

1. [Accueil](#)
2. Entity Print

Document de travail

[Covid-19](#)

# Données de prix immobiliers extraites d'Internet (web scraping) en temps réel : le cas de la crise de la Covid-19 au Royaume-Uni

31 Août 2021

Auteurs : [Jean-Charles Bricongne](#), Baptiste Meunier, Sylvain Pouget

**Document de travail n°827.** Alors que les données officielles sont disponibles avec un décalage et à un niveau souvent agrégé, de nombreuses informations peuvent être extraites en temps réel et publiquement des sites d'annonces immobilières. Cet article construit ainsi une large base de données en récupérant les annonces immobilières sur internet (web scraping). Cela permet de construire des indicateurs avancés et très granulaires du marché immobilier. L'originalité des données vient aussi de ce qu'elles donnent le point de vue des vendeurs – quand la plupart des statistiques et de la littérature se basent sur les transactions. En exploitant cette originalité, nous construisons des indicateurs innovants comme le nombre de nouvelles offres ou les ajustements de prix des offres déjà sur le marché. Cela permet également de comparer les prix demandés par les vendeurs – issus de notre base de données – avec les prix finaux de transaction – en utilisant la base de données des notaires : cela donne une mesure du pouvoir de négociation des acheteurs. Pendant la crise de la Covid-19, ces indicateurs documentent une baisse drastique de 80 % de l'activité sur le marché immobilier et le comportement attentiste des vendeurs. Par la suite, ils montrent la légère hausse des prix en zone rurale et leur baisse dans la région de Londres.

## [Weekly new offers during covid-19 period](#)

Les statistiques officielles sur le marché du logement résidentiel sont généralement disponibles avec un certain retard et la plupart sont fournies au niveau agrégé, alors que les divergences entre les zones urbaines et rurales ont été bien documentées dans la littérature (par exemple, Poon et Garratt, 2012 pour le Royaume-Uni). L'obtention d'informations fréquentes pourrait être encore plus critique lors d'épisodes de crise tels que la pandémie de Covid-19, car les délais de publication des statistiques officielles ne permettent pas de saisir les tournants importants et soudains de l'activité économique. Par ailleurs, de nombreuses informations sont disponibles publiquement et en temps réel sur les sites immobiliers, en particulier sur le segment résidentiel où 92 % des entreprises immobilières publient des annonces sur internet. L'utilisation de ces données alternatives permettrait alors de construire des indicateurs plus rapidement (en temps réel), à une fréquence plus élevée (quotidienne), et avec une granularité élevée (au niveau du code postal).

Notre approche se concentre sur le Royaume-Uni - mais pourrait être étendue de manière transparente à d'autres pays - et consiste à télécharger régulièrement ("web-scrapers") les offres sur les cinq principaux sites web immobiliers du Royaume-Uni. En moyenne, nous scrapons environ 1,5 million d'offres (à la vente et à la location) par jour avec de nombreuses informations sur le prix, l'emplacement, la surface, le nombre de

pièces, la description, le type d'offres et le type de logement. L'originalité des données recueillies sur le web est d'obtenir le point de vue des vendeurs à travers les offres qu'ils (ou les agences immobilières qu'ils mandatent) publient sur Internet - alors qu'une grande partie de la littérature et toutes les statistiques officielles se basent sur les transactions.

Cet ensemble de données permet d'abord de suivre le marché du logement en temps réel. Le prix de vente peut être suivi quotidiennement et à un niveau très granulaire, offrant une image précoce et plus fine des développements en cours sur le marché - complétant ainsi les statistiques officielles. Dans le même ordre d'idées, les indicateurs habituels du marché du logement (par exemple, le rapport loyer/prix) peuvent être publiés en temps réel. Ces indicateurs complètent les statistiques officielles en donnant un aperçu du point de vue des vendeurs. C'est là que réside l'originalité de notre ensemble de données extraites du web, et ce point de vue particulier permet de créer des indicateurs innovants. Un premier exemple est le nombre de nouvelles offres publiées chaque semaine, qui indique la volonté des vendeurs de mettre leurs biens sur le marché. Un deuxième exemple concerne les changements de prix d'une offre existante : l'analyse quotidienne du web permet de suivre une offre particulière dans le temps et d'observer comment le vendeur ajuste son prix. Il est intéressant de noter que cela donne un signal très précoce de la dynamique du marché du logement, car cela se produit avant même que toute transaction puisse être enregistrée dans les statistiques officielles.

En utilisant ces indicateurs quotidiennement, nous suivons le marché immobilier britannique pendant la crise de Covid-19 et documentons une baisse nette de 80 % du nombre de nouvelles offres pendant le premier confinement (voir la figure 1) tout en montrant que les vendeurs se sont abstenus de modifier leurs prix pendant cette période, ce qui suggère une approche attentiste. À la suite du confinement, les prix de vente moyens ont commencé à augmenter au niveau national. Toutefois, cela cache de grandes disparités entre les régions : alors que les prix ont augmenté régulièrement dans les zones rurales, ils ont diminué à Londres (avec toutefois des effets de composition potentiels) - la région qui a été la plus touchée par le virus, et où les données suggèrent que le marché immobilier est le plus tendu.

Cet ensemble de données permet également de faire correspondre les données extraites du Web avec les données notariales publiées sur les transactions, ce qui permet de calculer la différence entre les prix réalisés et les prix négociés. Il s'agit d'une indication directe de la marge de négociation des acheteurs dans la veine de Galesi et al. (2020), qui peut être calculée à un niveau très granulaire et suivie dans le temps. Dans le cas particulier du Royaume-Uni, cet indicateur montre de grandes disparités entre les régions, la marge de négociation des acheteurs étant la plus faible à Londres.

## **Télécharger l'intégralité de la publication**

[Document de travail n°827. Données de prix immobiliers extraites d'Int... \(PDF - 4 Mo\)](#)