XAI 3: Model-Agnostic methods

Code

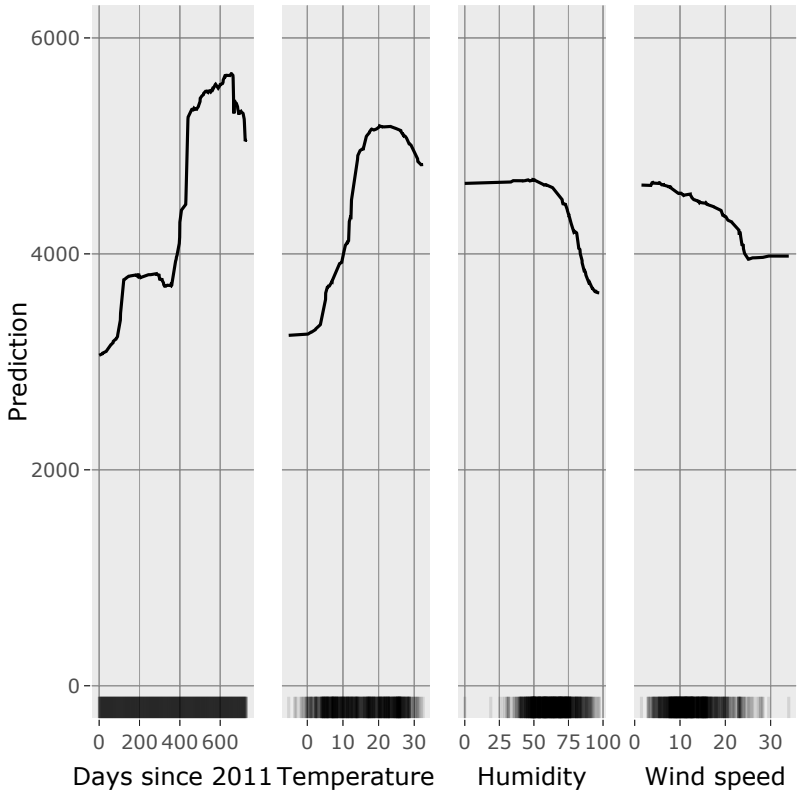
EXERCISE:

Apply PDP to the regression example of predicting bike rentals. Fit a random forest approximation for the prediction of bike rentals (**cnt**). Use the partial dependence plot to visualize the relationships the model learned. Use the slides shown in class as model.

QUESTION:

Analyse the influence of **days since 2011, temperature, humidity** and **wind speed** on the predicted bike counts.

```
Warning messages:
1: In readChar(file, size, TRUE) : truncating string with embedded nuls
2: In readChar(file, size, TRUE) : truncating string with embedded nuls
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.

Attaching package: 'randomForest'

The following object is masked from 'package:pre':

    importance

The following object is masked from 'package:ggplot2':

    margin

The following object is masked from 'package:dplyr':

    combine

The working directory was changed to /Users/cmonserr/OneDrive - UPV/Trabajo_2/Asignaturas/Evaluacion de modelos/P
racticas/Practica 3/Bike-Sharing-Dataset inside a notebook chunk. The working directory will be reset when the ch
unk is finished running. Use the knitr root.dir option in the setup chunk to change the working directory for not
ebook chunks.
```



EXERCISE:

Generate a 2D Partial Dependency Plot with humidity and temperature to predict the number of bikes rented depending of those parameters.

BE CAREFUL: due to the size, extract a set of random samples from the BBDD before generating the the data for the Partial Dependency Plot.

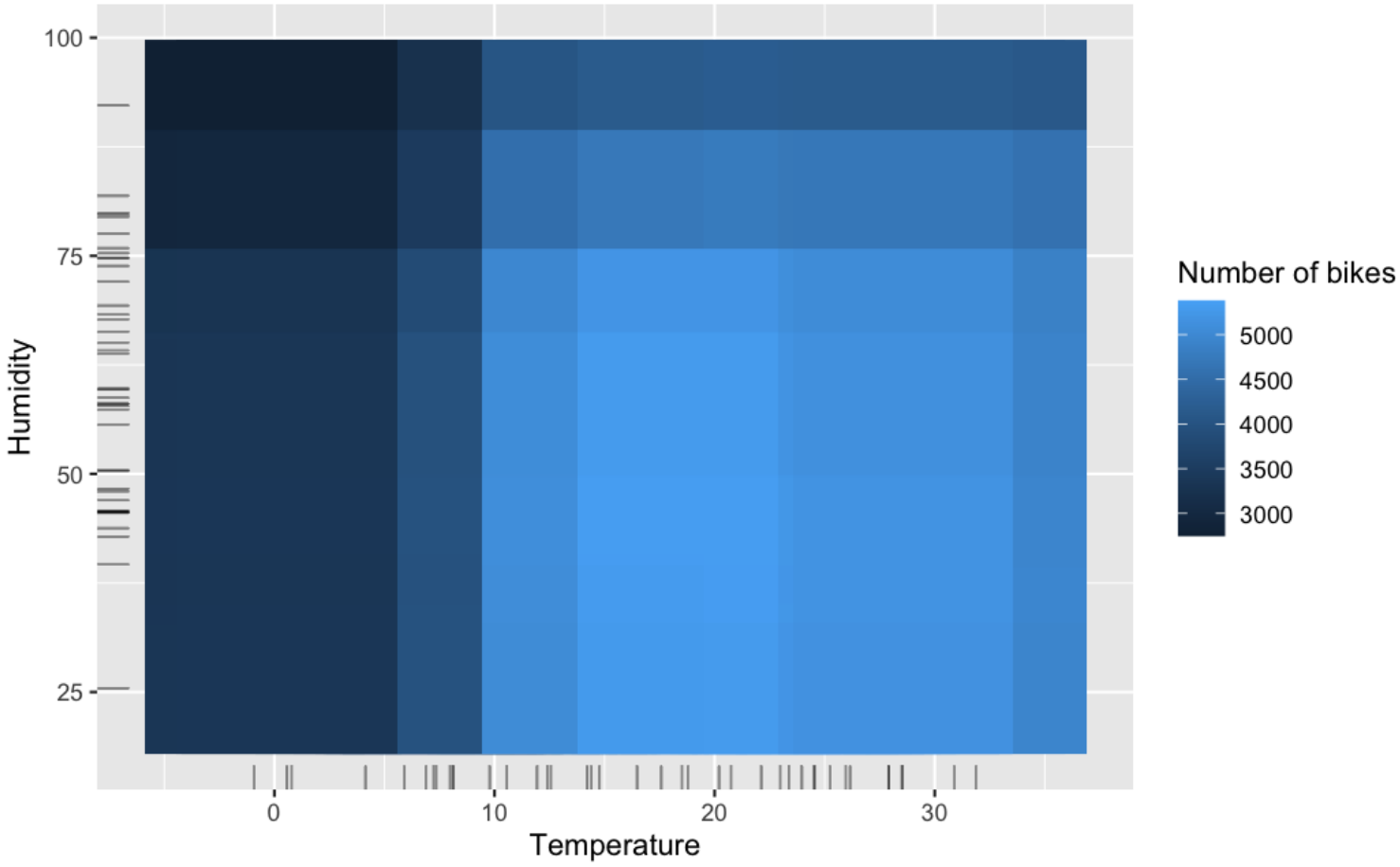
Show the density distribution of both input features with the 2D plot as shown in the class slides.

TIP: Use geom_tile() to generate the 2D plot. Set width and height to avoid holes.

QUESTION:

Interpret the results.

```
entra: 30.326 sec elapsed
```



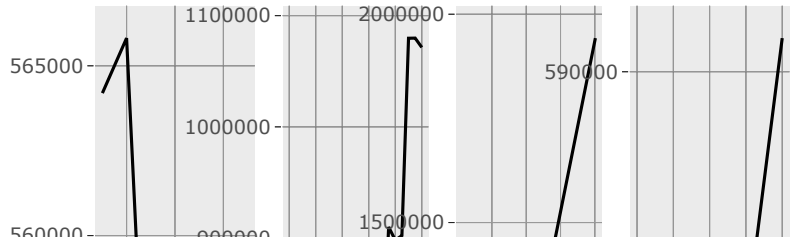
EXERCISE:

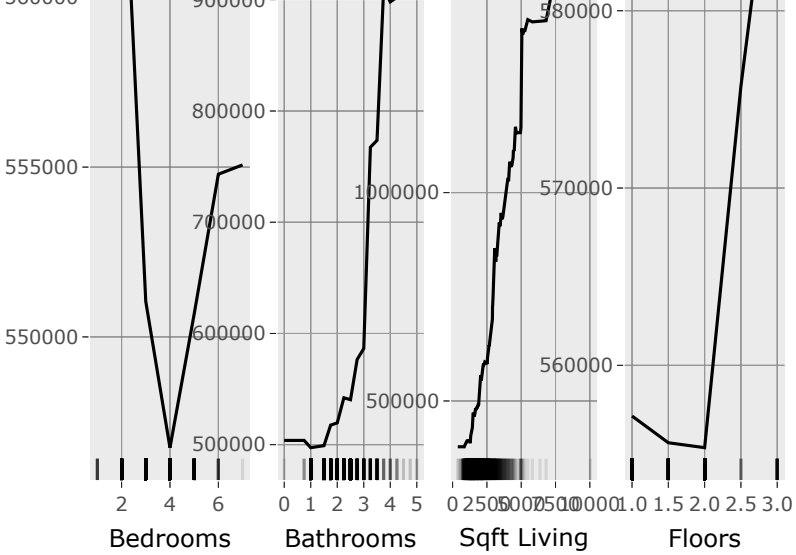
Apply the previous concepts to predict the **price** of a house from the database **kc_house_data.csv**. In this case, use again a random forest approximation for the prediction based on the features **bedrooms**, **bathrooms**, **sqft_living**, **sqft_lot**, **floors** and **yr_built**. Use the partial dependence plot to visualize the relationships the model learned.

BE CAREFUL: due to the size, extract a set of random samples from the BBDD before generating the data for the Partial Dependency Plot.

QUESTION:

Analyse the influence of **bedrooms, bathrooms, sqft_living** and **floors** on the predicted price.




900000
800000
700000
600000
500000
555000
550000
2
4
6
Bedrooms
1000000
500000
0
1
2
3
4
5
Bathrooms
580000
570000
560000
0
2500
5000
7500
10000
Sqft Living
1.0
1.5
2.0
2.5
3.0
Floors