie de 55 m<sup>2</sup>, 4 ml de vitrine, cave 8m2 située sous le magasin. Le bail inclus un appartement 2 pièces 24 m<sup>2</sup>, situé au 2eme étage de l'immeuble. Bail renouvelé 06/2017, loyer 1400 euros mensuel.

637

Boutique Local Commercial 122 m<sup>2</sup> de plain-pied à Levallois-Perret ! Contient une cave saine accessible par ascenseur de 47 m<sup>2</sup>. Très bon emplacement à 100 m du métro Louise Michel. Ancien Franprix Façade de 10 m rideau de fer, alarme, norme ERP. Endroit très commerçant ! Proche du marché couvert et de grandes enseigne : Nicolas, Au nom de la Rose, Subway... Bail tous commerces sauf restauration. Restauration sans extraction acceptée. Idéal pour magasin bio, primeur, épicerie, concessionnaire motos, audio, optique, parqueterie... Loyer : 3.300E HC/HT Cession : 99.000 euros.

1014

Levallois Perret 92300 Cession de Droit au Bail d'une boutique de 97 m<sup>2</sup>:(67 m<sup>2</sup> en rez-de-chaussée + 30 m<sup>2</sup> au 1er étage) Le local est équipé d'une toilette, de climatisation, et d'un rideau de fer. Vitrine de 4ml. Loyer de 2512 euros/mois TCC. Reprise: 56 000 euros honoraires inclus. Réf. FED-B111989.

1502

Levallois-Perret / Aristide Briand Location Pure - Un Local commercial d'une surface totale de 94m<sup>2</sup> comprenant une surface en rez-de-chaussée de 66 m<sup>2</sup> et un Sous-sol relié par un escalier intérieur d'une surface de 28m<sup>2</sup>. Façade de 8m avec rideaux métallique. Tout commerces sauf nuisances - Bail 3/6/9. Honoraires TTC en sus : 30% du loyer annuel €. Garantie : 6480€ €.

528

BOUTIQUE SANS PAS DE PORTE Levallois-perret - Métro Louise Michel - En face du marché, une boutique de 59m<sup>2</sup> décomposé d'un rez-de-chaussée de 37m<sup>2</sup> + 22m<sup>2</sup> de sous-sol en arrière boutique. Bail dérogatoire ou commercial selon locataire. Toute restauration interdite même sans extraction. Honoraires TTC en sus : 6653 €. Garantie : 3060 €.

951

L'équipe du Groupe Advisor vous propose un local commercial d'une surface totale de 97 m<sup>2</sup> à Levallois Perret. Celui-ci est composé d'un rez-de-chaussée de 67 m<sup>2</sup> ainsi que d'un 1er étage de 30 m<sup>2</sup> Idéal agence de voyage, salon de coiffure etc..

619

Très belle boutique avec belle devanture 6 mètres de façade. Arrière boutique bail neuf 3/6/9 toutes activités sauf restauration avec extraction a visiter !! Honoraires TTC en sus : 9000 €. Garantie : 7500 €.

967

porte d'asnières emplacement de parking 1er sous-sol immeuble neuf. Mise à jour le 25/06/2019 Demander des infos sur ce bien Référence annonce AVENDREALOVER :

1492891

1155

Dans une résidence de standing , et proche du métro Pont de Levallois, un emplacement de parking en sous-sol. Mise à jour le 06/06/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : 537

980

porte d'asnières emplacement de parking 1er sous-sol immeuble neuf

1267

Parking / box Levallois Perret\n\nDans une résidence de standing , et proche du métro Pont de Levallois, un emplacement de parking en sous-sol.\nRéférence annonce : 537\nNon soumis au DPE\nHonoraires à la charge du locataire : 90 € TTC\nDépôt de garantie : 90 €

766

emplacement de parking. Mise à jour le 25/06/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : 1533215

960

LA TOUR IMMO vous propose une boutique d'une surface de 35 m<sup>2</sup> de plain-pied. Tous commerces sauf restauration.

614

LA TOUR IMMO vous propose une boutique d'une surface totale de 40 m<sup>2</sup> de plain-pied, au coeur de Levallois-Perret. Tous commerces sauf restauration.

386

A louer bureaux rénovés et standing à Levallois Perret, dans un environnement calme et proche des commerces. Surface plug & play, filmé et climatisé. Baux dérogatoires si possibles.

859

Situés au sein d'un immeuble moderne à proximité de la Gare Clichy-Levallois, GROUPE BABYLONE vous propose à la location une surface de bureaux climatisée en excellent état. Parkings en sous-sol.

1466

Située au coeur de Levallois à toute proximité du métro et des commerces, GROUPE BABYLONE vous propose au sein d'un immeuble de qualité, une surface de bureaux climatisée alliant des espaces cloisonnés et en open space.

854

Parking situé 91-95 rue Anatole France au pied du Métro Anatole France.

Disponible de suite. - Dépôt de garantie : 100 EUR - Honoraires charge locataire : 100 EUR état des lieux inclus. Mise à jour le 17/08/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : 5888

287

Parking - Mo Anatole France\n\nParking situé 91-95 rue Anatole France au pied du Métro Anatole France. Disponible de suite.\n- Dépôt de garantie : 100 EUR\n- Honoraires charge locataire : 100 EUR état des lieux inclus\nRéférence annonce : 5888\nNon soumis au DPE\nHonoraires à la charge du locataire : 100 € TTC\nDépôt de garantie : 100 €

151

Parking situé 91-95 rue Anatole France au pied du Métro Anatole France.

Disponible de suite. - Dépôt de garantie : 100 € - Honoraires charge locataire :

100 € état des lieux inclus Réf. annonce : 5888

319

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de 50m2, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur. Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446

	PRICE	SURFACE	CRAWL_SOURCE_original
658	110.0	12.0	LE_BON_COIN
911	100.0	12.0	LE_BON_COIN
1475	100.0	11.0	LE_BON_COIN
551	75.0	11.0	LE_BON_COIN
654	100.0	10.0	LE_BON_COIN
27	110.0	11.0	LE_BON_COIN
1595	110.0	12.0	LE_BON_COIN
753	150.0	12.0	LE_BON_COIN
492	110.0	12.0	LE_BON_COIN
780	150.0	10.0	EXPLORIMMO
950	150.0	10.0	BIEN_ICI
1305	65000.0	73.0	A_VENDRE_A_LOUER
1580	132000.0	55.0	A_VENDRE_A_LOUER
684	220000.0	47.0	SE_LOGER
1530	3800.0	69.0	SE_LOGER
1255	220000.0	55.0	SE_LOGER
584	65000.0	73.0	SE_LOGER
1542	146.0	16.0	EXPLORIMMO
1636	60.0	14.0	PARU_VENDU
1008	60.0	17.0	LE_BON_COIN
850	201.0	19.0	EXPLORIMMO
78	60.0	19.0	A_VENDRE_A_LOUER
161	131.0	12.0	EXPLORIMMO
749	220000.0	35.0	SE_LOGER
655	60500.0	90.0	SE_LOGER
937	220000.0	34.0	SE_LOGER
1158	63500.0	125.0	SE_LOGER
83	220000.0	56.0	SE_LOGER
1027	220000.0	40.0	SE_LOGER
1004	65000.0	55.0	PAP
637	99000.0	122.0	PAP
1014	56000.0	97.0	SE_LOGER
1502	2300.0	94.0	SE_LOGER
528	1600.0	59.0	SE_LOGER

951	50000.0	97.0	SE_LOGGER
619	2590.0	84.0	SE_LOGGER
967	70.0	55.0	A_VENDRE_A_LOUER
1155	90.0	55.0	A_VENDRE_A_LOUER
980	70.0	55.0	EXPLORIMMO
1267	90.0	55.0	LE_BON_COIN
766	110.0	55.0	A_VENDRE_A_LOUER
960	220000.0	35.0	SE_LOGGER
614	220000.0	40.0	SE_LOGGER
386	3150.0	100.0	SE_LOGGER
859	4375.0	175.0	SE_LOGGER
1466	7548.0	283.0	SE_LOGGER
854	100.0	55.0	A_VENDRE_A_LOUER
287	100.0	55.0	LE_BON_COIN
151	100.0	55.0	LOGIC_IMMO
319	79900.0	50.0	EXPLORIMMO

Nous pouvons voir que notre algorithme a très bien regroupé les annonces similaires venant de différentes plateformes dans des groupes identiques mais qui a aussi fait un peu de confusion dans d'autres groupes car les structures des descriptions de ces annonces paraissent similaires.

```
[ ]: # Affiche des 10 premiers clusters ayant des plus d'elements
print(X_train_full_clustered['CLUSTER'].value_counts()[:10])
```

```
-1      1131
154      15
188      11
175       9
178       8
113       7
225       7
182       6
140       6
220       6
```

Name: CLUSTER, dtype: int64

```
[ ]: # Affichage des données du cluster le plus gros
X_train_full_clustered[X_train_full_clustered['CLUSTER'].value_counts().
↳ index[1] == X_train_full_clustered['CLUSTER']] [['CLUSTER',
↳ 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'SURFACE',
↳ 'CRAWL_SOURCE_original']].head(50)
```

```
[ ]:
CLUSTER  CLUSTER_DISTANCE  \
22      154              0.977470
196     154              0.900885
207     154              0.899579
288     154              0.981888
```

441	154	0.000000
534	154	0.991474
575	154	0.000000
667	154	0.883079
676	154	0.908855
1123	154	0.900885
1159	154	0.908855
1181	154	0.883079
1329	154	0.935591
1385	154	0.991474
1640	154	0.958237

## DESCRIPTION \

22

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

196

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

207

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

288

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec

cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

441

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

534 LEVALLOIS Un appartement traversant de 3 pièces d'environ 65m<sup>2</sup> dans une résidence sécurisée avec ascenseur et bien entretenu. L'appartement se compose d'une entrée avec rangement, d'un grand séjour exposé sud-ouest, une cuisine indépendante entièrement aménagée, de deux chambres sur jardin, d'une salle de bains et des wc séparés. Nombreux rangements. Calme et ensoleillé. Une cave complète ce bien. Charges annuelles : 3552 € Soumis au statut de la copropriété Prix de vente : 620.000 € dont 2.48% d'honoraires à la charge de l'acquéreur. Building Partners, fort d'une expertise de plus de 20 ans, est spécialisé dans la transaction de biens "Familiaux de Standing" et bénéficie d'implantations de choix dans l'ouest parisien. C'est avec plus de 450 appartements, maisons et hôtels particuliers que Building Partners offre le plus large choix de biens à la vente. 65, rue Président wilson 92300 Levallois (33) 1.47.31.01. 74.

575

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

667

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

676

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins

privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

1123

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

1159

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

1181

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

1329

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

1385 LEVALLOIS Un appartement traversant de 3 pièces d'environ 65m<sup>2</sup> dans une résidence sécurisée avec ascenseur et bien entretenu. L'appartement se compose d'une entrée avec rangement, d'un grand séjour exposé sud-ouest, une cuisine indépendante entièrement aménagée, de deux chambres sur jardin, d'une salle de bains et des WC séparés. Nombreux rangements. Calme et ensoleillé. Une cave complète ce bien. Charges annuelles: 3552 € Soumis au statut de la copropriété Prix de vente: 620.000 € dont 2.48% d'honoraires à la charge de l'acquéreur. Building Partners, fort d'une expertise de plus de 20 ans, est spécialisé dans

la transaction de biens "Familiaux de Standing" et bénéficie d'implantations de choix dans l'ouest parisien. C'est avec plus de 450 appartements, maisons et hôtels particuliers que Building Partners offre le plus large choix de biens à la vente. 65, rue Président wilson 92300 Levallois (33) 1.47.31.01. 74.

1640

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

	PRICE	SURFACE	CRAWL_SOURCE_original
22	947000.0	85.00	SE_LOGGER
196	516000.0	44.92	SE_LOGGER
207	661000.0	64.05	SE_LOGGER
288	1359000.0	111.86	SE_LOGGER
441	687000.0	61.85	SE_LOGGER
534	600000.0	65.00	MEILLEURS_AGENTS
575	687000.0	61.43	SE_LOGGER
667	704000.0	60.72	SE_LOGGER
676	486000.0	45.79	SE_LOGGER
1123	526000.0	43.74	SE_LOGGER
1159	520000.0	45.59	SE_LOGGER
1181	650000.0	58.14	SE_LOGGER
1329	515000.0	49.32	SE_LOGGER
1385	600000.0	65.00	SE_LOGGER
1640	687000.0	63.14	SE_LOGGER

Nous pouvons aussi remarqué qu'avec le gros cluster notre modèle à réussi à dectecter les annonces similaire venant de plusieurs plateforme differentes.

```
[ ]: # Affichage des données du cluster le plus petit
X_train_full_clustered[X_train_full_clustered['CLUSTER'].value_counts().
↳ index[-1] == X_train_full_clustered['CLUSTER']] [['CLUSTER',
↳ 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'SURFACE',
↳ 'CRAWL_SOURCE_original']].head()
```

```
[ ]: CLUSTER CLUSTER_DISTANCE \
1643      54              0.0
1651      54              0.0
```

```
DESCRIPTION \
1643 LEVALLOIS-PERRET 1 pièce 22.65 m²\n\nA deux pas du centre commercial So
Ouest dans résidence récente avec gardien & ascenseur, studio en bon état,
```



lumineux, vue dégagée, en tout collectif entrée, pièce principale avec coin cuisinette, salle d'eau avec wc, placard\nAssurance Multirisques Privilège \*: 10.01 € / mois\n\*Services facultatifs, voir conditions en agences\nRéférence annonce : 1558980\nHonoraires à la charge du locataire : 339 € TTC dont 67 € pour l'état des lieux\nDépôt de garantie : 590 €\nMontant des charges : 113 € / mois\nModalité de récupération des charges locatives : provision avec régularisation annuelle

1651 LEVALLOIS-PERRET 3 pièces 50.65 m<sup>2</sup>\n\nen plein centre ville, dans un quartier commerçant, immeuble ancien et sécurisé un 3 pièces bien aménagé avec belle cuisine et cheminée. à visiter Entrée, séjour, cuisine américaine, deux Salle de bains, wc et cave.\nAssurance Multirisques Privilège \*: 16.69 € / mois\n\*Services facultatifs, voir conditions en agences\nRéférence annonce : 1552939\nHonoraires à la charge du locataire : 759 € TTC dont 151 € pour l'état des lieux\nDépôt de garantie : 1211 €\nMontant des charges : 110 € / mois\nModalité de récupération des charges locatives : provision avec régularisation annuelle

	PRICE	SURFACE	CRAWL_SOURCE_original
1643	703.0	55.0	LE_BON_COIN
1651	1321.0	55.0	LE_BON_COIN

## 0.2.6 Prediction de nos données de test

```
[ ]: def pred_cluster(clusterer, df, scaled_cols):
    df = copy.deepcopy(df).reset_index(drop=True)
    labels, strengths = hdbscan.approximate_predict(clusterer, df.
↳ drop(col_del+['CRAWL_SOURCE_original'], axis=1, errors='ignore'))

    # undo scaling
    _df = pd.DataFrame(
        scaler_transformer.transformers_[0][1].inverse_transform(df[numeric_cols]),
        columns=numeric_cols
    )

    df[scaled_cols] = _df[scaled_cols] # update X_train with unscaled values
    # add cluster column
    df['CLUSTER'] = labels
    # make a column that contains distance from the cluster centroid. -1
↳ distance means that the point is an outlier
    df['CLUSTER_DISTANCE'] = strengths
    return df

X_test_full_clustered = pred_cluster(clusterer, X_test_full, numeric_cols)
```

## 0.2.7 Affichage de resultats de nos predicions de notre 1er modèle avec les données de test

```
[ ]: X_test_full_clustered.sort_values(by='CLUSTER', ascending=False)[['CLUSTER', 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'CRAWL_SOURCE_original']].head(30)
```

```
[ ]:      CLUSTER  CLUSTER_DISTANCE \
335      222      0.699928
281      219      0.847921
75       216      0.752165
149      216      0.693393
182      215      0.762483
85       213      0.450562
109      213      0.414517
212      211      0.675035
95       210      1.000000
63       209      1.000000
206      205      0.660064
322      203      1.000000
428      202      0.055299
32       202      0.139816
431      192      0.484970
128      188      1.000000
167      187      0.628491
251      185      0.660150
247      178      0.766182
243      178      0.652834
121      175      0.534492
102      173      0.263535
193      173      0.301629
265      173      0.937352
275      171      1.000000
404      167      1.000000
41       167      0.087138
272      165      1.000000
133      162      1.000000
191      156      0.241253
```

DESCRIPTION \

335

Cession de bail 92300 Levallois-Perret Le groupe POINT DE VENTE vous propose à proximité du métro Pont de Levallois - Bécon : un commerce idéal pour Restauration sans nuisance ou alimentaire d'une surface totale de 44 m<sup>2</sup> disposant d'une façade de 6 m - Au rez de chaussée, une surface de 44 m<sup>2</sup>. Prix : 55 000 € Loyer : 1 908, 50 € /HC /HT /MOIS Honoraires : 7 700 € H.T Ref annonce : P1-27014 Agence : Point de vente 01 77 39 27 91 Conseil en immobilier

commercial, Point de vente vous accompagne dans vos recherches de fonds de commerce, droit au bail, locations pures, murs de boutiques. Mise à jour le 19/08/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : P1-27014

281

Le groupe POINT DE VENTE vous propose un local commercial en angle d'une surface de 52 m<sup>2</sup> au rez-de-chaussée, Cave accessible par une trappe intérieure d'environ 45 m<sup>2</sup>. Le bien se situe dans un environnement très commerçant à quelques minutes à pied du métro Anatole France et des enseignes connues telles que La Poste, l'Atelier du Sourcil, Monoprix, Zara Home.. Restauration rapide sur dossier. Honoraires TTC en sus : 15500.00 €.

75

Le Groupe Babylone vous propose un local commercial en très bon état d'usage, avec une grande vitrine donnant sur deux angles d'environ 20 mètres linéaires, situé dans le centre de Levallois-Perret. Cession: 100 000 euros.

149

L'agence FLAGSHIP vous propose une boutique d'environ 97 m<sup>2</sup> composée d'un rez-de-chaussée de 67 m<sup>2</sup> et d'un sous-sol de 30m<sup>2</sup> environ, directement accessible par un escalier intérieur. Activités idéales: Tous commerces sauf nuisances, possibilité restauration rapide.

182

Dans une résidence de standing , et proche du métro Pont de Levallois, un emplacement de parking en sous-sol. Mise à jour le 02/06/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : 537

85

Surfaces de bureaux à Levallois Perret, dans un immeuble de standing. La surface est meublée, cloisonnée et climatisée. Espace lumineux, à proximité des transports en commun.

109

Dans un immeuble de qualité en plein coeur de Levallois, à toute proximité du métro et des commerces, nous vous proposons des surfaces de bureaux climatisées, cloisonnées et en open space.

212

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de 50m2, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur. Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446

95

Je vous propose un restaurant d'angle de 50 m2 en plein c?ur de Levallois Perret. Il y a 38 places assises à l'intérieur et une vingtaine en terrasse, une cuisine aménagée avec une hotte à charbon. Actuellement cuisine Française. La clientèle est une clientèle bureaux le midi et résident le soir avec une grande

demande de livraison. Affaire à saisir rapidement compte tenu de son emplacement et de son potentiel de développement. Loyer:2850 Euros par mois charges comprises.M. Yankel KOSKAS 0637729002 / 0180916809poste8326  
commerces.75017@maxihome.email Agent commercial indépendant du réseau national clairimmo Maxihome N° RSAC : 815 116 13200011 PRIUM - Ville du greffe : PARIS  
63

Marclay's Immobilier vous propose un local commercial de 40 m<sup>2</sup> de plain pied situé dans une rue prisée et très commerçante de Levallois à proximité direct du marché Henri Barbusse. Loyer mensuel: 1300 TTC Travaux de qualité refait à neuf. Point d'eau, WC. Il dispose d'une façade linéaire de 4 mètres. Bail tous commerces sauf nuisances. Emplacement idéal pour toutes activités d'optiques, audio-prothèse, barbier, coiffeur, créateur, dépôt-vente, accessoires de mode, prêt à porter, ameublement, décoration, salon de beauté, parfumerie, salon d'esthétique, bien être, vente cosmétiques, etc.. Honoraires: 10 % du montant de la cession. A visiter sans tarder ! N'hésitez pas à nous contacter pour de plus amples informations.

206

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce superbe restaurant sans extraction avec 22 places dont une dizaine de couverts en terrasse. Ce bien a une superficie totale de plus de 35 m2. Déco refait à neuf. Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC. Idéal vente à emporter et livraisons. A découvrir rapidement! Cession à 69000 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n°199606

322

92300 LEVALLOIS - BELLE AFFAIRE - Je vous propose un restaurant d'angle avec licence IV, offrant une surface de plus de 100 m<sup>2</sup> ouverte sur une vitrine de 10 mètres, ainsi qu'un sous-sol de 30 m2... Idéalement situé dans le quartier résidentiel et commerçant de Levallois-Perret, cet établissement de bonne réputation exploite actuellement près de 60 couverts, ainsi que 10 places en terrasse. Emplacement centre ville, exceptionnel et rare. Possibilité de transformation en brasserie. Excellent chiffre d'affaires. Loyer 2.890 euros. Prix 690.000 euros honoraires charge vendeur . Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389. Mandat n° 204762

428

Louez au mois un parking privé au pour €. Pas d'engagement, pas de caution, louez ce parking au mois et sans engagement. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-4018-1 Trouver un garage à louer n'aura jamais été aussi simple. Quelle que soit la taille de votre voiture neuve ou d'occasion, les emplacements sont suffisamment grands pour vous permettre de manoeuvrer sans difficulté. Tous nos parkings sont surveillés par un gardien et propres, il est donc très rassurant d'y laisser sa voiture.

32

Louez au mois un parking privé au pour €. Pas d'engagement, pas de caution, louez ce parking au mois et sans engagement. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-4270-1 Trouver un garage à louer n'aura jamais été aussi simple. Quelle que soit la taille de votre voiture neuve ou d'occasion, les emplacements sont suffisamment grands pour vous permettre de manoeuvrer sans difficulté. Tous nos parkings sont surveillés par un gardien et propres, il est donc très rassurant d'y laisser sa voiture.

431

A louer bureaux rénovés et lumineux à Levallois Perret, dans un environnement calme. Surface en cours de rénovation complète et disponible pour le mois de Juin 2019. Accès sécurisé et palier aménagé par le propriétaire.

128

Appartement, Studio Meublé, de 13m<sup>2</sup> clair et calme en bon état comprenant une pièce principale une cuisine équipée, Salle d'eau avec WC, Placards, Chauffage : Individuel électrique, A 2 pas de Neuilly et Paris, à 5 mn des métros Louise Michel et Porte de Champerret (M3), plusieurs lignes de bus au pied de l'immeuble, proche de la Défense.

167

EVOLIS vous propose une surface de bureaux rénovée avec terrasse technique dans un immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont très lumineux et climatisés.

251 RESTAURANT 150 M2 LEVALLOIS PERRET Je vous propose un restaurant de 150 m2 en plein c?ur de Levallois Perret.Situé dans une zone avec de nombreux bureaux mais aussi des habitations.Le restaurant a 80 places assises et une grande cuisine entièrement équipée d un matériel récent et en bon état de fonctionnement.Il y a aussi une grosse extraction aux normes. La clientèle est une clientèle bureaux le midi et le soir une clientèle de résidents fidèle et avec un bon pouvoir d achat. Belle affaire à saisir rapidement compte tenu sa superficie et de son emplacement. Loyer:3300 euros HT HC par mois. M. Yankel KOSKAS 06 37 72 90 02 / 01 809 16 809 poste 8326 commerces.75017@maxihome.email Agent commercial indépendant du réseau national clairimmo Maxihome N° RSAC : 815 116 13200011 PRIUM - Ville du greffe : PARIS Réf. annonce : 96663MX326

247

Marclay's Immobilier vous propose un local de 167 m<sup>2</sup> situé dans une rue prisée et très commerçante de Levallois. Ce local est composée d'un rez-de-chaussée d'une surface de 120 m<sup>2</sup>. Au sous-sol, une surface de 47 m<sup>2</sup>. Il dispose d'une façade linéaire de 6 mètres. Bail tous commerces sauf nuisances. Emplacement idéal pour toutes activités d'optiques, audio-prothèse, barbier, coiffeur, créateur, dépôt-vente, accessoires de mode, prêt à porter, ameublement, décoration, salon de beauté, salon d'esthétique, bien être, vente cosmétiques, micro-crèche etc.. Environnement: Subway, Planet Sushi, Nicolas, Au Nom de la Rose.. A visiter sans tarder ! Honoraires: 13 300 euros HT à la charge du preneur N'hésitez pas à nous contacter pour de plus amples informations.

243

Groupe Babylone vous propose un local de 90 m<sup>2</sup> de plain-pied avec une façade de

5 mètres, a proximité de la rue Président Wilson à Levallois, faible loyer. A visiter, cession 70 000 euros.

121

Bureaux à louer à Levallois Perret, excellent emplacement à proximité immédiate du métro Anatole France pour cet immeuble indépendant avec possibilité ERP: idéal centre de formation.

102

Le groupe POINT DE VENTE vous propose: un commerce d'une surface totale de 90 m<sup>2</sup> disposant d'une façade de 6 m: - Au rez-de-chaussée, une surface de 60 m<sup>2</sup>, point d'eau, WC, ainsi qu'un sous-sol d'une superficie de 30 m<sup>2</sup>, accessible par parties communes. Le local est Restauration sans nuisance ou alimentaire. Honoraires TTC en sus : 8000.00 €.

193

LA TOUR IMMO vous propose un local d'une surface de 90 m<sup>2</sup> comprenant un rez-de-chaussée de 60 m<sup>2</sup> et un sous-sol de 30 m<sup>2</sup> accessible par un escalier intérieur. Tous commerces sauf extraction. Pas de TVA. Charges: 200 € /Mois.

265

LA TOUR IMMO vous propose un local d'une surface totale de 56 m<sup>2</sup> de plain-pied; au coeur de Levallois. Refait à neuf. Tous commerces sauf restauration.

275

Appartement, 2 pièces, Meublé, de 36m<sup>2</sup>, clair et calme sans vis à vis, en très bon état avec parties communes de l'immeuble refaite à neuf, comprenant une pièce principale une cuisine US équipée , Salle d'eau, WC, Placards, Chauffage : Individuel électrique, Situé au 1er étage d'un immeuble sécurisé avec 2 porte d'accès. Proche des transports et des commodités

404

Vente de murs libres au jour de l'acquisition - Local d'activité - Ambiance société événementielle - utilisé à usage professionnel actuellement - possibilité bureaux; libéral, artistique, ou artisanal - Près Porte de Champerret. Surface atypique de 242 m<sup>2</sup>, dont 200 m<sup>2</sup> en rez-de-chaussée et 42 m<sup>2</sup> à l'étage avec entrée directe par la rue. Locaux entièrement rénovés en 2018.2 niveaux. Rez-de-chaussée: belle entrée. Espace cuisine. Salle de réunion. Grand open space. Salle de stockage. 2 WC. A l'étage: open-space + bureau fermé. Climatisation réversible, câblage informatique et baie de brassage. Très belle hauteur sous plafond. Très lumineux. Alarme et vidéosurveillance. Réf: MFA113305.

41

GIFOM - Vente Local commercial 92300 LEVALLOIS PERRET. Métro Louise Michel - Anatole France Surface atypique 242 m<sup>2</sup> avec entrée privative sur rue, locaux sur 2 niveaux entièrement rénovés en 2018 comprenant: rez-de-chaussée avec une belle entrée, espace cuisine, salle de réunion, grand open space, salle de stockage et 2 WC + 1er étage composé d'un open space et d'un bureau fermé. Très belle hauteur sous plafond, très lumineux. Alarme et vidéosurveillance. Charges: 14,35e/m<sup>2</sup>/an Foncier: 12,52e/m<sup>2</sup>/an Taxe bureaux: 17,55e/m<sup>2</sup>/an Prix de vente: 1 950 000 e NV + honoraires en sus de 7 %HT.

272

Le groupe POINT DE VENTE vous propose au centre ville de Levallois à proximité

des enseignes Monop - restaurants et de nombreux commerces. Un Local commercial d'une surface totale de 62 m<sup>2</sup> disposant d'une façade de 7 m: - Au rez-de-chaussée, une surface de 57 m<sup>2</sup>, hauteur sous plafond > 2m80, point d'eau, WC, ainsi qu'un sous-sol d'une superficie de 5 m<sup>2</sup>, accessible par parties communes. Cet emplacement correspond parfaitement à des activités comme Restauration rapide - traiteur - primeurs. Le local est Restauration sans nuisance ou alimentaire.

133

LEVALLOIS - BOX - 35 000 EUROS FAI Quartier Greffulhe, dans une résidence moderne et sécurisé, un box au 2ème sous-sol. Prix hors honoraires : 32 000 € Prix honoraires inclus : 35 000 € Honoraires d'agence : 3 000 € TTC Les honoraires sont à charge acquéreur. Imax, n°1 de l'immobilier à Levallois depuis 2003. N° mandat V18061. Réf. annonce : 25996191

191

Situé sur un axe principal de Levallois Perret, EVOLIS vous propose deux surfaces de bureaux rénovées et climatisées. ERP. Possible.

	PRICE	CRAWL_SOURCE_original
335	55000.0	A_VENDRE_A_LOUER
281	4509.0	SE_LOGGER
75	100000.0	SE_LOGGER
149	55000.0	SE_LOGGER
182	90.0	A_VENDRE_A_LOUER
85	4785.0	SE_LOGGER
109	6804.0	SE_LOGGER
212	79900.0	BIEN_ICI
95	150000.0	EXPLORIMMO
63	220000.0	SE_LOGGER
206	69000.0	BIEN_ICI
322	690000.0	EXPLORIMMO
428	100.0	EXPLORIMMO
32	65.0	EXPLORIMMO
431	6096.0	SE_LOGGER
128	620.0	EXPLORIMMO
167	11001.0	SE_LOGGER
251	190000.0	LOGIC_IMMO
247	100000.0	SE_LOGGER
243	70000.0	SE_LOGGER
121	25437.0	SE_LOGGER
102	2500.0	SE_LOGGER
193	60000.0	SE_LOGGER
265	220000.0	SE_LOGGER
275	1100.0	EXPLORIMMO
404	2067000.0	SE_LOGGER
41	2113800.0	SE_LOGGER
272	220000.0	SE_LOGGER
133	35000.0	LOGIC_IMMO

Nous avons pu prédire et affecter nos données de test dans des clusters dont les annonces correspondaient ou ils ont été considérés comme étant des annonces uniques

### 0.3 2ème Modèle

#### 0.3.1 2) Methode de similarité pour traiter les descriptions

Nous utilisons la similarité cosinus pour mesurer la similarité entre 2 paires des descriptions (la représentation vectorielle de la description). Nous allons utiliser cette methode et reprendre l'entraînement de notre modèle.

La méthode de similarité cosinus (cosine similarity) est couramment utilisée pour mesurer la similarité entre des textes, des documents ou des vecteurs de texte dans le traitement automatique du langage naturel (NLP). Elle repose sur une représentation vectorielle des textes, où chaque document est représenté sous forme d'un vecteur dans un espace multidimensionnel.

Pour appliquer la similarité cosinus, les textes sont tout d'abord représentés sous forme de vecteurs dans un espace vectoriel. Il existe différentes méthodes pour effectuer cette représentation, notamment la représentation en sac de mots (bag of words) ou des techniques de plongement de mots telles que Word2Vec, GloVe ou des modèles de langage pré-entraînés comme BERT. Chaque dimension de ces vecteurs correspond à un mot ou une caractéristique spécifique du texte.

Une fois que les textes sont représentés sous forme de vecteurs, vous obtenez deux vecteurs correspondant aux textes que vous souhaitez comparer. La similarité cosinus mesure l'angle entre les deux vecteurs dans l'espace vectoriel.

Plus la similarité cosinus est proche de 1, plus les textes sont similaires, car leurs vecteurs pointent dans des directions similaires. Plus elle est proche de -1, plus les textes sont différents. Une similarité cosinus proche de 0 signifie que les textes n'ont pas de similarité particulière.

```
[ ]: from sentence_transformers import SentenceTransformer, util
import umap

device = 'cuda' if torch.cuda.is_available() else 'cpu'
model = SentenceTransformer('dangvantuan/sentence-camembert-large')
model.to(device)
if device == 'cuda':
    model = model.half()

max_desc_len = 500

description_embeddings = df['DESCRIPTION'].apply(lambda x: f"query: {str(x)[:
    ↪max_desc_len]}").tolist()
description_embeddings = model.encode(description_embeddings,
    ↪show_progress_bar=True, convert_to_tensor=False)

# calculer du cosinus de similarités entre les paires des descriptions
```



```
description_embeddings_cos_sim = util.pytorch_cos_sim(description_embeddings,
↳description_embeddings)

df_cos_sim = pd.DataFrame(description_embeddings_cos_sim)
```

WARNING:sentence\_transformers.SentenceTransformer:No sentence-transformers model found with name /root/.cache/torch/sentence\_transformers/dangvantuan\_sentence-camembert-large. Creating a new one with MEAN pooling.

Batches: 0%| | 0/68 [00:00<?, ?it/s]

```
[ ]: df_cos_sim.shape
```

```
[ ]: (2164, 2164)
```

Le désavantage de cette technique est qu'il exige beaucoup de ressources et nous crée plus des variables car nous calculons la cosinus de similarité entre chaque valeur et tous les autres deux à deux c'est-à-dire si nous avons N valeurs dans une variable, le cosinus de similarité va nous donner N variables

### 0.3.2 Train et split pour le 2ème modèle

```
[ ]: from sklearn.model_selection import train_test_split

test_size = 0.2
df_with_all_features = pd.concat([df_scaled, df_cos_sim], axis=1)
df_with_all_features['CRAWL_SOURCE_original'] = df['CRAWL_SOURCE']
df_with_all_features = pd.concat([df_with_all_features, df['DESCRIPTION']],
↳axis=1)
X_train_full, X_test_full = train_test_split(df_with_all_features,
↳test_size=test_size, random_state=42)
X_train, X_test = X_train_full.drop(col_del+['CRAWL_SOURCE_original'], axis=1,
↳errors='ignore'), X_test_full.drop(col_del+['CRAWL_SOURCE_original'],
↳axis=1, errors='ignore')
```

```
[ ]: #affichage de la forme de nos données après division
X_train.shape, X_test.shape, X_train_full.shape, X_test_full.shape
```

```
[ ]: ((1731, 2195), (433, 2195), (1731, 2197), (433, 2197))
```

### 0.3.3 Entraînement du 2e modèle

```
[ ]: import hdbscan
seed_everything(42)
clusterer = hdbscan.HDBSCAN(min_cluster_size=2,
                             min_samples=1, # min_cluster_size by default
                             alpha=5.0, # 1.0 by default
                             # cluster_selection_epsilon=3, # 0.0 by default)
```

```

        prediction_data=True
    )
clusterer.fit(X_train)

```

```
[ ]: HDBSCAN(alpha=5.0, min_cluster_size=2, min_samples=1, prediction_data=True)
```

```
[ ]: # umap dimensionality reduction
seed_everything(42)
reducer = umap.UMAP(n_neighbors=6, n_components=2)
X_train_pca = reducer.fit_transform(X_train)

```

```
[ ]: # plot clusters

X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION'].str[:600]
X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION_wrapped'].str.
    ↪wrap(100)
X_train_full['DESCRIPTION_wrapped'] = X_train_full['DESCRIPTION_wrapped'].
    ↪apply(lambda x: str(x).replace('\n', '<br>'))

fig = go.Figure()

fig.add_trace(go.Scatter(
    x=X_train_pca[:, 0],
    y=X_train_pca[:, 1],
    mode='markers',
    marker=dict(
        size=5,
        color=clusterer.labels_,

    ),
    text=X_train.index,
    customdata=X_train_full['DESCRIPTION_wrapped'],
    hovertemplate='Description: %{customdata}<br>Index: %{text}<br>Cluster:␣
    ↪%{marker.color}<extra></extra>'
))

# increase size of the plot
fig.update_layout(
    title='Clusters',
    xaxis_title='PCA 1',
    yaxis_title='PCA 2',
    legend_title='Cluster',
    font=dict(
        family='Courier New, monospace',
        size=18,
        color='#7f7f7f'
    ),

```

```

        width=1200,
        height=800
    )
X_train_full.drop('DESCRIPTION_wrapped', axis=1, inplace=True)

fig.show()

```

### 0.3.4 Ajout de la colonne de regroupement pour le 2è modèle

```

[ ]: def add_cluster_column(clusterer, X_train_):
    X_train = copy.deepcopy(X_train_).reset_index(drop=True)
    labels = clusterer.labels_
    n_clusters_ = len(set(labels)) - (1 if -1 in labels else 0)
    print('Estimated number of clusters: %d' % n_clusters_)
    print('Number of outliers/unique ads: %d' % list(labels).count(-1))
    # print max amount of clusters, mean amount of clusters and min amount of
    ↪ clusters
    print('Max cluster: %d' % np.sort(np.unique(labels,
    ↪ return_counts=True)[1])[:-1][1]) # ignore -1
    print('Mean cluster: %d' % np.unique(labels, return_counts=True)[1].mean())
    print('Min cluster: %d' % np.unique(labels, return_counts=True)[1].min())

    # undo scaling
    _X_train = pd.DataFrame(scaler_transformer.named_transformers_['Scaler'].
    ↪ inverse_transform(X_train[numeric_cols]), columns=numeric_cols)
    X_train[numeric_cols] = _X_train[numeric_cols]
    # add cluster column
    X_train['CLUSTER'] = labels
    # make a column that contains distance from the cluster centroid. -1
    ↪ distance means that the point is an outlier
    X_train['CLUSTER_DISTANCE'] = clusterer.outlier_scores_
    return X_train

X_train_full_clustered = add_cluster_column(clusterer, X_train_full)

```

```

Estimated number of clusters: 52
Number of outliers/unique ads: 1566
Max cluster: 12
Mean cluster: 32
Min cluster: 2

```

Nous voyons que notre 2ème modèle a réussi de faire 52 groupements des annonces similaires et à classer 1566 annonces comme étant uniques.

Nous pouvons aussi remarquer que le plus gros groupe a 12 annonces similaires, le plus petit groupe en a 2 et la moyenne des annonces dans les groupes est de 32

### 0.3.5 Affichage de nos resultat de regroupement

```
[ ]: # Affichage de 5 premiers cluster ayant plus des annonces
print(X_train_full_clustered['CLUSTER'].value_counts()[:5])
```

```
-1      1566
39       12
50       10
36        8
26        7
Name: CLUSTER, dtype: int64
```

```
[ ]:
```

```
[ ]: # Affichage par cluster
display(X_train_full_clustered.sort_values(by='CLUSTER',
↪ascending=False)[['CLUSTER', 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE',
↪'SURFACE', 'CRAWL_SOURCE_original']].head(50))
```

	CLUSTER	CLUSTER_DISTANCE \
575	51	0.000000
667	51	0.878252
1086	51	0.752538
1331	51	0.878252
441	51	0.000000
207	51	0.752538
257	50	0.000000
1535	50	0.531232
787	50	0.000000
717	50	0.000000
148	50	0.339623
1159	50	0.502139
196	50	0.172156
1123	50	0.360203
676	50	0.000000
618	50	0.172156
50	49	0.000000
288	49	0.000000
22	49	0.438328
114	48	0.000000
325	48	0.000000
833	48	0.000000
294	48	0.000000
149	47	0.000000
168	47	0.000000
479	46	0.000000
1663	46	0.000000
953	45	0.000000

1213	45	0.000000
1380	44	0.000000
550	44	0.000000
1634	43	0.000000
132	43	0.000000
925	42	0.000000
215	42	0.000000
1173	41	0.000000
469	41	0.000000
390	40	0.000000
335	40	0.000000
503	39	0.000016
769	39	0.000000
1637	39	0.406992
204	39	0.000000
698	39	0.003959
858	39	0.000000
1679	39	0.321471
1370	39	0.003838
359	39	0.000000
1392	39	0.321471
713	39	0.000016

↳

↳

↳

↳

↳

↳

↳

↳

↳

↳DESCRIPTION \

575

↳

↳

↳Une adresse de choix pour devenir propriétaire. La station de métro, le lycée,  
 ↳et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
 ↳apprécierez la qualité résidentielle de la commune et choisirez parmi une  
 ↳typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
 ↳intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
 ↳terrasses ou jardins privés. Un cœur d'îlot paysager et des espaces  
 ↳communs végétalisés avec cheminements piétons desservent les différents halls  
 ↳d'entrée sécurisés. \*Source: Google Maps.

667

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

1086

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

1331

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

441

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

207

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

257

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

1535

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

787

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

717

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

148

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

1159

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

196

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.



1123

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

676

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

618

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

50

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

288

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

22

↳  
↳  
↳ Une adresse de choix pour devenir propriétaire. La station de métro, le lycée  
↳ et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous  
↳ apprécierez la qualité résidentielle de la commune et choisirez parmi une  
↳ typologie variée d'appartements du studio au 5 pièces duplex. Les espaces  
↳ intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons,  
↳ terrasses ou jardins privatifs. Un coeur d'ilot paysager et des espaces  
↳ communs végétalisés avec cheminements piétons desservent les différents halls  
↳ d'entrée sécurisés. \*Source: Google Maps.

114

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳ EVOLIS vous  
↳ propose une surface de bureaux rénovée avec terrasse technique dans un  
↳ immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont  
↳ très lumineux et climatisés.

325

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳ EVOLIS vous  
↳ propose une surface de bureaux rénovée avec terrasse technique dans un  
↳ immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont  
↳ très lumineux et climatisés.

833

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

EVOLIS vous

↳ propose une surface de bureaux rénovée avec terrasse technique dans un  
↳ immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont  
↳ très lumineux et climatisés.

294

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

EVOLIS vous

↳ propose une surface de bureaux rénovée avec terrasse technique dans un  
↳ immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont  
↳ très lumineux et climatisés.

149

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

A

↳ Levallois-Perret, à proximité des commerces et services, EVOLIS vous propose à  
↳ la location une surface de bureaux atypique de 330 m<sup>2</sup> en Rez-de-chaussée (dont  
↳ 175 m<sup>2</sup> en sous-sol).

168

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

A

↳ Levallois-Perret, à proximité des commerces et services, EVOLIS vous propose à  
↳ la location une surface de bureaux atypique de 330 m<sup>2</sup> en Rez-de-chaussée (dont  
↳ 175 m<sup>2</sup> en sous-sol).

479

↪  
↪  
↪  
↪ Local  
↪ commercial \*27 bis avenue Aristide Briand, 1 Vitrine sur rue, 2 Fenêtres sur  
↪ Jardin, le local de 57 m<sup>2</sup> en rez-de-chaussée comprend 1 surface en façade  
↪ communicante avec 1 surface sur jardin, un sous/sol direct sans fenêtre de 33  
↪ m<sup>2</sup> environ, un dégagement, un wc. Le loyer est présenté Hors Taxe. Bail Neuf 3.  
↪ 6.9 dans l'activité du preneur (pas de restauration). Les Commerces et  
↪ transports à proximité, le métro « Anatole France » Ligne 3 est à 3 minutes.  
↪ Honoraires Preneur 9770.40 TTC.

1663

↪  
↪  
↪  
↪ Local  
↪ commercial \*27 bis avenue Aristide Briand, 1 Vitrine sur rue, 2 Fenêtres sur  
↪ Jardin, le local de 57 m<sup>2</sup> en rez-de-chaussée comprend 1 surface en façade  
↪ communicante avec 1 surface sur jardin, un sous/sol direct sans fenêtre de 33  
↪ m<sup>2</sup> environ, un dégagement, un wc. Le loyer est présenté Hors Taxe. Bail Neuf 3.  
↪ 6.9 dans l'activité du preneur (pas de restauration). Les Commerces et  
↪ transports à proximité, le métro « Anatole France » Ligne 3 est à 3 minutes.  
↪ Honoraires Preneur 9770.40 TTC.

953

↪  
↪  
↪  
↪ Local commercial \*27 bis avenue Aristide Briand, 1 Vitrine sur rue, 2  
↪ Fenêtres sur Jardin, le local de 57 m<sup>2</sup> en rez-de-chaussée comprend 1 surface en  
↪ façade communicante avec 1 surface sur jardin, un sous/sol direct sans fenêtre  
↪ de 33 m<sup>2</sup> environ, un dégagement, un WC. Le loyer est présenté Hors Taxe. Bail  
↪ Neuf 3.6.9 dans l'activité du preneur (pas de restauration). Les Commerces et  
↪ transports à proximité, le métro " Anatole France " Ligne 3 est à 3 minutes.  
↪ Honoraires Preneur 9770.40 euros TTC. Honoraires TTC en sus : 29.50 €.  
↪ Garantie : 12858.00 €.

1213

↪  
↪  
↪  
↪ Local commercial \*27 bis avenue  
↪ Aristide Briand, 1 Vitrine sur rue, 2 Fenêtres sur Jardin, le local de 57 m<sup>2</sup>  
↪ en rez-de-chaussée comprend 1 surface en façade communicante avec 1 surface sur  
↪ jardin, un sous/sol direct sans fenêtre de 33 m<sup>2</sup> environ, un dégagement, un WC.  
↪ Le loyer est présenté Hors Taxe. Bail Neuf 3.6.9 dans l'activité du preneur  
↪ (pas de restauration). Les Commerces et transports à proximité, le métro "  
↪ Anatole France " Ligne 3 est à 3 minutes. Honoraires Preneur 9770.40 euros TTC.  
↪ Garantie : 12858.00 €.

1380

↳  
↳  
↳  
↳  
↳  
↳  
↳

□  
□  
□  
□  
□  
□  
□

↳ EVOLIS vous propose à LEVALLOIS Perret (Métro ANATOLE France) une location  
↳ pure d'un local commercial sur deux niveaux donnant sur rue passante. Surface:  
↳ 265 m<sup>2</sup> en rez-de-chaussée + 55 en sous sol environ. Loyer 150 000 euros HT HC  
↳ Beau linéaire.

550

↳  
↳  
↳  
↳  
↳  
↳  
↳

□  
□  
□  
□  
□  
□  
□

↳ EVOLIS vous propose à LEVALLOIS Perret (Métro ANATOLE France) une location  
↳ pure d'un local commercial sur deux niveaux donnant sur rue passante. Surface:  
↳ 265 m<sup>2</sup> en rez-de-chaussée + 55 en sous sol environ. Loyer 150 000 euros HT HC  
↳ Beau linéaire.

1634

↳  
↳  
↳  
↳

□  
□  
□  
□

↳ EVOLIS vous propose un local commercial de 340 m<sup>2</sup> au coeur de Levallois  
↳ réparti comme suit: 140 m<sup>2</sup> de rez-de-chaussée, sous-sol 60 m<sup>2</sup>, 1er étage 119  
↳ m<sup>2</sup>, 7 m de façade, extraction. Droit d'achalandage de 30 m<sup>2</sup> Environnement  
↳ commercial: Enseigne Kryss, Monoprix Carrefour city, Kayzer, nombreux  
↳ commerces de bouche. Loyer de 8 500 € HT/HC/mois Cession 600 000 € + 5% HT  
↳ d'honoraires de commercialisation.

132

↳  
↳  
↳  
↳  
↳

□  
□  
□  
□  
□

↳ EVOLIS vous propose un local commercial de 340 m<sup>2</sup> au coeur de Levallois  
↳ réparti comme suit: - 140 m<sup>2</sup> de rez-de-chaussée - Sous-sol 60 m<sup>2</sup> - 1er étage  
↳ 119 m<sup>2</sup> - 7 m de façade - extraction Droit d'achalandage de 30 m<sup>2</sup> Environnement  
↳ commercial: Enseigne Kryss, Monoprix Carrefour city, Kayzer, nombreux  
↳ commerces de bouche. Loyer de 8 500 € HT/HC/mois Cession 600 000 € + 5% HT  
↳ d'honoraires de commercialisation.

925

↳  
↳  
↳  
↳ EVOLIS vous propose, au sein d'un immeuble de grand standing, une  
↳ surface de bureaux avec accès à un très beau jardin. Porte de Champerret à  
↳ toute proximité. Métro "Louise Michel" ligne 3. . Immeuble de très bon  
↳ standing . Ascenseur . Hôtesse d'accueil - Climatisation - Petite terrasse au  
↳ 3ème étage - Câblage informatique - Flexibilité d'aménagement Situation/  
↳ Transports : Anatole France (3) Louise Michel (3) Dépôt de garantie : 3 mois  
↳ de loyer HT HC. Mise à jour le 22/08/2019 Demander des infos sur ce bien  
↳ Référence annonce AVENDREALOUE : 823641F

215

↳  
↳  
↳  
↳ EVOLIS vous propose, au sein d'un immeuble de grand standing, plusieurs  
↳ surfaces de bureaux avec accès à un très beau jardin. Porte de Champerret à  
↳ toute proximité. Métro "Louise Michel" ligne 3. . Immeuble de très bon  
↳ standing . Ascenseur . Hôtesse d'accueil - Climatisation - Petite terrasse au  
↳ 3ème étage - Câblage informatique - Flexibilité d'aménagement Situation/  
↳ Transports : Anatole France (3) Louise Michel (3) Dépôt de garantie : 3 mois  
↳ de loyer HT HC. Mise à jour le 22/08/2019 Demander des infos sur ce bien  
↳ Référence annonce AVENDREALOUE : 817892F

1173

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳ SUCCÈS COMMERCIAL ! Idéalement située,  
↳ cette nouvelle adresse regroupe tous les services pour profiter d'une vraie  
↳ vie de quartier: commerces, écoles, restaurants, bars.. Vous bénéficiez de la  
↳ proximité du Parc Collange.La résidence à taille humaine propose des  
↳ appartements aux prestations de gamme, du studio au 5 pièces. Caves et  
↳ parkings.(Réf. 402/726294-1220600).

469

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳ SUCCÈS COMMERCIAL ! Idéalement située,  
↳ cette nouvelle adresse regroupe tous les services pour profiter d'une vraie  
↳ vie de quartier: commerces, écoles, restaurants, bars.. Vous bénéficiez de la  
↳ proximité du Parc Collange.La résidence à taille humaine propose des  
↳ appartements aux prestations de gamme, du studio au 5 pièces. Caves et  
↳ parkings.(Réf. 302/726280-1220969).

390

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

┐  
┐  
┐  
┐  
┐  
┐  
┐  
┐

↳A proximité de la gare "Clichy-Levallois" sur l'axe Victor Hugo, dans un┐  
↳immeuble moderne et de bon standing, EVOLIS vous propose des surfaces de┐  
↳bureaux livrés rénovés.

335

↳  
↳  
↳  
↳  
↳  
↳  
↳  
↳

┐  
┐  
┐  
┐  
┐  
┐  
┐  
┐

↳A proximité de la gare "Clichy-Levallois" sur l'axe Victor Hugo, dans un┐  
↳immeuble moderne et de bon standing, EVOLIS vous propose des surfaces de┐  
↳bureaux livrés rénovés.

503

↳  
↳

┐  
┐

92300┐

↳LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION - Je vous propose ce┐  
↳restaurant de restauration rapide sandwicherie et saladerie idéalement situé┐  
↳dans un quartier de bureaux, écoles et commerçant. Superficie de 20 m2 avec 6┐  
↳places debout. Pas d'extraction. Loyer mensuel très faible à 998 euros.┐  
↳Cession à 127900 euros honoraires inclus à la charge du vendeur. Contactez Mme┐  
↳Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce┐  
↳et agissant en tant que chargée d'affaires indépendante auprès de la SAS┐  
↳PROPRIETES PRIVEES - Mandat n° 188156

769

↳  
↳  
↳

┐  
┐  
┐

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS┐

↳EXTRACTION : Je vous propose ce restaurant de restauration rapide idéalement┐  
↳situé dans un quartier de bureaux, écoles et commerçants, avec 10/12 places┐  
↳intérieur et d'une superficie de 25 m2. Loyer mensuel 940 euros. Cession à┐  
↳99200 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique┐  
↳Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et┐  
↳agissant en tant que chargée d'affaires indépendante auprès de la SAS┐  
↳PROPRIETES PRIVEES - Mandat n° 188597

1637

↳ 92300 LEVALLOIS PERRET - RESTAURATION RAPIDE.↳  
↳ SANS EXTRACTION - Je vous propose ce restaurant de restauration rapide.↳  
↳ sandwicherie et saladerie idéalement situé dans un quartier de bureaux, écoles.↳  
↳ et commerçant. Superficie de 20 m2 avec 6 places debout. Pas d'extraction.↳  
↳ Loyer mensuel très faible à 998 euros. Cession à 127900 euros honoraires.↳  
↳ inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96.↳  
↳ 11, spécialisée en cession de fonds de commerce et agissant en tant que.↳  
↳ chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat.↳  
↳ n° 188156 Réf. annonce : 2807FR188156B-FRZ

204

↳ 92300 LEVALLOIS PERRET - VRAI COUP DE COEUR :↳  
↳ Je vous propose ce superbe restaurant sans extraction avec 22 places dont une.↳  
↳ dizaine de couverts en terrasse. Ce bien a une superficie totale de plus de 35.↳  
↳ m2. Déco refait à neuf. Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC.↳  
↳ Idéal vente à emporter et livraisons. A découvrir rapidement! Cession à 69000.↳  
↳ euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé.↳  
↳ au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en.↳  
↳ tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES -↳  
↳ Mandat n°199606

698

↳ 92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION : Je vous propose.↳  
↳ ce restaurant de restauration rapide idéalement situé dans un quartier de.↳  
↳ bureaux, écoles et commerçants, avec 10/12 places intérieur et d'une.↳  
↳ superficie de 25 m2. Loyer mensuel 940 euros. Cession à 99200 euros honoraires.↳  
↳ inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96.↳  
↳ 11, spécialisée en cession de fonds de commerce et agissant en tant que.↳  
↳ chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat.↳  
↳ n° 188597 Réf. annonce : 2807FR188597B-FRZ

858

↳ 92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce superbe.↳  
↳ restaurant sans extraction avec 22 places dont une dizaine de couverts en.↳  
↳ terrasse. Ce bien a une superficie totale de plus de 35 m2. Déco refait à neuf.↳  
↳ Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC. Idéal vente à emporter.↳  
↳ et livraisons. A découvrir rapidement! Cession à 69000 euros honoraires inclus.↳  
↳ à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11,↳  
↳ spécialisée en cession de fonds de commerce et agissant en tant que chargée.↳  
↳ d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n°199606.↳  
↳ Réf. annonce : 2807FR199606B-FRZ



1679 92300 LEVALLOIS - BELLE AFFAIRE - Je vous  
propose un restaurant d'angle avec licence IV, offrant une surface de plus de  
100 m<sup>2</sup> ouverte sur une vitrine de 10 mètres, ainsi qu'un sous-sol de 30 m<sup>2</sup>...  
Idéalement situé dans le quartier résidentiel et commerçant de  
Levallois-Perret, cet établissement de bonne réputation exploite actuellement  
près de 60 couverts, ainsi que 10 places en terrasse. Emplacement centre  
ville, exceptionnel et rare. Possibilité de transformation en brasserie.  
Excellent chiffre d'affaires. Loyer 2.890 euros. Prix 690.000 euros honoraires  
charge vendeur . Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant  
en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES  
PRIVEES. CPI : 75012016000006389. Mandat n° 204762

1370 92300 LEVALLOIS PERRET  
- VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide  
avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de  
50m<sup>2</sup>, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail  
échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A  
découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur.  
Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11,  
spécialisée en cession de fonds de commerce et agissant en tant que chargée  
d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446  
Réf. annonce : 2807FR187446B-FRZ

359 92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS  
EXTRACTION : Je vous propose ce restaurant de restauration rapide idéalement  
situé dans un quartier de bureaux, écoles et commerçants, avec 10/12 places  
intérieur et d'une superficie de 25 m<sup>2</sup>. Loyer mensuel 940 euros. Cession à  
99200 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique  
Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et  
agissant en tant que chargée d'affaires indépendante auprès de la SAS  
PROPRIETES PRIVEES - Mandat n° 188597

1392 92300 LEVALLOIS - BELLE AFFAIRE - Je vous propose un restaurant d'angle  
avec licence IV, offrant une surface de plus de 100 m<sup>2</sup> ouverte sur une vitrine  
de 10 mètres, ainsi qu'un sous-sol de 30 m<sup>2</sup>... Idéalement situé dans le quartier  
résidentiel et commerçant de Levallois-Perret, cet établissement de bonne  
réputation exploite actuellement près de 60 couverts, ainsi que 10 places en  
terrasse. Emplacement centre ville, exceptionnel et rare. Possibilité de  
transformation en brasserie. Excellent chiffre d'affaires. Loyer 2.890 euros.  
Prix 690.000 euros honoraires charge vendeur . Contactez Dominique  
GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que chargée d'affaires  
indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389.  
Mandat n° 204762 Réf. annonce : 307DG204762B-DOR

↳

↳

92300\_

↳ LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION - Je vous propose ce\_

↳ restaurant de restauration rapide sandwicherie et saladerie idéalement situé\_

↳ dans un quartier de bureaux, écoles et commerçant. Superficie de 20 m2 avec 6\_

↳ places debout. Pas d'extraction. Loyer mensuel très faible à 998 euros.\_

↳ Cession à 127900 euros honoraires inclus à la charge du vendeur. Contactez Mme\_

↳ Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce\_

↳ et agissant en tant que chargée d'affaires indépendante auprès de la SAS\_

↳ PROPRIETES PRIVEES - Mandat n° 188156

	PRICE	SURFACE	CRAWL_SOURCE_original
575	687000.0	61.43	SE_LOGER
667	704000.0	60.72	SE_LOGER
1086	626000.0	60.72	SE_LOGER
1331	742000.0	65.53	SE_LOGER
441	687000.0	61.85	SE_LOGER
207	661000.0	64.05	SE_LOGER
257	487000.0	44.84	SE_LOGER
1535	527000.0	49.27	SE_LOGER
787	480000.0	41.23	SE_LOGER
717	481000.0	42.32	SE_LOGER
148	556000.0	46.98	SE_LOGER
1159	520000.0	45.59	SE_LOGER
196	516000.0	44.92	SE_LOGER
1123	526000.0	43.74	SE_LOGER
676	486000.0	45.79	SE_LOGER
618	558000.0	49.25	SE_LOGER
50	1331000.0	111.86	SE_LOGER
288	1359000.0	111.86	SE_LOGER
22	947000.0	85.00	SE_LOGER
114	8028.0	301.00	SE_LOGER
325	10999.0	412.40	SE_LOGER
833	5006.0	187.70	SE_LOGER
294	3019.0	113.20	SE_LOGER
149	5676.0	330.00	SE_LOGER
168	5676.0	330.00	SE_LOGER
479	2440.0	90.00	BIEN_ICI
1663	2440.0	90.00	BIEN_ICI
953	2440.0	90.00	SE_LOGER
1213	2440.0	90.00	SE_LOGER
1380	12499.0	320.00	SE_LOGER
550	12499.0	320.00	SE_LOGER
1634	642000.0	369.00	SE_LOGER
132	642000.0	369.00	SE_LOGER
925	4376.0	150.00	A_VENDRE_A_LOUER

215	12310.0	422.00	A_VENDRE_A_LOUER
1173	985000.0	87.11	SE_LOGER
469	965000.0	87.11	SE_LOGER
390	29265.0	1133.00	SE_LOGER
335	29265.0	1133.00	SE_LOGER
503	127900.0	20.00	BIEN_ICI
769	99200.0	25.00	EXPLORIMMO
1637	127900.0	20.00	LOGIC_IMMO
204	69000.0	35.00	EXPLORIMMO
698	99200.0	25.00	LOGIC_IMMO
858	69000.0	35.00	LOGIC_IMMO
1679	690000.0	100.00	BIEN_ICI
1370	79900.0	50.00	LOGIC_IMMO
359	99200.0	25.00	BIEN_ICI
1392	690000.0	100.00	LOGIC_IMMO
713	127900.0	20.00	EXPLORIMMO

Nous remarquons que notre modèle à tres bien pu detecter les annonces similaires. Surtout quand dans le groupe il y'a moins de 3 annonces similaires mais en cas de plus de 3 le modèle n'arrive plus à bien sonder la difference les descriptions de ses annonces sont presque similaires

```
[ ]: # Affichage du groupe le plus gros
X_train_full_clustered[X_train_full_clustered['CLUSTER'].value_counts().
↳index[1] == X_train_full_clustered['CLUSTER']] [['CLUSTER',
↳'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'SURFACE',
↳'CRAWL_SOURCE_original']].head(50)
```

```
[ ]: CLUSTER CLUSTER_DISTANCE \
204      39      0.000000
319      39      0.003838
359      39      0.000000
503      39      0.000016
698      39      0.003959
713      39      0.000016
769      39      0.000000
858      39      0.000000
1370     39      0.003838
1392     39      0.321471
1637     39      0.406992
1679     39      0.321471
```

```
DESCRIPTION \
204
92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce superbe
restaurant sans extraction avec 22 places dont une dizaine de couverts en
terrasse. Ce bien a une superficie totale de plus de 35 m2. Déco refait à neuf.
Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC. Idéal vente à emporter et
livraisons. A découvrir rapidement! Cession à 69000 euros honoraires inclus à la
```

charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n°199606

319

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de 50m2, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur. Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446

359

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION : Je vous propose ce restaurant de restauration rapide idéalement situé dans un quartier de bureaux, écoles et commerçants, avec 10/12 places intérieur et d'une superficie de 25 m2. Loyer mensuel 940 euros. Cession à 99200 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 188597

503

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION - Je vous propose ce restaurant de restauration rapide sandwicherie et saladerie idéalement situé dans un quartier de bureaux, écoles et commerçant. Superficie de 20 m2 avec 6 places debout. Pas d'extraction. Loyer mensuel très faible à 998 euros. Cession à 127900 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 188156

698

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION : Je vous propose ce restaurant de restauration rapide idéalement situé dans un quartier de bureaux, écoles et commerçants, avec 10/12 places intérieur et d'une superficie de 25 m2. Loyer mensuel 940 euros. Cession à 99200 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 188597 Réf. annonce : 2807FR188597B-FRZ

713

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION - Je vous propose ce restaurant de restauration rapide sandwicherie et saladerie idéalement situé dans un quartier de bureaux, écoles et commerçant. Superficie de 20 m2 avec 6 places debout. Pas d'extraction. Loyer mensuel très faible à 998 euros. Cession à 127900 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS

PROPRIETES PRIVEES - Mandat n° 188156

769

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION : Je vous propose ce restaurant de restauration rapide idéalement situé dans un quartier de bureaux, écoles et commerçants, avec 10/12 places intérieur et d'une superficie de 25 m2. Loyer mensuel 940 euros. Cession à 99200 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 188597

858

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce superbe restaurant sans extraction avec 22 places dont une dizaine de couverts en terrasse. Ce bien a une superficie totale de plus de 35 m2. Déco refait à neuf. Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC. Idéal vente à emporter et livraisons. A découvrir rapidement! Cession à 69000 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n°199606 Réf. annonce : 2807FR199606B-FRZ

1370

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de 50m2, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur. Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446 Réf. annonce : 2807FR187446B-FRZ

1392 92300 LEVALLOIS - BELLE AFFAIRE - Je vous propose un restaurant d'angle avec licence IV, offrant une surface de plus de 100 m<sup>2</sup> ouverte sur une vitrine de 10 mètres, ainsi qu'un sous-sol de 30 m2... Idéalement situé dans le quartier résidentiel et commerçant de Levallois-Perret, cet établissement de bonne réputation exploite actuellement près de 60 couverts, ainsi que 10 places en terrasse. Emplacement centre ville, exceptionnel et rare. Possibilité de transformation en brasserie. Excellent chiffre d'affaires. Loyer 2.890 euros. Prix 690.000 euros honoraires charge vendeur . Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389. Mandat n° 204762 Réf. annonce : 307DG204762B-DOR

1637

92300 LEVALLOIS PERRET - RESTAURATION RAPIDE SANS EXTRACTION - Je vous propose ce restaurant de restauration rapide sandwicherie et saladerie idéalement situé dans un quartier de bureaux, écoles et commerçant. Superficie de 20 m2 avec 6 places debout. Pas d'extraction. Loyer mensuel très faible à 998 euros. Cession à 127900 euros honoraires inclus à la charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce

et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 188156 Réf. annonce : 2807FR188156B-FRZ 1679 92300 LEVALLOIS - BELLE AFFAIRE - Je vous propose un restaurant d'angle avec licence IV, offrant une surface de plus de 100 m<sup>2</sup> ouverte sur une vitrine de 10 mètres, ainsi qu'un sous-sol de 30 m<sup>2</sup>... Idéalement situé dans le quartier résidentiel et commerçant de Levallois-Perret, cet établissement de bonne réputation exploite actuellement près de 60 couverts, ainsi que 10 places en terrasse. Emplacement centre ville, exceptionnel et rare. Possibilité de transformation en brasserie. Excellent chiffre d'affaires. Loyer 2.890 euros. Prix 690.000 euros honoraires charge vendeur . Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389. Mandat n° 204762

	PRICE	SURFACE	CRAWL_SOURCE_original
204	69000.0	35.0	EXPLORIMMO
319	79900.0	50.0	EXPLORIMMO
359	99200.0	25.0	BIEN_ICI
503	127900.0	20.0	BIEN_ICI
698	99200.0	25.0	LOGIC_IMMO
713	127900.0	20.0	EXPLORIMMO
769	99200.0	25.0	EXPLORIMMO
858	69000.0	35.0	LOGIC_IMMO
1370	79900.0	50.0	LOGIC_IMMO
1392	690000.0	100.0	LOGIC_IMMO
1637	127900.0	20.0	LOGIC_IMMO
1679	690000.0	100.0	BIEN_ICI

Nous pouvons aussi remarqué qu'avec le gros cluster notre 2e modèle à réussi à detecter les annonces similaire venant de plusieurs plateforme differentes mais aaussi failli en ajoutant les annonces qui sont presque similaires mais qui ne sont pas les memes.

```
[ ]:
```

```
[ ]: # Affichage du groupe le plus petit
X_train_full_clustered[X_train_full_clustered['CLUSTER'].value_counts().
↳ index[-1] == X_train_full_clustered['CLUSTER']][['CLUSTER',
↳ 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'SURFACE',
↳ 'CRAWL_SOURCE_original']].head(50)
```

```
[ ]:
    CLUSTER  CLUSTER_DISTANCE  \
1462      15                0.0
1605      15                0.0

DESCRIPTION
1462 RUE RIVAY - Au sein d'une copropriété bien tenue, au 3ème sous-sol, un BOX
```

mesurant 2,49m largeur, 4,6m longueur, 1,93 hauteur. A visiter sans tarder !!!!  
 1605 RUE RIVAY - Au sein d'une copropriété bien tenue, au 3ème sous-sol, un BOX  
 mesurant 2,49m largeur, 4,6m longueur, 1,93 hauteur. A visiter sans tarder !!!!

	PRICE	SURFACE	CRAWL_SOURCE_original
1462	44000.0	12.0	BIEN_ICI
1605	44000.0	12.0	EXPLORIMMO

Pour le groupe le plus petit, notre cluster a correctement détecté ces deux annonces similaires venant de deux plateformes différentes.

### 0.3.6 Prediction de regroupement les données de test

```
[ ]: def pred_cluster(clusterer, df, scaled_cols):
    df = copy.deepcopy(df).reset_index(drop=True)
    labels, strengths = hdbscan.approximate_predict(clusterer, df.
↳ drop(col_del+['CRAWL_SOURCE_original'], axis=1, errors='ignore'))

    # undo scaling
    _df = pd.DataFrame(
    scaler_transformer.transformers_[0][1].inverse_transform(df[numeric_cols]),
    columns=numeric_cols
    )

    df[scaled_cols] = _df[scaled_cols] # update X_train with unscaled values
    # add cluster column
    df['CLUSTER'] = labels
    # make a column that contains distance from the cluster centroid. -1
↳ distance means that the point is an outlier
    df['CLUSTER_DISTANCE'] = strengths
    return df
X_test_full_clustered = pred_cluster(clusterer, X_test_full, numeric_cols)
```

### 0.3.7 Affichage de resultats de la prediction du 2è modèle avec nos données de test

```
[ ]: # Affichage de resultats de la prediction
X_test_full_clustered.sort_values(by='CLUSTER', ascending=False)[['CLUSTER',
↳ 'CLUSTER_DISTANCE', 'DESCRIPTION', 'PRICE', 'CRAWL_SOURCE_original']].
↳ head(40)
```

```
[ ]:   CLUSTER  CLUSTER_DISTANCE  \
286      50          1.000000
188      50          0.495147
348      50          1.000000
167      48          0.771735
322      39          1.000000
212      39          0.662596
```

206	39	0.662603
404	31	1.000000
32	7	0.921642
428	7	0.162496
60	3	1.000000
317	-1	0.000000
285	-1	0.000000
293	-1	0.000000
292	-1	0.000000
291	-1	0.000000
290	-1	0.000000
289	-1	0.000000
288	-1	0.000000
287	-1	0.000000
320	-1	0.000000
284	-1	0.000000
295	-1	0.000000
283	-1	0.000000
282	-1	0.000000
281	-1	0.000000
280	-1	0.000000
279	-1	0.000000
278	-1	0.000000
277	-1	0.000000
276	-1	0.000000
321	-1	0.000000
275	-1	0.000000
274	-1	0.000000
294	-1	0.000000
296	-1	0.000000
319	-1	0.000000
307	-1	0.000000
315	-1	0.000000
314	-1	0.000000

#### DESCRIPTION \

286

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

188

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et



les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

348

Une adresse de choix pour devenir propriétaire. La station de métro, le lycée et les commerces sont accessibles à pied dans un rayon de 500 m\*. Vous apprécierez la qualité résidentielle de la commune et choisirez parmi une typologie variée d'appartements du studio au 5 pièces duplex. Les espaces intérieurs sont pour la plupart prolongés d'agréables extérieurs, balcons, terrasses ou jardins privatifs. Un coeur d'îlot paysager et des espaces communs végétalisés avec cheminements piétons desservent les différents halls d'entrée sécurisés.

\*Source: Google Maps.

167

EVOLIS vous propose une surface de bureaux rénovée avec terrasse technique dans un immeuble situé à toute proximité de la Porte de Champerret. Les locaux sont très lumineux et climatisés.

322

92300 LEVALLOIS - BELLE AFFAIRE - Je vous propose un restaurant d'angle avec licence IV, offrant une surface de plus de 100 m<sup>2</sup> ouverte sur une vitrine de 10 mètres, ainsi qu'un sous-sol de 30 m<sup>2</sup>... Idéalement situé dans le quartier résidentiel et commerçant de Levallois-Perret, cet établissement de bonne réputation exploite actuellement près de 60 couverts, ainsi que 10 places en terrasse. Emplacement centre ville, exceptionnel et rare. Possibilité de transformation en brasserie. Excellent chiffre d'affaires. Loyer 2.890 euros. Prix 690.000 euros honoraires charge vendeur . Contactez Dominique GIRAUD-DESERT au 06 62 23 01 49 agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES. CPI : 75012016000006389. Mandat n° 204762

212

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce restaurant de restauration rapide avec extraction dans un quartier d'affaires. Ce restaurant d'une superficie de 50m<sup>2</sup>, dispose de 6/8 places assises et terrasse. Matériel très récent. Bail échéance en 2028. Aucun travaux à prévoir! Loyer mensuel 1990 euros. A découvrir! Cession à 79900 euros honoraires inclus à la charge du vendeur. Affaire à saisir rapidement! Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n° 187446

206

92300 LEVALLOIS PERRET - VRAI COUP DE COEUR : Je vous propose ce superbe restaurant sans extraction avec 22 places dont une dizaine de couverts en terrasse. Ce bien a une superficie totale de plus de 35 m<sup>2</sup>. Déco refait à neuf. Aucun travaux à prévoir! Loyer mensuel 1540 euros TTC. Idéal vente à emporter et livraisons. A découvrir rapidement! Cession à 69000 euros honoraires inclus à la

charge du vendeur. Contactez Mme Frédérique Rézé au 06 72 37 96 11, spécialisée en cession de fonds de commerce et agissant en tant que chargée d'affaires indépendante auprès de la SAS PROPRIETES PRIVEES - Mandat n°199606

404

Vente de murs libres au jour de l'acquisition - Local d'activité - Ambiance société événementielle - utilisé à usage professionnel actuellement - possibilité bureaux; libéral, artistique, ou artisanal - Près Porte de Champerret. Surface atypique de 242 m<sup>2</sup>, dont 200 m<sup>2</sup> en rez-de-chaussée et 42 m<sup>2</sup> à l'étage avec entrée directe par la rue. Locaux entièrement rénovés en 2018.2 niveaux. Rez-de-chaussée: belle entrée. Espace cuisine. Salle de réunion. Grand open space. Salle de stockage. 2 WC. A l'étage: open-space + bureau fermé. Climatisation réversible, câblage informatique et baie de brassage. Très belle hauteur sous plafond. Très lumineux. Alarme et vidéosurveillance. Réf: MFA113305.

32

Louez au mois un parking privé au pour €. Pas d'engagement, pas de caution, louez ce parking au mois et sans engagement. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-4270-1 Trouver un garage à louer n'aura jamais été aussi simple. Quelle que soit la taille de votre voiture neuve ou d'occasion, les emplacements sont suffisamment grands pour vous permettre de manoeuvrer sans difficulté. Tous nos parkings sont surveillés par un gardien et propres, il est donc très rassurant d'y laisser sa voiture.

428

Louez au mois un parking privé au pour €. Pas d'engagement, pas de caution, louez ce parking au mois et sans engagement. 2 journées d'essai gratuit pour tester le parking avant l'abonnement. Envie de tester gratuitement ce parking ? Retrouvez toutes les informations sur le site Yespark avec la référence 92300-4018-1 Trouver un garage à louer n'aura jamais été aussi simple. Quelle que soit la taille de votre voiture neuve ou d'occasion, les emplacements sont suffisamment grands pour vous permettre de manoeuvrer sans difficulté. Tous nos parkings sont surveillés par un gardien et propres, il est donc très rassurant d'y laisser sa voiture.

60

Grand 2 pi?ces haut standing + Parking + Cave R?sidence de haut standing en front de seine pr?s de la place Georges Pompidou. Metro Pont de Levallois ? 50 m. Magnifique 2 Pi?ces (NON MEUBLE) d'angle de 57 m? en ?tage ?lev? avec 2 grands balcons (16 m2). Ascenseur. Tr?s lumineux car ?tage ?lev? et double exposition (Sud-Est) et calme. PARKING et CAVE (4 m2) en sous-sol. Libre Photos suppl?mentaires sur demande par mail

317

Levallois, quartier Mairie. EngelVolkers vous propose un appartement traversant de 2/3 pièces avec balcon filant d'une superficie d'environ 68 m2 situé au 4ème étage, avec ascenseur d'un immeuble ancien. Cet appartement se compose d'un double séjour, d'une cuisine équipée, d'une chambre au calme sur cour, d'une salle de bains et wc indépendants. Ce bien, calme et lumineux avec une

exposition Sud possède une hauteur sous plafond de 2.92m. Une cave complète cet appartement pour lequel quelques travaux sont à prévoir. Les parties communes de l'immeuble, sécurisé par vigik et digicode ont été entièrement rénovées. Honoraires de transaction à la charge de l'acquéreur: 4%. Prix net vendeur: 760 000€. Levallois- proximité de toutes les commodités. Quartier Mairie, proche du métro Louise Michel (ligne 3) et à 10 minutes de la gare Levallois/Clichy. Réf. annonce : W-02FZ90

285

Place de Parking à vendre.\n\nSous un immeuble d'habitation bien entretenu à l'angle de la rue d'Alsace et de la rue Victor Hugo à Levallois-Perret 92300,\nà 150 mètres de la Porte d'Asnières/Paris,\nà coté du centre commercial So-Ouest.\n\nAccès sécurisé 24h/24 7j/7 par clé-bip électronique,\nascenseur piéton,\ngardien,\ncaméra de surveillance. \n\nL'accès et les parkings sont très larges (pas de murs « trop » près pour rayer les peintures)\nLa circulation facile\nParking accessible de la berline au 4X4.\nPossibilité de moto ou scooter + berline de moyenne taille sur une place.\n\nTous les gros travaux de copropriété ont été effectué il y a 1 an environ:\n\_Ravalement de l'immeuble.\n\_Cabine d'ascenseur refaite.\n\_Electricité du parking toutes refaite au normes actuelles et passé en basse tension.\n\_Jardin extérieur refait et clôturé.\n\_Caméra de surveillances\n\nIl n'y a aucun gros travaux à prévoir pour de nombreuses années.\n\nBonne rentabilité locative dans un quartier avec très forte de demande de location de parking.\n\n23 000 €\n\nTaxe foncière: 128€ par an\nCharges de copropriété: 10€ par mois\n\nContact téléphonique de préférence. (Pas de SMS)\n\n(AGENCE, EXPERTISE ET DÉMARCHAGE DIVERS S'ABSTENIR)\n\nparking/pas un box/place de voiture/moto/

293

CE QUE NOUS AIMONS: La localisation des bureaux en plein coeur de Levallois-Perret Le charme de la cours commune dotée d'un jardin La proximité du métro.

292

dans un immeuble ancien un appartement meublé de 13.32 m<sup>2</sup>, comprenant : une pièce à vivre avec coin cuisine et une salle d'eau avec wc Assurance Multirisques Privilège \*: 10.1 € / mois \*Services facultatifs, voir conditions en agences

291

Au calme complet, au sein d'une résidence de standing, superbe 2 pièces d'environ 55m2 en parfait état. Exclusivement sur cour arborée, ce bien vous offre une belle pièce de vie (exposée Est) donnant sur balcon d'environ 6m2 avec une cuisine semi ouverte intégralement équipée. Une chambre de belle taille (13 m2) avec rangements, une salle d'eau et un WC séparé. Une cave et une place de parking au sein de l'immeuble complètent ce bien. A noter que l'appartement a été refait par ses propriétaires récemment, sans aucun travaux à prévoir. Disponible début Septembre. Contactez nous au 06 95 75 55 87 pour visiter ce bien.

290

Idéalement placé , rue Aristide Briand face à la mairie de Levallois Perret - Local professionnel de 44,56 m2 situé en rez de chaussée sur rue comprenant : Une entrée , deux pièces, une cuisine /débaras, un wc, une cave . Chauffage et eau chaude individuels gaz. Digicode et interphone. Accès au local par le hall

de l'immeuble. Actuellement occupé par un stomatologue. Disponible le 1<sup>er</sup> août 2019.

289

Au calme complet, au sein d'une résidence de standing, superbe 2 pièces d'environ 55 m<sup>2</sup> en parfait état. Exclusivement sur cour arborée, ce bien vous offre une belle pièce de vie (exposée Est) donnant sur balcon d'environ 6 m<sup>2</sup> avec une cuisine semi ouverte intégralement équipée. Une chambre de belle taille (13 m<sup>2</sup>) avec rangements, une salle d'eau et un WC séparé. Une cave et une place de parking au sein de l'immeuble complètent ce bien. A noter que l'appartement a été refait par ses propriétaires récemment, sans aucun travaux à prévoir.

Contactez nous pour visiter ce bien. Honoraires TTC en sus : 806.70 €.

Disponible le 01/09/2019.

288

Appartement de 3 pièces , 72.60 m<sup>2</sup>, 5ème étage, avec balcon offrant une vue dégagée sur la Défense et l'île de la Jatte . SISLEY s'organise autour d'un bel espace arboré. L'immeuble est situé à proximité de toutes commodités : transports, commerces, écoles. Balcon : 9,10 m<sup>2</sup>. Réf. annonce : 198378\_082

287

3 pièces 65 m<sup>2</sup> avec balcon et parking Au pied des commerces et à quelques minutes du métro Louise Michel, dans une résidence semi-récente avec gardien, au 4<sup>ème</sup> étage avec ascenseur, appartement de 65 m<sup>2</sup> comprenant une entrée, 27 m<sup>2</sup> de pièce de vie (séjour avec cuisine semi-ouverte), un balcon, un wc séparé, un dressing, 2 chambres, une salle de bains, un emplacement de parking et une cave. DPE : NC. A VISITER RAPIDEMENT ! Jonathan Aisenberg et toute l'équipe de L'AGENCE DE LEVALLOIS sont à votre écoute tous les jours, samedi, dimanche et jours fériés de 10h à 21h !

320

Levallois-Perret (92), A proximité du métro Pont de Levallois, vous recherchez une très belle affaire avec un fort potentiel de croissance, cette affaire est pour vous ! Belle clientèle d'affaires le midi et mixte le soir. Carnet de réservation rempli un mois à l'avance pour la clientèle d'affaires. Excellente notoriété ! Plus de 100 places assises. Surface utile de 260 m<sup>2</sup> dont deux salles de 103 m<sup>2</sup> et une terrasse extérieure chauffée (plus de 30 places). Extraction avec gros diamètre. Licence IV. Loyer mensuel 4100 euros hors charge. CA de plus de 850 000 euros (ouvert 232 jours en moyenne). Restaurant fermé actuellement tous les week-ends et jours fériés et 3 semaines annuelles. Facilité de stationnement ! Possibilité d'acheter les murs à moyen ou long terme. Excellent emplacement !

284

Levallois-Perret • Louise Michel • maison de ville 127 m<sup>2</sup> • 4 chambres. Idéalement située dans une rue calme à proximité du métro Louise Michel, maison de ville en parfait état comprenant : grande pièce à vivre de 60 m<sup>2</sup>, cuisine américaine avec verrière atelier, chambre parentale avec salle de bains et dressing, 3 chambres enfants, 2 salles de douche, 2 wc. Maison entièrement rénovée avec des matériaux de qualité. Nombreux rangements. Proche commerces, transports et écoles. Pour plus d'informations ou pour organiser une visite, merci de contacter Alain ETJEMESIAN au 06 49 35 21 13 ou votre interlocuteur

habituel chez BOSETTA IMMOBILIER. Nous sommes également à votre disposition si vous souhaitez vendre ou faire estimer un bien immobilier à Paris 17ème, 16ème, 8ème, 9ème, Neuilly ou Levallois.

295

Je vous propose un local commercial de 110 M2 en plein coeur de Levallois Perret. Situé sur la rue principale, dans un environnement très commerçant avec énormément de sociétés et de résidents. Actuellement l'activité est vente de Meubles et décorations mais il y a la possibilité de faire tous commerces sauf nuisance et notamment restauration rapide sans extraction. Le magasin a une belle visibilité avec une façade de 10m. Affaire à saisir rapidement compte tenu de son emplacement, de sa clientèle et de sa superficie. Loyer: 3200 euros charges comprises. M. Yankel KOSKAS 0637729002 / 0180916809 poste 8326

commerces.75017@maxihome.email Agent commercial indépendant du réseau national clairimmo Maxihome N° RSAC : 815 116 13200011 PRIUM - Ville du greffe : PARIS 283

Dans un Appartement, 2 pièces, Chambres : 1, vide, de 32m<sup>2</sup>, Clair, Calme, Bon état, Pièce principale, Cuisine, Salle d'eau, WC dans la salle d'eau, Carrelage, Parquet, Placards, Chauffage : collectif, Avec ascenseur, Situé au 7ème étage, Dans un immeuble Récent Proches des transports et toutes commodités. Réf. annonce : 25060

282

Au pied du métro Anatole France et proche de toutes commodités, un appartement familial de 5 pièces d'une superficie de 99.90m<sup>2</sup> situé au 1er étage, comprenant une entrée avec placard, une cuisine séparée meublée et équipée (plaque de cuisson, hotte, four), un séjour, 4 chambres avec placards, une salle de bains, une salle de douche avec wc ainsi qu'un wc séparé. Ce bien est proposé avec une cave et un emplacement parking.

281

Le groupe POINT DE VENTE vous propose un local commercial en angle d'une surface de 52 m<sup>2</sup> au rez-de-chaussée, Cave accessible par une trappe intérieure d'environ 45 m<sup>2</sup>. Le bien se situe dans un environnement très commerçant à quelques minutes à pied du métro Anatole France et des enseignes connues telles que La Poste, l'Atelier du Sourcil, Monoprix, Zara Home.. Restauration rapide sur dossier. Honoraires TTC en sus : 15500.00 €.

280

Le Groupe Babylone vous propose une cession de bail proche du centre commerciale SO OUEST, 50 m<sup>2</sup> rez-de-chaussée, loyer: 1 500 euros/ mois CC. Cession de bail 120 000 euros. Pas d'extraction Activités autorisées: prêt à porter, accessoire de mode..

279

Dans un immeuble de bon standing et rénové, doté d'un jardin/patio, des surfaces de bureaux rénovées à louer. A proximité immédiate du métro, et dans un quartier très commerçant; ant, nous vous proposons des surfaces de bureaux rénovées au sein d'un immeuble de bon standing. Les parties communes ainsi que le RIE viennent d'être rénovés. Grand hall d'accueil avec hôtesses d'accueil et PC sécurité.

278

Au pied du square Edith de Villepin ! Appartement familial de 5 pièces (dont 3 chambres) Dans une petite copropriété, charmante et bien entretenue (les 4 propriétaires y habitent) se trouve au 3e et dernier étage (sans ascenseur), cet appartement de 83.01 m<sup>2</sup> loi Carrez et 88.64m<sup>2</sup> au sol Lumineux, esprit scandinave, traversant, calme, très bon état avec son plan idéal (petite entrée avec rangements, cuisine équipée avec table, double séjour (salle à manger, salon spacieux), 2 chambres enfant, chambre parents avec rangement, salle de bains). Faibles charges Appartement fonctionnel et rempli de " hygge ", grand palier au dernier étage et jardinet commun au rez-de-chaussée. Bien soumis au statut de la copropriété. Dont 2.53 % honoraires TTC à la charge de l'acquéreur. Copropriété de 5 lots. Charges annuelles: 2150.00 euros.

277

En exclusivité à Levallois, vente d'un fonds de commerce en restauration d'environ 89 m<sup>2</sup> de plain-pied. 60 couverts en salle et 60 couverts en terrasse. Licence IV et extraction. Changement d'activité possible. Bail 3/6/9, courant jusqu'en septembre 2020. Honoraires d'agence à la charge du vendeur.

276

Grande place de parking située au niveau -1. Large rampe et manoeuvre facile. Sécurisé. Entrée Rue Pierre Brossolette.\nPossibilité de louer la place mitoyenne.

321

Bureaux à louer à Levallois-Perret (92300) dans les Hauts de Seine. Situés entre le métro "Pont de Levallois" et la gare SNCF "Clichy-Levallois", immeuble de standing proposant des plateaux rénovés, lumineux, rationnels et fonctionnels de plus de 1 500 m<sup>2</sup>. Twogether 2 bénéficie de nombreux services dont un restaurant inter-entreprises, une cafétéria, un business center ainsi qu'un vaste hall d'accueil avec une hôtesse. "#@##@#.

275

Appartement, 2 pièces, Meublé, de 36m<sup>2</sup>, clair et calme sans vis à vis, en très bon état avec parties communes de l'immeuble refaite à neuf, comprenant une pièce principale une cuisine US équipée , Salle d'eau, WC, Placards, Chauffage : Individuel électrique, Situé au 1er étage d'un immeuble sécurisé avec 2 porte d'accès. Proche des transports et des commodités

274

A deux pas de la Gare et de ses commerces, joli T2 non meublé, entièrement rénové au sein d'un immeuble sécurisé avec ascenseur. Ce grand 2 pièces vous offre une entrée avec rangements, un séjour de belle taille exposé plein Ouest, une cuisine aménagée dinatoire, un cellier, une chambre avec armoire faite sur mesure, un balcon et une salle de bains avec WC. Appartement lumineux qui profite de nombreux rangements. Disponible immédiatement.

294 LEVALLOIS PERRET (92300) RESIDENCE SERVICES SENIORS - 3 P\n\nRESIDENCE SERVICES Levallois Perret, à l'orée de Neuilly , dans une résidence services seniors non médicalisée, pour seniors autonomes et valides (sécurité 24h/24 toute l'année, restauration, homme d'entretien pour petit bricolage et autres services), en étage élevé avec ascenseurs, magnifique 3 pièces de 77 m<sup>2</sup> en parfait état , avec double exposition sur terrasse et balcon : entrée avec placards/penderie, cuisine équipée, séjour avec balcon de 6,60 m<sup>2</sup>, une chambre

de 13 m<sup>2</sup> avec terrasse de 9 m<sup>2</sup>, une chambre de 13 m<sup>2</sup> avec placard/penderie et accès au balcon du séjour (cette pièce peut être utilisée comme salle à manger, une porte coulissante la séparant du séjour), dégagement avec placards/penderies, salle de bains avec WC, salle d'eau avec WC . Double vitrage. Volets roulants électriques . Cave au sous-sol.\nPossibilité d'acquérir une place de parking située au sous-sol, au prix de 23.500 €.\nCompte tenu des services proposés, les charges mensuelles s'élèvent à 1.370 € (charges de copropriété et charges de services).\nRésidence seniors bien située à l'orée de Neuilly, proche commerces et transports (métro Anatole France, ligne 3).\nContacter Janine GRANDAUBERT au 06 89 22 30 73 ou jgrandaubert@orange.fr\nRéférence annonce : RSIM762\nLe prix indiqué comprend les honoraires à la charge de l'acheteur : 4,71% TTC du prix du bien hors honoraires\nPrix hors honoraires : 425 000 €

296

Dans résidence de standing, appartement de type 5 pièces composé d'une entrée avec WC séparés, d'un séjour de 46m<sup>2</sup> offrant accès à balcons exposé S/O, cuisine dînatoire de 18m<sup>2</sup>, 3 chambres avec rangements dont une suite parentale, 2 salle de bains avec WC. Cave et parking en sous-sol. David TEIXEIRA 06.75.72.43.45. Mise à jour le 08/06/2019 Demander des infos sur ce bien Référence annonce AVENDREALOUE : 1308680

319

LEVALLOIS | MAIRIE | 2 PIÈCES Appartement 2 pièces de 50 m<sup>2</sup> situé dans un immeuble ancien de standing avec ascenseur au pied de la Mairie et des commerces. Ce bien vous offre: entrée, séjour, cuisine dînatoire, 1 chambre (poss. 2), une salle de bains et des toilettes séparées. Très calme, beaux volumes ! Une cave complète ce bien. Charges annuelles.. € Soumis au statut de la copropriété (Aucune procédure en cours).. Lots de copropriété Prix de vente 480.000 Euros dont 2.13 % d'honoraires à la charge de l'acquéreur.. Building Partners, fort d'une expertise de plus de 20 ans, est spécialisé dans la transaction de biens "Familiaux de Standing" et bénéficie d'implantations de choix dans l'ouest parisien. C'est avec plus de 450 appartements, maisons et hôtels particuliers que Building Partners offre le plus large choix de biens à la vente. 65, rue Président wilson 92300 Levallois (33) 1.47.31.01. 74.

307

Levallois / Square Edith de Villepin Dans un immeuble récent de bon standing avec gardien, un appartement de 52.86 m<sup>2</sup> situé au 2ème étage avec ascenseur. Il se compose de : une entrée avec rangement, un séjour lumineux, une grande cuisine indépendante, une chambre, une salle de bains et des wc séparés. Une cave complète ce bien. Parking en SUS du prix : 20.000 € Charges annuelles : - € Soumis au statut de la copropriété Prix de vente : 449.000 € dont 3.08 % d'honoraires à la charge de l'acquéreur. Building Partners, fort d'une expertise de plus de 20 ans, est spécialisé dans la transaction de biens "Familiaux de Standing" et bénéficie d'implantations de choix dans l'ouest parisien. C'est avec plus de 450 appartements, maisons et hôtels particuliers que Building Partners offre le plus large choix de biens à la vente. 65, rue Président wilson 92300 Levallois (33) 1.47.31.01. 74.

315

LEVALLOIS Jean Jaures / Pdt Wilson: au 2ème étage, un appartement 3 pièces de 43m<sup>2</sup> traversant comprenant: entrée, SDB, WC, séjour avec cuisine, 2 chambres, cave. travaux à prévoir.

314

Location parking en sous sol à Levallois Perret, à côté de la boulangerie Kayser place de la Liberation.

	PRICE	CRAWL_SOURCE	original
286	483000.0	SE	LOGGER
188	576000.0	SE	LOGGER
348	477000.0	SE	LOGGER
167	11001.0	SE	LOGGER
322	690000.0	EXPLORIMMO	
212	79900.0	BIEN_ICI	
206	69000.0	BIEN_ICI	
404	2067000.0	SE	LOGGER
32	65.0	EXPLORIMMO	
428	100.0	EXPLORIMMO	
60	1690.0	LE_BON_COIN	
317	791000.0	LOGIC_IMMO	
285	23000.0	LE_BON_COIN	
293	2256.0	SE	LOGGER
292	712.0	EXPLORIMMO	
291	1850.0	MEILLEURS_AGENTS	
290	1211.0	EXPLORIMMO	
289	1850.0	SE	LOGGER
288	755000.0	LOGIC_IMMO	
287	560000.0	MEILLEURS_AGENTS	
320	585000.0	SE	LOGGER
284	1290000.0	MEILLEURS_AGENTS	
295	65000.0	BIEN_ICI	
283	920.0	LOGIC_IMMO	
282	2927.0	MEILLEURS_AGENTS	
281	4509.0	SE	LOGGER
280	220000.0	SE	LOGGER
279	92925.0	SE	LOGGER
278	769000.0	SE	LOGGER
277	550000.0	SE	LOGGER
276	150.0	LE_BON_COIN	
321	95577.0	SE	LOGGER
275	1100.0	EXPLORIMMO	
274	1390.0	BIEN_ICI	
294	445000.0	LE_BON_COIN	
296	1350000.0	A_VENDRE_A_LOUER	
319	480000.0	SE	LOGGER
307	449000.0	MEILLEURS_AGENTS	
315	415000.0	MEILLEURS_AGENTS	



Nous avons pu prédire et affecter nos données de test dans des clusters dont les annonces correspondaient ou ils ont été considérés comme étant des annonces uniques.

## 0.4 Conclusion et Perspectives d'Amélioration

Le premier modèle, qui utilise l'encodage par "embedding" exclusivement pour la colonne "DESCRIPTION", a réussi à regrouper efficacement des annonces similaires, bien que parfois il ait associé des annonces qui ne l'étaient pas, mais dont les descriptions étaient très semblables. Selon nos résultats, c'est le modèle qui s'en est le mieux sorti.

Le deuxième modèle, qui se base sur la similarité cosinus, a bien regroupé les annonces similaires, mais a échoué à détecter d'autres annonces similaires s'il n'y avait pas de proximité évidente dans les descriptions. Il a également eu du mal à distinguer des annonces avec des descriptions lexicalement proches mais non identiques.

Pour une utilisation fréquente, le premier modèle s'avère plus adapté si l'objectif est de prédire le groupe de nouvelles annonces inconnues, car il n'y a pas besoin de réentraîner l'encodage par "embedding", et le modèle généralisera mieux vers de nouvelles données.

Le deuxième modèle est plus pertinent si l'objectif est de détecter les similarités dans les descriptions des données connues pour lesquelles la similarité cosinus a été calculée. Toutefois, il n'est pas adéquat pour des données inconnues, car il ne sera pas en mesure de détecter des similitudes approximatives.

Ainsi, le premier modèle est préférable pour des prédictions, en raison de sa capacité à généraliser, tandis que le deuxième modèle est utile pour analyser des données connues.

### Perspectives d'Amélioration :

- L'exploration de la variable IMAGE, qui n'a pas été abordée dans ce travail, pourrait s'avérer très utile. Si nous pouvons établir des liens de similarité et détecter des images similaires, cela pourrait être une extension précieuse de cette analyse.
- Un traitement plus approfondi de la colonne DESCRIPTION est recommandé. Il est évident que des erreurs d'orthographe subsistent, probablement dues à la récupération des données par l'API. Par exemple, des mots mal orthographiés, des caractères inattendus comme des caractères spéciaux mélangés avec des mots, et d'autres petites erreurs de structure nécessitent une correction.

Je suis ouvert aux appréciations et aux critiques constructives, car je suis conscient que mes compétences dans ce domaine sont en constante évolution. J'aspire à travailler sur des projets stimulants afin de continuer à progresser.

Je reste à votre disposition pour discuter en détail de ce travail si vous le souhaitez.

```
[ ]: ! pip install nbconvert >/dev/null 2>/dev/nul
! apt-get install texlive-xetex texlive-fonts-recommended texlive-plain-generic
↪>/dev/null 2>/dev/nul
```

```
[ ]: !jupyter nbconvert --to pdf "/content/drive/MyDrive/Colab Notebooks/  
↳ Ads_clustering_Joel_KabanguKabobo.ipynb" >/dev/null 2>/dev/nul
```