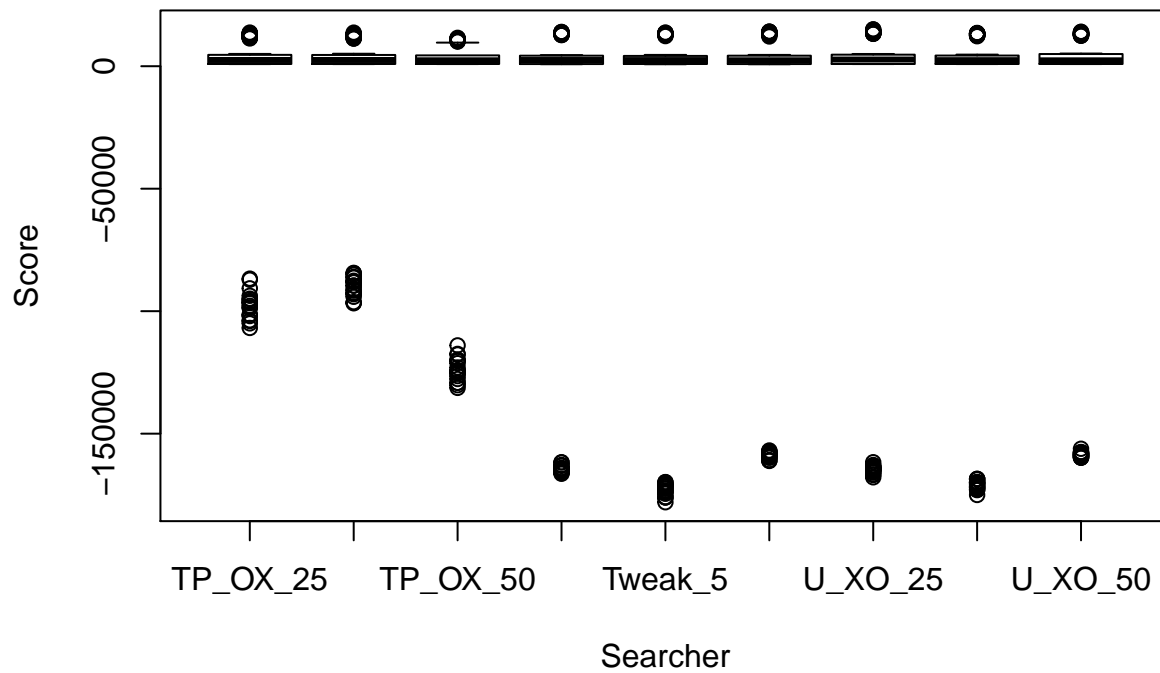


# Population\_Search\_Analysis

*Maggie Casale, Lenny Scott, & Zach Copic*

*March 4, 2016*

```
pop_data_20_runs <- read.csv("/home/casal033/ECAI/new-simple-search-the-second/population-result.txt",  
plot(pop_data_20_runs$Score ~ pop_data_20_runs$Searcher,  
      xlab="Searcher", ylab="Score")
```



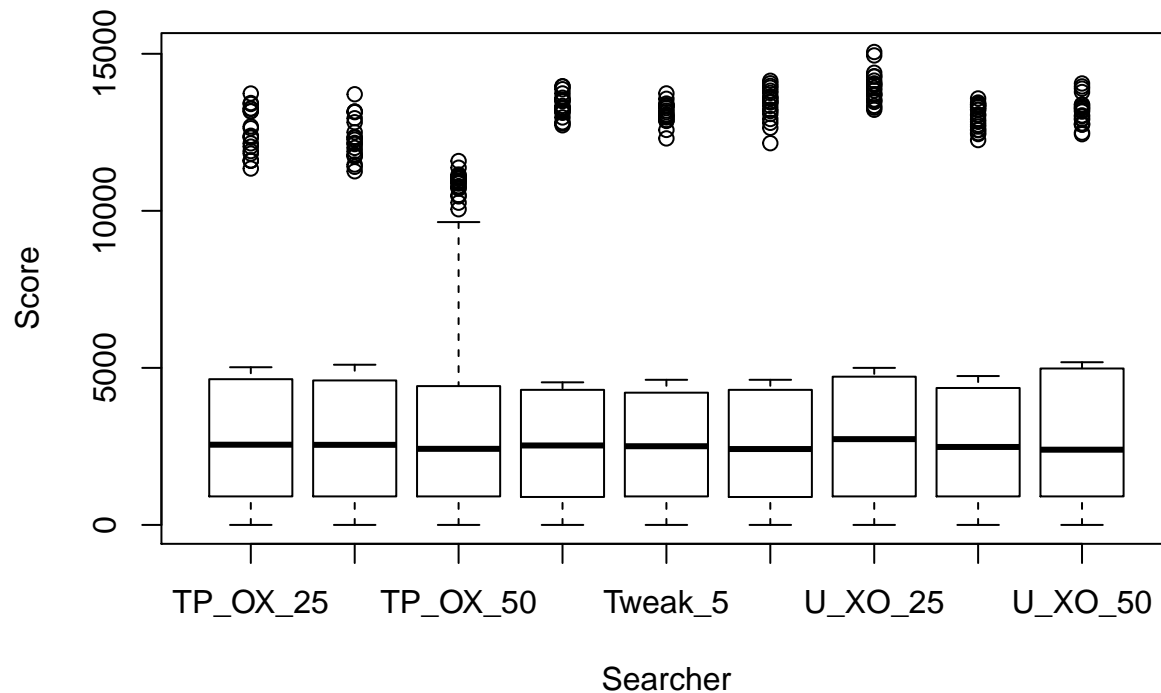
```
negs <- subset(pop_data_20_runs, Score<0)  
nrow(negs)
```

```
## [1] 180
```

```
unique(negs$Problem)
```

```
## [1] K_16_1000_3  
## 7 Levels: K_11_20 K_11_200 K_13_20 K_13_200 K_16_1000_3 ... K_16_200
```

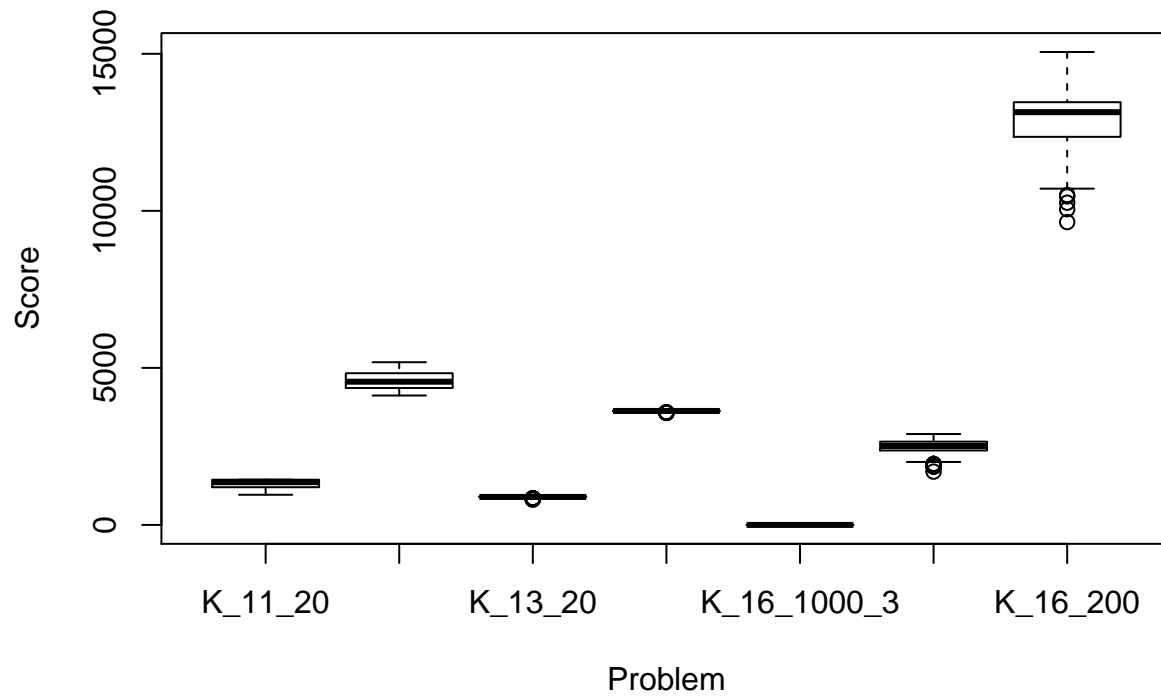
```
pop_data_20_runs$Non_negative_score = ifelse(pop_data_20_runs$Score<0, 0, pop_data_20_runs$Score)  
plot(pop_data_20_runs$Non_negative_score ~ pop_data_20_runs$Searcher,  
      xlab="Searcher", ylab="Score")
```



```
pairwise.wilcox.test(pop_data_20_runs$Non_negative_score, pop_data_20_runs$Search)
```

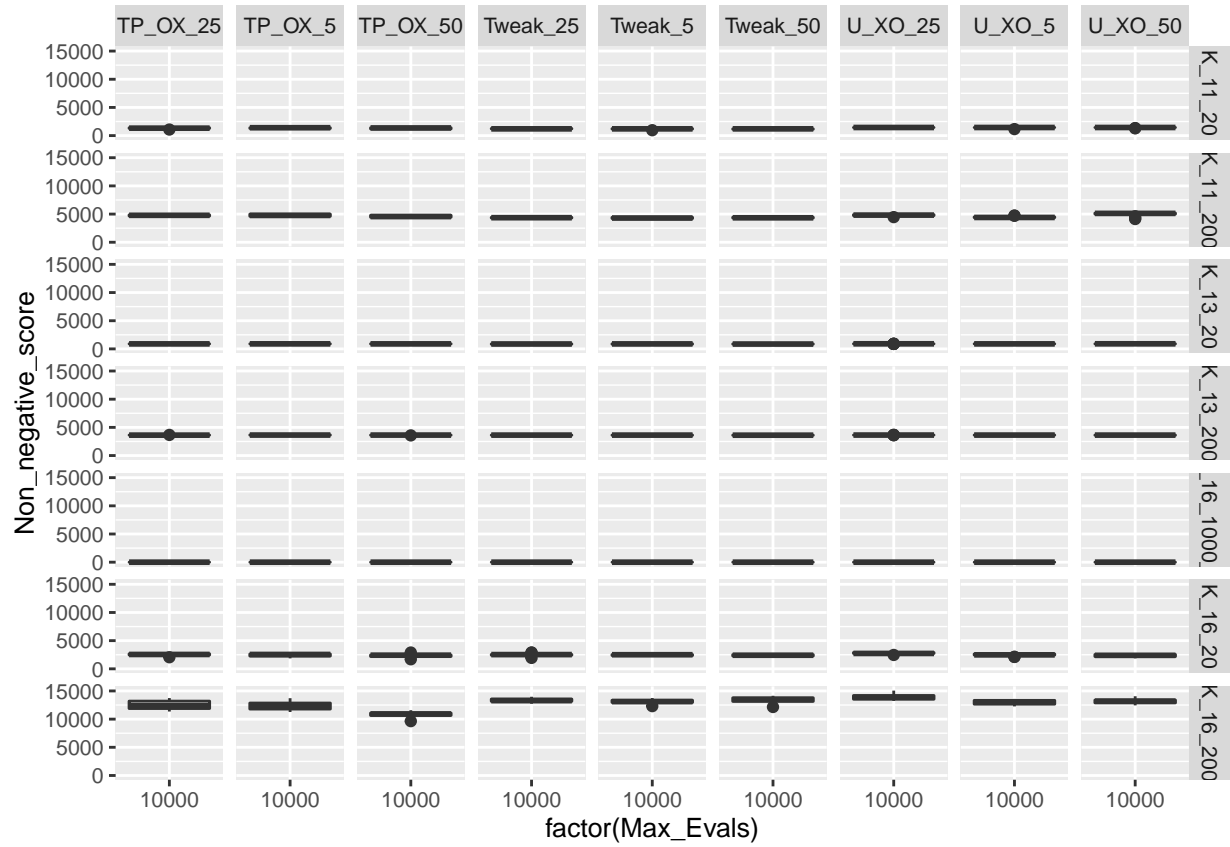
```
##
## Pairwise comparisons using Wilcoxon rank sum test
##
## data: pop_data_20_runs$Non_negative_score and pop_data_20_runs$Search
##
##      TP_OX_25 TP_OX_5 Tweak_25 Tweak_5 Tweak_50 U_XO_25
## TP_OX_5  1      -      -      -      -      -
## TP_OX_50 1      1      -      -      -      -
## Tweak_25 1      1      1      -      -      -
## Tweak_5  1      1      1      -      -      -
## Tweak_50 1      1      1      1      -      -
## U_XO_25  1      1      1      1      1      -
## U_XO_5   1      1      1      1      1      1
## U_XO_50  1      1      1      1      1      1
##
##      U_XO_5
## TP_OX_5   -
## TP_OX_50  -
## Tweak_25  -
## Tweak_5   -
## Tweak_50  -
## U_XO_25   -
## U_XO_5    -
## U_XO_50   1
##
## P value adjustment method: holm
```

```
plot(pop_data_20_runs$Non_negative_score ~ pop_data_20_runs$Problem,
      xlab="Problem", ylab="Score")
```

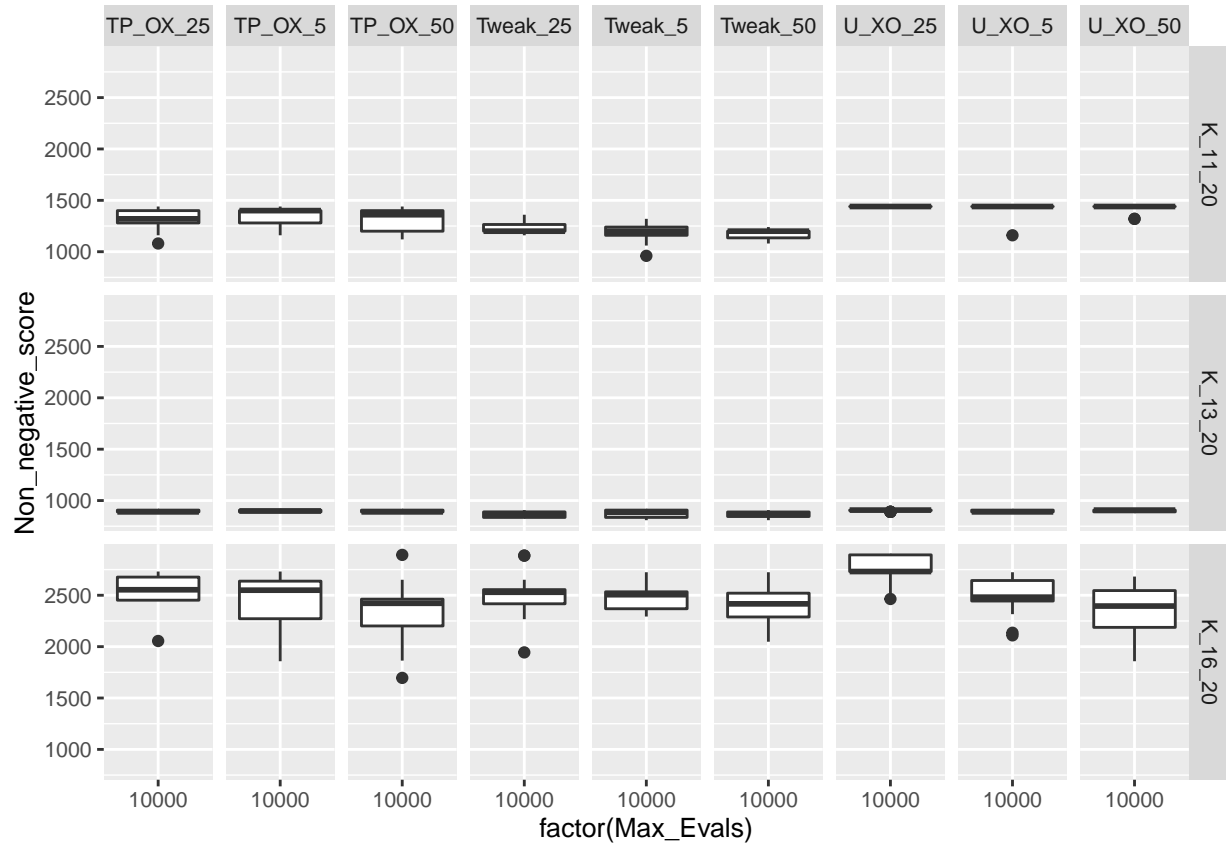


```
library("ggplot2")

ggplot(pop_data_20_runs,
  aes(x=factor(Max_Evals), y=Non_negative_score, group=Max_Evals)) +
  geom_boxplot() + facet_grid(Problem ~ Search) + theme_grey(base_size = 10)
```



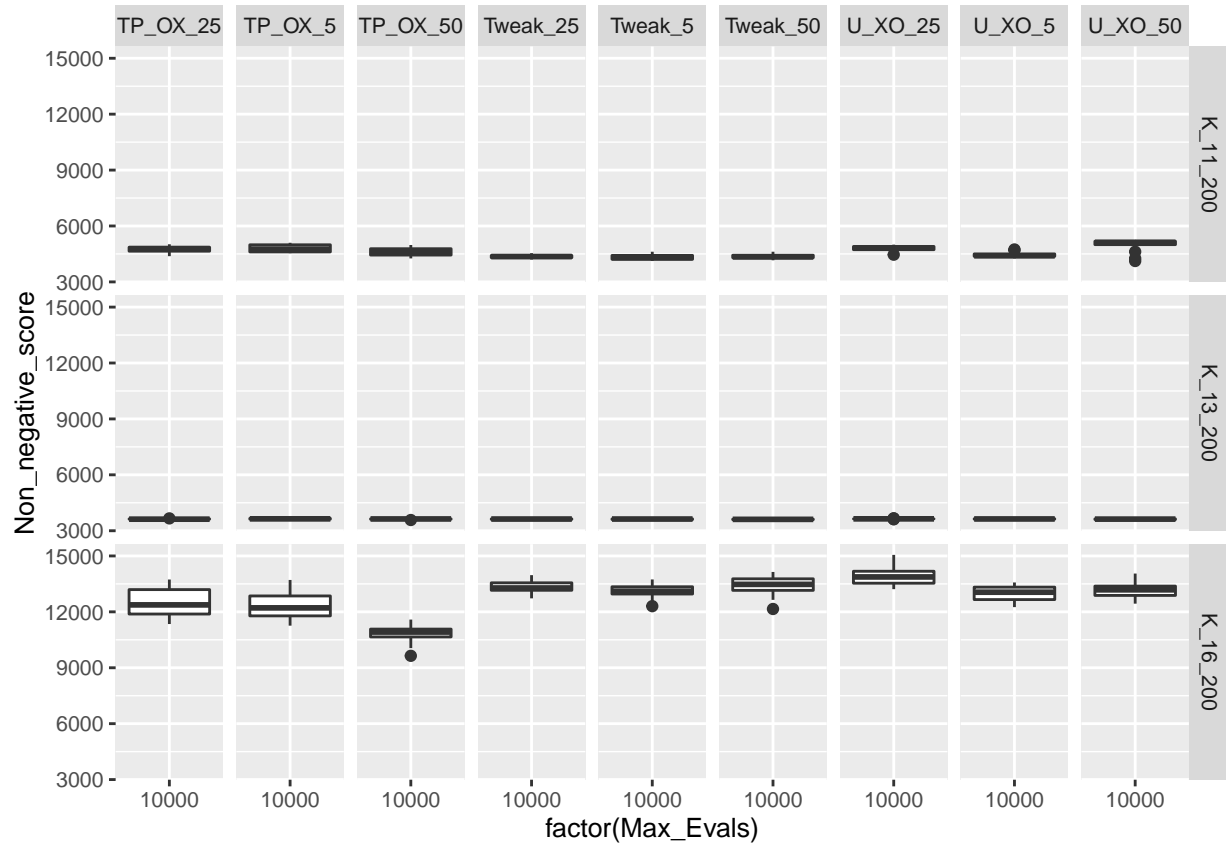
```
twenty_item_problems_11_16 = subset(pop_data_20_runs, Problem=="K_11_20" | Problem=="K_13_20" | Problem=="K_16_20")
ggplot(twenty_item_problems_11_16, aes(factor(Max_Evals), Non_negative_score)) + geom_boxplot() + facet_grid(. ~ .)
```



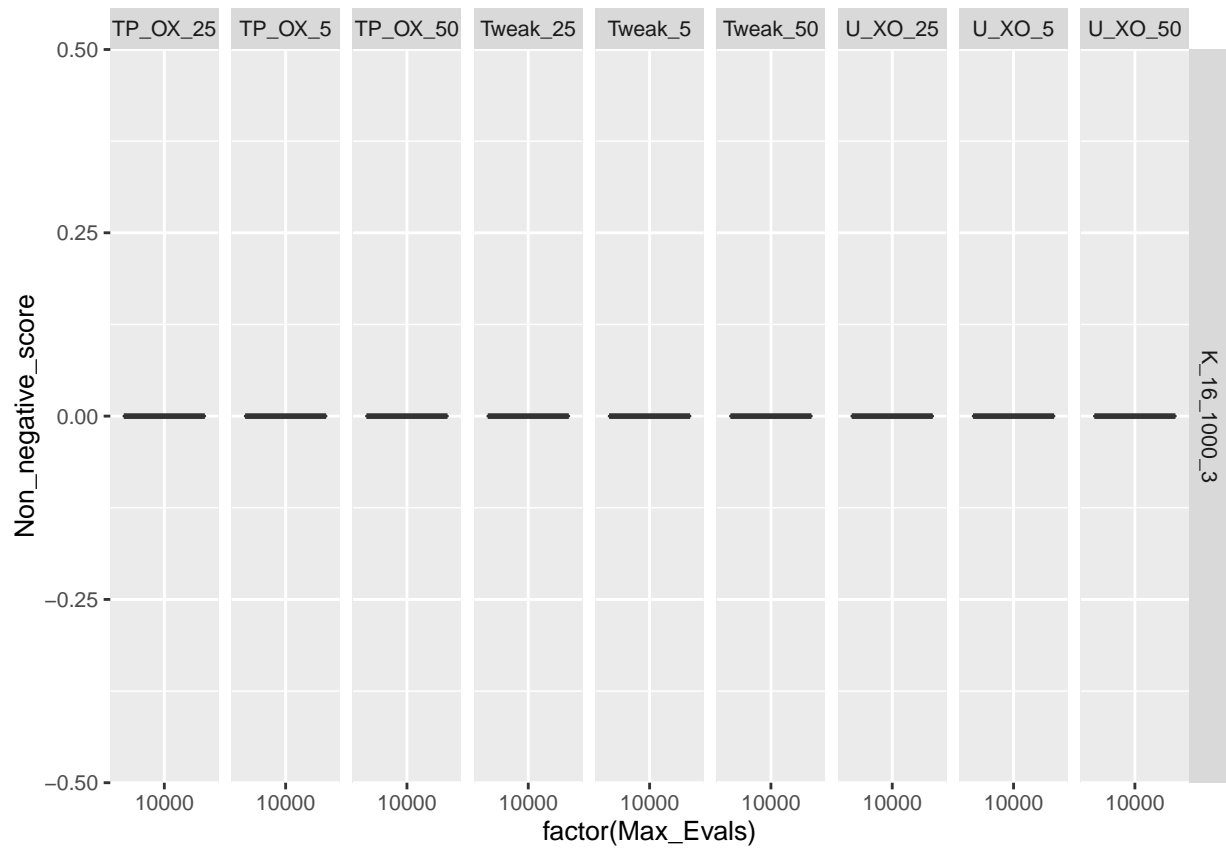
```

twohundred_item_problems_11_16 = subset(pop_data_20_runs, Problem=="K_11_200" | Problem=="K_13_200" | Problem=="K_16_200")
ggplot(twohundred_item_problems_11_16, aes(factor(Max_Evals), Non_negative_score)) + geom_boxplot() + facet_grid(Problem ~ .)

```



```
thousand_item_problem_16 = subset(pop_data_20_runs, Problem=="K_16_1000_3")
ggplot(thousand_item_problem_16, aes(factor(Max_Evals), Non_negative_score)) + geom_boxplot() + facet_g
```



```
# https://cran.r-project.org/web/packages/rpart.plot/rpart.plot.pdf
library("rpart")
library("rpart.plot")

rp <- rpart(Non_negative_score ~ Search + Problem + Max_Evals, data=pop_data_20_runs)
rp

## n= 1260
##
## node), split, n, deviance, yval
##      * denotes terminal node
##
## 1) root 1260 205706000000 3679.2240
##    2) Problem=K_11_20,K_11_200,K_13_20,K_13_200,K_16_1000_3,K_16_20 1080 2764299000 2152.0390
##      4) Problem=K_11_20,K_13_20,K_16_1000_3 540 165251300 735.3222 *
##      5) Problem=K_11_200,K_13_200,K_16_20 540 431394500 3568.7560
##      10) Problem=K_16_20 180 10107390 2478.5560 *
##      11) Problem=K_11_200,K_13_200 360 100382400 4113.8560 *
##    3) Problem=K_16_200 180 174165000 12842.3300 *
```

```
rpart.plot(rp, type=4, extra=100, Margin=0.0001)
```

