

Rapport - ProjetR Utilisation de jeu de donnée sur les appartements à Seattle

Amine Halhali Ismail Taaissat Mohammed-Ibrahim Benahmida Anas Imarrighen

2025-11-02

Table des matières

1	Introduction	3
1.1	Contexte de l'étude	3
1.2	Objectifs	3
2	Description du jeu de données	3
2.1	Variables disponibles	3
2.2	Description des données	4
3	Analyse descriptive	5
3.1	Statistiques de base	5
3.2	Distribution des variables clés	5
3.3	Visualisations simples	5
4	Analyse exploratoire bivariée	5
4.1	Corrélations entre variables numériques	5
4.2	Comparaison des prix selon des catégories (t-test, ANOVA)	5
4.3	Visualisations : scatterplots, boxplots, heatmaps	5
5	Modélisation	5
5.1	Régression linéaire multiple	5
5.2	Régression logarithmique	5
6	Analyse géographique	5
6.1	Cartographie des prix par quartier ou coordonnées	5
6.2	Cartes de densité	5
6.3	Comparaison géographique des prix	5

7	Interprétation et conclusions	5
7.1	Facteurs influençant le prix	5
7.2	Quartiers ou caractéristiques les plus chers/abordables	5
7.3	Limites de l'analyse et recommandations	5

1 Introduction

1.1 Contexte de l'étude

Le marché des locations de courte durée via des plateformes comme Airbnb connaît un développement rapide dans les grandes villes. Comprendre les facteurs qui influencent le prix d'une location est essentiel pour les propriétaires souhaitant optimiser leurs revenus, mais aussi pour les voyageurs souhaitant comparer les offres.

1.2 Objectifs

L'objectif de cette analyse est donc de pouvoir observer le lien de corrélation et/ou de causalité entre les paramètres mis à disposition dans le jeu de données. Nous nous attarderons sur les points importants comme le prix et les avis comme base pour notre étude afin de bien répondre à la problématique: Quels facteurs influencent le prix d'une location Airbnb ?

2 Description du jeu de données

2.1 Variables disponibles

On compte pas moins de 17 paramètres dans le jeu de données:

- room_id
- host_id
- room_type
- adress
- reviews
- overall_satisfaction
- accommodates
- bedrooms
- bathrooms
- price
- last_modified
- latitude
- longitude
- location
- name
- currency
- rate_type

2.2 Description des données

room_id : cet élément représente l'identifiant du rbnb à louer

host_id : cet élément représente l'identifiant de la personne qui loue le rbnb

room_type : cet élément représente le type de rbnb

adress : cet élément représente l'adresse du rbnb

reviews : cet élément représente le nombre d'avis sur le rbnb

overall_satisfaction : cet élément représente la note moyenne du rbnb

accommodates :

bedrooms : cet élément représente le nombre de chambre du rbnb

bathrooms : cet élément représente le nombre de salle de bain du rbnb

price : cet élément représente le prix en fonction du rate_type et de currency

last_modified : cet élément représente la dernière modification du rbnb

latitude : cet élément représente la latitude du rbnb

longitude : cet élément représente la longitude du rbnb

location :

name : cet élément représente le nom du rbnb

currency : cet élément représente la devise utilisé pour payer

rate_type : cet élément représente le mode de facturation du rbnb

3 Analyse descriptive

3.1 Statistiques de base

3.2 Distribution des variables clés

3.3 Visualisations simples

4 Analyse exploratoire bivariée

4.1 Corrélations entre variables numériques

4.2 Comparaison des prix selon des catégories (t-test, ANOVA)

4.3 Visualisations : scatterplots, boxplots, heatmaps

5 Modélisation

5.1 Régression linéaire multiple

5.2 Régression logarithmique

6 Analyse géographique

6.1 Cartographie des prix par quartier ou coordonnées

6.2 Cartes de densité

6.3 Comparaison géographique des prix

7 Interprétation et conclusions

7.1 Facteurs influençant le prix

7.2 Quartiers ou caractéristiques les plus chers/abordables

7.3 Limites de l'analyse et recommandations