# STAT361 Laboratory for Advanced R for Data Science

## Lab 3

Sidi Wu

Feburary 2/3, 2023

Department of Statistics and Actuarial Science
Simon Fraser University

# Weekly Office Hours

- 10am-11am on Fridays
- (Hybrid) Room K10504 + Zoom
- Start this week

# Lab 3 Exercise

# Lecture 4 Exercise

In order to test your Lab 3 code, you need complete lecture 4 exercise, including:

- split_points()
- LOF()
- split.region()

In other words, you need the "tree" output from **recpart()** function in lecture 4 exercise.

# Lab 3 Exercise: Methods for Recursive Partitioning Tree

1. Write functions, including:
    - **print.region()**
    - **plot_regions.tree()**
    - **plot_regions.node()**
2. Test functions

# print.region()

**Objective**: to print important information from a region.

- What should be the input?
- What information do we have in a region?
- Print the information that you think most useful (be creative).

# plot_regions.tree()

**Objective**: to generate a plot for a tree

- Take a “tree” as its input
- Which function produces a tree output?
- Do the scatterplot of the covariates
- Call the **plot_regions.node()** function for each child

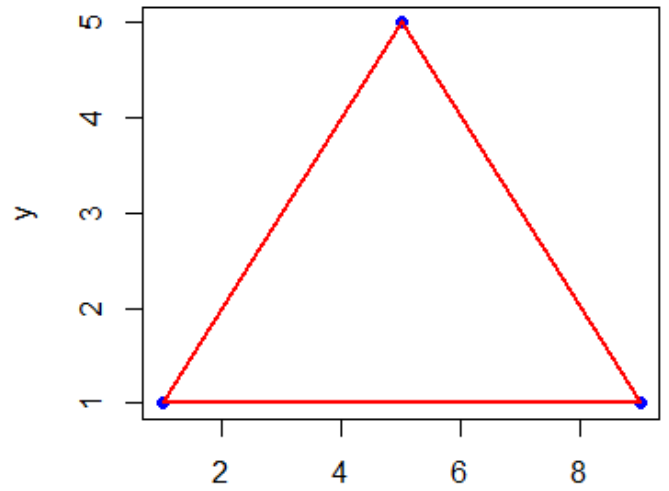
# plot_regions.node()

**Objective**: to draw boxes based on the coordinates matrix in a node and those in its child nodes.

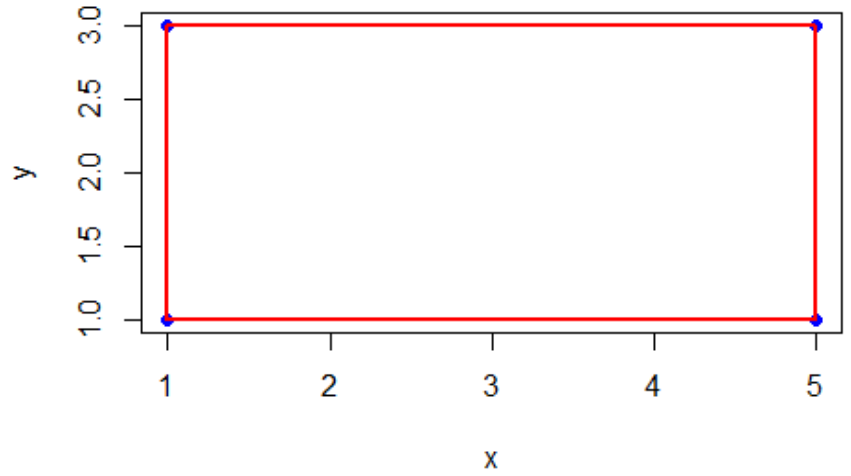
- Take a “node” input
- Which function produces a node output?
- Exit criteria: if node is NULL (refer **recpart_recursive()**)
- Use **lines()** or **rect()** to draw a box based on the coordinate matrix
- Call the **plot_regions.node()** function for each child within the same function recursively until you meet the exit criteria

# Example: Draw a Box using lines() function

```
x = c(1, 5, 9)
y = c(1, 5, 1)
plot(x, y, col='blue', pch=16)
lines(c(x[1], x[2], x[3], x[1]),
      c(y[1], y[2], y[3], y[1]),
      col='red', lwd=2)
```

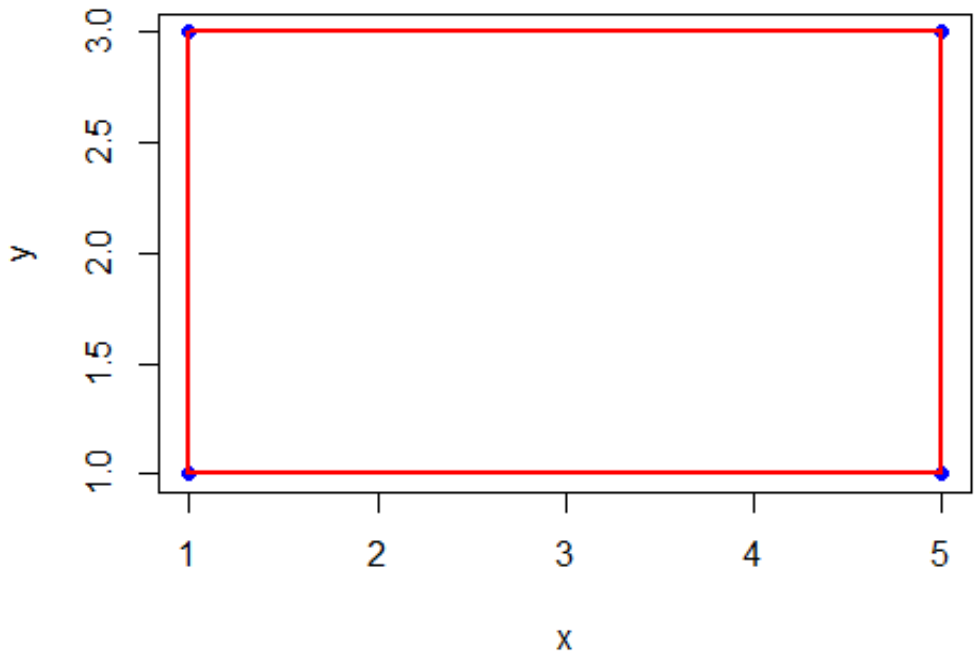



```
x = c(1, 5, 5, 1)
y = c(1, 1, 3, 3)
plot(x, y, col='blue', pch=16)
lines(c(x[1], x[2], x[3], x[4], x[1]),
      c(y[1], y[2], y[3], y[4], y[1]),
      col='red', lwd=2)
```



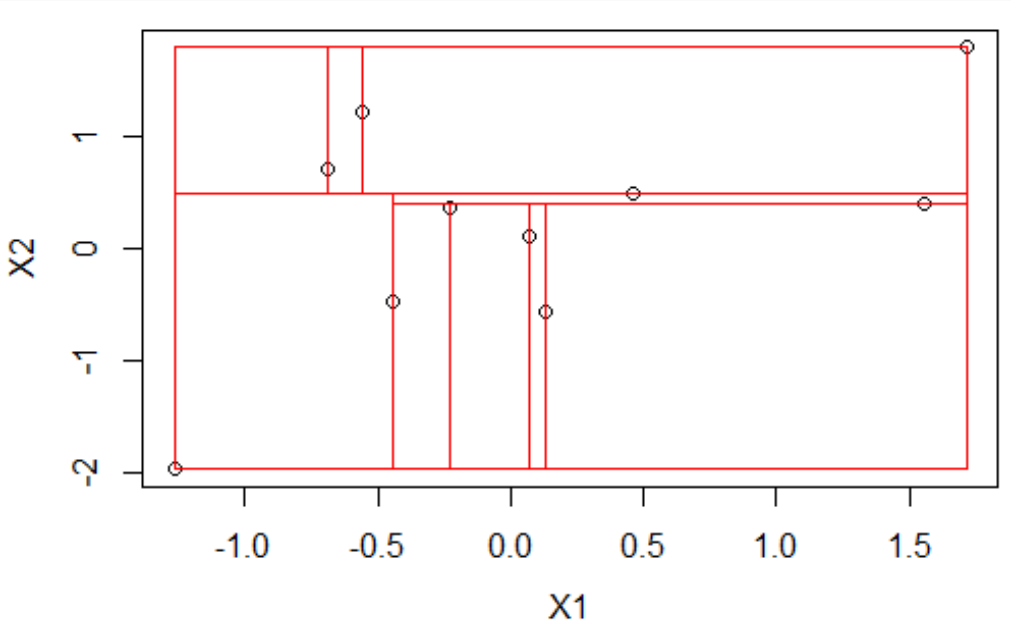
# Example: Draw a Box using rect() function

```
x = c(1, 5, 5, 1); x_range = sort(range(x))
y = c(1, 1, 3, 3); y_range = sort(range(y))
plot(x, y, col='blue', pch=16)
rect(x_range[1],y_range[1],x_range[2],y_range[2],border='red', lwd=2)
```



# Test Functions

```
# Test:
set.seed(123); n <- 10
x <- data.frame(x1=rnorm(n),x2=rnorm(n))
y <- rnorm(n)
mytree <- recpart(x,y)
plot_regions.tree(mytree)
```



# About Lab Quiz 2

- Feb. 9/10
- Canvas quiz during the lab session
- 25 minutes
- About 10 multiple-choice questions
- Based on materials we covered in previous labs