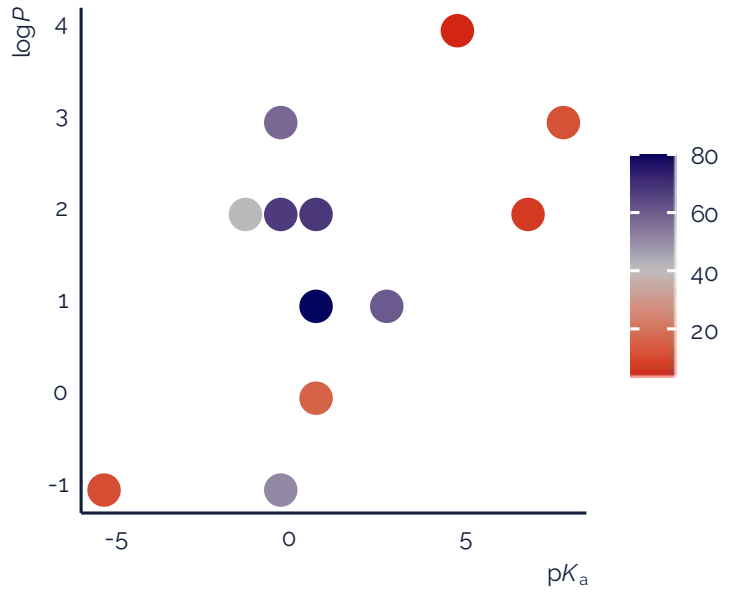
Random Forest Regression

You have now developed an extraction method for the analysis of pesticides from the previous seminar. For some of the pesticides, the recovery of extraction is good while for others it is much lower. You would now like to predict quantitiatively the extraction recovery from the pK_a and log*P* values. You will develop a Random Forest Regression for this task, where each group developes one single Regression Tree.

(1) Your first task is to develop a Regression Tree. Do not forger to: (1) create a bootstrap sample and (2) feature subspace sampling. Your accuracy metric is RMSE and smallest allowed leaf size is 3.

(2) Draw your Regression Tree to the board next to the trees fro other groups and add your predictions here:

https://docs.google.com/spreadsheets/d/1g2iblxZ_LiwuAkKJfyb7xdGR0FSYIVzIk4cEttG98OI/edit?usp=sharing



Pesticide	pK_a	log*P*	Recovery
Aldicarb sulphoxide	1	0	14
Aldicarb sulphone	0	-1	51
Demeton-*S*-methyl sulphoxide	-5	-1	9
Carbendazim	7	2	5
Methomyl	1	1	81
Thiabendazole	8	3	10
Methiocarb sulphoxide	-1	2	41
Methiocarb sulphone	0	2	68
Aldicarb	3	1	61
Imazalil	5	4	2
Thiodicarb	1	2	69
Phorate sulphoxide			57
Phorate sulphone			64
Methiocarb	0	3	58