

Application de la méthode des prix hédonique au marché de l'immobilier à Lyon

Adrien Casse - Victor Ouledi - Robin Tordjeman - Yoann Tregouet

31-01-2023

English abstract :

Hedonic pricing is a widely used technique for estimating the prices of housing attributes or characteristics. This technique is based on the idea that the price of a dwelling is determined by its characteristics, such as its location, size, and amenities.

In this study, we applied the hedonic pricing method to the housing market in Lyon in order to understand how housing prices are set and how they evolve.

We used several regression models that allow us to show that the number of rooms, the distance to the city center and the distance to Bellecour are the main factors that determine housing prices in Lyon.

We found that the number of rooms has a positive and significant effect on housing prices, indicating that larger homes are more expensive than smaller ones. The distance to the city center also has a positive and significant effect on housing prices, indicating that housing located closer to the city center is more expensive than housing located further away. The results obtained are therefore consistent with our initial intuition.

This study provides valuable information about the housing market in Lyon. It can be used by policy makers, developers, and real estate agents to understand the factors that determine housing prices in the city and to make informed decisions.

Abstract en Français :

La méthode des prix hédoniques est une technique largement utilisée pour estimer les prix des attributs ou caractéristiques du logement. Cette technique repose sur l'idée que le prix d'un logement est déterminé par les caractéristiques de ce logement, telles que son emplacement, sa taille et ses équipements.

Dans cette étude, nous avons appliqué la méthode des prix hédoniques au marché du logement à Lyon afin de comprendre comment les prix des logements sont fixés et comment ils évoluent.

Nous avons utilisés plusieurs modèles de regression qui nous permettent de mettre en évidence que le nombre de pièces, la distance au centre-ville et la distance à Bellecour sont les principaux facteurs qui déterminent les prix des logements à Lyon.

Nous avons constaté que le nombre de pièces a un effet positif et significatif sur les prix des logements, indiquant que les grands logements sont plus chers que les petits. La distance au centre ville a également un effet positif et significatif sur les prix des logements, indiquant que les logements situés plus près du centre ville sont plus chers que les logements situés plus loin. Les résultats obtenus sont donc conforme à notre intuition de départ.

Cette étude fournit des informations précieuses sur le marché du logement à Lyon. Ils peuvent être utilisés par les décideurs politiques, les promoteurs et les agents immobiliers pour comprendre les facteurs qui déterminent les prix des logements dans la ville et pour prendre des décisions éclairées.

Problématique :

En quoi la méthode des prix hédoniques nous permet-elle d'attester de l'impact des caractéristiques des logements sur leur valeur immobilière ?

PLAN :

I.Comment situer la problématique ? Quel intérêt au travail économétrique ?

II.Présentation de la littérature et modèle économétrique pris pour référence

III.Appropriation de la base dans le cadre de l'étude

IV.Conclusion

V.Bibliographie complète

I.COMMENT SITUER LA PROBLÉMATIQUE ? QUEL INTÉRÊT AU TRAVAIL ÉCONOMÉTRIQUE ?

1) Introduction :

Lyon est une ville en croissance continue qui attire de plus en plus de nouveaux habitants chaque année. Les biens les plus recherchés sont les appartements anciens dans le centre-ville, ainsi que les maisons individuelles dans les quartiers périphériques. Cependant, cette croissance démographique a eu un impact sur le marché immobilier local, qui est devenu de plus en plus tendu. Les prix de l'immobilier à Lyon ont considérablement augmenté ces dernières années, en particulier dans les quartiers les plus prisés de la ville. Les logements sont de

plus en plus rares et les acheteurs doivent faire face à une forte concurrence pour trouver un logement qui correspond à leurs besoins et à leur budget. Les prix élevés et la pénurie de logements ont également rendu l'accès à la propriété plus difficile pour les ménages à revenus modestes. Cependant, malgré ces défis, le marché immobilier à Lyon reste attractif pour les investisseurs et les acheteurs

L'étude du marché de l'immobilier à Lyon permet de comprendre les tendances et les dynamiques de l'offre et de la demande de logements dans cette ville. Elle permet également d'identifier les facteurs qui influencent les prix de l'immobilier, les types de biens les plus recherchés et les tendances en matière de location et d'achat.

2) Enjeux du travail :

Notre travail qui concerne avant tout l'études des prix de l'immobilier à Lyon peut avoir plusieurs enjeux :

- Comprendre comment les prix des logements sont fixés.
- Comment les prix des logements évoluent.
- Comment la distance au centre ville (Bellecour) influence le prix des logements
- Orienter les décideurs publics et privés dans leurs politiques de gestion des logements

3) Situation de la problématique par rapport à la littérature :

Afin de mieux appréhender le sujet nous avons eu recours à la littérature scientifique déjà existante. Les différents articles que nous avons alors choisi d'utiliser sont caractérisés par une approche commune d'analyse spatiale du processus de valorisation immobilière. En effet le lien entre les caractéristiques spatiales des biens et leur prix nous intéresse principalement tout au long de l'étude. Ainsi de la littérature scientifique nous avons tiré plusieurs choses.

Hypothèses :

- Les caractéristiques des logements n'ont pas les mêmes effets sur leur valeur immobilière en fonction de leur situation spatiale
- La situation spatiale d'un logement est le critère qui joue le plus sur sa valeur immobilière

Méthodologie employée :

- Les différents articles utilisent la méthode des prix hédoniques

Notre problématique nous permet donc d'interroger les hypothèses de valorisation immobilière en fonction de différents quartiers de Lyon grâce à la méthode des prix hédoniques

II. PRÉSENTATION DE LA LITTÉRATURE ET MODÈLE ÉCONOMÉTRIQUE PRIS POUR RÉFÉRENCE

1) Le Modèle hédonique

Notre étude se sera intéressée avant tout à la dynamique des prix des logements au sein de l'agglomération Lyonnaise. Dans le cadre de nombreuses études similaires à la notre, il est monnaie courante d'avoir recours à la méthode des prix hédoniques pour mieux comprendre ces dernières dynamiques concernant les prix des logements. Ainsi, comme notre problématique l'indique, nous avons nous aussi eu recours à cette méthode. Il est donc intéressant d'entrée de jeu de savoir à quoi elle consiste. Celle-ci repose sur l'analyse de la relation entre le prix d'un bien ou d'un service et les caractéristiques qui le composent. Elle permet de décomposer le prix d'un bien en associant à chacune des caractéristique un prix, que l'on nomme « prix hédonique ».

Cette méthode aura été démocratisée et aura pris forme grâce au cadre théorique mis en place par 3 auteurs ayant réalisé des travaux majeurs sur la dynamique des prix de l'immobilier.

On peut déjà citer Kevin J. Lancaster (1966) qui adopte quant à lui une approche plus traditionnelle de la théorie du consommateur, où l'utilité perçue pour un bien immobilier est dérivée des caractéristiques des biens. Sachant qu'un bien est une combinaison de plusieurs caractéristiques. Son approche peut être récapitulée en trois points :

- L'utilité du consommateur n'est pas liée au bien en soi, ce sont ses caractéristiques qui donnent lieu à cette utilité

- Un bien possède en général plusieurs caractéristiques, de même ces caractéristiques sont communes à plusieurs biens.
- Les biens combinés peuvent avoir des attributs qui diffèrent des biens pris séparément.

Le cadre théorique mis en place par Rosen (1974) aura lui aussi participé au développement du modèle hédonique. Il combine l'approche de Lancaster et la théorie de l'utilité de la concurrence appliquée au marché des biens immobiliers hétérogènes. Dans la fonction hédonique de Rosen, les consommateurs cherchent à maximiser leur utilité et les producteurs quant à eux cherchent à maximiser leur profit, de ce fait, le prix d'achat du bien est la somme des prix implicites des caractéristiques du logement.

Le dernier auteur que nous pouvons citer est d'Alonso (1964). Son modèle théorique se concentre sur le marché foncier en supposant que l'unité de logement est un produit homogène. En considérant que les variations de prix des logements étaient dues principalement à la distance au centre-ville, selon lui, la concurrence pour l'emplacement central ferait grimper le prix des logements situés plus près des Central Business District (CBD). La théorie des prix hédoniques d'Alonso est une théorie de l'économie spatiale qui décrit comment les prix des biens immobiliers varient en fonction de leur emplacement. Dans ce cadre, le ménage est considéré comme un individu disposant d'un certain revenu qu'il peut dépenser entre des terrains et d'autres biens. C'est sur le cadre théorique d'Alonso que nous nous sommes le plus appuyé durant notre étude.

En effet grâce aux hypothèses émises par l’auteur, il nous aura été possible d’introduire les deux concepts que sont le monocentrisme et le polycentrisme au sein d’une ville. Le monocentrisme impliquant qu’il existe un unique centre d’emplois et d’activités comme noyau autour du quel la population se répartie et pour lequel les prix de l’immobilier sont déterminés en fonction de leur plus ou moins proximité ou éloignement autour de ce noyau. Le polycentrisme implique lui alors l’existence de plusieurs noyaux au sein du même ville autour duquel se produisent les effets précédemment décrits. Pour une ville polycentrique, on pourrait alors retrouver les biens les plus chers concentré autours des 2 centres névralgiques de la ville, et les biens les moins chers étant à la fois éloignés d’un noyau et de l’autre au sein de la même agglomération.

Il nous sera alors possible de vérifier si Lyon présente les caractéristiques d’une ville poly ou monocentrique.

Les hypothèses de base retenues pour le modèle des prix hédonique

Grâce aux travaux des 3 auteurs précédemment cités, nous pouvons à présent définir un cadre théorique commun pour l’application de la méthode des prix hédoniques. Rappelons que le modèle prévoit qu’un bien se caractérise par la somme de ses composantes individuelles, il est alors dérivé de la valeur des caractéristiques connexes (nombre de chambres, balcons, etc) ou des services accessibles grâce à la possession de celui-ci (accès au transport en commun, centres commerciaux). La valeur de ces caractéristiques ne pouvant pas toujours être observées directement, mais ayant une grande influence sur le prix du marché d’un bien.

Pour que ce modèle soit cohérent il faut ainsi considérer plusieurs hypothèses

- Les biens doivent nécessairement être hétérogènes, ce qui signifie que leurs caractéristiques doivent différer.
- Le marché fonctionne dans un cadre de la concurrence parfaite, l'impact d'un individu peut être ignoré, car le nombre d'acheteurs & de vendeurs est trop important.
- Les acheteurs et les vendeurs disposent d'une information parfaite sur les biens.
- Le marché est en équilibre, il ne doit pas y avoir d'intercorrélation entre chacune des caractéristiques.

2) Présentation de la Littérature :

Afin de mener à bien notre application de la méthode des prix hédoniques à Lyon, cette dernière centrée sur l'intérêt particulier porté aux caractéristiques spatiales des biens, il nous a fallu nous intéresser à des applications abordant la même approche mais sur d'autres villes. Nous repasserons ici en revue les articles sur lesquels nous nous sommes appuyés majoritairement et relèverons les éléments de méthodes apportés par ces derniers qui nous ont permis de mener à bien notre application à la ville de Lyon.

a) 1er article

Arslanlı, K. Y. (2020). Analysis of house prices : a hedonic model proposal for Istanbul metropolitan area. Journal of Design for Resilience in Architecture and Planning1(1), 57-68**

Ce document, par son étude sur le marché des biens immobiliers à Istanbul, nous a permis de porter attention au problème de l'autocorrélation spatiale dans le cadre de l'application de la méthode hédonique. La compréhension de la distribution des logements étant nécessaire à l'élaboration de leurs prix ; à Istanbul, les mêmes caractéristiques de colonies se construisent autour des mêmes, zones cela signifie que les propriétés des maisons sont similaires aux caractéristiques environnantes. Ces logements sont ainsi vendus en fonction des commodités entourées, les prix des maisons à Istanbul sont de ce fait autocorrélés. Le fait qu'il y ait de l'autocorrélation spatiale est contraignant, car cela peut entraîner des erreurs dans l'estimation des coefficients du modèle ainsi que dans les résultats de la modélisation.

Il y a de plus des facteurs externes tels que des contraintes physiques, les revenus, et les variables démographiques qui peuvent affecter les prix des logements sur le marché, et de ce fait être source de biais puis peuvent en outre apporter une inefficacité des prix ou du marché.

Nous rendre compte de ces potentiels problèmes nous permettent d'être conscient des potentielles limites de notre application.

Concernant la spécification du prix hédonique des maisons, dans ce document, le prix des maisons est déterminé sur le marché par ses caractéristiques peut

être représenté de la manière suivante :

$$P = f(H, N, t)$$

Où :

P = Le prix de vente

H = Les variables décrivant les caractéristiques de l'unité en m².

N = La distance par rapport à un point de références.

T = La période pendant laquelle le bien est vendu.

Puis une forme fonctionnelle semi-logarithmique est utilisé pour la modélisation :

$$V = e^{x\beta+u}$$

Où :

V = La valeur de la propriété

X = Le vecteur de caractéristiques du logement

β = Le vecteur de coefficients hédoniques inconnus

u = Le résidu

Les formes de spécification du modèles nous auront aidé dans la spécification du modèle mais adapté à lyon cette fois-ci.

b) 2eme article

Xiaochen Hu, Luechai Chulasai & Sasipen Phuangsaichai. (2011). Hedonic Pricing Model for Housing Market in City of Kunming, the People's Republic of China. (*Journal of Economics Chiang Mai University*), 15(1), 144-165

Cette revue traite de nombreuses études utilisant la méthode des prix hédoniques dans plusieurs villes chinoise. Pour la sélection des variables expliquant les différents prix, il a ainsi été recensé plusieurs catégories :

L'accessibilité au CBD est déterminée par la mesure de la distance entre le logement et le CBD. Plus les commodités de circulation sont nombreuses et proches d'une unité de logement, plus elle a de valeur. L'emplacement est considéré comme un attribut essentiel du logement, il détermine l'accessibilité au CBD et les conditions de circulation de l'unité de logement.

Les attributs structurels faisant référence aux caractéristiques propres au logement, tels que la surface de plancher, d'un garage, d'une cheminée, l'âge du bâtiment.

Les attributs de voisinage concernent la proximité à des installations de divertissement, à des lieux de culte, à des services éducatifs... La proximité des établissements de vie (supermarché, restaurant, banque...), de lieux de divertissements (parc, cinéma...) a une influence directe sur la commodité de la vie quotidienne des ménages.

L'externalité de la qualité environnementale fait référence au bruit, à la pollution, la sécurité, la propreté.

Pour cette étude, les 3 formes fonctionnelles ont été utilisées pour estimer le modèle hédonique soit :

- La forme linéaire qui représente le fait qu'une variation d'une unité de chaque caractéristique entraîne la variation du prix du logement par cette caractéristique.

$$p = \beta_0 + \sum_{i=1}^{n=12} \beta_i z_i + \varepsilon \quad (i = 1, 2, 3 \dots n)$$

- La forme logarithme, qui a toutes les variables de la fonction sous forme logarithmique, le coefficient de régression correspond à l'élasticité du prix de la caractéristique.

$$\ln P = \beta_0 + \sum_{i=1}^{n=12} \ln \beta_i z_i + \varepsilon \quad (i = 1, 2, 3 \dots n)$$

- La forme fonctionnelle semi-logarithme est une variante de l'équation double-log, où certaines variables, dépendantes ainsi que indépendantes, sont exprimées en termes de leur logarithme naturel.

$$\ln P = \beta_0 + \sum_{i=1}^{n=12} \beta_i z_i + \varepsilon \quad (i = 1, 2, 3 \dots n)$$

Ainsi, grâce à l'article, on se rend compte des différentes manières de pouvoir exprimer notre modèle pour ensuite pour estimer des paramètres avec des interprétations différentes selon la spécification du modèle choisie.

III. APPROPRIATION DE LA BASE DANS LE CADRE DE L'ÉTUDE

Nous arrivons à présent sur l'application du modèle des prix hédoniques à Lyon. Nous avons pour cela importé 6 bases de données retraçant les transactions financières pour différents biens immobiliers ainsi que leurs caractéristiques pour différents quartiers de la métropole de Lyon.(2 quartiers par arrondissement avec un quartier proche et un quartier loin du centre par arrondissement) Nous nous sommes assurés que les variables d'intérêt soient considérées sous la bonne forme. Exemple : la latitude et longitude sous forme numérique

1) Méthodologie et traitement de la base

Nous avons importé 6 bases de données retraçant les transactions financières pour différents biens immobiliers ainsi que leurs caractéristiques pour différents quartiers de la métropole de Lyon.(2 quartiers par arrondissement avec un quartier proche et un quartier loin du centre par arrondissement) Nous nous sommes assurés que les variables d'intérêt soient considérées sous la bonne forme. Exemple : la latitude et longitude sous forme numérique Puis nous supprimons ensuite les variables du dataframe qui ne nous intéressent pas, puis supprimons les lignes avec des valeurs manquantes pour les variables nous intéressant car les NA étant distribués aléatoirement en fonction des différents types de biens, de surfaces et de nombre de pièces + nous avons assez de valeurs à notre disposition + corriger les valeurs en vue de la nature des variables n'a pas de sens.

Ensuite il s'agit de nous débarrasser des valeurs abérantes à première vue et

d'une catégorie de bien immobilier dont la distribution des valeurs foncières ne semblait pas être expliquée par les variables à notre disposition.

Nous avons ensuite créé une nouvelle variable distance par rapport à Part-Dieu et distance par rapport à Bellecour ainsi que transformé les variables catégorielles en dummy variables. À savoir si les valeurs sont dans tel ou tel arrondissement (0 ou 1) et si les valeurs font parti du quartier loin ou non de centre pour chaque arrondissement concerné (0 ou 1). Nous avons pu également créer une variable prix au mètre carré et log prix au mètre carré

Une fois la base bien construite, dans un premier temps nous pouvons nous intéresser aux outliers plus en détail

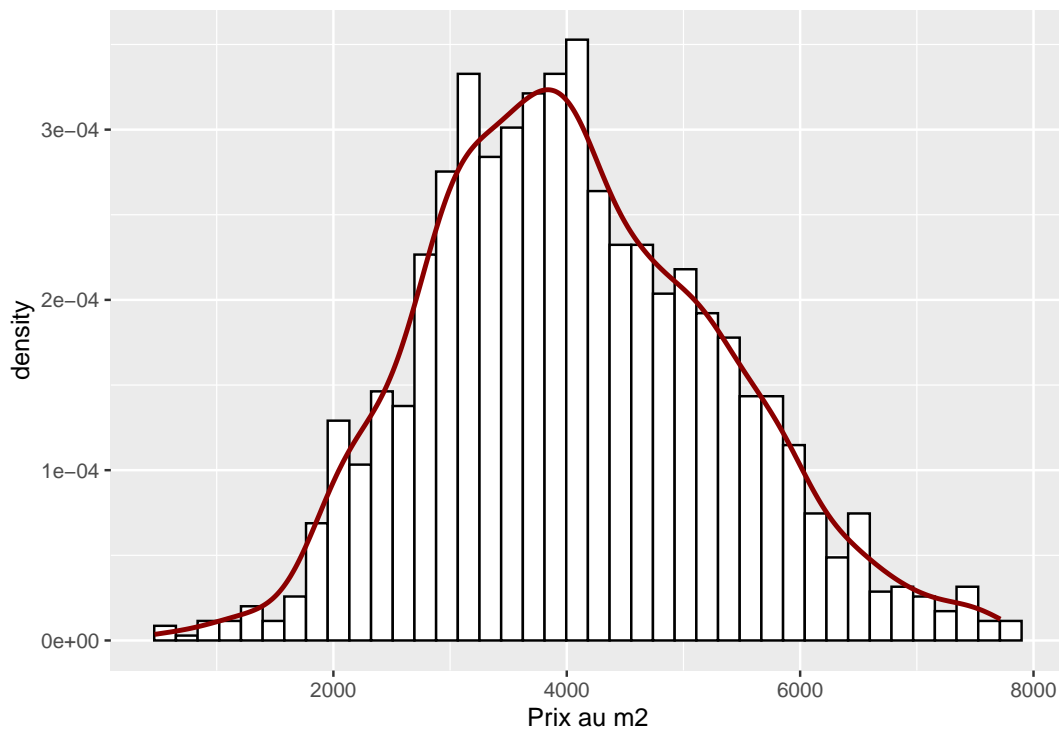
Nous supprimons ceux nous paraissant être des biens immobiliers particuliers et dont les variables à notre disposition ne semblent pas expliquer intuitivement la valeur foncière

2) Statistiques descriptives de la base

Nous nous intéressons de nouveau cette fois-ci aux statistiques descriptives de la base pour en vérifier la cohérence

Table 1: Summary Statistics

VF_au_m2	valeur_fonciere	surface_reelle_bati	distancePD	distanceBellecourt
Min. : 468.2	Min. : 17450	Min. : 12.00	Min. : 1.527	Min. : 0.6613
1st Qu.: 3149.5	1st Qu.: 150000	1st Qu.: 41.00	1st Qu.: 2.253	1st Qu.: 2.8101
Median : 3958.7	Median : 215000	Median : 60.00	Median : 3.606	Median : 4.9246
Mean : 4062.5	Mean : 225704	Mean : 58.68	Mean : 3.657	Mean : 4.3623
3rd Qu.: 4923.5	3rd Qu.: 286004	3rd Qu.: 74.00	3rd Qu.: 4.249	3rd Qu.: 5.2528
Max. : 7715.6	Max. : 598400	Max. : 182.00	Max. : 7.912	Max. : 7.4897

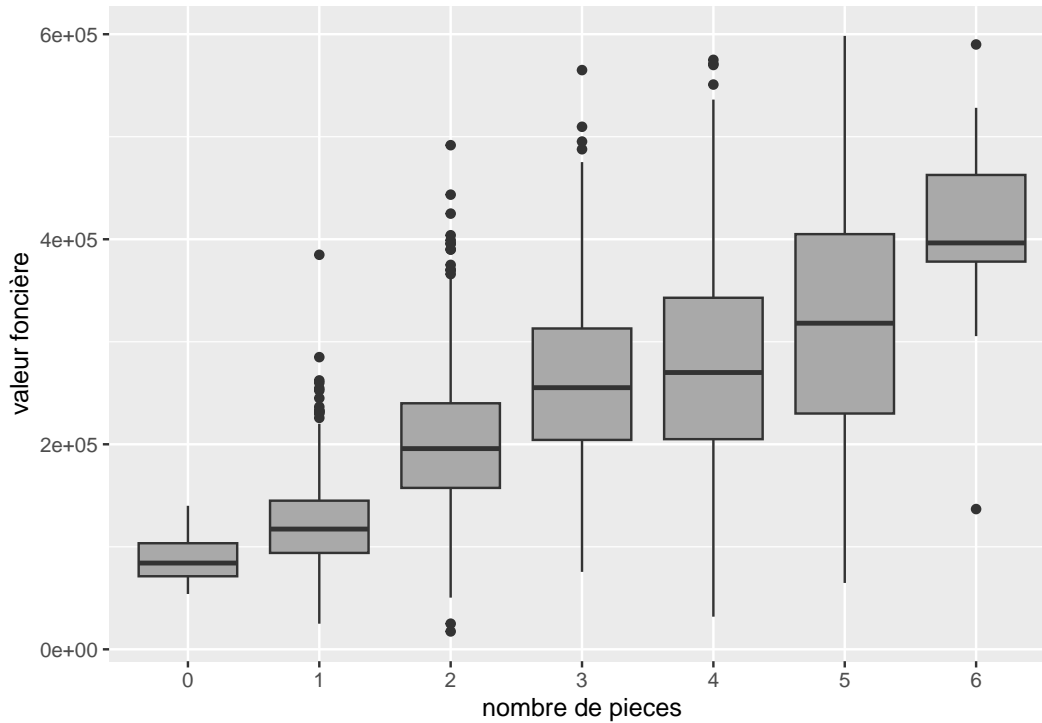


on se rend compte, vu la “courbe en cloche” que celle-ci suit une loi normale. En effet les valeurs sont concentrées autour de la moyenne et plus on s’en éloigne plus le nombre d’observations diminue. L’hypothèse de normalité est d’autant plus essentielle pour la suite de notre travail et la création de nos

modèles.

3) Statistiques descriptives approfondies

Boxplot des valeurs foncières en fonction du nombre de pièces



Pour ce boxplot ainsi que les suivants nous retrouvons des outliers que nous avons décider de garder et de considérer. En effet il ne s'agit pas de valeurs complètement aberrantes comme celle que nous avons pu déjà éliminer mais elles correspondent plutôt à des biens se distinguant un peu de la majorité mais pas de manière démesurée. Les outliers déjà éliminés résultaient eux d'erreurs de frappe lors de la constitution des bases que nous avons utilisées pour notre étude (nous supposons)

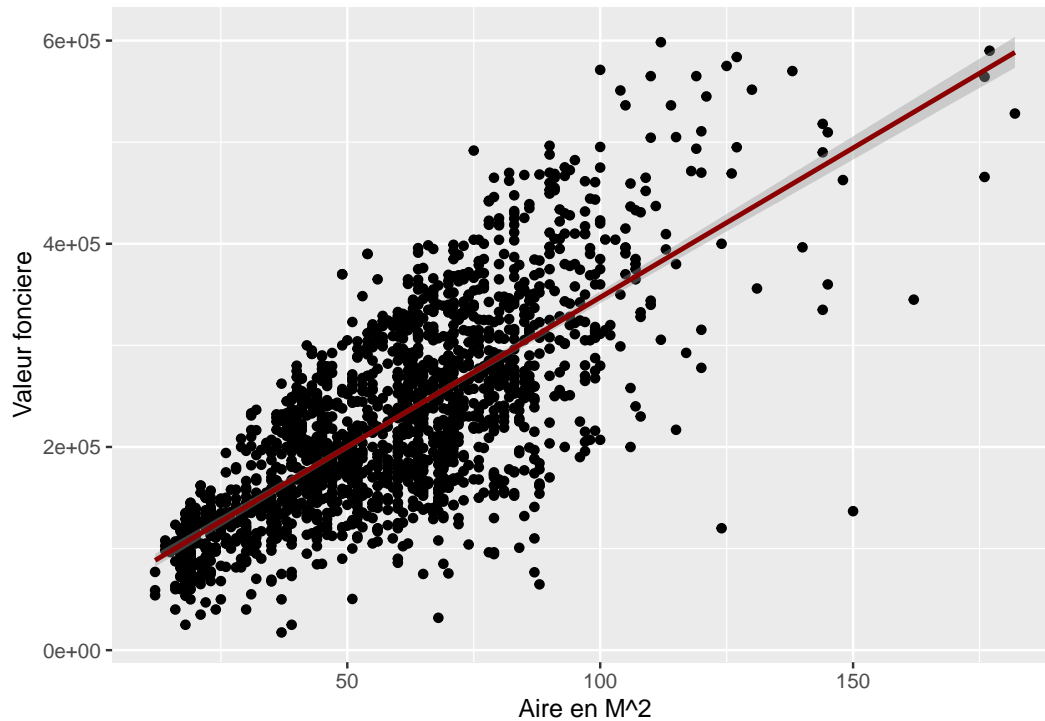
Calcul moyenne valeur foncière en fonction du nombre de pièces des biens

Table 2: Moyenne de Prix en fonction du nombre de pieces

nombre_pieces_principales	valeur_fonciere
0	90570.0
1	124201.9
2	202172.5
3	259346.0
4	280700.1
5	325476.7
6	403998.1

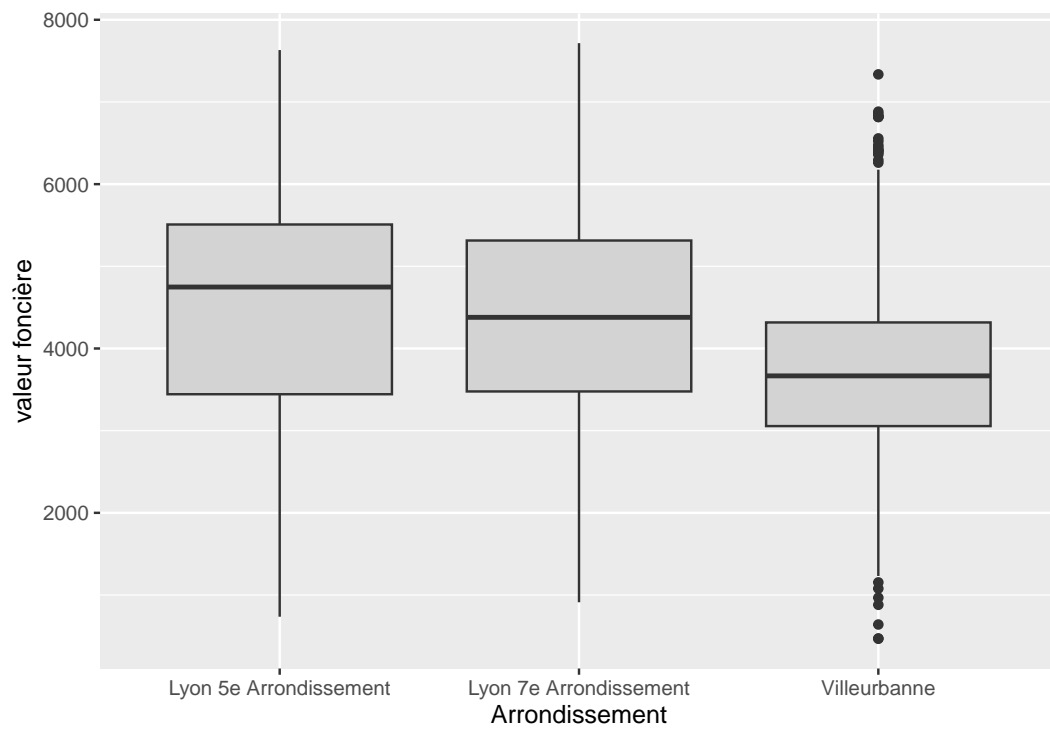
On remarque une corrélation positive et cohérente entre le nombre de pièces et la valeur foncière. En effet plus il y a de pièces plus sa valeur foncière sera élevée.

Graphique des valeurs foncières en fonction de a surfaces des biens immobiliers



Ici nous retrouvons une corrélation intuitive et attendue : plus un bien est grand plus sa valeur foncière est grande. Nous retrouvons de l'hétéroscédasticité cependant du au manque de variables explicatives liées à la détermination de la valeur foncière des biens comme la proximité au centre, aux transports, proximités d'endroits produisant des externalités négatives sur les logements.

Boxplot des valeurs foncières en fonction des quartiers :



Calcul moyenne valeur foncière au m2 en fonction des différents quartiers :

Table 3: Moyenne de Prix au m2 par Arrondissements

nom_commune	VF_au_m2
Lyon 5e Arrondissement	4481.99
Lyon 7e Arrondissement	4425.39
Villeurbanne	3738.32

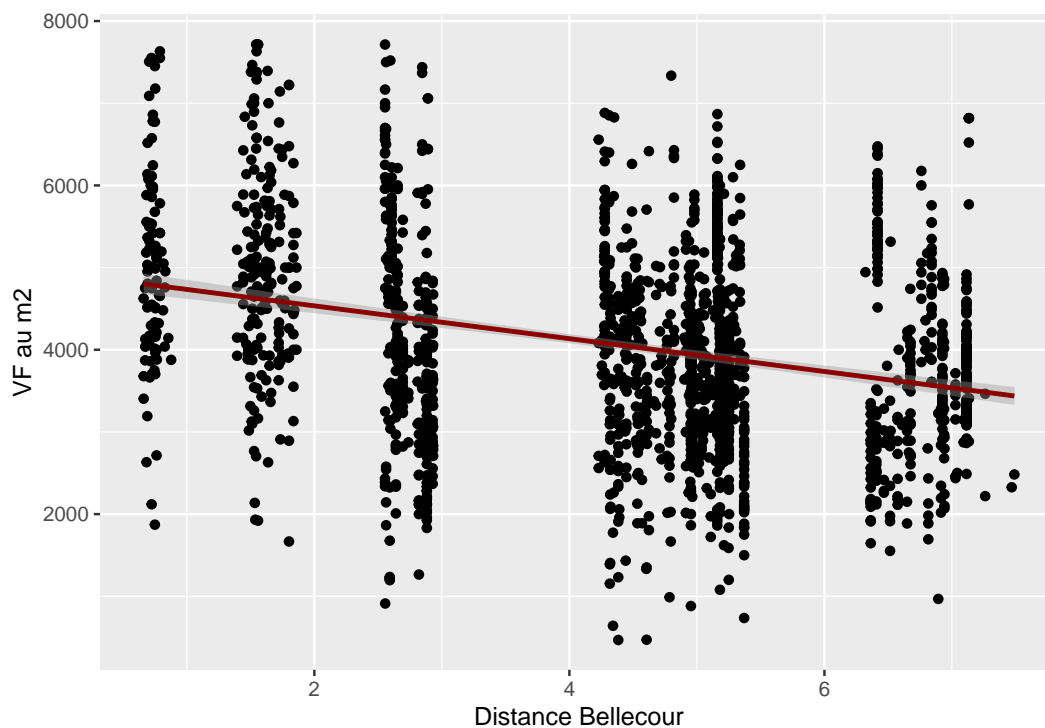
Nous pouvons ici faire un classement hypothétique concernant le prestige et la qualité de vie des arrondissements. Ainsi le 5ème arrondissement serait légèrement plus prestigieux que le 7ème arrondissement avec des prix au mètre carré du logement sensiblement plus élevés. En suivant ce raisonnement Villeurbanne serait beaucoup moins prestigieux que le 5ème et le 7ème arrondissement. Le but principale de notre étude étant d’attester de l’effet de l’éloignement des biens par rapport au centre sur leur prix, il faut dans un premier temps savoir quel point de lyon est à consider comme le centre. En effet selon la théorie d’Alonso sur laquelle nous nous appuyons, cette dernière prévoit une corrélation négative et significative entre le prix et l’éloignement au centre des affaires.

On se rend compte (voir annexe correlation) qu’il n’y a aucune significativé dans la corrélation entre valeurs foncières des biens et leur éloignement au quartier des affaires (ici considéré comme Part-Dieu)

Nous avons alors décidé de tester de l’hypothèses d’Alonso en considérant le coeur historique de la ville (Bellecour).

Ainsi nous avons trouvé plus forte corrélation entre le prix et l'éloignement des biens à Bellecour plutôt qu'à Part-Dieu. La Théorie d'Alonso prévoit que les prix des biens soient négativement corrélés à leur proximité du centre ville ou centre d'affaires. Dans un premier temps nous pouvons donc émettre l'hypothèse que cette théorie est réalisée à Lyon si nous considérons le centre historique comme le centre ville et non le centre d'affaire.

Scatter plot des Valeurs foncières au m2 en fonction de la distance



On remarque que 4 catégories de biens peuvent se créer en fonction de leur distance. **Etant donné le choix de nos lots, la variable distance n'est pas vraiment continue, il serait plus pertinent dans le cas de notre étude de la considérer comme une variable catégorielle**

Catégories observables :

1-> Très près: entre 0 et 2 km par rapport à Bellecour

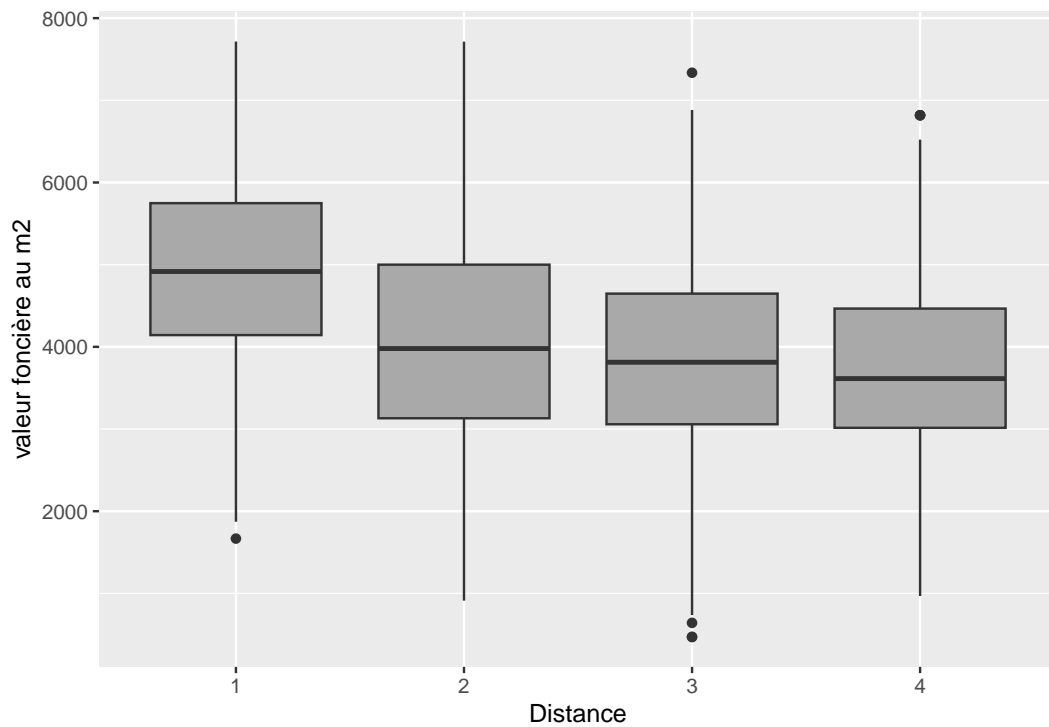
2-> Pres : entre 2 et 4 km par rapport à Bellecour

3-> Loin : entre 4 et 6 km par rapport à Bellecour

4-> Tres Loin : plus de 6 km par rapport à Bellecour

Nous pouvons ainsi produire de nouveau des statistiques descriptives en considérant ces catégories en les reliant au prix.

Boxplot des Valeurs foncières au m2 en fonction de la distance



Ainsi en considérant les 4 catégories de distances on se rend compte une fois de plus que l'éloignement est corrélé négativement au prix

IV .REGRESSIONS ET RESULTATS DE L'ETUDE

1) Regression avec les MCO :

Nous allons à présent établir les régressions pour tester les effets de la distance en abordant cette dernière de 2 manières différentes:

1ere approche - La distance comme variable continue 2eme approche - La distance comme variables ordinal

Les regressions qui suivent nous permettent de tester l'effet de la distance absolue au centre (distance-inter arrondissement) soit nous poser la question : l'écart de distance absolue par rapport au centre entre chaque arrondissement a-t-il un effet sur le prix ?

1) Distance comme variable Continue Modèle pour déterminer la
Valeur foncière avec la distance considérée comme valeur continue

$$VF = \beta_0(Villeurbanne) + \beta_1 SurfaceReelle + \beta_2 5Arr + \beta_3 7Arr + \beta_4 DistanceBellecourt$$

Table 4 : Régression des prix sur la distance
considérée comme continue

	(1)	(2)
(Intercept)	5458.279 (133.741)	8.603 (0.037)
Lyon 5eme	582.879 (77.819)	0.127 (0.022)
Lyon 7eme	232.060 (92.030)	0.046 (0.026)
distanceBellecourt	-103.078 (21.937)	-0.026 (0.006)
surface_reelle_bati	-19.017 (1.035)	-0.005 (0.000)
Num.Obs.	1876	1876
R2	0.238	0.185
R2 Adj.	0.236	0.183
AIC	31 609.6	907.8
BIC	31 642.8	941.1
Log.Lik.	-15 798.787	-447.924
F	145.863	105.876
RMSE	1099.48	0.31

(1): Prix au m2 / (2): Log(prix au m2)

Les résultats obtenus nous permettent d'attribuer une valeur moyenne du prix du mètre carré à Villeurbanne de 5458.279 euros et ainsi dire que le prix du mètre carré Dans le 5ème arrondissement et dans le 7ème arrondisse-

ment sont respectivement 582.879 et 232.060 plus cher. En considérant ensuite l'éloignement à Bellecour, on peut présumer que le prix moyen du mètre carré diminue de 103.078 euros quand on s'éloigne de 1km de Bellecour. cet effet est-il significatif ? **on s'aperçoit que cet effet d'éloignement à Bellecour d'1km, (-103,078 euros) est bien statistiquement significatif au seuil de 99% (P-value<0,01)** Le R-squared est quant à lui égale à 23,8% ainsi , la variance des variables explicatives explique 23,8% de la variance du prix des logements à Villeurbanne. Le F-test test de la nullité globale de tout les coefficients est lui aussi est statistiquement très significatif (p-value<0,01) les coefficients sont donc tous significativement différents.

2) Distance en inter-arrondissement comme variable ordinale (1 à 4)
Modèle pour déterminer la Valeur foncière avec la distance considérée comme variable ordinale

$$VF = \beta_0(TrèsPrès) + \beta_1SurfaceReelle + \beta_2Près + \beta_3Loin + \beta_4TrèsLoin$$

Table 5 : Régression des prix sur la distance considérée comme ordinale

	(1)	(2)
(Intercept)	5949.921 (84.080)	8.716 (0.023)
distanceB_Cat2	-804.025 (89.235)	-0.202 (0.025)
distanceB_Cat3	-963.107 (75.039)	-0.234 (0.021)
distanceB_Cat4	-1003.033 (86.889)	-0.240 (0.024)
surface_reelle_bati	-18.612 (1.030)	-0.005 (0.000)
Num.Obs.	1876	1876
R2	0.237	0.191
R2 Adj.	0.235	0.189
AIC	31 611.5	893.8
BIC	31 644.7	927.0
Log.Lik.	-15 799.744	-440.876
F	145.237	110.202
RMSE	1100.04	0.31

(1): prix au m2 / (2): Log(prix au m2)

En considérant l'éloignement inter-arrondissement comme une variable catégorielle (cat1,cat2,cat3,cat4) on peut fixer le prix moyen du mètre carré à

5949.921euros, pour tous les biens se trouvant “très près” (cat1) de Bellecour. En se déplaçant dans la catégorie d'éloignement suivante “pres” (cat2), le prix du mètre carré diminue de 804.025euros.

De plus, cette diminution est statistiquement significative au seuil de 99% (*). Le R-squared permet de dire que la variance des variables explicatives de cette régression (distance a bellecour & Surface du logement) explique 23,7% de la variables à expliquer (valeur fonciere des logements en cat1) **

La P-value du Test de significativité global est également très faible, les coefficients ne sont pas donc pas tous égaux et nulles. Ici la variable arrondissement a été écartée car celle-ci biaisait les résultats. En effet du fait que les biens que nous avons sélectionnés de Villeurbanne soient uniquement situés loin ou très loin de Bellecour ou que les biens du 7ème soient près ou loin, il nous était impossible de spécifier un modèle avec la prise en compte de l'effet de l'arrondissement et des catégories. Ainsi en prenant comme référence les biens de Villeurbanne la régression ne produit pas de résultats pour l'effet de la catégorie près et très près. Il en est de même en considérant les autres arrondissements comme référence, la régression ne produit pas de résultat pour l'effet des catégories de distances n'incluant pas les biens des catégories d'arrondissements de référence.

2) ANOVA

L'ANOVA (Analyse de la Variance) est une technique statistique utilisée pour comparer les moyennes de plusieurs groupes ou catégories. L'objectif de l'ANOVA est de déterminer si les moyennes des groupes étudiés sont similaires ou différentes, et si ces différences sont statistiquement significatives. Dans le cadre de notre étude, pour faciliter notre analyse on a considéré la distance au centre ville (très près, près, loin, très loin). Ainsi avec le test de Tukey qui va comparer chacune des différentes catégories entre elles, permettront de dire si les différences entre les groupes sont statistiquement significatives et donc si la distance au centre-ville impacte le prix des logements.

a) Anova de la distance

Il y a-t-il une différence significative entre les différents "groupes" de distance que nous avons établis ?

Table 6 : Anova des prix en fonction des catégories de distances

	diff	lwr	upr	p adj
2-1	-884.3040	-1132.5936	-636.01444	0.0000000
3-1	-1104.8895	-1312.7937	-896.98537	0.0000000
4-1	-1217.9454	-1457.7281	-978.16276	0.0000000
3-2	-220.5855	-419.8449	-21.32614	0.0231397
4-2	-333.6414	-565.9685	-101.31430	0.0013043
4-3	-113.0559	-301.6091	75.49728	0.4126669

Il s'avère que oui, il y a une différence au niveau du prix des logements en fonction des catégories de distances. On observe en effet, qu'en prenant la catégorie 1 (Très près) comme catégorie de référence, plus on avance dans

les categories plus le prix du logement diminue ,**en d'autres termes plus on s'éloigne du centre ville plus le prix du logement diminue** Avec le test de tukey , on peut s'apercevoir de la difference significative qu'il existe entre les categories de distance par rapport au centre ville. Par exemple ici on s'aperçoit qu'en passant de la categorie 1 (tres pres du centre) à la categorie 3 (loin du centre),**le prix du logement diminue significativement de 1104euros le m²**. Il est intéressant ici de s'intéresser à la significativité de passer d'une catégorie à une autre. On voit par exemple que le fait de passer d'un logement loin à très loin n'est pas significatif sur les prix des biens. On peut alors émettre l'hypothèse que passé un certain éloignement, la proximité au centre n'est plus impactante dans la détermination des prix.

b) Anova du nombre de pieces

Il y'a t-il une difference significative du prix au m2 en fonction du nombre de pieces ?

Table 7 : Tukey test des prix en fonction du nombre de pièces des biens

	diff	lwr	upr	p adj
1-0	-489.85244	-2201.4943	1221.78947	0.9801527
2-0	-856.38757	-2566.5803	853.80519	0.7580508
3-0	-1348.70614	-3057.5716	360.15933	0.2303633
4-0	-1839.52269	-3551.3533	-127.69208	0.0257948
5-0	-2040.27089	-3782.6488	-297.89298	0.0100543
6-0	-2010.69997	-3957.8957	-63.50427	0.0377158
2-1	-366.53513	-602.3400	-130.73031	0.0000973
3-1	-858.85370	-1084.8312	-632.87623	0.0000000
4-1	-1349.67025	-1597.0741	-1102.26643	0.0000000
5-1	-1550.41845	-1958.7387	-1142.09823	0.0000000
6-1	-1520.84753	-2481.2723	-560.42278	0.0000648
3-2	-492.31856	-707.0442	-277.59296	0.0000000
4-2	-983.13512	-1220.3058	-745.96447	0.0000000
5-2	-1183.88332	-1586.0856	-781.68104	0.0000000
6-2	-1154.31239	-2112.1521	-196.47266	0.0070388
4-3	-490.81655	-718.2189	-263.41423	0.0000000
5-3	-691.56476	-1088.0853	-295.04417	0.0000060
6-3	-661.99383	-1617.4617	293.47405	0.3865222
5-4	-200.74820	-609.8587	208.36230	0.7752556
6-4	-171.17728	-1131.9383	789.58372	0.9984828
6-5	29.57093	-984.6176	1043.75949	1.0000000

Tout comme pour le premier tukey test concernant les différentes catégories de distance, pour les différents biens avec différents nombres de pièces, il est aussi intéressant de s'intéresser à la **significativité** lorsqu'on gagne une pièce en plus. Ici on peut émettre l'hypothèse qu'un bien d'une pièce et un autre de 0 pièce sont à rapprocher. En effet il n'y a pas un écart significatif dans leur prix. Lorsqu'on passe d'un bien de 4 à 5 pièces et de 5 à 6 pièces on

remarque une fois de plus qu'il n'y a pas de significativité au niveau de l'écart de leur prix. On peut alors émettre la nouvelle hypothèse qui est que passé un certain nombre de pièces l'apport d'une nouvelle pièce n'a plus beaucoup d'effet sur les prix. Grâce à ces tests, on peut alors se rendre compte de la décroissance de l'effet marginal du nombre de pièce ainsi que de la distance. Il reste tout de même des limites aux interprétations précédentes, en effet ces tests sont faits sans prendre en compte de l'arrondissement et l'éloignement. Ainsi si par exemple les biens de 0 pièce sont majoritairement situés plus près du centre que les biens de 1 pièce, ces 2 types de biens peuvent avoir un prix assez similaire, l'un car ayant une pièce en plus et les autres car étant plus près du centre. C'est un biais potentiellement existant qui représente une limite à la possibilité d'interprétation des résultats.

IV. CONCLUSION

En conclusion, l'étude a eu pour but d'appliquer les méthode de prix hédoniques pour expliquer le prix des logements à Lyon dans 3 Arrondissement de la ville :

- Lyon 5eme
- Lyon 7eme
- Villeurbanne

Dans notre étude nous avons utilisé la méthode des MCO afin d'effectuer une regression lineaire multiple, dans le but d'identifier les facteurs qui influencent le plus fortement les prix des logements dans la ville (dans notre cas spécifiquement la distance). Cependant, il est important de noter que l'utilisation des MCO implique une relation linéaire entre les variables indépendantes et dépendantes, ainsi que la normalité et l'homoscédasticité des erreurs. Notre étude a révélé que les données ne répondaient pas à toutes ces hypothèses, ce qui a pu avoir un impact sur les résultats , ce qui est une des principale limite de notre modèle ainsi que le peu de variables explicatives que nous disposions. Afin d'en tenir compte et de mieux comprendre l'importance de la distance au centre ville (Bellecour) sur les prix des logements, une ANOVA a également été réalisée, pour realiser cela nous avons donc creer des categories de distance, ainsi grace aux résultats de l'ANOVA nous avons reussi a trouver l'effet statisquement significatif de la distance au centre ville sur le prix des logements, en effet plus on s'éloigne du centre , plus le prix des logements diminuent significativement ,ce qui confirme l'importance de l'emplacement comme facteur déterminant des prix des logements à Lyon et confirme les hypothèses du théoreme d'Alonso et

fait de Lyon une ville dite monocentrique. Dans l'ensemble, malgré les nombreuses limites présentes (manque de variables explicatives et non respect de toutes les hypothèses MCO) notre étude fournit des informations utiles sur les facteurs qui influencent les prix des logements à Lyon ainsi que les limites de l'utilisation de la régression MCO dans ce contexte. Elle souligne l'importance de l'emplacement, en particulier la proximité au centre-ville, dans la détermination des prix des logements ainsi que la surface habitable du logement ainsi au sein d'un même arrondissement l'éloignement joue également sur les prix (voir annexe 2). Cependant, des recherches supplémentaires sont nécessaires pour explorer davantage les autres facteurs qui contribuent au prix des maisons à Lyon, tel que la proximité des espaces verts, proximité des arrêts de transport ou même le taux de criminalité, en effet, tous ces facteurs peuvent varier entre les différents quartiers de la ville et ainsi influencer la dynamique du marché immobilier à Lyon. En outre, les résultats de cette étude qui prouvent que Lyon est une ville monocentrique, peuvent servir de point de départ à d'autres recherches sur le marché du logement à Lyon et dans d'autres régions similaires.

V. BIBLIOGRAPHIE

[1] **Anderson, S. T. & Bates, M. D. (2017). HEDONIC PRICES AND EQUILIBRIUM**

SORTING IN HOUSING MARKETS : A CLASSROOM SIMULATION. National Tax Journal, 70(1), 171-183.

[2] Arslanlı, K. Y. (2020). Analysis of house prices : a hedonic model proposal for Istanbul metropolitan area. Journal of Design for Resilience in Architecture and Planning, 1(1), 57-68.

[3] Ayoub, K., Breuillé, M. L., Grivault, C., Le Gallo, J. & Nappi-Choulet, I. (2019). Hétérogénéité spatiale des prix hédoniques des appartements du marché locatif privé en France. Revue française d'économie, Vol. XXXIV(2), 203-247.

[4] Deutsch, E., Quettier, A. & Simon, A. (2017). L'influence des arrondissements dans la formation des prix immobiliers à Paris : diffusion ou localisation ? Revue d'Économie Régionale & ; Urbaine, Octobre(4), 673-710.

[5] Napoléone Claude. Prix fonciers, immobiliers et localisation des ménages au sein d'une agglomération urbaine. In: Les Annales de la recherche urbaine, N°107, 2012. La ville en thèse. pp. 18-27;

[6] Sedoarisoa, N., Desponds, D., Dalmas, L. & Lavandier, C. (2017). Prix de l'immobilier et proximité de la plate-forme aéroportuaire de Paris

[7] Charles-de-Gaulle (CDG). L'Espace géographique, Tome 46(1), 61-78. Thierry Brossard, Daniel Joly, François-Pierre Tourneux, Jean Cavaillès, Mohamed Hilal, Pierre Wavresky, Julie Le Gallo, Ghislain Geniaux, Claude Napoleone, Hubert Jayet, Nicolas Ovtracht & Pierre-Yves Péguy. (2007). La

valeur économique des paysages des villes périurbanisées. Research Papers in Economics.

[8]Xiaochen Hu, Luechai Chulasai & Sasipen Phuangsai. (2011). Hedonic Pricing Model for Housing Market in City of Kunming, the People's Republic of China. (Journal of Economics Chiang Mai University), 15(1), 144-165.

Table des Annexes

Annexe 1 : Corrélation des distances aux différents centres aux prix.

Annexe 2 : Boxplots des valeurs foncières au mètre carré en fonction de l'éloignement d'un quartier au sein du même arrondissement**

Annexe 3 : Test des Hypotheses MCO

Annexe 1: Corrélation des distances aux différents centres aux prix.

Corrélation de la distance à Part-Dieu et le prix

Pearson's product-moment correlation

```
data: DVF_LYON$VF_au_m2 and DVF_LYON$distancePD
t = -0.097104, df = 1874, p-value = 0.9227
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.04749493  0.04301792
sample estimates:
      cor
-0.002243103
```

Corrélation de la distance à Bellecourt et le prix

Pearson's product-moment correlation

```
data: DVF_LYON$VF_au_m2 and DVF_LYON$distanceBellecourt
t = -13.056, df = 1874, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.3297012 -0.2467205
sample estimates:
```

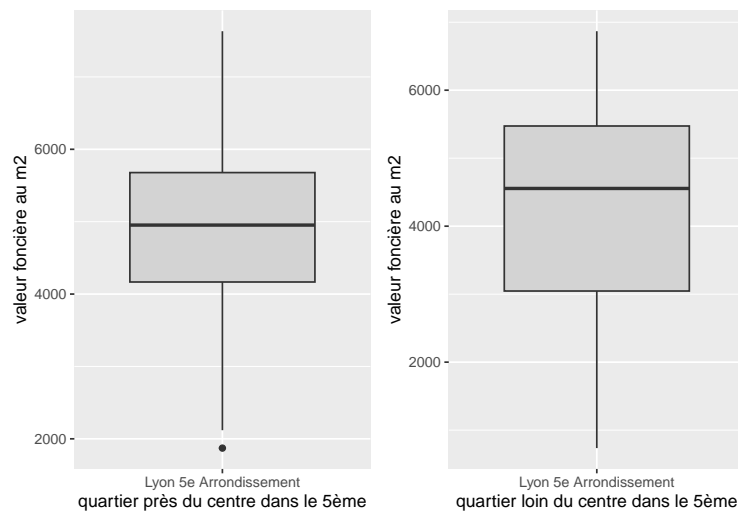
cor

-0.288753

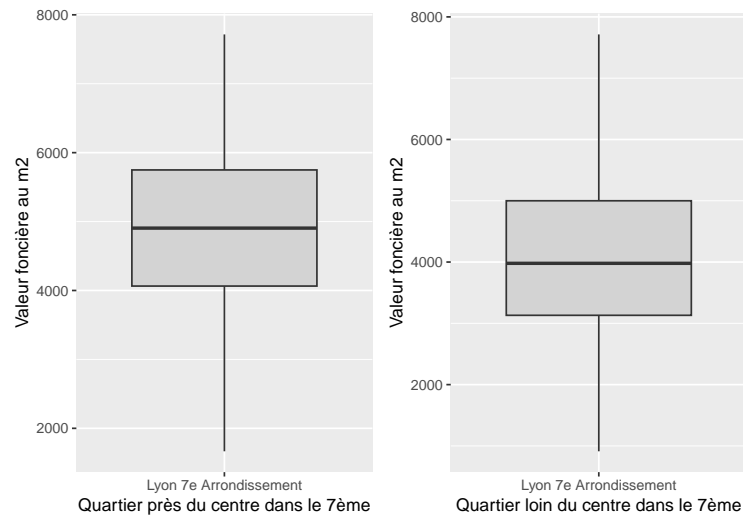
Annexe 2 :

Boxplots des valeurs foncières au mètre carré en fonction de l'éloignement d'un quartier au sein du même arrondissement

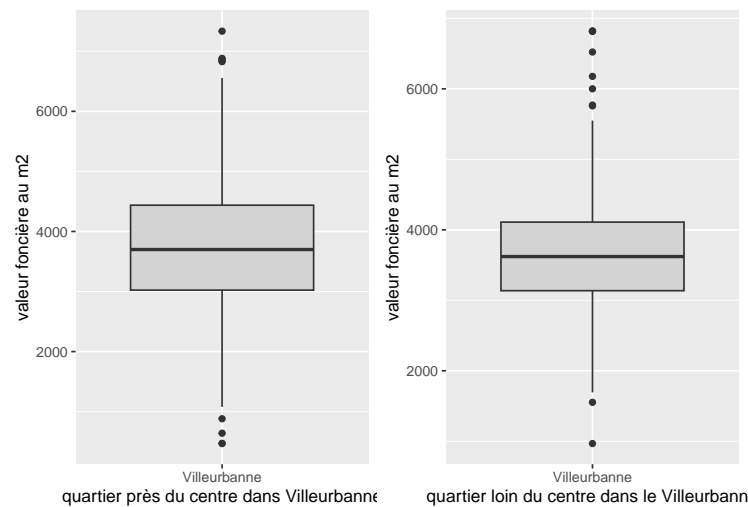
Nous voulons comparer ici l'effet de la distance sur des biens du même arrondissement mais ayant un éloignement au centre différent de par leur position au sein de deux différents quartiers.



On voit que globalement les prix au m2 sont plus élevés pour un quartier du 5ème arrondissement près du centre que loins du centre.



Il en est de même pour le 7e arrondissement. Les prix au mètre carré semblent plus élevé pour les biens proches du centre que pour les biens loin du centre au sein du même arrondissement.



Pour la commune de Villeurbanne on remarque peu de différence dans la répartition des prix au mètre carré pour les biens proches et loin du centre.

Annexe 3 : Test des Hypotheses MCO

Independance des erreurs :

Box-Ljung test

```
data: residuals(Reg5)
```

```
X-squared = 49.013, df = 1, p-value = 2.542e-12
```

Normalité des erreurs :

Shapiro-Wilk normality test

```
data: residuals(Reg5)
```

```
W = 0.99103, p-value = 2.412e-09
```

Homosedasticity

studentized Breusch-Pagan test

```
data: Reg5
```

```
BP = 245.17, df = 4, p-value < 2.2e-16
```