Practical No. 10

AIM - Practical of Decision Tree.

Source Code - *(CSV → shorturl.at/fBIS0)

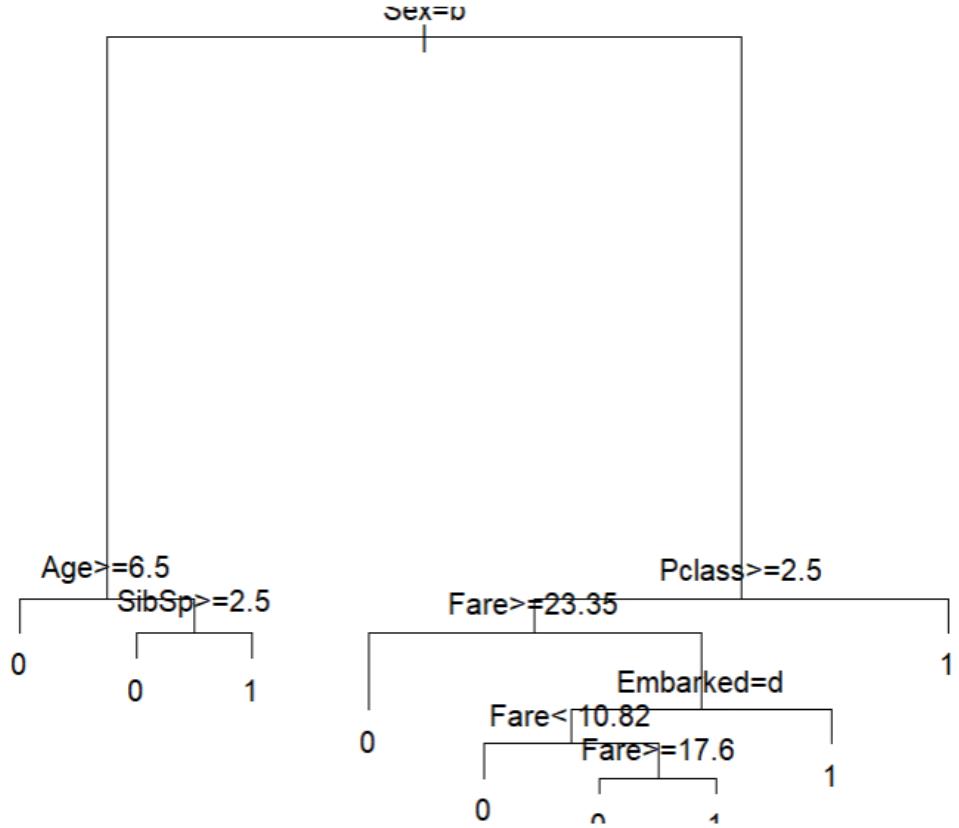
```
"PassengerId:type should be integers
Survived:Survived or Not
Pclass:Class of Travel
Name:Name of Passenger
Sex:Gender
Age:Age of Passengers
SibSp:Number of Sibling/Spouse aboard
Parch:Number of Parent/Child aboard
Ticket
Fare
Cabin
Embarked:The port in which a passenger has embarked. C - Cherbourg, S - Southampton,
Q = Queenstown"
titanic<-read.csv(file.choose(), header=TRUE, sep=",")
summary(titanic)
names(titanic)
install.packages("partykit")
install.packages("CHAID",repos="http://R-Forge.R-project.org", type="source")
library(CHAID)
library(partykit)
titanic$Survived<-as.factor(titanic$Survived)
summary(titanic$Survived)
names(titanic)
tree<-chaid(formula=Survived~Pclass+Sex+Age+SibSp+Parch+Fare+Embarked, data=titanic)
class(titanic$Survived)
library(rpart)
fit<-rpart(Survived~Pclass+Sex+Age+SibSp+Parch+Fare+Embarked, data=titanic,
method="class")
plot(fit)
text(fit)
install.packages('rattle')
install.packages('rpart.plot')
install.packages('RColorBrewer')
library(rattle)
library(rpart.plot)
```

```
library(RColorBrewer)
fancyRpartPlot(fit)
Prediction<-predict(fit, titanic, type="class")
Prediction

hitters<-read.csv(file.choose(), sep=",", header=TRUE)
summary(hitters)
reg.tree<-rpart(Salary~Years+Hits, data=hitters)
rpart.plot(reg.tree, type=4)
reg.tree$variable.importance
install.packages("MASS")
library(MASS)
set.seed(1984)
library(rpart)
train<-sample(1:nrow(hitters), nrow(hitters)/2)
tree_baseball<-rpart(Salary~Hits+HmRun+Runs+RBI+Walks+Years+Errors, subset=train,
data=hitters)
library(rpart.plot)
rpart.plot(tree_baseball)
tree_baseball$variable.importance
```

OUTPUT -

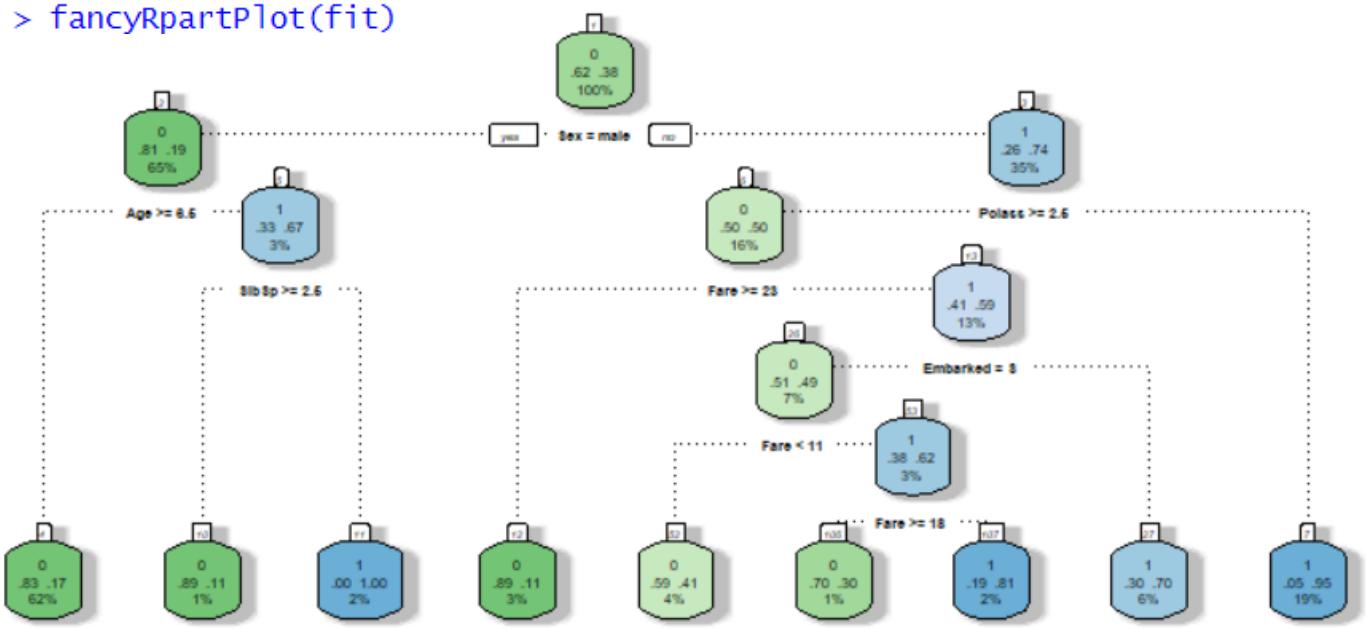
```
> "PassengerId:type should be integers
+ Survived:Survived or Not
+ Pclass:Class of Travel
+ Name:Name of Passenger
+ Sex:Gender
+ Age:Age of Passengers
+ SibSp:Number of Sibling/Spouse aboard
+ Parch:Number of Parent/Child aboard
+ Ticket
+ Fare
+ Cabin
+ Embarked:The port in which a passenger has embarked. C - Cherbourg, S - Southampton, Q = Queenstown"
[1] "PassengerId:type should be integers\nSurvived:Survived or Not\nPclass:Class of Travel\nName:Name
of Passenger\nSex:Gender\nAge:Age of Passengers\nSibSp:Number of Sibling/Spouse aboard\nParch:Number o
f Parent/Child aboard\nTicket\nFare\nCabin\nEmbarked:The port in which a passenger has embarked. C - C
herbourg, S - Southampton, Q = Queenstown"
> titanic<-read.csv(file.choose(), header=TRUE, sep=",")
> summary(titanic)
  PassengerId       Survived          Pclass          Name              Sex
 Min.   :  1.0   Min.   :0.0000   Min.   :1.000   Length:891         Length:891
 1st Qu.:223.5   1st Qu.:0.0000   1st Qu.:2.000   Class :character   Class :character
 Median :446.0   Median :0.0000   Median :3.000   Mode  :character   Mode  :character
 Mean   :446.0   Mean   :0.3838   Mean   :2.309
 3rd Qu.:668.5   3rd Qu.:1.0000   3rd Qu.:3.000
 Max.   :891.0   Max.   :1.0000   Max.   :3.000

      Age            SibSp           Parch           Ticket             Fare
 Min.   : 0.42   Min.   :0.000   Min.   :0.0000   Length:891         Min.   :  0.00
 1st Qu.:20.12   1st Qu.:0.000   1st Qu.:0.0000   Class :character   1st Qu.:  7.91
 Median :28.00   Median :0.000   Median :0.0000   Mode  :character   Median : 14.45
 Mean   :29.70   Mean   :0.523   Mean   :0.3816                      Mean   : 32.20
 3rd Qu.:38.00   3rd Qu.:1.000   3rd Qu.:0.0000                      3rd Qu.: 31.00
 Max.   :80.00   Max.   :8.000   Max.   :6.0000                      Max.   :512.33
 NA's   :177
```

```
    Cabin             Embarked
 Length:891         Length:891
 Class :character   Class :character
 Mode  :character   Mode  :character
> names(titanic)
 [1] "PassengerId" "Survived"    "Pclass"      "Name"        "Sex"         "Age"         "SibSp"
 [8] "Parch"       "Ticket"      "Fare"        "Cabin"       "Embarked"
> install.packages("partykit")
 package 'partykit' successfully unpacked and MD5 sums checked

 The downloaded binary packages are in
        C:\Users\Kunal\AppData\Local\Temp\RtmpmKwA7v\downloaded_packages
 > install.packages("CHAID",repos="http://R-Forge.R-project.org", type="source")

* installing *source* package 'CHAID' ...
** using staged installation
** R
** data
*** moving datasets to lazyload DB
** inst
** byte-compile and prepare package for lazy loading
** help
*** installing help indices
** building package indices
** testing if installed package can be loaded from temporary location
** testing if installed package can be loaded from final location
** testing if installed package keeps a record of temporary installation path
* DONE (CHAID)

The downloaded source packages are in
        'C:\Users\Kunal\AppData\Local\Temp\RtmpmKwA7v\downloaded_packages'
> library(CHAID)
Loading required package: partykit
Loading required package: grid
Loading required package: libcoin
Loading required package: mvtnorm
> library(partykit)
> titanic$Survived<-as.factor(titanic$Survived)
> summary(titanic$Survived)
  0   1
549 342
> names(titanic)
 [1] "PassengerId" "Survived"    "Pclass"      "Name"        "Sex"         "Age"         "SibSp"
 [8] "Parch"       "Ticket"      "Fare"        "Cabin"       "Embarked"
> tree<-chaid(formula=Survived~Pclass+Sex+Age+SibSp+Parch+Fare+Embarked, data=titanic)
Error in step1internal(response, xvars[[i]], weights, indices[[i]], ctrl) :
  is.factor(x) is not TRUE
> class(titanic$Survived)
[1] "factor"
> library(rpart)
> fit<-rpart(Survived~Pclass+Sex+Age+SibSp+Parch+Fare+Embarked, data=titanic, method="class")
> plot(fit)
> text(fit)
```




```
> install.packages('rattle')
package 'rattle' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Kunal\AppData\Local\Temp\RtmpmKwA7v\downloaded_packages
> install.packages('rpart.plot')
package 'rpart.plot' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Kunal\AppData\Local\Temp\RtmpmKwA7v\downloaded_packages
> install.packages('RColorBrewer')
package 'RColorBrewer' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\Kunal\AppData\Local\Temp\RtmpmKwA7v\downloaded_packages
> library(rattle)
Loading required package: tibble
Loading required package: bitops
Rattle: A free graphical interface for data science with R.
Version 5.5.1 Copyright (c) 2006-2021 Togaware Pty Ltd.
Type 'rattle()' to shake, rattle, and roll your data.
> library(rpart.plot)
> library(RColorBrewer)
```

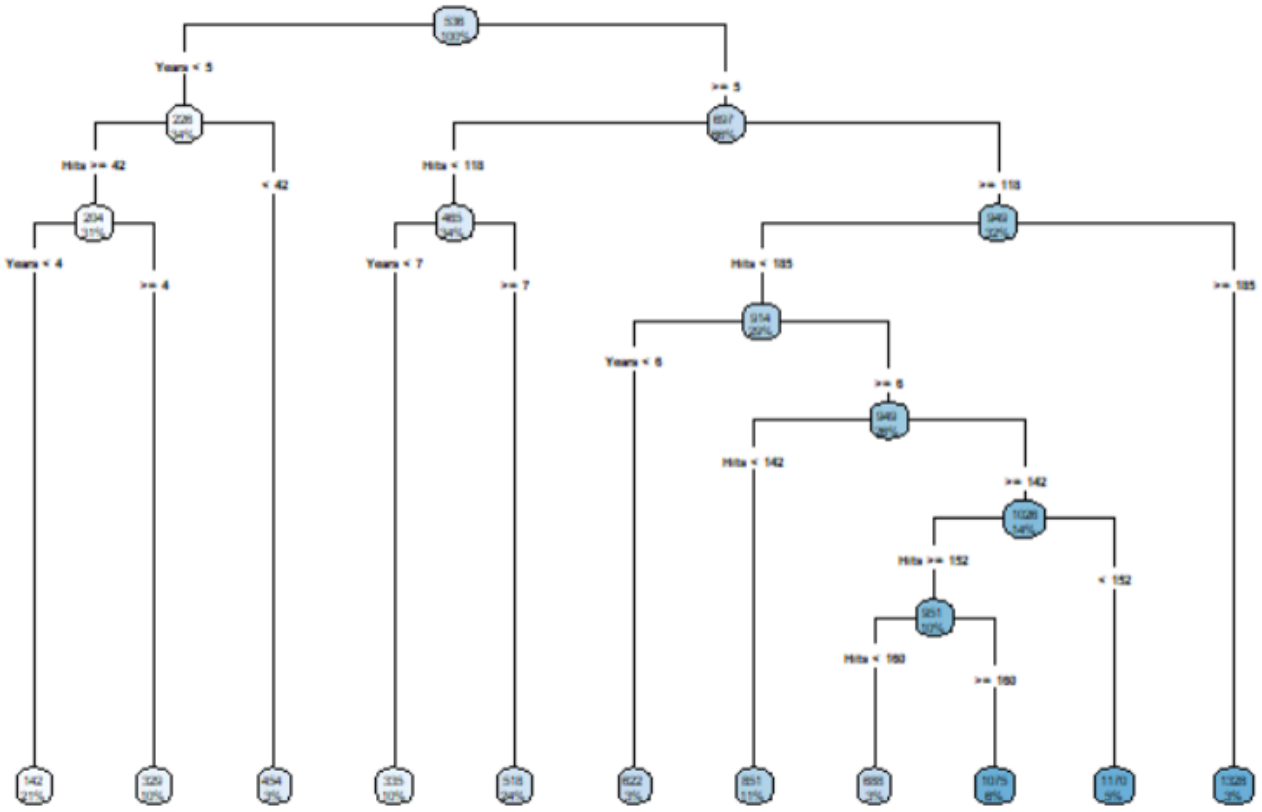
```
> fancyRpartPlot(fit)
```



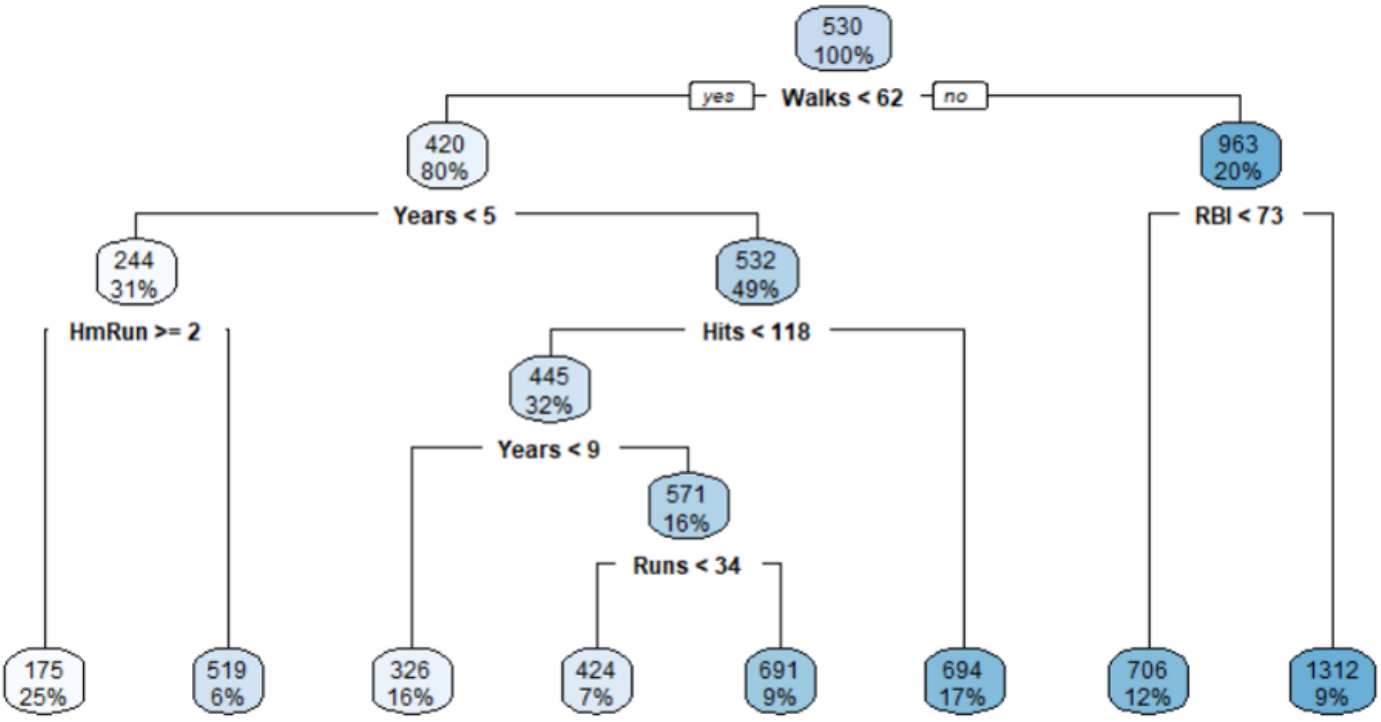

```
> Prediction<-predict(fit, titanic, type="class")
> Prediction
  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25
  0   1   0   1   0   0   0   0   1   1   1   1   0   0   0   1   0   0   0   1   0   0   1   0   0
 26  27  28  29  30  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50
  0   0   0   1   0   0   1   1   0   0   0   0   0   0   1   0   1   0   1   1   0   0   1   0   0
 51  52  53  54  55  56  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
  0   0   1   1   0   0   1   0   1   0   0   1   0   0   0   0   1   0   0   0   0   0   0   0   0
 76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93  94  95  96  97  98  99 100
  0   0   0   1   1   0   0   1   0   1   1   0   0   1   0   0   0   0   0   0   0   0   0   1   0
101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125
  0   0   0   0   0   0   0   0   0   0   0   1   0   0   1   0   0   0   0   0   0   0   0   1   0
126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
  0   0   0   1   0   0   0   1   1   0   0   1   0   0   0   1   0   1   0   0   0   0   0   0   0
151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175
  0   1   0   0   0   0   1   0   0   0   0   1   0   0   0   0   1   0   0   0   0   0   1   0   0
176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
  0   0   1   0   0   0   0   0   1   0   0   1   0   0   0   1   0   0   1   1   1   0   0   1   1
201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225
  0   0   0   0   0   0   0   0   1   0   0   1   0   0   0   1   0   0   1   0   0   0   0   0   0
226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250
  0   0   0   0   0   1   0   0   0   0   0   0   1   0   0   1   1   0   0   0   0   0   1   0   0
251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275
  0   0   0   0   0   1   1   1   1   1   0   0   0   0   1   0   0   0   1   1   0   0   1   0   1
276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
  1   0   0   0   0   0   0   0   0   0   0   0   0   0   1   1   1   0   0   0   0   0   1   0   1
301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325
  1   0   0   1   0   1   1   1   0   1   1   1   1   0   0   0   1   0   1   1   0   0   1   1   0
326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350
  1   0   1   0   1   1   0   0   0   1   0   0   1   0   0   1   1   0   0   0   1   1   1   1   0
351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375
```

```
  0   0   0   0   0   0   1   1   1   1   0   0   1   0   0   0   1   1   1   1   0   0   0   0   0
376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
  1   0   0   0   0   1   1   0   1   0   0   0   1   0   1   0   0   0   1   1   0   0   0   0   1
401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425
  0   0   0   0   0   0   0   1   0   0   0   0   1   0   0   0   1   1   0   0   0   0   0   1   0
426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450
  0   1   1   0   0   0   1   1   0   0   1   0   1   0   0   1   0   0   1   0   1   1   0   1   0
451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475
  0   0   0   0   0   0   0   1   1   0   0   0   0   0   0   0   0   0   0   1   0   0   1   1   0
476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
  0   0   0   0   1   0   0   0   0   0   0   1   0   0   0   0   0   0   0   0   0   1   0   1   0
501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525
  0   1   1   0   1   0   1   0   0   0   0   0   0   1   0   0   1   0   1   0   1   0   0   1   0
526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550
  0   1   0   0   0   1   0   0   1   0   1   0   1   0   1   1   0   0   0   0   0   1   0   0   0
551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
  0   0   0   0   0   0   1   0   1   1   0   0   0   0   0   0   0   0   0   0   1   0   1   0
576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
  0   1   1   1   0   1   1   0   0   0   1   0   0   0   0   0   1   0   1   0   0   1   0   0   0
601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625
  1   0   0   0   0   0   0   0   1   1   0   0   1   0   0   1   0   1   1   0   0   0   0   0   0
626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650
  0   0   1   0   0   0   0   0   0   0   1   0   0   0   0   0   1   0   0   1   0   0   0   0   0
651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675
  0   1   0   1   1   0   0   1   0   0   0   0   0   0   0   0   0   0   0   1   1   0   0   0   0
676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700
  0   0   0   0   0   1   0   0   0   0   0   0   0   0   1   0   1   0   0   0   0   0   1   0   0
701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725
  1   0   1   0   0   0   1   0   1   0   1   0   0   0   0   0   1   1   0   0   1   0   0   0   0
726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750
  0   1   1   0   0   1   0   0   0   0   0   0   0   0   0   0   0   1   0   0   0   0   1   0   0
751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775
  1   1   0   0   1   1   0   0   0   1   0   0   0   1   0   1   0   1   0   0   0   0   1   0   1
776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800
  0   0   1   0   1   1   1   0   0   0   0   0   0   1   0   0   0   0   0   0   0   0   1   0   0   0
801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825
  0   1   0   1   0   0   0   0   0   1   0   0   0   0   0   0   0   0   0   0   1   0   0   1   0
826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850
  0   0   1   0   1   1   1   0   0   0   1   0   0   0   0   0   0   1   0   0   0   0   0   0   1
851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875
  0   0   1   1   1   0   1   0   1   0   0   0   1   0   0   1   1   0   0   1   0   1   0   0   1
876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891
  1   0   0   0   1   1   0   0   0   0   0   0   1   0   0   0
Levels: 0 1
> hitters<-read.csv(file.choose(), sep=",", header=TRUE)
> summary(hitters)
       X                  AtBat            Hits         HmRun            Runs             RBI        
 Length:322         Min.   : 16.0   Min.   :  1   Min.   : 0.00   Min.   :  0.00   Min.   :  0.00  
 Class :character   1st Qu.:255.2   1st Qu.: 64   1st Qu.: 4.00   1st Qu.: 30.25   1st Qu.: 28.00  
 Mode  :character   Median :379.5   Median : 96   Median : 8.00   Median : 48.00   Median : 44.00  
                    Mean   :380.9   Mean   :101   Mean   :10.77   Mean   : 50.91   Mean   : 48.03  
                    3rd Qu.:512.0   3rd Qu.:137   3rd Qu.:16.00   3rd Qu.: 69.00   3rd Qu.: 64.75  
                    Max.   :687.0   Max.   :238   Max.   :40.00   Max.   :130.00   Max.   :121.00  
```

```
     Walks             Years            CAtBat            CHits            CHmRun      
 Min.   :  0.00   Min.   : 1.000   Min.   :   19.0   Min.   :   4.0   Min.   :  0.00  
 1st Qu.: 22.00   1st Qu.: 4.000   1st Qu.:  816.8   1st Qu.: 209.0   1st Qu.: 14.00  
 Median : 35.00   Median : 6.000   Median : 1928.0   Median : 508.0   Median : 37.50  
 Mean   : 38.74   Mean   : 7.444   Mean   : 2648.7   Mean   : 717.6   Mean   : 69.49  
 3rd Qu.: 53.00   3rd Qu.:11.000   3rd Qu.: 3924.2   3rd Qu.:1059.2   3rd Qu.: 90.00  
 Max.   :105.00   Max.   :24.000   Max.   :14053.0   Max.   :4256.0   Max.   :548.00  

     CRuns             CRBI            CWalks        League            Division        
 Min.   :   1.0   Min.   :   0.00   Min.   :   0.00   Length:322         Length:322        
 1st Qu.: 100.2   1st Qu.:  88.75   1st Qu.:  67.25   Class :character   Class :character  
 Median : 247.0   Median : 220.50   Median : 170.50   Mode  :character   Mode  :character  
 Mean   : 358.8   Mean   : 330.12   Mean   : 260.24                                        
 3rd Qu.: 526.2   3rd Qu.: 426.25   3rd Qu.: 339.25                                        
 Max.   :2165.0   Max.   :1659.00   Max.   :1566.00                                        

    PutOuts          Assists          Errors           Salary       NewLeague        
 Min.   :   0.0   Min.   :  0.0   Min.   : 0.00   Min.   :  67.5   Length:322        
 1st Qu.: 109.2   1st Qu.:  7.0   1st Qu.: 3.00   1st Qu.: 190.0   Class :character  
 Median : 212.0   Median : 39.5   Median : 6.00   Median : 425.0   Mode  :character  
 Mean   : 288.9   Mean   :106.9   Mean   : 8.04   Mean   : 535.9                     
 3rd Qu.: 325.0   3rd Qu.:166.0   3rd Qu.:11.00   3rd Qu.: 750.0                     
 Max.   :1378.0   Max.   :492.0   Max.   :32.00   Max.   :2460.0                     
                                                  NA's   :59                        
> reg.tree<-rpart(Salary~Years+Hits, data=hitters)
> rpart.plot(reg.tree, type=4)
```



```
> reg.tree$variable.importance
   Years     Hits
15696202 14259635
> install.packages("MASS")
The downloaded binary packages are in
        C:\Users\Kunal\AppData\Local\Temp\RtmpcJYP4W\downloaded_packages
> library(MASS)
> set.seed(1984)
> library(rpart)
> train<-sample(1:nrow(hitters), nrow(hitters)/2)
> tree_baseball<-rpart(Salary~Hits+HmRun+Runs+RBI+Walks+Years+Errors, subset=train,
data=hitters)
> library(rpart.plot)
> rpart.plot(tree_baseball)
```




```
> tree_baseball$variable.importance
    Walks       RBI      Runs     HmRun     Years      Hits    Errors
7572989.1 4580168.5 4351984.7 4327446.6 3409560.6 2243492.5  538739.9
```