Using Data Analytics to Retain Human Resources

Import the data

```
> View(HR_comma_sep)
> attach(HR_comma_sep)
> hrdata<-HR_comma_sep
> View(hrdata)
```

```
Analyze and check data quality
> summary(hrdata)
```

```
 satisfaction_level last_evaluation  number_project  average_montly_hours time_spend_comp
 Min.   :0.0900     Min.   :0.3600   Min.   :2.000   Min.   : 96.0        Min.   : 2.000
 1st Qu.:0.4400     1st Qu.:0.5600   1st Qu.:3.000   1st Qu.:156.0        1st Qu.: 3.000
 Median :0.6400     Median :0.7200   Median :4.000   Median :200.0        Median : 3.000
 Mean   :0.6128     Mean   :0.7161   Mean   :3.803   Mean   :201.1        Mean   : 3.498
 3rd Qu.:0.8200     3rd Qu.:0.8700   3rd Qu.:5.000   3rd Qu.:245.0        3rd Qu.: 4.000
 Max.   :1.0000     Max.   :1.0000   Max.   :7.000   Max.   :310.0        Max.   :10.000
 Work_accident          left         promotion_last_5years     Dept
 Min.   :0.0000   Min.   :0.0000   Min.   :0.00000       Length:14999
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000       Class :character
 Median :0.0000   Median :0.0000   Median :0.00000       Mode  :character
 Mean   :0.1446   Mean   :0.2381   Mean   :0.02127
 3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.00000
 Max.   :1.0000   Max.   :1.0000   Max.   :1.00000
    salary
 Length:14999
 Class :character
 Mode  :character
```

```
Splitting the data into training and testing dataset

> library(caTools)

> split<-sample.split(hrdata,SplitRatio = 0.8)
> split
 [1]  TRUE FALSE  TRUE  TRUE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE

> training<- subset(hrdata,split=="TRUE")
> testing<- subset(hrdata,split=="FALSE")

View training data
> View(training)
View testing Data
> View(testing)
Create a logistic regression model on training dataset
> model<- glm(left~.,training, family="binomial")
> summary(model)

Call:
glm(formula = left ~ ., family = "binomial", data = training)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2397  -0.6633  -0.4032  -0.1148   3.1293

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)
```

```
(Intercept)            -1.4412900  0.2147004  -6.713 1.91e-11 ***
satisfaction_level     -4.1361302  0.1098570 -37.650  < 2e-16 ***
last_evaluation         0.6674703  0.1670034   3.997 6.42e-05 ***
number_project         -0.3094176  0.0236712 -13.072  < 2e-16 ***
average_montly_hours    0.0043747  0.0005772   7.579 3.49e-14 ***
time_spend_company      0.2707020  0.0173423  15.609  < 2e-16 ***
Work_accident          -1.5121485  0.0998091 -15.150  < 2e-16 ***
promotion_last_5years -1.6329003  0.3184333  -5.128 2.93e-07 ***
Depthr                  0.1990695  0.1445491   1.377 0.168458    
DeptIT                 -0.2513547  0.1338220  -1.878 0.060343 .  
Deptmanagement         -0.4469871  0.1767391  -2.529 0.011436 *  
Deptmarketing          -0.0614583  0.1438805  -0.427 0.669271    
Deptproduct_mng        -0.1565207  0.1425095  -1.098 0.272066    
DeptRandD              -0.5733537  0.1590852  -3.604 0.000313 ***
Deptsales              -0.0689375  0.1112058  -0.620 0.535318    
Deptsupport             0.0111727  0.1190611   0.094 0.925236    
Depttechnical           0.0366984  0.1159269   0.317 0.751574    
salarylow               1.9760651  0.1438844  13.734  < 2e-16 ***
salarymedium            1.4529393  0.1447543  10.037  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 13170  on 11998  degrees of freedom
Residual deviance: 10298  on 11980  degrees of freedom
AIC: 10336

Number of Fisher Scoring iterations: 5
```

```
# removed variable Department in model one but this AIC increased and also
residual deviance increased
therefore we continue with the previous model having all variables
> model1<- glm(left~.-Dept,training, family="binomial")
> summary(model1)

Call:
glm(formula = left ~ . - Dept, family = "binomial", data = training)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-2.2206  -0.6652  -0.4079  -0.1190   3.1503  

Coefficients:
                        Estimate Std. Error z value Pr(>|z|)    
(Intercept)            -1.5398795  0.1914579  -8.043 8.77e-16 ***
satisfaction_level     -4.1344256  0.1095259 -37.748  < 2e-16 ***
last_evaluation         0.6693286  0.1665202   4.020 5.83e-05 ***
number_project         -0.3090173  0.0235950 -13.097  < 2e-16 ***
average_montly_hours    0.0043464  0.0005751   7.557 4.12e-14 ***
time_spend_company      0.2641438  0.0171475  15.404  < 2e-16 ***
Work_accident          -1.5186137  0.0997111 -15.230  < 2e-16 ***
promotion_last_5years -1.6816057  0.3167320  -5.309 1.10e-07 ***
salarylow               2.0306533  0.1426226  14.238  < 2e-16 ***
salarymedium            1.5065106  0.1435910  10.492  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 13170  on 11998  degrees of freedom
Residual deviance: 10340  on 11989  degrees of freedom
AIC: 10360

Number of Fisher Scoring iterations: 5
```

Test the model on the testing dataset and print the response

```
> test<- predict(model,testing,type="response")
> test
          1           2           3           4           5           6 
7 
0.172375966 0.137929828 0.752518562 0.354923520 0.524371596 0.480583988 0.508
129355 
          8           9          10          11          12          13 
14 
0.447303656 0.478779972 0.633166706 0.469291225 0.781868064 0.547444291 0.158
182481 
         15          16          17          18          19          20 
21 
0.676029096 0.365579801 0.545040085 0.429725513 0.159057458 0.742639739 0.410
266543 
         22          23          24          25          26          27 
28 
0.346651371 0.364348110 0.731356941 0.501977150 0.370639720 0.540014408 0.360
914294 
         29          30          31          32          33          34 
35 
0.412719411 0.725935106 0.463269069 0.784341383 0.189246857 0.481995128 0.159
410314 
         36          37          38          39          40          41 
42 
0.379997629 0.628736878 0.421200115 0.035123659 0.776795768 0.512608101 0.271
649290 
         43          44          45          46          47          48 
49 
0.538102375 0.129345235 0.491368913 0.529833890 0.184531533 0.444972131 0.529
521793 
         50          51          52          53          54          55 
56 
0.433152545 0.269070605 0.363938579 0.171291407 0.103912576 0.617368154 0.301
152218 
         57          58          59          60          61          62 
63 
0.731766581 0.380979117 0.125937324 0.350716308 0.727136583 0.792944062 0.493
267792 
         64          65          66          67          68          69 
70 
0.526251010 0.183730690 0.215874050 0.693439162 0.386263848 0.411392808 0.642
747261 
         71          72          73          74          75          76 
77 
0.702846520 0.555053350 0.325104594 0.442531062 0.479679758 0.365865242 0.443
646255 
         78          79          80          81          82          83 
84 
0.526310804 0.453239360 0.802131409 0.294933772 0.744901037 0.219394224 0.639
987784 
         85          86          87          88          89          90 
91 
```

```
0.677867432 0.292339962 0.530913042 0.503597105 0.674287403 0.780717613 0.107
635730
         92          93          94          95          96          97
98
0.136146159 0.347651454 0.670017021 0.461276655 0.740817772 0.731939921 0.750
532510
         99         100         101         102         103         104
105
0.448157835 0.119467080 0.037570704 0.544323161 0.592248618 0.739217049 0.272
617390
        106         107         108         109         110         111
112
0.687377978 0.153811839 0.101264409 0.255546798 0.519815489 0.257715378 0.486
966270
        113         114         115         116         117         118
119
0.513891656 0.711875517 0.252504102 0.440509932 0.132966764 0.251123275 0.400
746312
        120         121         122         123         124         125
126
0.691258425 0.194603153 0.549854279 0.361752433 0.163171244 0.601080304 0.173
737206
        127         128         129         130         131         132
133
0.360618677 0.467728197 0.362118313 0.339238064 0.660731934 0.111098717 0.378
714791
        134         135         136         137         138         139
140
0.276926580 0.697850426 0.281400225 0.472945004 0.552946502 0.362697074 0.752
089967
        141         142         143         144         145         146
147
0.117924332 0.231057347 0.516827917 0.748163900 0.562015676 0.387813440 0.188
312792
        148         149         150         151         152         153
154
0.372449878 0.277551586 0.403060712 0.103886947 0.528733898 0.736901893 0.241
230052
        155         156         157         158         159         160
161
0.664005971 0.650432163 0.178505333 0.508874587 0.769160519 0.822393017 0.377
426758
        162         163         164         165         166         167
168
0.114885348 0.616520494 0.629427451 0.269989731 0.071267016 0.275165705 0.686
459122
        169         170         171         172         173         174
175
0.661396856 0.164209427 0.388417262 0.636496693 0.457828000 0.481161921 0.272
115944
        176         177         178         179         180         181
182
0.284991130 0.113837828 0.368191937 0.042493869 0.425214457 0.422197523 0.746
372395
        183         184         185         186         187         188
189
0.458454935 0.542131007 0.357346346 0.501379657 0.080676693 0.735419729 0.462
398765
        190         191         192         193         194         195
196
0.569150992 0.576798543 0.399107926 0.048933626 0.173654842 0.214020042 0.780
809385
        197         198         199         200         201         202
203
```

```
0.138870241 0.596452270 0.269663280 0.100370790 0.610502690 0.128000042 0.336
927423
        204         205         206         207         208         209
210
0.202561277 0.293798984 0.782752089 0.832326883 0.595635462 0.138037991 0.147
536627
        211         212         213         214         215         216
217
0.591508756 0.463176216 0.106620281 0.443557267 0.073605459 0.134906603 0.132
413879
        218         219         220         221         222         223
224
0.347659979 0.334020710 0.565122239 0.712941945 0.294005882 0.475945775 0.186
623069
        225         226         227         228         229         230
231
0.401588787 0.493176467 0.491664647 0.210496549 0.411251624 0.215458415 0.266
208939
        232         233         234         235         236         237
238
0.104384693 0.587391851 0.094030710 0.774631498 0.287957482 0.510241423 0.788
003988
        239         240         241         242         243         244
245
0.685990738 0.136596609 0.687148740 0.109815028 0.692892270 0.214322870 0.230
180577
        246         247         248         249         250         251
252
0.496753093 0.613447630 0.195794868 0.344942935 0.548672807 0.210396632 0.259
530677
        253         254         255         256         257         258
259
0.488241451 0.511747996 0.668520344 0.226796729 0.038427878 0.646925458 0.697
505132
        260         261         262         263         264         265
266
0.497755359 0.471738049 0.441155552 0.374877976 0.377918405 0.314993160 0.497
851267
        267         268         269         270         271         272
273
0.811715505 0.517066293 0.162920485 0.679981254 0.624046871 0.252982614 0.147
867799
        274         275         276         277         278         279
280
0.434148299 0.343019130 0.394947106 0.128386384 0.355682250 0.113641039 0.342
968650
        281         282         283         284         285         286
287
0.618844944 0.314575084 0.792041233 0.827933223 0.045572446 0.245904044 0.432
765412
        288         289         290         291         292         293
294
0.106977492 0.563967907 0.603594478 0.276017529 0.713322254 0.474246749 0.557
233440
        295         296         297         298         299         300
301
0.369440448 0.518114247 0.752380071 0.473436621 0.526791225 0.523375022 0.526
746535
        302         303         304         305         306         307
308
0.356581626 0.287607021 0.300525372 0.156960141 0.508761152 0.571052737 0.512
513346
        309         310         311         312         313         314
315
```

```
0.222350890 0.609314010 0.110256098 0.212331864 0.795954654 0.502033501 0.326
929918
        316         317         318         319         320         321
322
0.574893028 0.064713008 0.571693951 0.336357890 0.097439745 0.036373689 0.190
355960
        323         324         325         326         327         328
329
0.720873241 0.295074409 0.088963800 0.397333650 0.636633394 0.705826735 0.529
070436
        330         331         332         333         334         335
336
0.512716848 0.655347015 0.723740807 0.435080641 0.070951383 0.103799478 0.390
279108
        337         338         339         340         341         342
343
0.255597752 0.484005443 0.529901221 0.702207914 0.693189426 0.325400577 0.537
244205
        344         345         346         347         348         349
350
0.792443863 0.741744650 0.751823131 0.499754844 0.645787201 0.273563369 0.366
158275
        351         352         353         354         355         356
357
0.394776696 0.535058251 0.632462040 0.294631680 0.724262824 0.386926825 0.320
184350
        358         359         360         361         362         363
364
0.383881156 0.569263492 0.463959472 0.735948349 0.756002958 0.467021768 0.341
981393
        365         366         367         368         369         370
371
0.599924035 0.267511492 0.319134692 0.652673637 0.186039930 0.742183569 0.438
641787
        372         373         374         375         376         377
378
0.603855030 0.639647718 0.489109267 0.156385032 0.522832308 0.506588383 0.485
471230
        379         380         381         382         383         384
385
0.541191594 0.511473111 0.300685481 0.079736713 0.246271314 0.520220793 0.188
540200
        386         387         388         389         390         391
392
0.757591472 0.445452885 0.422766998 0.252067879 0.191811913 0.095616071 0.270
117431
        393         394         395         396         397         398
399
0.765478554 0.181840240 0.085179425 0.333501439 0.155943644 0.114735246 0.524
977188
        400         401         402         403         404         405
406
0.500969263 0.197557505 0.052607311 0.397705804 0.420395105 0.038620199 0.065
879483
        407         408         409         410         411         412
413
0.336064537 0.228663808 0.095333899 0.090342744 0.084618474 0.081665109 0.014
117636
        414         415         416         417         418         419
420
0.112561297 0.071714780 0.213575202 0.052361468 0.256945507 0.340535899 0.021
984698
        421         422         423         424         425         426
427
```

```
0.124145392 0.071316873 0.010783257 0.037684268 0.160674250 0.053794388 0.038
437101
        428         429         430         431         432         433
434
0.244542180 0.334034153 0.339536586 0.242748592 0.196938657 0.051034858 0.056
243289
        435         436         437         438         439         440
441
0.155925444 0.522440876 0.006926426 0.464428138 0.190949984 0.049648598 0.779
224425
        442         443         444         445         446         447
448
0.030024798 0.191840591 0.020842932 0.104008298 0.306448598 0.181555924 0.145
549403
        449         450         451         452         453         454
455
0.057609863 0.017922960 0.221961564 0.099904506 0.141747182 0.471022007 0.269
044933
        456         457         458         459         460         461
462
0.231030475 0.565385056 0.425624327 0.059761976 0.046731900 0.144500099 0.307
882119
        463         464         465         466         467         468
469
0.152718604 0.198646650 0.022937936 0.325906229 0.120194158 0.149868928 0.179
339109
        470         471         472         473         474         475
476
0.164725197 0.048464140 0.145081320 0.156040822 0.127358215 0.050385324 0.113
523819
        477         478         479         480         481         482
483
0.425478104 0.105738622 0.021227809 0.059294608 0.054646559 0.134547228 0.031
933495
        484         485         486         487         488         489
490
0.058128895 0.358354159 0.113576752 0.163141433 0.090852263 0.198443190 0.329
602696
        491         492         493         494         495         496
497
0.024815518 0.123605741 0.776045099 0.085743099 0.321019991 0.147517518 0.132
120338
        498         499         500         501         502         503
504
0.148606053 0.059615533 0.065610228 0.052322403 0.026597430 0.273797208 0.173
142762
        505         506         507         508         509         510
511
0.133202061 0.131400763 0.246975645 0.175773937 0.116387821 0.085519953 0.243
057872
        512         513         514         515         516         517
518
0.153793624 0.138569515 0.082749033 0.078999366 0.172134024 0.147845152 0.096
420906
        519         520         521         522         523         524
525
0.467535625 0.164769263 0.065181478 0.056213158 0.054711759 0.290917470 0.022
892979
        526         527         528         529         530         531
532
0.140717168 0.168043224 0.294708385 0.060549733 0.014806764 0.110830961 0.349
840866
        533         534         535         536         537         538
539
```

```
0.115261424 0.201603488 0.084588037 0.140717446 0.081756970 0.146990311 0.061
112354
        540         541         542         543         544         545
546
0.207253202 0.209478274 0.078552574 0.138891611 0.123679836 0.430213809 0.471
442046
        547         548         549         550         551         552
553
0.594934361 0.282164291 0.112372569 0.047148343 0.031817804 0.198974439 0.177
094304
        554         555         556         557         558         559
560
0.159688559 0.227176354 0.089396962 0.218465965 0.453058778 0.685967644 0.069
215295
        561         562         563         564         565         566
567
0.390863496 0.061009594 0.093622409 0.013627981 0.159213324 0.179522015 0.281
014536
        568         569         570         571         572         573
574
0.132154651 0.233858701 0.152253950 0.040825378 0.699467187 0.019721694 0.425
988425
        575         576         577         578         579         580
581
0.070312332 0.288736662 0.596470555 0.037143813 0.397449739 0.004074694 0.488
750112
        582         583         584         585         586         587
588
0.548595161 0.438057258 0.090525382 0.108473736 0.229948410 0.577037602 0.071
491067
        589         590         591         592         593         594
595
0.401058607 0.070957655 0.142061392 0.102454684 0.073746358 0.051378372 0.054
102368
        596         597         598         599         600         601
602
0.239258613 0.079194773 0.193402899 0.062999798 0.033489442 0.028178510 0.073
671056
        603         604         605         606         607         608
609
0.126946983 0.014472214 0.193498472 0.266594740 0.005947501 0.179180981 0.086
962002
        610         611         612         613         614         615
616
0.284913899 0.076010568 0.066456390 0.380513967 0.082325247 0.416531782 0.016
721356
        617         618         619         620         621         622
623
0.090865348 0.867464760 0.138150612 0.236590396 0.154540680 0.132583516 0.051
851123
        624         625         626         627         628         629
630
0.076974946 0.074204126 0.197181535 0.506583754 0.143555721 0.020097126 0.210
280050
        631         632         633         634         635         636
637
0.093802643 0.208747282 0.064594523 0.039572365 0.455288142 0.172394497 0.158
894982
        638         639         640         641         642         643
644
0.169434602 0.043755202 0.008259291 0.020708435 0.342146378 0.053020921 0.078
422952
        645         646         647         648         649         650
651
```

```
0.105777800 0.147501329 0.005867490 0.007965451 0.221535823 0.139569668 0.104
592031
        652         653         654         655         656         657
658
0.133437271 0.054555467 0.028241159 0.198049102 0.052125776 0.052821104 0.097
320846
        659         660         661         662         663         664
665
0.156066257 0.388066891 0.257731511 0.030298787 0.038023172 0.083267756 0.015
509658
        666         667         668         669         670         671
672
0.379675308 0.044893395 0.069640965 0.198157458 0.256701917 0.153611362 0.014
948615
        673         674         675         676         677         678
679
0.049101684 0.029446696 0.041055456 0.721721351 0.218546977 0.002913202 0.029
502827
        680         681         682         683         684         685
686
0.129328173 0.155897975 0.279395320 0.274500347 0.185119984 0.226349795 0.133
626035
        687         688         689         690         691         692
693
0.018001109 0.117243161 0.237135688 0.064404620 0.204449117 0.019093759 0.058
026694
        694         695         696         697         698         699
700
0.068058754 0.098887590 0.303744632 0.300315639 0.718127134 0.129170756 0.168
199382
        701         702         703         704         705         706
707
0.237553430 0.068244164 0.203919656 0.460490098 0.166307179 0.269580565 0.010
554265
        708         709         710         711         712         713
714
0.075218751 0.210074832 0.227959269 0.333593791 0.140735810 0.088708027 0.132
640137
        715         716         717         718         719         720
721
0.016586589 0.182191260 0.149514657 0.056707102 0.032397543 0.205348019 0.214
198530
        722         723         724         725         726         727
728
0.216379936 0.210912301 0.537233933 0.043220542 0.651606599 0.431215509 0.057
473530
        729         730         731         732         733         734
735
0.589283221 0.040204577 0.114170662 0.018359007 0.002621570 0.066195390 0.250
031237
        736         737         738         739         740         741
742
0.019754711 0.476996408 0.077315275 0.019615439 0.113290845 0.052956394 0.422
034445
        743         744         745         746         747         748
749
0.218329563 0.112360123 0.028576798 0.154192629 0.040275599 0.553213077 0.073
139399
        750         751         752         753         754         755
756
0.066891931 0.097767270 0.032981484 0.080020664 0.050449485 0.177881618 0.078
582533
        757         758         759         760         761         762
763
```

```
0.339746234 0.039452874 0.045020540 0.212682360 0.079303369 0.013740590 0.029
812182
        764         765         766         767         768         769
770
0.323472948 0.104341512 0.232124085 0.011029826 0.198215398 0.071517389 0.072
149958
        771         772         773         774         775         776
777
0.701233415 0.164862318 0.243352258 0.030159273 0.241865578 0.218257979 0.023
515091
        778         779         780         781         782         783
784
0.096979214 0.073250849 0.258322336 0.107965551 0.057238878 0.117264603 0.161
791600
        785         786         787         788         789         790
791
0.529144584 0.170432871 0.560426160 0.343316840 0.109171149 0.058973580 0.079
593417
        792         793         794         795         796         797
798
0.226784799 0.009463697 0.503883449 0.232869902 0.268820696 0.278377581 0.302
103659
        799         800         801         802         803         804
805
0.387085491 0.211244286 0.752055652 0.065809236 0.201495559 0.212075425 0.066
307017
        806         807         808         809         810         811
812
0.187639997 0.291866225 0.039646185 0.541448347 0.068516821 0.890996181 0.044
699452
        813         814         815         816         817         818
819
0.242342164 0.744102712 0.098810009 0.166971469 0.125107115 0.036875964 0.306
552951
        820         821         822         823         824         825
826
0.211870100 0.007079116 0.221614203 0.481209445 0.399520080 0.062400102 0.027
495649
        827         828         829         830         831         832
833
0.019683937 0.123925162 0.112008641 0.084489363 0.176599846 0.029067663 0.034
184116
        834         835         836         837         838         839
840
0.437073866 0.119048036 0.089010058 0.626913661 0.440341898 0.174252260 0.036
550281
        841         842         843         844         845         846
847
0.141670752 0.194744516 0.097946805 0.233435006 0.011986392 0.164113363 0.714
408178
        848         849         850         851         852         853
854
0.044976962 0.016900883 0.180015159 0.132890823 0.154189837 0.051917676 0.010
839990
        855         856         857         858         859         860
861
0.105410229 0.044231449 0.181258513 0.094438928 0.078077056 0.023779744 0.115
504904
        862         863         864         865         866         867
868
0.105138830 0.055579221 0.097118940 0.042011327 0.170152701 0.102163801 0.119
405027
        869         870         871         872         873         874
875
```

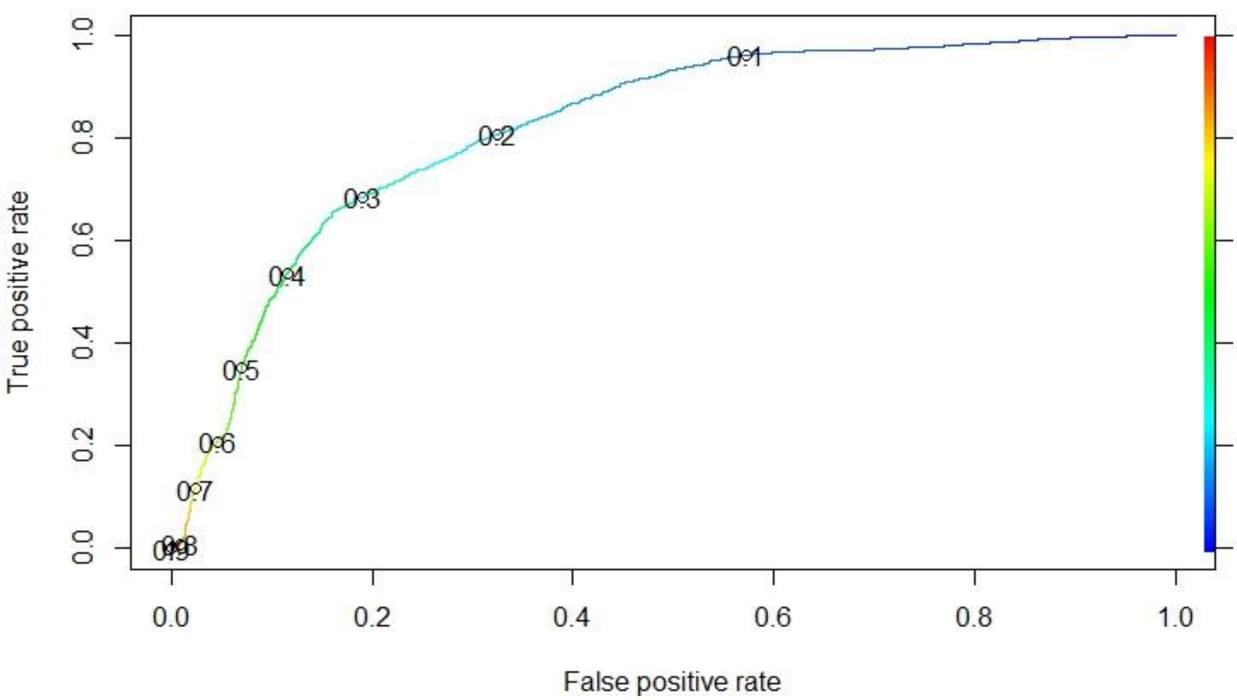
```
0.225711648 0.666629455 0.082225532 0.158683342 0.127163444 0.045306938 0.803
512984
        876         877         878         879         880         881
882
0.154703964 0.227651989 0.234305221 0.154496397 0.009498935 0.391269764 0.166
400762
        883         884         885         886         887         888
889
0.692971606 0.016385068 0.275414624 0.086258438 0.574478631 0.039948502 0.312
953522
        890         891         892         893         894         895
896
0.062412901 0.069419865 0.443700816 0.342672013 0.014746905 0.300354628 0.175
172410
        897         898         899         900         901         902
903
0.102375001 0.253998731 0.814374074 0.223719260 0.043726015 0.043148760 0.037
039376
        904         905         906         907         908         909
910
0.238953963 0.441449532 0.071135373 0.012169931 0.058782527 0.111972159 0.015
431422
        911         912         913         914         915         916
917
0.036362146 0.076602211 0.049358690 0.204838797 0.080975237 0.205974626 0.185
596401
        918         919         920         921         922         923
924
0.087030700 0.010176772 0.003253672 0.826632409 0.125272617 0.073796093 0.016
200046
        925         926         927         928         929         930
931
0.431882460 0.226322434 0.002753454 0.162354882 0.127540266 0.196784058 0.027
717472
        932         933         934         935         936         937
938
0.012793899 0.171695094 0.015845128 0.033500036 0.264639784 0.268108059 0.014
255005
        939         940         941         942         943         944
945
0.047113483 0.135991496 0.417476782 0.265072526 0.055894205 0.203485070 0.398
204574
        946         947         948         949         950         951
952
0.003984705 0.005323905 0.029378900 0.047481444 0.090853296 0.246243509 0.080
090753
        953         954         955         956         957         958
959
0.080633282 0.134659864 0.062541035 0.029564621 0.124304535 0.049355841 0.240
621349
        960         961         962         963         964         965
966
0.017828108 0.822847987 0.192490421 0.058188103 0.370399792 0.042232740 0.119
890086
        967         968         969         970         971         972
973
0.036607283 0.676623951 0.257493017 0.037699842 0.453571490 0.351253771 0.238
882632
        974         975         976         977         978         979
980
0.308209231 0.055438584 0.167156380 0.292975739 0.257167995 0.011373043 0.078
005143
        981         982         983         984         985         986
987
```

```
0.011169971 0.187375565 0.172286988 0.125668965 0.168455551 0.201699602 0.233
990101
        988         989         990         991         992         993
994
0.220010102 0.383127512 0.322209428 0.353186856 0.065709235 0.027587530 0.044
488113
        995         996         997         998         999        1000
0.100895382 0.210538186 0.083568677 0.327669991 0.039697865 0.169904372
 [ reached getOption("max.print") -- omitted 2000 entries ]
```

Confusion matrix For threshold 0.5

```
> table(ActualValue=testing$left,PredictedValue=test>0.5)
           PredictedValue
ActualValue FALSE TRUE
          0    2102  183
          1   434  281
Calculate accurace from confusion matrix
> accuracy=(2102+281)/(2102+281+183+434)
> accuracy
[1] 0.7943333

Plotting the the ROC curve
> test1<- predict(model,training,type="response")
> library(ROCR)
> res<- predict(model,training,type="response")
> library(ROCR)
> ROCRPred= prediction(res,training$left)
> ROCRPref<- performance(ROCRPred,"tpr","fpr")
> plot(ROCRPref,colorize=TRUE,print.cutoffs.at=seq(0.1,by=0.1))
```




```
Confusion matrix and accuracy for Threshold 0.4
> table(ActualValue=testing$left,PredictedValue=test>0.4)
           PredictedValue
ActualValue FALSE TRUE
```

```
           0  2014  271
           1   321  394
> accuracy1=(2014+394)/(2014+394+271+321)
> accuracy1
[1] 0.8026667
```

Confusion matrix and accuracy for Threshold 0.2

```
> table(ActualValue=testing$left,PredictedValue=test>0.2)
           PredictedValue
ActualValue FALSE TRUE
          0  1529  756
          1   135  580
> acc4=(1529+580)/(1529+580+135+756)
> acc4
[1] 0.703
```

Confusion matrix and accuracy for Threshold 0.35

```
> table(ActualValue=testing$left,PredictedValue=test>0.35)
           PredictedValue
ActualValue FALSE TRUE
          0  1945  340
          1   254  461
> accuracy2=(1945+461)/(1945+461+254+340)
> accuracy2
[1] 0.802
```

Confusion matrix and accuracy for Threshold 0.3

```
> table(ActualValue=testing$left,PredictedValue=test>0.3)
           PredictedValue
ActualValue FALSE TRUE
          0  1846  439
          1   219  496
> acc3=(1846+496)/(1846+496+219+439)
> acc3
[1] 0.7806667
```

Principal Component Analysis

```
> my_data <-HR_comma_sep
> new_hr_data<-my_data[c(-9,-10)]
> View(new_hr_data)
> library(caTools)

#Splitting data into train and test data
> split<-sample.split(new_hr_data,SplitRatio = 0.8)
> split
[1]  TRUE  TRUE FALSE  TRUE  TRUE FALSE  TRUE  TRUE

> pca.train<- subset(new_hr_data,split=="TRUE")


> View(pca.train)
> pca.test<- subset(new_hr_data,split=="FALSE")
> View(pca.test)
#Principal component analysis and normalization of the variables
> prin_comp <- prcomp(pca.train, scale. = T)
```

Tprcomp() function results in 5 useful measures:

```
> names(prin_comp)
[1] "sdev"     "rotation" "center"   "scale"    "x"
```

```
#output of mean of variables
> prin_comp$center
   satisfaction_level       last_evaluation        number_project  average_mo
ntly_hours
           0.61119388            0.71628234            3.81402791           20
1.43888346
   time_spend_company         Work_accident                  left promotion_l
ast_5years
           3.50644502            0.14650191            0.23806561
0.02115744

#output the standard deviation of variables
> prin_comp$scale
   satisfaction_level       last_evaluation        number_project  average_mo
ntly_hours
            0.2497568             0.1713073             1.2387096
49.8909677
   time_spend_company         Work_accident                  left promotion_l
ast_5years
            1.4610210             0.3536244             0.4259184
0.1439154


#rotation measure provides the principal component loading
> prin_comp$rotation
                              PC1         PC2          PC3          PC4
PC5
satisfaction_level     0.21981730 -0.59683577  0.144938610 -0.259385721  3.31
5146e-01
last_evaluation       -0.45024694 -0.33450961  0.152266384 -0.093595870  6.78
4173e-02
number_project        -0.55098945 -0.14716969  0.008783725  0.074795059 -1.91
9125e-01
average_montly_hours  -0.51928533 -0.20122380  0.114143857 -0.002466779 -2.45
3380e-01
time_spend_company    -0.33008364  0.09776795 -0.463184963 -0.039883094  7.85
9095e-01
Work_accident          0.07022193 -0.28400759 -0.411381956  0.819537875 -7.74
9069e-02
left                  -0.24726698  0.60639451  0.014723828  0.011978125 -2.73
1029e-05
promotion_last_5years  0.02615574 -0.10866123 -0.747468768 -0.494956573 -4.05
9753e-01
                              PC6         PC7         PC8
satisfaction_level    -0.23589511  0.30867249 -0.49633876
last_evaluation       -0.53068033 -0.55172267  0.25062671
number_project         0.47556202 -0.22350068 -0.59677304
average_montly_hours  -0.01382189  0.71156366  0.33183163
time_spend_company     0.16545099  0.08010518  0.11794331
Work_accident         -0.24153859  0.04724174 -0.08338854
left                  -0.57742389  0.18724000 -0.44978584
promotion_last_5years -0.13133290 -0.01845405 -0.03757960

#principal component score vectors
> dim(prin_comp$x)
[1] 11249     8

#Computing the standard deviation of each principal component
> std_dev <- prin_comp$sdev

#computing the variance
> pr_var <- std_dev^2

#varaince of components
```

```
> pr_var[1:8]
[1] 1.8809638 1.4406816 1.0647661 0.9575586 0.8439379 0.7098382 0.6266033 0.4
756506

#explaining the proportion of variance
> prop_varex <- pr_var/sum(pr_var)
> prop_varex[1:8]
[1] 0.23512048 0.18008519 0.13309576 0.11969482 0.10549223 0.08872978 0.07832
541 0.05945632

#scree plot used to access components with variability in data, shown in desc
ending value
> plot(prop_varex, xlab = "Principal Component",
+      ylab = "Proportion of Variance Explained",
+      type = "b")

#Cumulative variance plot for the principal components
> plot(cumsum(prop_varex), xlab = "Principal Component",
+      ylab = "Cumulative Proportion of Variance Explained",
+      type = "b")

#Adding a training set with principal components
> training<- subset(new_hr_data,split=="TRUE")
> train.data <- data.frame(left = training$left, prin_comp$x)
#Selecting the components
> train.data <- train.data[,1:7]

Run the logistic regressionmodel
> lrmodel<- glm(left~.,data = train.data, family="binomial")

> summary(lrmodel)

Call:
glm(formula = left ~ ., family = "binomial", data = train.data)

Deviance Residuals:
      Min         1Q     Median         3Q        Max
-0.005429   0.000000   0.000000   0.000000   0.004387

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -833.483   6102.722  -0.137    0.891
PC1         -357.504   2652.817  -0.135    0.893
PC2          795.578   5822.952   0.137    0.891
PC3            2.286    193.640   0.012    0.991
PC4          137.348   1020.115   0.135    0.893
PC5          -93.917    724.979  -0.130    0.897
PC6         -489.328   3581.104  -0.137    0.891

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1.2348e+04  on 11248  degrees of freedom
Residual deviance: 1.0074e-04  on 11242  degrees of freedom
AIC: 14

Number of Fisher Scoring iterations: 25

#transforming test to PCA
> test.data <- predict(prin_comp, newdata = pca.test)
> test.data <- as.data.frame(test.data)
> testpredict<- predict(lrmodel,test.data,type="response")
# Select the components
> test.data <- test.data[,1:7]
```

```
#making prediction on test data
> testpredict<- predict(lrmodel,test.data,type="response")
```