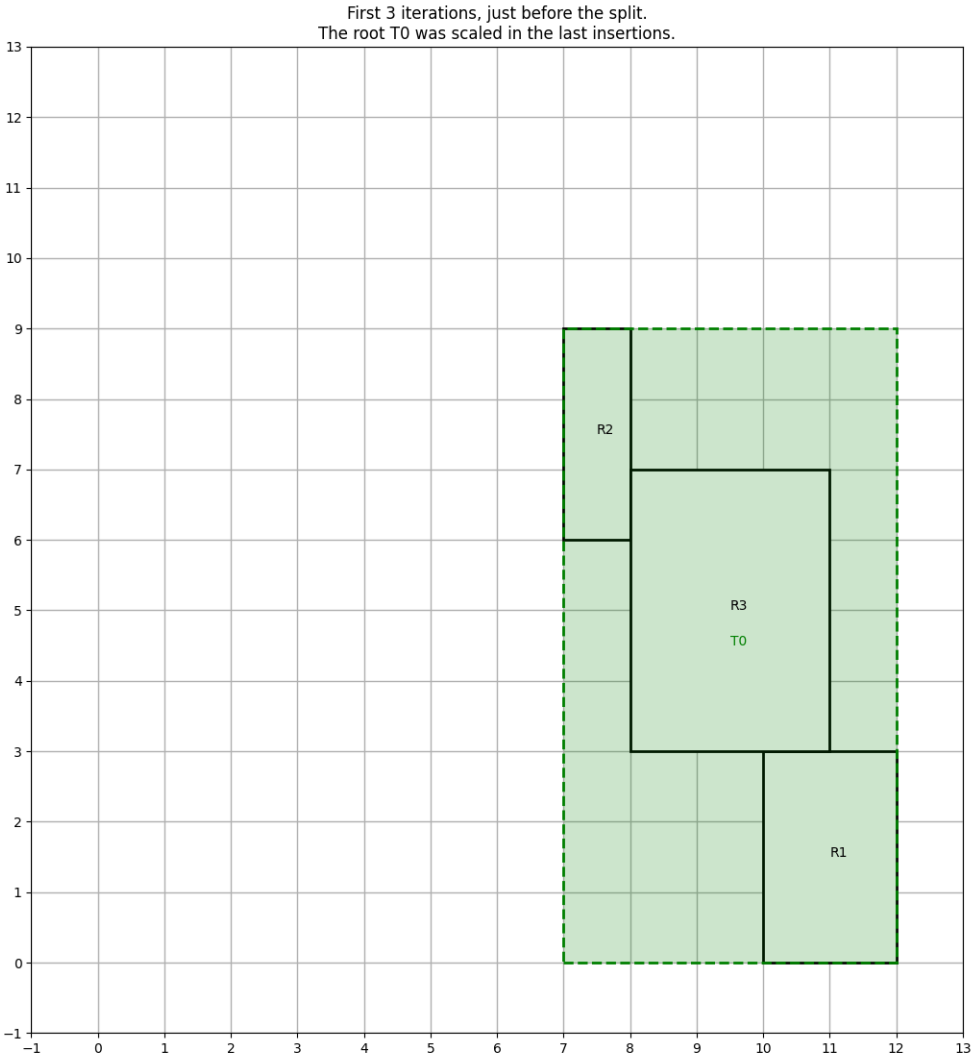
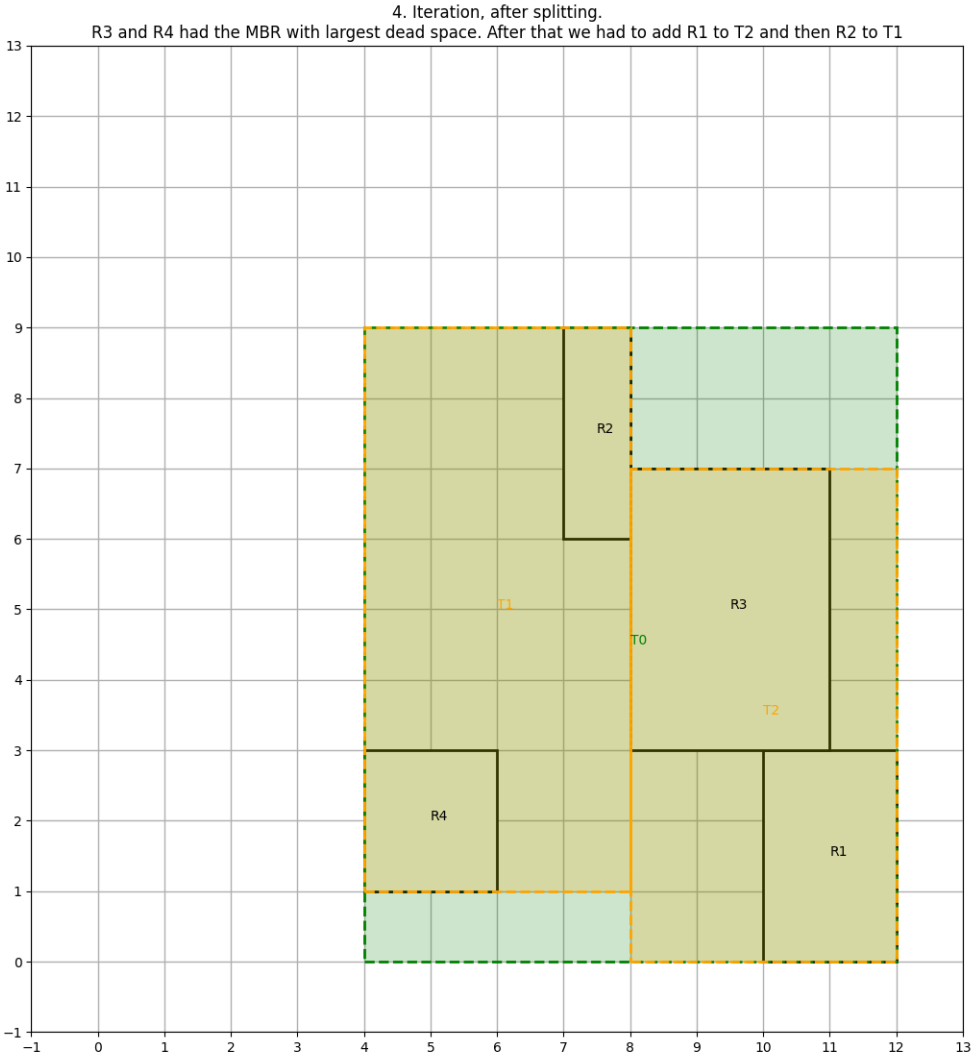
Data Science Assignment 5

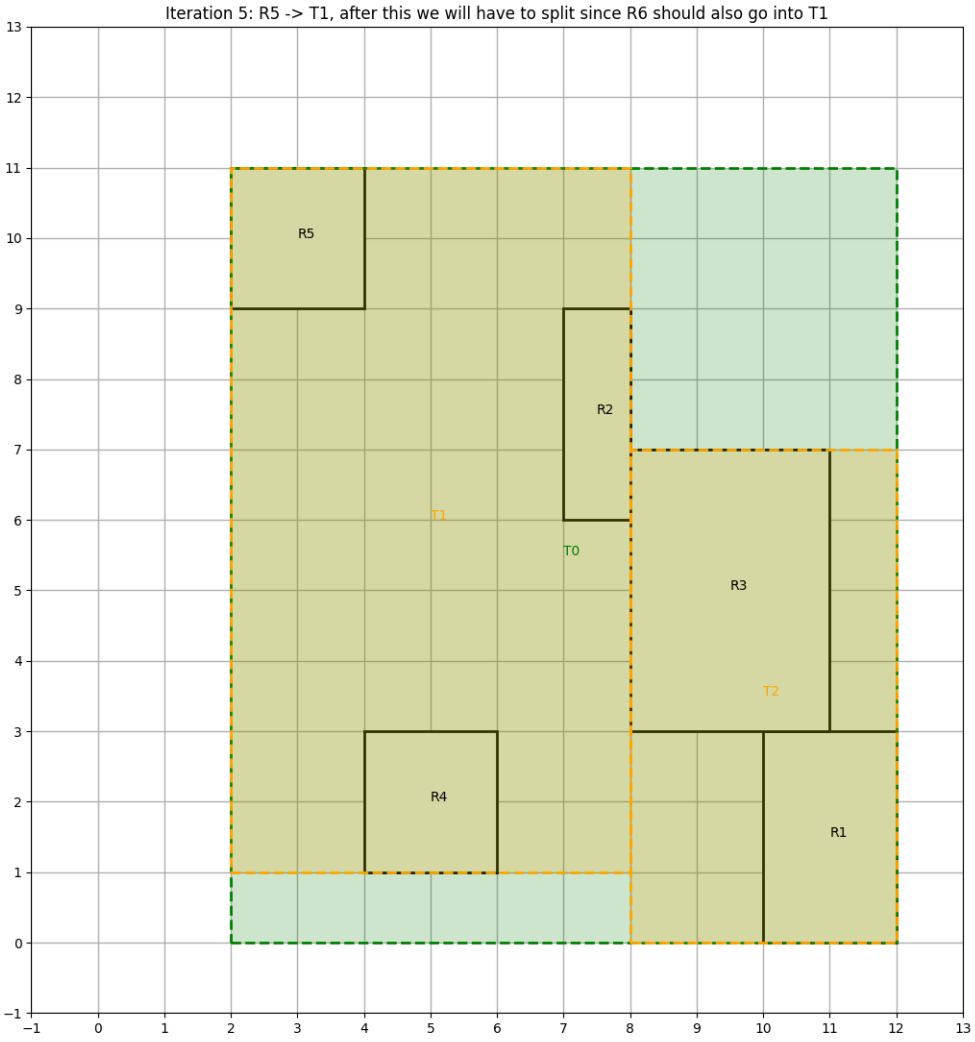
Nike Marie Pulow – Henri Paul Heyden
stu239549 – stu240825

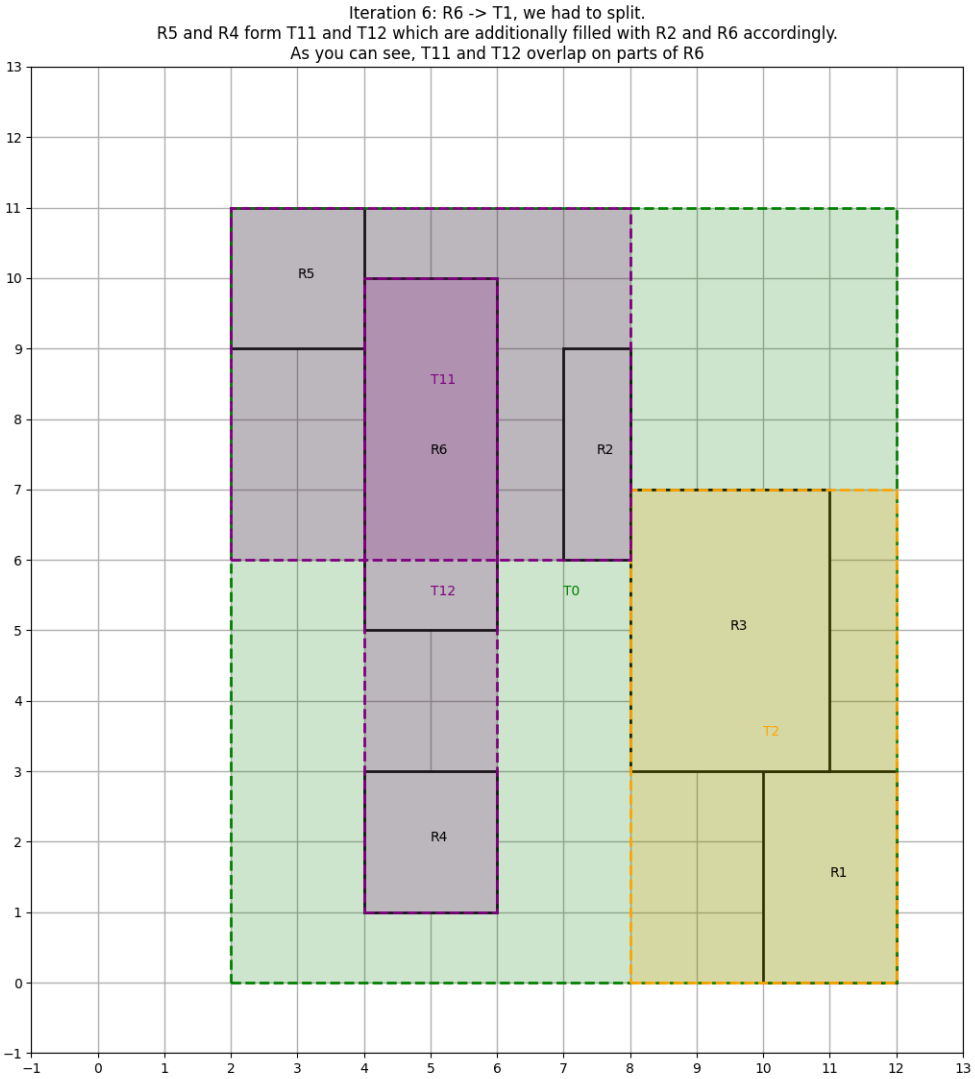
1 R-Tree

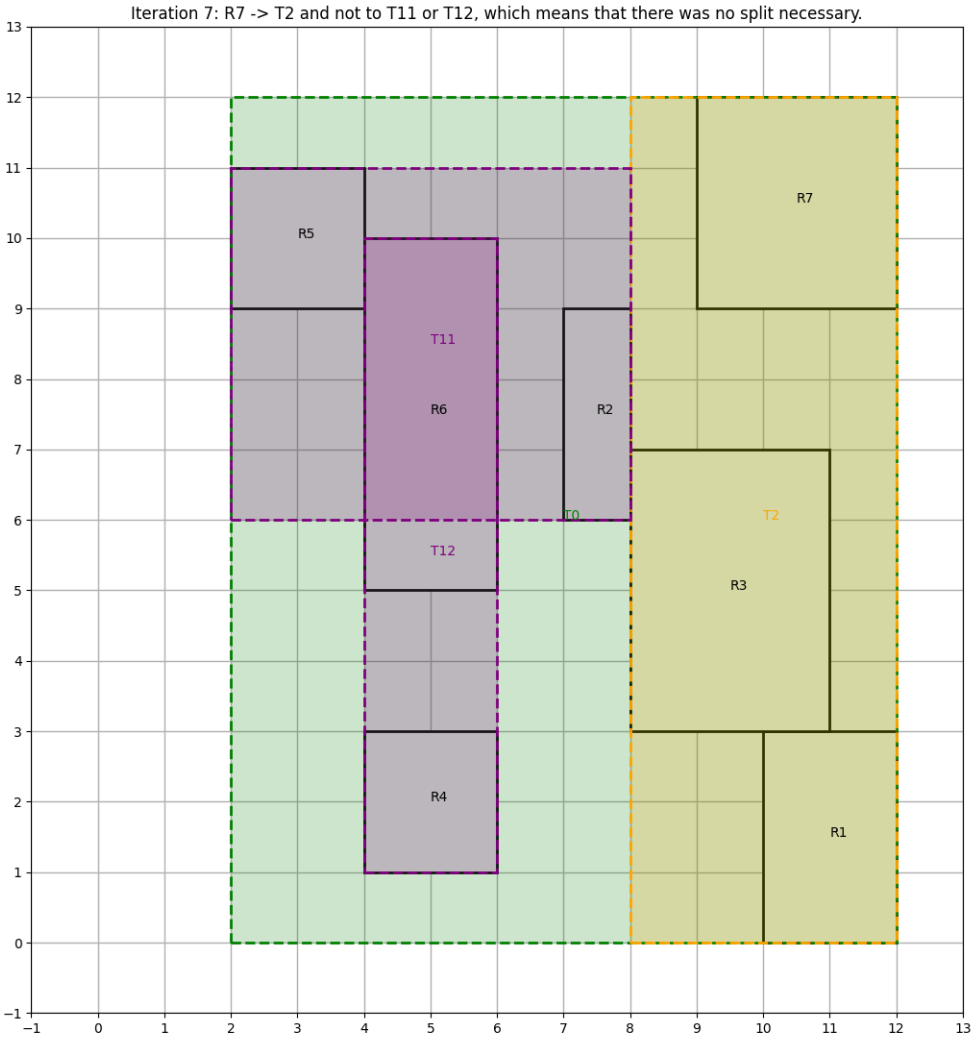
The following is an illustrated computation of the 7 insertions and needed splits.
The first 2 insertions were not included since they are trivial.




4. Iteration, after splitting.
R3 and R4 had the MBR with largest dead space. After that we had to add R1 to T2 and then R2 to T1
R2
R3
R4
R1
T0
T1
T2


Iteration 5: R5 -> T1, after this we will have to split since R6 should also go into T1
R5
R2
T1
T0
R3
T2
R4
R1


Iteration 6: R6 -> T1, we had to split.
R5 and R4 form T11 and T12 which are additionally filled with R2 and R6 accordingly.
As you can see, T11 and T12 overlap on parts of R6
R5
T11
R6
R2
T12
T0
R3
T2
R4
R1


Iteration 7: R7 -> T2 and not to T11 or T12, which means that there was no split necessary.
R7
R5
T11
R6
R2
T0
T2
T12
R3
R4
R1

2 kNN-Index-APL

1: $q = (9, 9)$, $k = 3$

Iteration	APL	*result*	*pruningdist*
1.	[(0,R1), (0,R2)]	[$(\infty, null)$, $(\infty, null)$, $(\infty, null)$]	∞
2.	[(0,R2), (0,A), (0,B)]	[$(\infty, null)$, $(\infty, null)$, $(\infty, null)$]	∞
3.	[(0,A), (0,B), (4,R21), (4,R22)]	[$(\infty, null)$, $(\infty, null)$, $(\infty, null)$]	∞
4.	[(0,B), (4,R21), (4,R22)]	[(3,A), $(\infty, null)$, $(\infty, null)$]	∞
5.	[(4,R21), (4,R22)]	[(3,A), (13,B), $(\infty, null)$]	∞
6.	[(4,R22), (6,C), (6,D)]	[(3,A), (13,B), $(\infty, null)$]	∞
7.	[(4,E), (4,F), (6,C), (6,C)]	[(3,A), (13,B), $(\infty, null)$]	∞
8.	[(4,F), (6,C), (6,D)]	[(3,A), (6,E), (13,B)]	**13**
9.	[(6,C), (6,D)]	[(3,A), (6,E), (6,F)]	**6**
10.	[(6,D)]	[(3,A), (6,E), (6,F)]	**6**
11.	[]	[(3,A), (6,E), (6,F)]	**6**

2: $q = (4, 6)$, $k = 2$

Iteration	APL	*result*	*pruningdist*
1.	[(0,R1), (0,R2)]	[$(\infty, null)$, $(\infty, null)$]	∞
2.	[(0,R2), (2,A), (2,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
3.	[(1,R21), (1,R22), (2,A), (2,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
4.	[(1,R22), (2,A), (2,B), (4,C), (4,D)]	[$(\infty, null)$, $(\infty, null)$]	∞
5.	[(2,A), (2,B), (2,E), (2,F), (4,C), (4,D)]	[$(\infty, null)$, $(\infty, null)$]	∞
6.	[(2,B), (2,E), (2,F), (4,C), (4,D)]	[(6,A), $(\infty, null)$]	∞
7.	[(2,E), (2,F), (4,C), (4,D)]	[(6,A), (14,B)]	**14**
8.	[(2,F), (4,C), (4,D)]	[(2,E), (6,A)]	**6**
9.	[(4,C), (4,D)]	[(2,E), (2,F)]	**2**

3: $q = (7, 5)$, $k = 2$

Iteration	APL	*result*	*pruningdist*
1.	[(0,R1), (0,R2)]	[$(\infty, null)$, $(\infty, null)$]	∞
2.	[(0,R2), (3,A), (3,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
3.	[(0,R21), (0,R22), (3,A), (3,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
4.	[(0,R22), (2,C), (2,D), (3,A), (3,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
5.	[(0,E), (0,F), (2,C), (2,D), (3,A), (3,B)]	[$(\infty, null)$, $(\infty, null)$]	∞
6.	[(0,F), (2,C), (2,D), (3,A), (3,B)]	[(2,E), $(\infty, null)$]	∞
7.	[(2,C), (2,D), (3,A), (3,B)]	[(2,E), (2,F)]	2
8.	[(2,D), (3,A), (3,B)]	[(2,E), (2,F)]	2
9.	[(3,A), (3,B)]	[(2,E), (2,F)]	2

4: TODO