projet

groupe 5 (Wilson FOTSING, Laurent Fabrice NGWE, Gladisse Diane KOUMBO)

2025-05-14

## Description des données

Le jeu de données contient les prix de ventes des maisons dans une compté de la ville de Seattle aux USA. ces maisons ont été vendues entre mai 2014 et mai 2015.

Tirées sur Kagle <https://www.kaggle.com/datasets/harlfoxem/housesalesprediction>

licence : CC0: Public Domain

Le jeu a 21613 enregistrement et 21 variables.

Bien que le set se fait un peu vieux, il reste d’actualité et vivement apprécié, compte tenu de ses 8789 vues et ses 2399 teléchargement (y compris le notre) durant les 30 derniers jours.

### Description des variables :

## nettoyage de la base des données :

#sum(is.na(house))  
# on va aussi modifier la colonne date pour utiliser le format ymd  
  
if (!inherits(house$date, "Date")) {  
 house <- house %>%  
 mutate(  
 date = ymd(substr(date, 1, 8))   
 )  
}  
  
head(house)%>%  
 kable()

| id | date | price | bedrooms | bathrooms | sqft\_living | sqft\_lot | floors | waterfront | view | condition | grade | sqft\_above | sqft\_basement | yr\_built | yr\_renovated | zipcode | lat | long | sqft\_living15 | sqft\_lot15 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 7129300520 | 2014-10-13 | 221900 | 3 | 1.00 | 1180 | 5650 | 1 | 0 | 0 | 3 | 7 | 1180 | 0 | 1955 | 0 | 98178 | 47.5112 | -122.257 | 1340 | 5650 |
| 6414100192 | 2014-12-09 | 538000 | 3 | 2.25 | 2570 | 7242 | 2 | 0 | 0 | 3 | 7 | 2170 | 400 | 1951 | 1991 | 98125 | 47.7210 | -122.319 | 1690 | 7639 |
| 5631500400 | 2015-02-25 | 180000 | 2 | 1.00 | 770 | 10000 | 1 | 0 | 0 | 3 | 6 | 770 | 0 | 1933 | 0 | 98028 | 47.7379 | -122.233 | 2720 | 8062 |
| 2487200875 | 2014-12-09 | 604000 | 4 | 3.00 | 1960 | 5000 | 1 | 0 | 0 | 5 | 7 | 1050 | 910 | 1965 | 0 | 98136 | 47.5208 | -122.393 | 1360 | 5000 |
| 1954400510 | 2015-02-18 | 510000 | 3 | 2.00 | 1680 | 8080 | 1 | 0 | 0 | 3 | 8 | 1680 | 0 | 1987 | 0 | 98074 | 47.6168 | -122.045 | 1800 | 7503 |
| 7237550310 | 2014-05-12 | 1225000 | 4 | 4.50 | 5420 | 101930 | 1 | 0 | 0 | 3 | 11 | 3890 | 1530 | 2001 | 0 | 98053 | 47.6561 | -122.005 | 4760 | 101930 |

**Nombre de valeurs manquantes : 0**

## Description du site

Initialement, le visiteur est sur la page d’acceuil présentée précédament, Il a la possibilité de choisir son profil : **acheteur** ou **Propriétaire**.

S’il fait le choix de rester en mode visite (il ne choisit aucun profil), alors, il a la possibilité de naviguer à travers la carte du Compté. La carte quant à elle, est divisée en plusieurs zones géographiques (le nombre de régions n’est pas encore déterminé) séparés par des grilles.

À mesure qu’il navigue à travers les grilles, si le curseur est dans une grille, ce qui se passe :

* on affiche au dessus du curseur, le nom du quartier où il se trouve (on a les coordonnées géographiques) et le prix moyen des propriétés dans ce quartier.
* Les graphiques G1,G2, G3,G4, D1,D2,D3,D4 se mettent à jour localement lorsqu’on navigue sur la carte.

## Description des graphiques dynamiques :

G1 : Distribution des prix dans la localité selecionnée : (count = f(price)) –> histogramme

G2 : relation prix ~ surface –> (geom\_plot+geom\_Smooth) coloré selon view

G3 : Tendance des prix moyens dans le temps –> serie temporelle

G4 : Distribution du nombre de chambres –> Diagramme à bandes

D1 : Nombre de maisons par categories –> diagramme à bandes

D2 : Nombre de maison par grade (Qualité de la construction et du design ) –> diagramme à bandes

D3 : Vue sur l’eau oui ou non (waterfront) –> Diagramme circulaire

D4 : Nombre de maison selon l’état general (condition) –> diagramme à bandes

Pour résumer, les graphiques de gauche () donnent une idée générale du coût de vie, alors que les graphiques de droite donne une idée générale sur la qualité de vie

* Si le visiteur pense qu’un graphe est trop petit, il a la possibilité de l’agrandir en le mettant en avant plan. Il peut par exemple double cliquer sur la zone correspondante pour stopper l’effet dynamique, puis aller cliquer sur le graphe qui l’interesse pour le mettre en avant plan. pour revenir à l’effet dynamique, il doit double cliquer à nouveu n’importe où dans la grille.

## La barre de recherche :

Si il connaît le code postal ou le nom du quartier, il peut simplement entrer les informations dans le menu rechercher et les graphes se mettrons à jour automatiquement.

## Les différents profils :

### profil acheteur :

Si il clique dans la case correspondante, un menu deroulant se déploie. Les options sont :

* localite : En choisissant localité, on lui déroule une liste des localités disponibles. il lui suffira alors de cliquer dessus et il sera redirrigé vers une nouvelle page contenant uniquement les propriétés dans cette localité.

il pourra cliquer sur la propriété qui l’interesse pour voir le contenu (photos,prix,… )

* Tranches des prix : Si il clique dessus, un menu déroulant contenant les différentes tranches des prix (tranches de 25000$) apparaît. Il lui suffira faire un choix pour être redirigé vers la page correspondant aux propriétés désirées
* Année de Construction –> idème

Pour faire court, le comportement précédant sera analogue pour tout ce qui suit

* Nombre de chambres
* Nombre de salles de bain
* Superficie habitable
* Superficie du terrain
* Nombre d’étages
* Vue sur l’eau
* Qualité de la vue
* État général de la maison
* Qualité de la construction et du design
* Année de rénovation

**Notons aussi qu’il peut faire sa recherche en rajoutant des filtres.**

finalement, une fois qu’il a été redirigé sur une page, il pourra toujours retourner vers le site en cliquant sur retour ou en appuyant la touche echap du clavier.

### profil Propriétaire (ou vendeur, pas encore décidé) :

Si il clique dessus, le menu deroulant s’affiche

* s’identifier : il doit rentrer ses paramètres de connexion. Si l’identification est réussie, il est redirigé vers la page Vendeur (voir le fichier powerpoint)
* S’il n’est pas inscrit, il pourra s’inscrire, puis s’identifier.

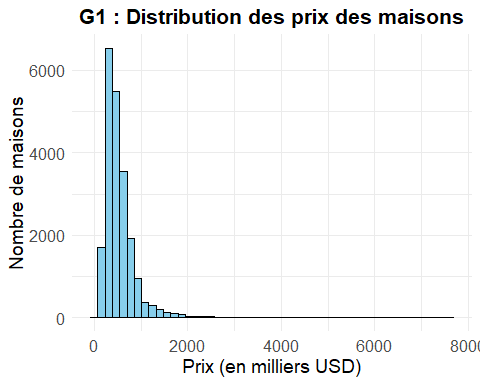
Le site a exactement le même fonctionnement que pour le profil visiteur, sauf qu’en plus, il peut choisir de :

* Publier une annonce : il devra alors rentrer les caractéristiques de sa maison en remplissant un formulaire. le formulaire est simple ; on lui donne la liste des variables, et il doit compléter. À la fin, il a la possibilité de télécharger quelques photos de sa propriété s’il le souhaite.

Remarque : Dans le formulaire, le pris de la maison sera la dernière case à remplir. On souhaite qu’il entre toutes les informations d’abord ; à partir de ces informations, on va faire une estimation (prédiction) du prix de sa propriété, il pourra alors s’en inspirer pour fixer son prix.

* S’il clique sur Historique de mes ventes, on lui affiche la liste des annonces qu’il a déjà faites. il pourra alors consulter chaque annonce en tant que acheteur (pas le droit de modifier).
* S’il clique sur acheter, le site se comporte exactement comme si il était dans le profil acheteur, à la seule différence qu’il reste connecté.
* Il a aussi l’option paramètre où il pourra changer son mot de passe…

## Les graphiques :



## `geom\_smooth()` using formula = 'y ~ x'

