

# LES CRITÈRES QUI IMPACTENT LE PRIX DE L'IMMOBILIER

Licence 3 MIAGE Projet Commun

Seohyun PARK



# CONTEXTE

On sait qu'à Paris le prix de l'immobilier peut fluctuer selon ses arrondissements :

- Mais alors **quels sont critères qui influent le prix au m2?**
- Quelle pourrait être l'évolution de ces prix dans le temps ?



# LA PROBLÉMATIQUE :

Comment les **nouvelles technologies** et l'analyse des données nous permettent de découvrir les **critères** et les **corrélations** qui **influent** sur le **prix** de l'immobilier dans les arrondissements **parisiens** ?



# PLAN



**01**

## ANALYSE DU TERRAIN

- Enquête du terrain
- Enjeux et objectifs

**02**

## METHODOLOGIES

- Technologies
- CRISP-DM
- Vue statique/dynamique

**03**

## ARCHITECTURES

- Structure des données
- Démo

**04**

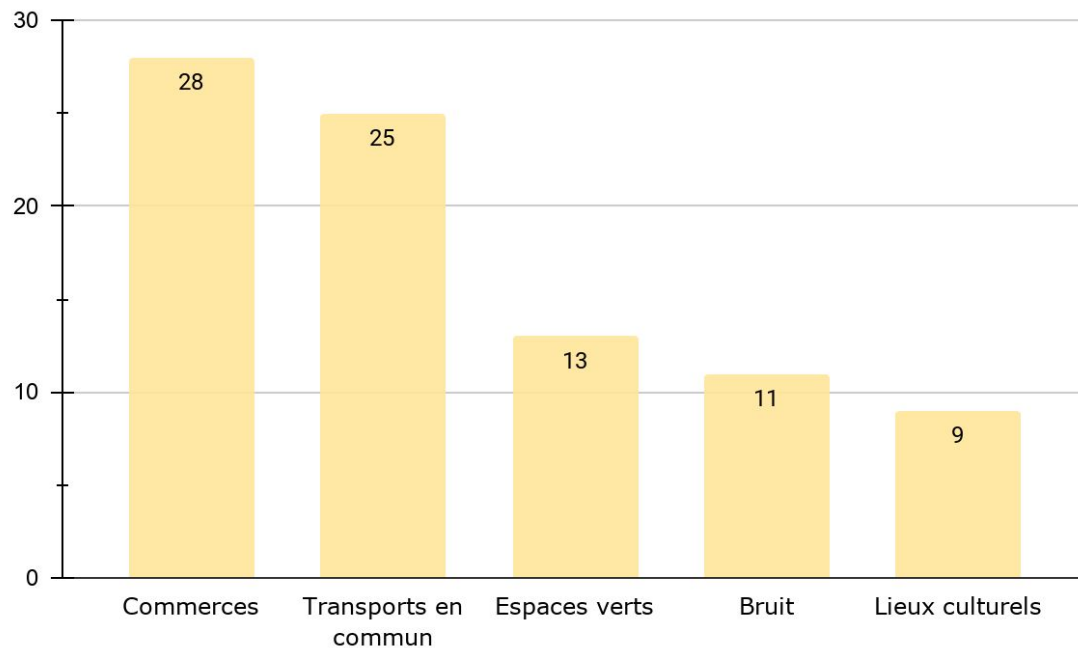
## BILAN

- Résultats
- Retours

# **01. ANALYSE DU TERRAIN**



## RÉSULTATS D'ANALYSE DU TERRAIN



*TOP 5 Critères de choix d'un nouveau logement (plusieurs réponses possibles)*

## ENJEUX & OBJECTIFS



### ENJEUX

Découvrir / confirmer les critères et les corrélations qui influent sur le prix de prix immobilier.



### OBJECTIFS

- Définir des critères objectifs et subjectifs dans la recherche d'un bien chez une personne
- Collecter des données et les transformer en jeu de donnée
- Afficher différentes données sur un tableau du bord selon différents arrondissements
- Calculer les corrélations entre le prix de l'immobilier et les critères sélectionnés

## 02. MÉTHODOLOGIE





## TECHNOLOGIES UTILISÉES



### LANGAGE DE PROGRAMMATION

Data Scraping, Nettoyage et analyse  
des données



### INDEXATION

Indexation des documents  
orientés textes



### BASE DE DONNÉES

Gestion d'un volume important de  
données structurées

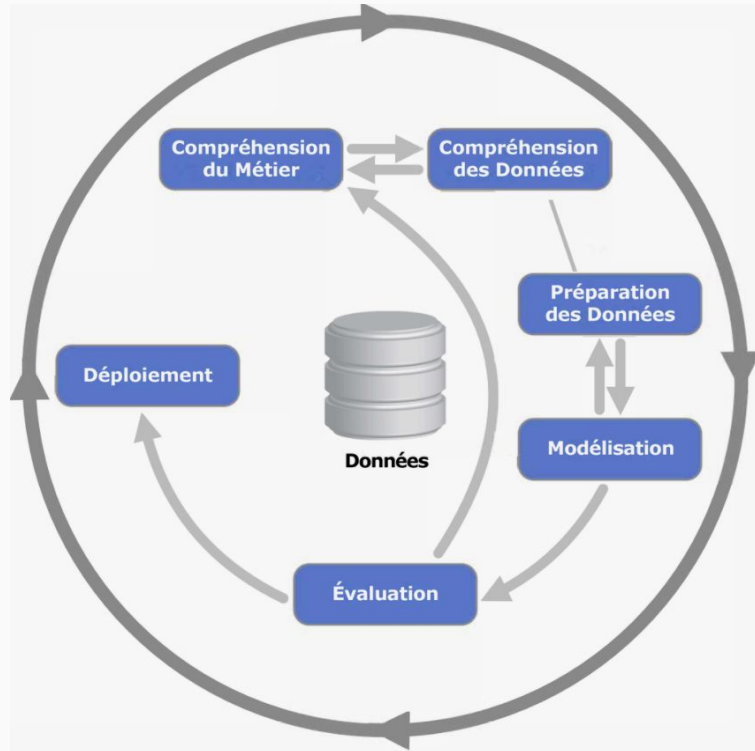


### VISUALISATION

Agrégation des résultats dans des  
graphes et des diagrammes

# Méthodologie data

## CROSS INDUSTRY STANDARD PROCESS FOR DATA MINING



I. COMPRÉHENSION DU MÉTIER

II. COMPRÉHENSION DES DONNÉES

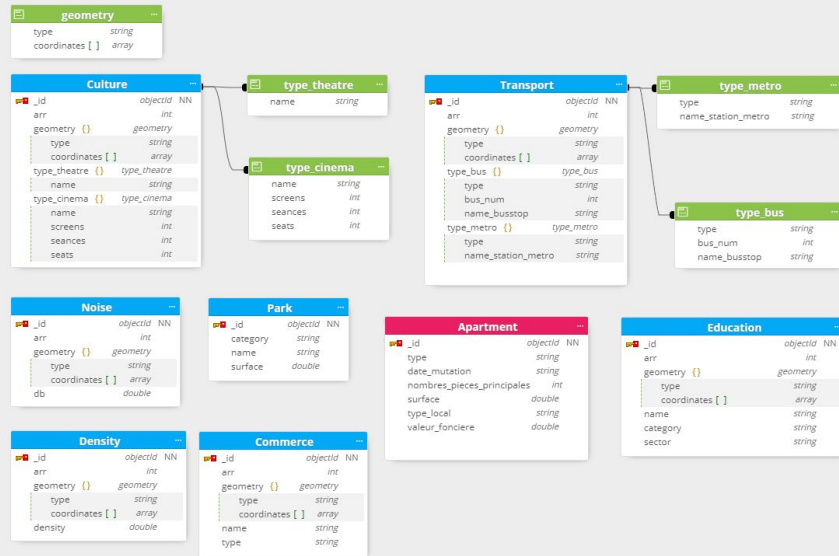
III. PRÉPARATION DES DONNÉES

IV. MODÉLISATION

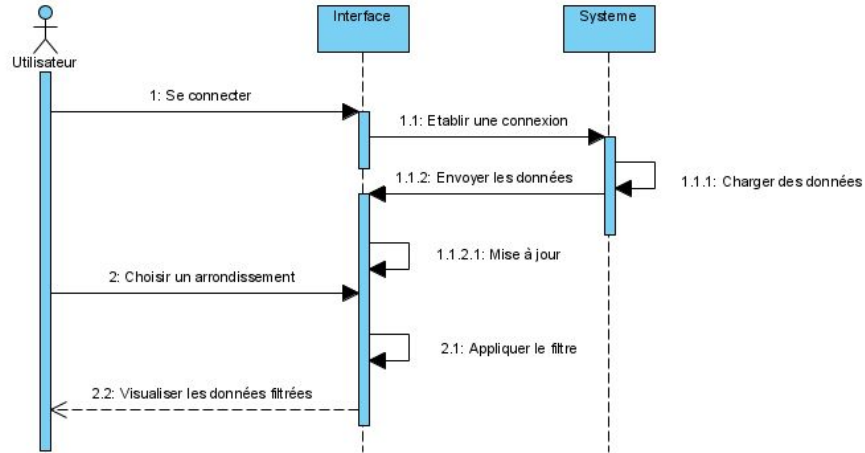
V. ÉVALUATION

VI. DÉPLOIEMENT

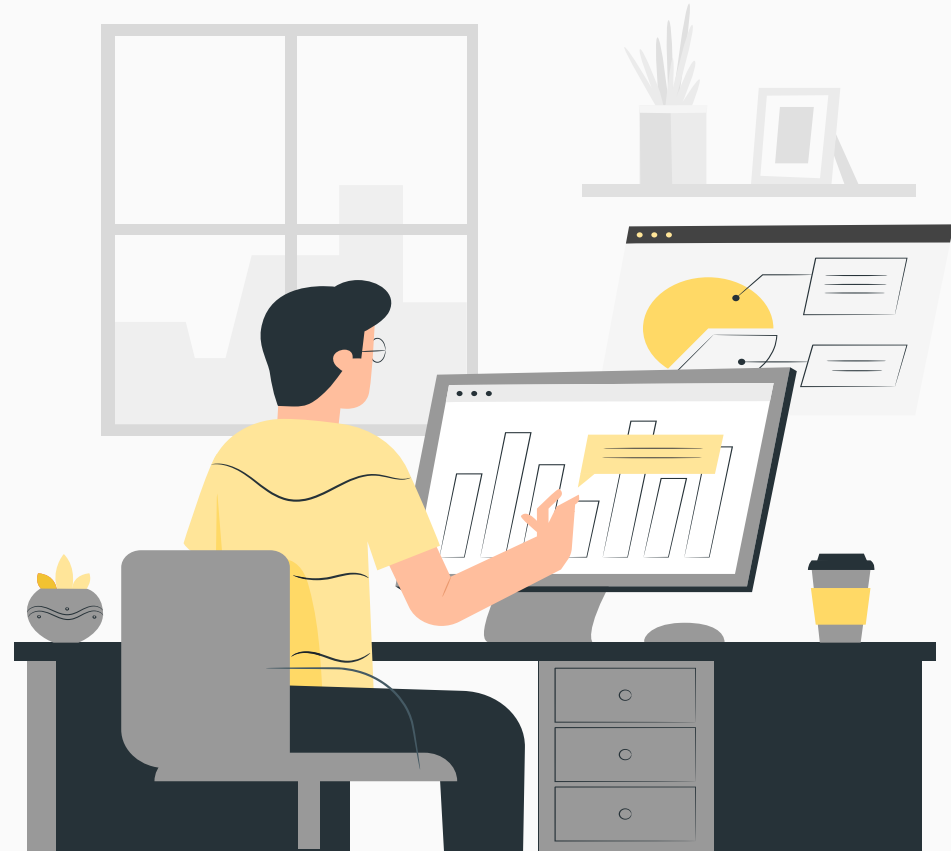
# SCHÉMA DE BASE DE DONNÉES



# FONCTIONNALITÉS



# 03. ARCHITECTURE



## PRÉPARATION DES DONNÉES

```
{
  "_id": {
    "$oid": "603d0edf6dfbfc34309f2ed5"
  },
  "datasetid": "accessibilite-des-arrets-de-bus-ratp",
  "recordid": "883d22613d71765c4e35b9e19421e85c1e500c1e",
  "fields": {
    "annonce_visuelle_situations_perturbees": 1,
    "accessibilite_ufr": 1,
    "idamivif": "IDAMIVIF",
    "doublon": "Ok",
    "annonce_sonore_situations_perturbees": 1,
    "stif": 100100021,
    "gipa": 11053,
    "ligne": 21,
    "annonce_sonore_prochain_passage": 1,
    "coord": [48.87108938611164, 2.3315871332906513],
    "x": 599639,
    "idptar": 7921092,
    "annonce_visuelle_prochain_passage": 1,
    "y": 2430308,
    "sens": "A",
    "codeinsee": 75109,
    "nomptar": "OPERA"
  },
  "geometry": {
    "type": "Point",
    "coordinates": [2.3315871332906513, 48.87108938611164]
  },
  "record_timestamp": "2020-04-16T15:35:27.175+02:00"
}
```



```
{
  "_id": {
    "$oid": "603d0edf6dfbfc34309f2ecb"
  },
  "arr": "75011",
  "bus_num": 20,
  "geometry": {
    "type": "Point",
    "coordinates": [2.3677846082656013, 48.86897358519814]
  },
  "name_busstop": "JULES FERRY",
  "type": "bus"
}
```

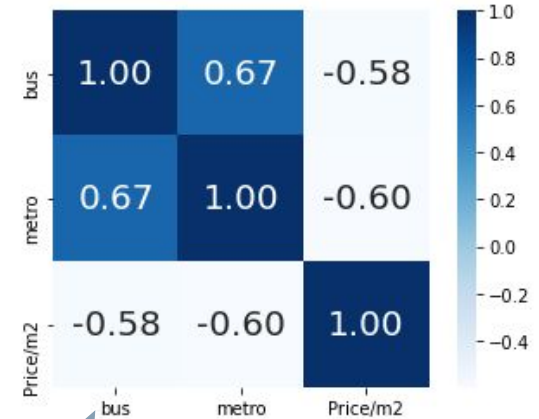
## CALCUL DES CORRÉLATIONS

	bus	metro	Price/m2
75001	159.0	30.0	13940
75002	76.0	18.0	12250
75003	68.0	12.0	12650
75004	165.0	20.0	13660
75005	201.0	22.0	12750

$$X_{normalise} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Normalisation

	bus	metro	Price/m2
75001	0.208238	0.450	0.882883
75002	0.018307	0.150	0.578378
75003	0.000000	0.000	0.650450
75004	0.221968	0.200	0.832432
75005	0.304348	0.250	0.668468



$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

Coefficient de Pearson

## RÉSULTATS



MAGASINS ALIMENTAIRES

ESPACES VERTS

ECOLES



## ANALYSES

1

### Arrondissement résidentiel

- Ecole : ++
- Transports en commun : ++
- Commerces : ++

2

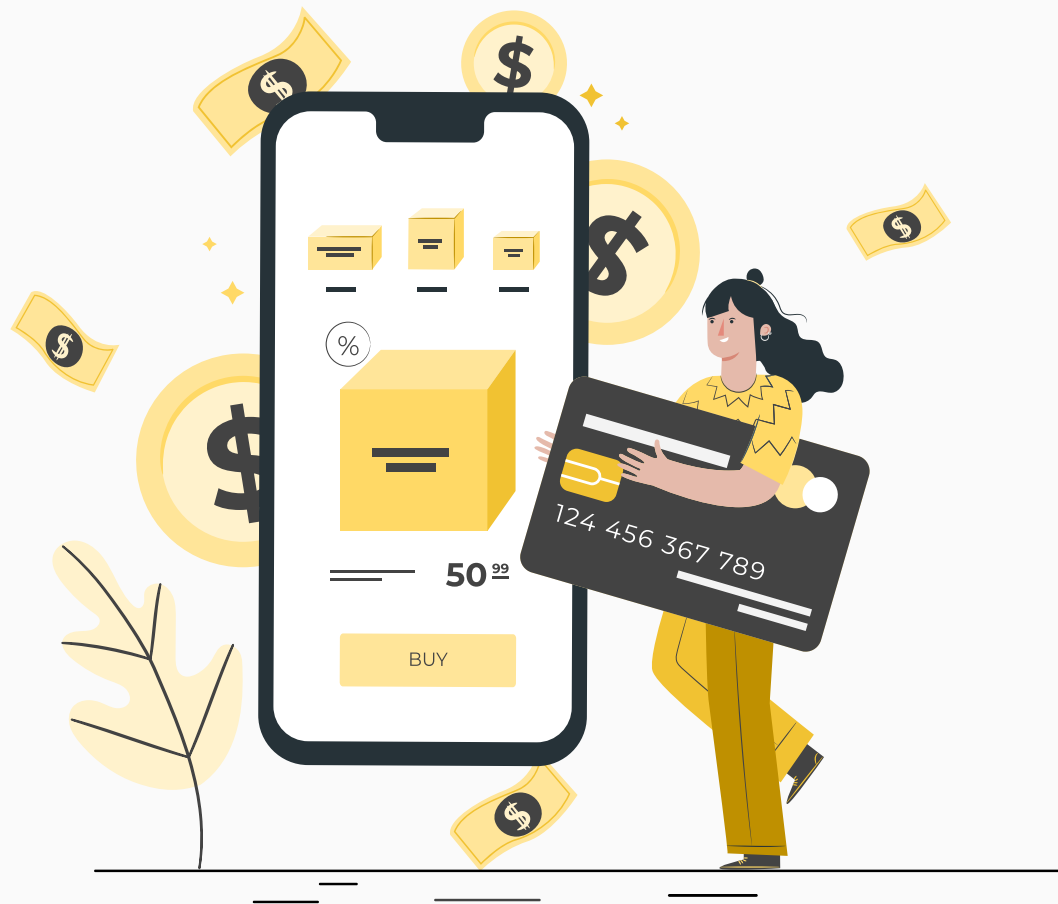
### Arrondissement non résidentiel

- Magasins alimentaires : --
- Transports : --
- Restaurant : ++
- Magasins non alimentaires : ++

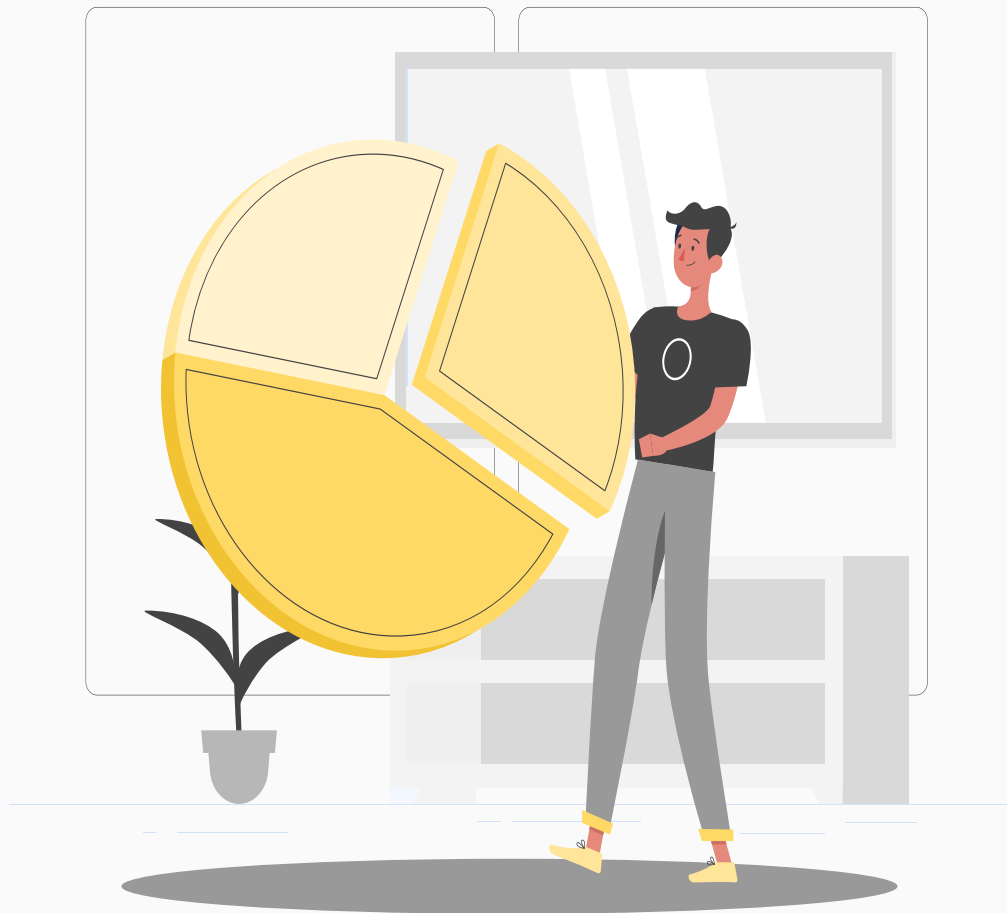
3

Les critères subjectifs  
recherchés ne participent  
pas à l'augmentation du  
prix de l'immobilier

# DÉMO



## 04. BILAN



## RETOURS

### DIFFICULTÉS RENCONTRÉES

- Technologies Open Source avec ses limites fonctionnelles
- Organisation et Communication

### AXE D'AMÉLIORATION

- Automatisation en préparation des données
- Gestion du temps

# MERCI

Avez-vous des question ?

[https://github.com/soyamimi/Invisible\\_immobilier](https://github.com/soyamimi/Invisible_immobilier)

