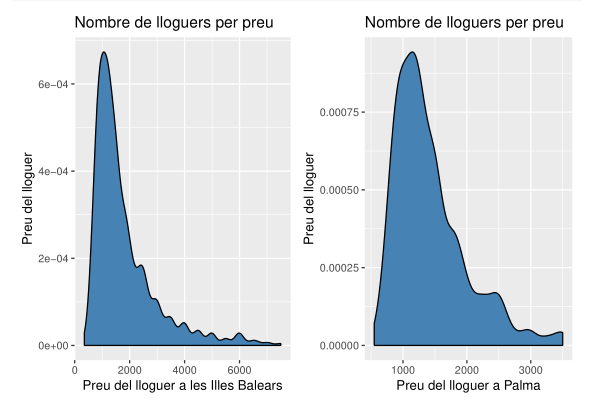
## Resum del projecte

### Descripció del conjunt de dades

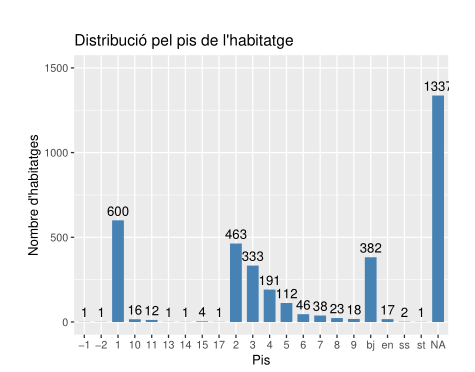
El conjunt de dades es correspon amb les característiques dels lloguers d’habitatges particulars publicats al portal immobiliari Idealista per a les Illes Balears[[1]](#footnote-1). El dataset és públic[[2]](#footnote-2) i de lliure accés[[3]](#footnote-3).

El conjunt de dades es correspon amb 3.600 registres i 25 variables. No totes les variables són rellevants i són propícies per a ser visualitzades, com la referència externa de l’habitatge o la direcció postal. Les variables que es tindran en compte son:

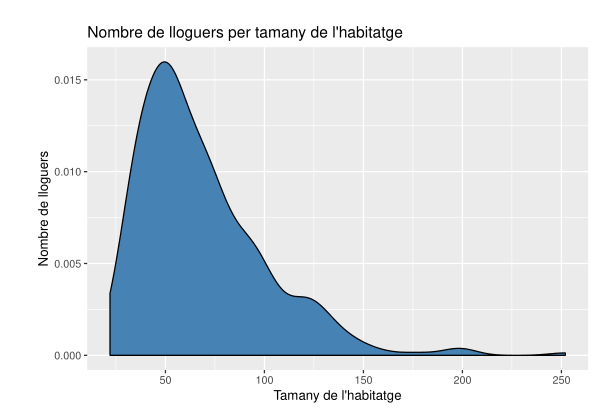
1. **price.** Variable numèrica.Preu mensual de l’habitatge en lloguer en euros. Serà la variable objectiu i totes les demés es mostraran en funció d‘aquesta. És a dir: com influeix una característica determinada en el preu de l’habitatge? Llevant els outliers, podem veure la distribució de preus tant a Palma com a les Illes Balears.



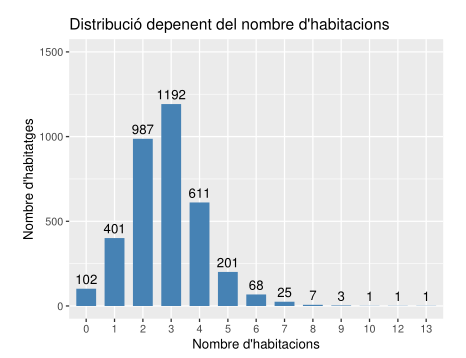
1. **numPhotos.** Variable numèrica. Nombre de fotos publicades a l’anunci de l’habitatge. Pot ser interessant mostrar l’afectació que té sobre el preu el nombre de fotos; sembla clar que més fotos pot ser un indicatiu de més interès de llogar per part del propietari. El mínim és 0 i el màxim és 90, amb un percentil 50 de 16.
2. **floor.** Variable categòrica. Pis de l’habitatge. Combinat amb la variable **hasLift** pot afectar de forma important sobre el preu. NA es correspon amb valors desconeguts o no introduïts.



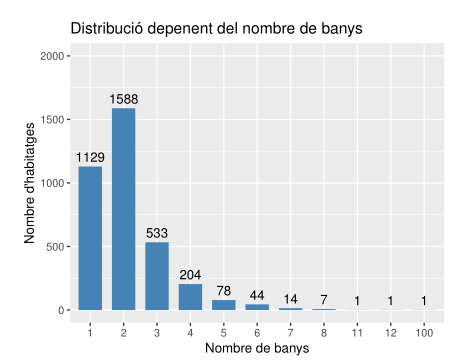
1. **size.** Variable numèrica. Tamany de la propietat en metres quadrats. La correlació amb el preu és molt clara. S’ha de dir que no és una dada homogènia: a alguns habitatges es correspon amb la superfície total; a d’altres, amb el total de la propietat. Llevant els outliers, podem veure la distribució de tamanys tant a les Illes Balears



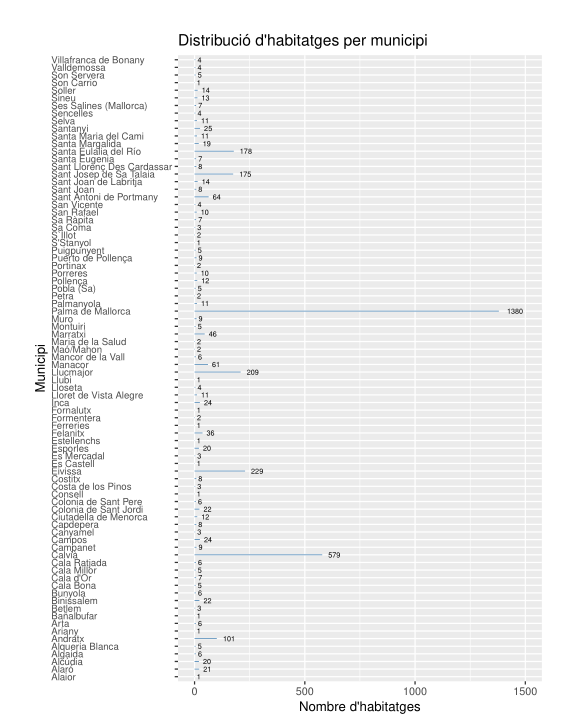
1. **rooms.** Variable numèrica. Nombre d’habitacions. És la variable amb una major correlació amb el preu.



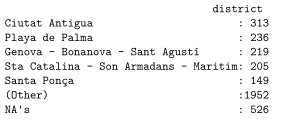
1. **bathrooms.** Variable numèrica. Nombre de banys de l’habitatge.



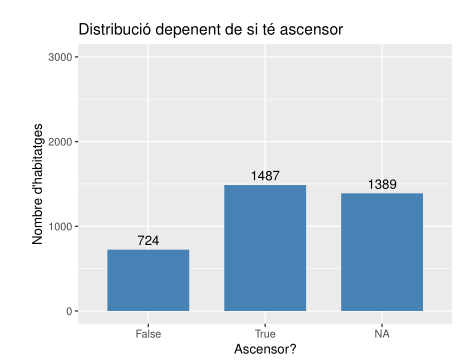
1. **municipality.** Variable categòrica. Nucli on és l’habitatge. Pot ser de molta utilitat per a la visualització de les dades. Com a exemple es mostraran les dades a Palma. S’ha de dir que, malgrat el nom, no es correspon totalment amb el municipi[[4]](#footnote-4). Podem veure la distribució per nucli (municipi si seguim la nomenclatura d’Idealista):



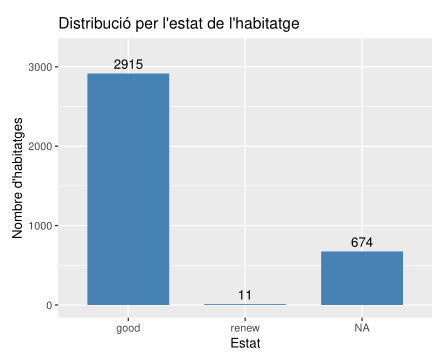
1. **district.** Variable categòrica. Districte o barri de l’habitatge. Pot ser de molta utilitat per a la visualització de les dades. No mostrem totes les dades, simplement com a exemple:



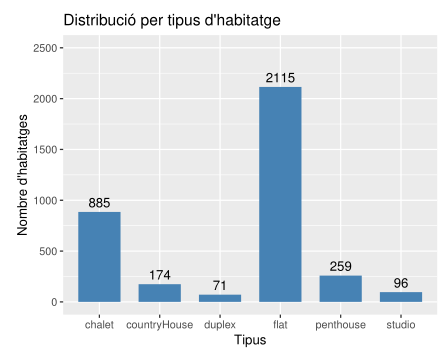
1. **latitud.** Latitud de la coordenada física on es situa l’habitatge. Molt útil per a la visualització de dades.
2. **longitud.** Longitud de la coordenada física on es situa l’habitatge. Molt útil per a la visualització de dades.
3. **hasLift.** Valor booleà que indica si l’habitatge té ascensor o no. Evidentment, és només rellevant depenent del tipus i pis de l’habitatge.



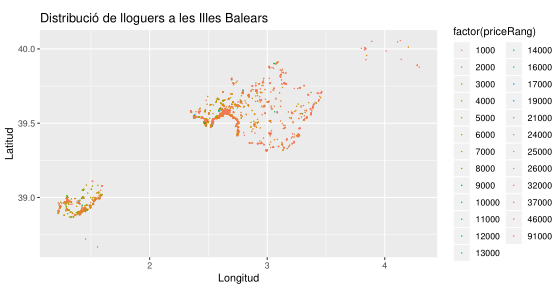
1. **status.** Variable categòrica. Estat de l’habitatge. Útil per a la visualització de dades.

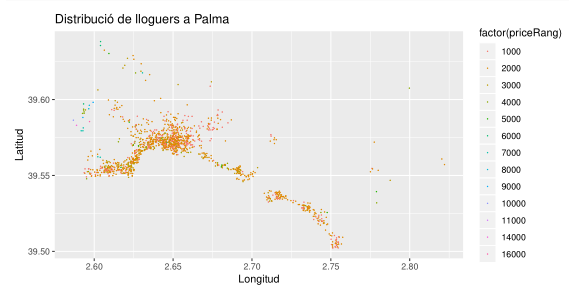


1. **propertyType.** Variable categòrica. Tipus de propietat. Té una afectació directa sobre el preu.



Com a exemple, mostrem la distribució per preu tant a les Illes Balears com a Palma.





1. El conjunt de dades original es corresponia amb un dataset trobat a una de les webs que se’ns varen proposar. Per a una altra assignatura he obtingut el dataset al que faig referencia. Em sembla més interessant per dues raons: és més actual i està més relacionat amb el meu entorn. A més, és un tema igualment actual. [↑](#footnote-ref-1)
2. <https://github.com/caran77/M2.960---Periodisme-de-dades/blob/master/PAC2/dataOutput.csv> [↑](#footnote-ref-2)
3. Les dades s’han obtingut accedint a l’API privada d’Idealista. S’han fet 72 crides, obtenint fitxers JSON de 50 registres cada un. Posteriorment, mitjançant un programa fet en Python s’ha creat el dataset. Tot el codi font és accessible a la mateixa carpeta de GitHub. [↑](#footnote-ref-3)
4. Com a exemple: Cala Rajada és una població però no un municipi. [↑](#footnote-ref-4)