Regresión Logística

jerf

19/6/2021

```
dataset = read.csv("Social_Network_Ads.csv")
dataset = dataset[,3:5]
```

Dividir dataset en conjunto de entranmiento y conjunto de test

```
library(caTools)
```

```
## Warning: package 'caTools' was built under R version 4.0.5
```

```
set.seed(123)
split = sample.split(dataset$Purchased, SplitRatio = 0.75)
```

```
training_set = subset(dataset, split == TRUE)
training_set
```

```
##     Age EstimatedSalary Purchased
## 1    19           19000         0
## 3    26           43000         0
## 6    27           58000         0
## 7    27           84000         0
## 8    32          150000         1
## 10   35           65000         0
## 11   26           80000         0
## 13   20           86000         0
## 14   32           18000         0
## 15   18           82000         0
## 16   29           80000         0
## 17   47           25000         1
## 21   45           22000         1
## 23   48           41000         1
## 24   45           22000         1
## 25   46           23000         1
## 26   47           20000         1
## 27   49           28000         1
## 28   47           30000         1
## 30   31           18000         0
## 31   31           74000         0
## 33   21           16000         0
## 36   35           27000         0
## 37   33           28000         0
## 39   26           72000         0
## 40   27           31000         0
## 41   27           17000         0
```

```
## 42   33           51000          0
## 43   35          108000          0
## 44   30           15000          0
## 47   25           79000          0
## 49   30          135000          1
## 50   31           89000          0
## 51   24           32000          0
## 53   29           83000          0
## 54   35           23000          0
## 55   27           58000          0
## 56   24           55000          0
## 57   23           48000          0
## 58   28           79000          0
## 59   22           18000          0
## 60   32          117000          0
## 61   27           20000          0
## 62   25           87000          0
## 63   23           66000          0
## 64   32          120000          1
## 65   59           83000          0
## 67   24           19000          0
## 68   23           82000          0
## 70   31           68000          0
## 71   25           80000          0
## 72   24           27000          0
## 73   20           23000          0
## 76   34          112000          1
## 77   18           52000          0
## 78   22           27000          0
## 79   28           87000          0
## 80   26           17000          0
## 81   30           80000          0
## 83   20           49000          0
## 88   28           85000          0
## 90   35           50000          0
## 91   22           81000          0
## 92   30          116000          0
## 93   26           15000          0
## 94   29           28000          0
## 95   29           83000          0
## 96   35           44000          0
## 97   35           25000          0
## 98   28          123000          1
## 99   35           73000          0
## 100  28           37000          0
## 101  27           88000          0
## 102  28           59000          0
## 105  19           21000          0
## 106  21           72000          0
## 110  38           80000          0
## 111  39           71000          0
## 112  37           71000          0
## 113  38           61000          0
## 114  37           55000          0
```

```
## 115  42           80000         0
## 116  40           57000         0
## 118  36           52000         0
## 119  40           59000         0
## 120  41           59000         0
## 121  36           75000         0
## 122  37           72000         0
## 123  40           75000         0
## 125  41           51000         0
## 128  26           32000         0
## 129  30           17000         0
## 130  26           84000         0
## 132  33           31000         0
## 133  30           87000         0
## 135  28           55000         0
## 136  23           63000         0
## 137  20           82000         0
## 138  30          107000         1
## 140  19           25000         0
## 141  19           85000         0
## 142  18           68000         0
## 143  35           59000         0
## 144  30           89000         0
## 145  34           25000         0
## 146  24           89000         0
## 147  27           96000         1
## 149  29           61000         0
## 150  20           74000         0
## 151  26           15000         0
## 152  41           45000         0
## 153  31           76000         0
## 155  40           47000         0
## 157  46           59000         0
## 158  29           75000         0
## 160  32          135000         1
## 161  32          100000         1
## 164  35           38000         0
## 165  33           69000         0
## 166  18           86000         0
## 167  22           55000         0
## 168  35           71000         0
## 169  29          148000         1
## 171  21           88000         0
## 172  34          115000         0
## 173  26          118000         0
## 174  34           43000         0
## 177  35           47000         0
## 178  25           22000         0
## 179  24           23000         0
## 180  31           34000         0
## 181  26           16000         0
## 182  31           71000         0
## 183  32          117000         1
## 184  33           43000         0
```

```
## 185  33           60000         0
## 186  31           66000         0
## 187  20           82000         0
## 188  33           41000         0
## 189  35           72000         0
## 190  28           32000         0
## 191  24           84000         0
## 192  19           26000         0
## 194  19           70000         0
## 195  28           89000         0
## 196  34           43000         0
## 197  30           79000         0
## 198  20           36000         0
## 201  35           39000         0
## 202  49           74000         0
## 203  39          134000         1
## 204  41           71000         0
## 205  58          101000         1
## 206  47           47000         0
## 207  55          130000         1
## 209  40          142000         1
## 210  46           22000         0
## 211  48           96000         1
## 212  52          150000         1
## 214  35           58000         0
## 215  47           43000         0
## 216  60          108000         1
## 217  49           65000         0
## 218  40           78000         0
## 219  46           96000         0
## 220  59          143000         1
## 221  41           80000         0
## 222  35           91000         1
## 223  37          144000         1
## 225  35           60000         0
## 227  36          126000         1
## 231  35          147000         1
## 232  39           42000         0
## 233  40          107000         1
## 235  38          112000         0
## 238  37           80000         0
## 240  53          143000         1
## 242  38           59000         0
## 243  50           88000         1
## 244  56          104000         1
## 245  41           72000         0
## 246  51          146000         1
## 247  35           50000         0
## 248  57          122000         1
## 249  41           52000         0
## 250  35           97000         1
## 251  44           39000         0
## 252  37           52000         0
## 253  48          134000         1
```

```
## 254  37          146000         1
## 256  52           90000         1
## 257  41           72000         0
## 258  40           57000         0
## 259  58           95000         1
## 260  45          131000         1
## 261  35           77000         0
## 262  36          144000         1
## 263  55          125000         1
## 267  40           75000         0
## 268  37           74000         0
## 269  47          144000         1
## 270  40           61000         0
## 271  43          133000         0
## 272  59           76000         1
## 275  57           26000         1
## 276  57           74000         1
## 277  38           71000         0
## 278  49           88000         1
## 279  52           38000         1
## 280  50           36000         1
## 282  35           61000         0
## 283  37           70000         1
## 284  52           21000         1
## 285  48          141000         0
## 287  37           62000         0
## 288  48          138000         1
## 289  41           79000         0
## 290  37           78000         1
## 291  39          134000         1
## 293  55           39000         1
## 294  37           77000         0
## 295  35           57000         0
## 296  36           63000         0
## 297  42           73000         1
## 298  43          112000         1
## 300  46          117000         1
## 301  58           38000         1
## 303  37          137000         1
## 304  37           79000         1
## 306  42           54000         0
## 308  47          113000         1
## 309  36          125000         1
## 311  42           70000         0
## 312  39           96000         1
## 313  38           50000         0
## 314  49          141000         1
## 315  39           79000         0
## 317  54          104000         1
## 318  35           55000         0
## 319  45           32000         1
## 320  36           60000         0
## 321  52          138000         1
## 322  53           82000         1
```

```
## 323  41            52000          0
## 325  48           131000          1
## 327  41            72000          0
## 328  42            75000          0
## 329  36           118000          1
## 330  47           107000          1
## 331  38            51000          0
## 333  42            65000          0
## 334  40            65000          0
## 335  57            60000          1
## 336  36            54000          0
## 337  58           144000          1
## 338  35            79000          0
## 340  39           122000          1
## 342  35            75000          0
## 344  47            51000          1
## 345  47           105000          1
## 346  41            63000          0
## 348  54           108000          1
## 349  39            77000          0
## 350  38            61000          0
## 351  38           113000          1
## 352  37            75000          0
## 354  37            57000          0
## 355  36            99000          1
## 356  60            34000          1
## 357  54            70000          1
## 358  41            72000          0
## 359  40            71000          1
## 360  42            54000          0
## 361  43           129000          1
## 362  53            34000          1
## 365  42           104000          1
## 366  59            29000          1
## 370  54            26000          1
## 371  60            46000          1
## 374  59           130000          1
## 375  37            80000          0
## 376  46            32000          1
## 377  46            74000          0
## 378  42            53000          0
## 379  41            87000          1
## 381  42            64000          0
## 382  48            33000          1
## 384  49            28000          1
## 385  57            33000          1
## 386  56            60000          1
## 387  49            39000          1
## 388  39            71000          0
## 390  48            35000          1
## 391  48            33000          1
## 393  45            45000          1
## 394  60            42000          1
## 396  46            41000          1
```

```
## 397  51           23000         1
## 398  50           20000         1
## 399  36           33000         0
```

```
testing_set = subset(dataset, split == FALSE)
testing_set
```

```
##     Age EstimatedSalary Purchased
## 2    35           20000         0
## 4    27           57000         0
## 5    19           76000         0
## 9    25           33000         0
## 12   26           52000         0
## 18   45           26000         1
## 19   46           28000         1
## 20   48           29000         1
## 22   47           49000         1
## 29   29           43000         0
## 32   27          137000         1
## 34   28           44000         0
## 35   27           90000         0
## 38   30           49000         0
## 45   28           84000         0
## 46   23           20000         0
## 48   27           54000         0
## 52   18           44000         0
## 66   24           58000         0
## 69   22           63000         0
## 74   33          113000         0
## 75   32           18000         0
## 82   39           42000         0
## 84   35           88000         0
## 85   30           62000         0
## 86   31          118000         1
## 87   24           55000         0
## 89   26           81000         0
## 103  32           86000         0
## 104  33          149000         1
## 107  26           35000         0
## 108  27           89000         0
## 109  26           86000         0
## 117  35           75000         0
## 124  35           53000         0
## 126  39           61000         0
## 127  42           65000         0
## 131  31           58000         0
## 134  21           68000         0
## 139  28           59000         0
## 148  41           30000         0
## 154  36           50000         0
## 156  31           15000         0
## 159  26           30000         0
## 162  25           90000         0
## 163  37           33000         0
## 170  29           47000         0
```

```
## 175  34           72000         0
## 176  23           28000         0
## 193  29           43000         0
## 199  26           80000         0
## 200  35           22000         0
## 208  52          114000         0
## 213  59           42000         0
## 224  60          102000         1
## 226  37           53000         0
## 228  56          133000         1
## 229  40           72000         0
## 230  42           80000         1
## 234  49           86000         1
## 236  46           79000         1
## 237  40           57000         0
## 239  46           82000         0
## 241  42          149000         1
## 255  50           44000         0
## 264  35           72000         0
## 265  48           90000         1
## 266  42          108000         1
## 273  60           42000         1
## 274  39          106000         1
## 281  59           88000         1
## 286  37           93000         1
## 292  49           89000         1
## 299  45           79000         0
## 302  48           74000         1
## 305  40           60000         0
## 307  51          134000         0
## 310  38           50000         0
## 316  39           75000         1
## 324  48           30000         1
## 326  41           60000         0
## 332  48          119000         1
## 339  38           55000         0
## 341  53          104000         1
## 343  38           65000         0
## 347  53           72000         1
## 353  42           90000         1
## 363  47           50000         1
## 364  42           79000         0
## 367  58           47000         1
## 368  46           88000         1
## 369  38           71000         0
## 372  60           83000         1
## 373  39           73000         0
## 380  58           23000         1
## 383  44          139000         1
## 389  47           34000         1
## 392  47           23000         1
## 395  39           59000         0
## 400  49           36000         1
```

Escalado de datos

Standardisation

$$x_{stand} = \frac{x - mean(x)}{sd(x)}$$

```
training_set[,1:2] = scale(training_set[,1:2])
training_set
```

```
##             Age EstimatedSalary Purchased
## 1   -1.76554750     -1.47334137         0
## 3   -1.09629664     -0.78837605         0
## 6   -1.00068938     -0.36027273         0
## 7   -1.00068938      0.38177303         0
## 8   -0.52265305      2.26542765         1
## 10  -0.23583125     -0.16049118         0
## 11  -1.09629664      0.26761214         0
## 13  -1.66994024      0.43885347         0
## 14  -0.52265305     -1.50188159         0
## 15  -1.86115477      0.32469259         0
## 16  -0.80947485      0.26761214         0
## 17   0.91145593     -1.30210004         1
## 21   0.72024140     -1.38772071         1
## 23   1.00706320     -0.84545650         1
## 24   0.72024140     -1.38772071         1
## 25   0.81584866     -1.35918049         1
## 26   0.91145593     -1.44480115         1
## 27   1.10267046     -1.21647938         1
## 28   0.91145593     -1.15939893         1
## 30  -0.61826032     -1.50188159         0
## 31  -0.61826032      0.09637081         0
## 33  -1.57433297     -1.55896204         0
## 36  -0.23583125     -1.24501960         0
## 37  -0.42704579     -1.21647938         0
## 39  -1.09629664      0.03929037         0
## 40  -1.00068938     -1.13085871         0
## 41  -1.00068938     -1.53042182         0
## 42  -0.42704579     -0.56005428         0
## 43  -0.23583125      1.06673835         0
## 44  -0.71386758     -1.58750226         0
## 47  -1.19190391      0.23907192         0
## 49  -0.71386758      1.83732433         1
## 50  -0.61826032      0.52447414         0
## 51  -1.28751117     -1.10231849         0
## 53  -0.80947485      0.35323281         0
## 54  -0.23583125     -1.35918049         0
## 55  -1.00068938     -0.36027273         0
## 56  -1.28751117     -0.44589340         0
## 57  -1.38311844     -0.64567495         0
## 58  -0.90508211      0.23907192         0
## 59  -1.47872570     -1.50188159         0
## 60  -0.52265305      1.32360034         0
## 61  -1.00068938     -1.44480115         0
## 62  -1.19190391      0.46739370         0
## 63  -1.38311844     -0.13195096         0
```

```
## 64  -0.52265305      1.40922101         1
## 65   2.05874311      0.35323281         0
## 67  -1.28751117     -1.47334137         0
## 68  -1.38311844      0.32469259         0
## 70  -0.61826032     -0.07487051         0
## 71  -1.19190391      0.26761214         0
## 72  -1.28751117     -1.24501960         0
## 73  -1.66994024     -1.35918049         0
## 76  -0.33143852      1.18089923         1
## 77  -1.86115477     -0.53151406         0
## 78  -1.47872570     -1.24501960         0
## 79  -0.90508211      0.46739370         0
## 80  -1.09629664     -1.53042182         0
## 81  -0.71386758      0.26761214         0
## 83  -1.66994024     -0.61713472         0
## 88  -0.90508211      0.41031325         0
## 90  -0.23583125     -0.58859450         0
## 91  -1.47872570      0.29615237         0
## 92  -0.71386758      1.29506012         0
## 93  -1.09629664     -1.58750226         0
## 94  -0.80947485     -1.21647938         0
## 95  -0.80947485      0.35323281         0
## 96  -0.23583125     -0.75983583         0
## 97  -0.23583125     -1.30210004         0
## 98  -0.90508211      1.49484167         1
## 99  -0.23583125      0.06783059         0
## 100 -0.90508211     -0.95961738         0
## 101 -1.00068938      0.49593392         0
## 102 -0.90508211     -0.33173251         0
## 105 -1.76554750     -1.41626093         0
## 106 -1.57433297      0.03929037         0
## 110  0.05099054      0.26761214         0
## 111  0.14659781      0.01075015         0
## 112 -0.04461672      0.01075015         0
## 113  0.05099054     -0.27465207         0
## 114 -0.04461672     -0.44589340         0
## 115  0.43341960      0.26761214         0
## 116  0.24220507     -0.38881295         0
## 118 -0.14022399     -0.53151406         0
## 119  0.24220507     -0.33173251         0
## 120  0.33781234     -0.33173251         0
## 121 -0.14022399      0.12491104         0
## 122 -0.04461672      0.03929037         0
## 123  0.24220507      0.12491104         0
## 125  0.33781234     -0.56005428         0
## 128 -1.09629664     -1.10231849         0
## 129 -0.71386758     -1.53042182         0
## 130 -1.09629664      0.38177303         0
## 132 -0.42704579     -1.13085871         0
## 133 -0.71386758      0.46739370         0
## 135 -0.90508211     -0.44589340         0
## 136 -1.38311844     -0.21757162         0
## 137 -1.66994024      0.32469259         0
## 138 -0.71386758      1.03819813         1
```

```
## 140 -1.76554750    -1.30210004        0
## 141 -1.76554750     0.41031325        0
## 142 -1.86115477    -0.07487051        0
## 143 -0.23583125    -0.33173251        0
## 144 -0.71386758     0.52447414        0
## 145 -0.33143852    -1.30210004        0
## 146 -1.28751117     0.52447414        0
## 147 -1.00068938     0.72425569        1
## 149 -0.80947485    -0.27465207        0
## 150 -1.66994024     0.09637081        0
## 151 -1.09629664    -1.58750226        0
## 152  0.33781234    -0.73129561        0
## 153 -0.61826032     0.15345126        0
## 155  0.24220507    -0.67421517        0
## 157  0.81584866    -0.33173251        0
## 158 -0.80947485     0.12491104        0
## 160 -0.52265305     1.83732433        1
## 161 -0.52265305     0.83841658        1
## 164 -0.23583125    -0.93107716        0
## 165 -0.42704579    -0.04633029        0
## 166 -1.86115477     0.43885347        0
## 167 -1.47872570    -0.44589340        0
## 168 -0.23583125     0.01075015        0
## 169 -0.80947485     2.20834721        1
## 171 -1.57433297     0.49593392        0
## 172 -0.33143852     1.26651990        0
## 173 -1.09629664     1.35214056        0
## 174 -0.33143852    -0.78837605        0
## 177 -0.23583125    -0.67421517        0
## 178 -1.19190391    -1.38772071        0
## 179 -1.28751117    -1.35918049        0
## 180 -0.61826032    -1.04523805        0
## 181 -1.09629664    -1.55896204        0
## 182 -0.61826032     0.01075015        0
## 183 -0.52265305     1.32360034        1
## 184 -0.42704579    -0.78837605        0
## 185 -0.42704579    -0.30319229        0
## 186 -0.61826032    -0.13195096        0
## 187 -1.66994024     0.32469259        0
## 188 -0.42704579    -0.84545650        0
## 189 -0.23583125     0.03929037        0
## 190 -0.90508211    -1.10231849        0
## 191 -1.28751117     0.38177303        0
## 192 -1.76554750    -1.27355982        0
## 194 -1.76554750    -0.01779007        0
## 195 -0.90508211     0.52447414        0
## 196 -0.33143852    -0.78837605        0
## 197 -0.71386758     0.23907192        0
## 198 -1.66994024    -0.98815761        0
## 201 -0.23583125    -0.90253694        0
## 202  1.10267046     0.09637081        0
## 203  0.14659781     1.80878411        1
## 204  0.33781234     0.01075015        0
## 205  1.96313585     0.86695680        1
```

```
## 206  0.91145593     -0.67421517         0
## 207  1.67631405      1.69462322         1
## 209  0.24220507      2.03710588         1
## 210  0.81584866     -1.38772071         0
## 211  1.00706320      0.72425569         1
## 212  1.38949226      2.26542765         1
## 214 -0.23583125     -0.36027273         0
## 215  0.91145593     -0.78837605         0
## 216  2.15435038      1.06673835         1
## 217  1.10267046     -0.16049118         0
## 218  0.24220507      0.21053170         0
## 219  0.81584866      0.72425569         0
## 220  2.05874311      2.06564610         1
## 221  0.33781234      0.26761214         0
## 222 -0.23583125      0.58155458         1
## 223 -0.04461672      2.09418633         1
## 225 -0.23583125     -0.30319229         0
## 227 -0.14022399      1.58046234         1
## 231 -0.23583125      2.17980699         1
## 232  0.14659781     -0.81691628         0
## 233  0.24220507      1.03819813         1
## 235  0.05099054      1.18089923         0
## 238 -0.04461672      0.26761214         0
## 240  1.48509952      2.06564610         1
## 242  0.05099054     -0.33173251         0
## 243  1.19827773      0.49593392         1
## 244  1.77192132      0.95257746         1
## 245  0.33781234      0.03929037         0
## 246  1.29388499      2.15126677         1
## 247 -0.23583125     -0.58859450         0
## 248  1.86752858      1.46630145         1
## 249  0.33781234     -0.53151406         0
## 250 -0.23583125      0.75279591         1
## 251  0.62463413     -0.90253694         0
## 252 -0.04461672     -0.53151406         0
## 253  1.00706320      1.80878411         1
## 254 -0.04461672      2.15126677         1
## 256  1.38949226      0.55301436         1
## 257  0.33781234      0.03929037         0
## 258  0.24220507     -0.38881295         0
## 259  1.96313585      0.69571547         1
## 260  0.72024140      1.72316344         1
## 261 -0.23583125      0.18199148         0
## 262 -0.14022399      2.09418633         1
## 263  1.67631405      1.55192212         1
## 267  0.24220507      0.12491104         0
## 268 -0.04461672      0.09637081         0
## 269  0.91145593      2.09418633         1
## 270  0.24220507     -0.27465207         0
## 271  0.52902687      1.78024389         0
## 272  2.05874311      0.15345126         1
## 275  1.86752858     -1.27355982         1
## 276  1.86752858      0.09637081         1
## 277  0.05099054      0.01075015         0
```

```
## 278  1.10267046      0.49593392         1
## 279  1.38949226     -0.93107716         1
## 280  1.19827773     -0.98815761         1
## 282 -0.23583125     -0.27465207         0
## 283 -0.04461672     -0.01779007         1
## 284  1.38949226     -1.41626093         1
## 285  1.00706320      2.00856566         0
## 287 -0.04461672     -0.24611184         0
## 288  1.00706320      1.92294500         1
## 289  0.33781234      0.23907192         0
## 290 -0.04461672      0.21053170         1
## 291  0.14659781      1.80878411         1
## 293  1.67631405     -0.90253694         1
## 294 -0.04461672      0.18199148         0
## 295 -0.23583125     -0.38881295         0
## 296 -0.14022399     -0.21757162         0
## 297  0.43341960      0.06783059         1
## 298  0.52902687      1.18089923         1
## 300  0.81584866      1.32360034         1
## 301  1.96313585     -0.93107716         1
## 303 -0.04461672      1.89440477         1
## 304 -0.04461672      0.23907192         1
## 306  0.43341960     -0.47443362         0
## 308  0.91145593      1.20943946         1
## 309 -0.14022399      1.55192212         1
## 311  0.43341960     -0.01779007         0
## 312  0.14659781      0.72425569         1
## 313  0.05099054     -0.58859450         0
## 314  1.10267046      2.00856566         1
## 315  0.14659781      0.23907192         0
## 317  1.58070679      0.95257746         1
## 318 -0.23583125     -0.44589340         0
## 319  0.72024140     -1.10231849         1
## 320 -0.14022399     -0.30319229         0
## 321  1.38949226      1.92294500         1
## 322  1.48509952      0.32469259         1
## 323  0.33781234     -0.53151406         0
## 325  1.00706320      1.72316344         1
## 327  0.33781234      0.03929037         0
## 328  0.43341960      0.12491104         0
## 329 -0.14022399      1.35214056         1
## 330  0.91145593      1.03819813         1
## 331  0.05099054     -0.56005428         0
## 333  0.43341960     -0.16049118         0
## 334  0.24220507     -0.16049118         0
## 335  1.86752858     -0.30319229         1
## 336 -0.14022399     -0.47443362         0
## 337  1.96313585      2.09418633         1
## 338 -0.23583125      0.23907192         0
## 340  0.14659781      1.46630145         1
## 342 -0.23583125      0.12491104         0
## 344  0.91145593     -0.56005428         1
## 345  0.91145593      0.98111768         1
## 346  0.33781234     -0.21757162         0
```

```
## 348  1.58070679      1.06673835         1
## 349  0.14659781      0.18199148         0
## 350  0.05099054     -0.27465207         0
## 351  0.05099054      1.20943946         1
## 352 -0.04461672      0.12491104         0
## 354 -0.04461672     -0.38881295         0
## 355 -0.14022399      0.80987635         1
## 356  2.15435038     -1.04523805         1
## 357  1.58070679     -0.01779007         1
## 358  0.33781234      0.03929037         0
## 359  0.24220507      0.01075015         1
## 360  0.43341960     -0.47443362         0
## 361  0.52902687      1.66608300         1
## 362  1.48509952     -1.04523805         1
## 365  0.43341960      0.95257746         1
## 366  2.05874311     -1.18793916         1
## 370  1.58070679     -1.27355982         1
## 371  2.15435038     -0.70275539         1
## 374  2.05874311      1.69462322         1
## 375 -0.04461672      0.26761214         0
## 376  0.81584866     -1.10231849         1
## 377  0.81584866      0.09637081         0
## 378  0.43341960     -0.50297384         0
## 379  0.33781234      0.46739370         1
## 381  0.43341960     -0.18903140         0
## 382  1.00706320     -1.07377827         1
## 384  1.10267046     -1.21647938         1
## 385  1.86752858     -1.07377827         1
## 386  1.77192132     -0.30319229         1
## 387  1.10267046     -0.90253694         1
## 388  0.14659781      0.01075015         0
## 390  1.00706320     -1.01669783         1
## 391  1.00706320     -1.07377827         1
## 393  0.72024140     -0.73129561         1
## 394  2.15435038     -0.81691628         1
## 396  0.81584866     -0.84545650         1
## 397  1.29388499     -1.35918049         1
## 398  1.19827773     -1.44480115         1
## 399 -0.14022399     -1.07377827         0
```

```
testing_set[,1:2] = scale(testing_set[,1:2])
testing_set
```

```
##             Age EstimatedSalary Purchased
## 2   -0.30419063     -1.51354339         0
## 4   -1.05994374     -0.32456026         0
## 5   -1.81569686      0.28599864         0
## 9   -1.24888202     -1.09579256         0
## 12  -1.15441288     -0.48523366         0
## 18   0.64050076     -1.32073531         1
## 19   0.73496990     -1.25646596         1
## 20   0.92390818     -1.22433128         1
## 22   0.82943904     -0.58163769         1
## 29  -0.87100546     -0.77444577         0
## 32  -1.05994374      2.24621408         1
```

```
## 34  -0.96547460     -0.74231109         0
## 35  -1.05994374      0.73588415         0
## 38  -0.77653633     -0.58163769         0
## 45  -0.96547460      0.54307608         0
## 46  -1.43782030     -1.51354339         0
## 48  -1.05994374     -0.42096430         0
## 52  -1.91016600     -0.74231109         0
## 66  -1.34335116     -0.29242558         0
## 69  -1.53228944     -0.13175218         0
## 74  -0.49312891      1.47498177         0
## 75  -0.58759805     -1.57781275         0
## 82   0.07368593     -0.80658045         0
## 84  -0.30419063      0.67161480         0
## 85  -0.77653633     -0.16388686         0
## 86  -0.68206719      1.63565517         1
## 87  -1.34335116     -0.38882962         0
## 89  -1.15441288      0.44667204         0
## 103 -0.58759805      0.60734544         0
## 104 -0.49312891      2.63183023         1
## 107 -1.15441288     -1.03152320         0
## 108 -1.05994374      0.70374947         0
## 109 -1.15441288      0.60734544         0
## 117 -0.30419063      0.25386397         0
## 124 -0.30419063     -0.45309898         0
## 126  0.07368593     -0.19602154         0
## 127  0.35709335     -0.06748283         0
## 131 -0.68206719     -0.29242558         0
## 134 -1.62675858      0.02892121         0
## 139 -0.96547460     -0.26029090         0
## 148  0.26262421     -1.19219660         0
## 154 -0.20972149     -0.54950301         0
## 156 -0.68206719     -1.67421679         0
## 159 -1.15441288     -1.19219660         0
## 162 -1.24888202      0.73588415         0
## 163 -0.11525235     -1.09579256         0
## 170 -0.87100546     -0.64590705         0
## 175 -0.39865977      0.15745993         0
## 176 -1.43782030     -1.25646596         0
## 193 -0.87100546     -0.77444577         0
## 199 -1.15441288      0.41453736         0
## 200 -0.30419063     -1.44927403         0
## 208  1.30178474      1.50711645         0
## 213  1.96306872     -0.80658045         0
## 224  2.05753786      1.12150030         1
## 226 -0.11525235     -0.45309898         0
## 228  1.67966130      2.11767536         1
## 229  0.16815507      0.15745993         0
## 230  0.35709335      0.41453736         1
## 234  1.01837732      0.60734544         1
## 236  0.73496990      0.38240268         1
## 237  0.16815507     -0.32456026         0
## 239  0.73496990      0.47880672         0
## 241  0.35709335      2.63183023         1
## 255  1.11284646     -0.74231109         0
```

```
## 264 -0.30419063      0.15745993         0
## 265  0.92390818      0.73588415         1
## 266  0.35709335      1.31430838         1
## 273  2.05753786     -0.80658045         1
## 274  0.07368593      1.25003902         1
## 281  1.96306872      0.67161480         1
## 286 -0.11525235      0.83228819         1
## 292  1.01837732      0.70374947         1
## 299  0.64050076      0.38240268         0
## 302  0.92390818      0.22172929         1
## 305  0.16815507     -0.22815622         0
## 307  1.20731560      2.14981004         0
## 310 -0.02078321     -0.54950301         0
## 316  0.07368593      0.25386397         1
## 324  0.92390818     -1.19219660         1
## 326  0.26262421     -0.22815622         0
## 332  0.92390818      1.66778985         1
## 339 -0.02078321     -0.38882962         0
## 341  1.39625388      1.18576966         1
## 343 -0.02078321     -0.06748283         0
## 347  1.39625388      0.15745993         1
## 353  0.35709335      0.73588415         1
## 363  0.82943904     -0.54950301         1
## 364  0.35709335      0.38240268         0
## 367  1.86859958     -0.64590705         1
## 368  0.73496990      0.67161480         1
## 369 -0.02078321      0.12532525         0
## 372  2.05753786      0.51094140         1
## 373  0.07368593      0.18959461         0
## 380  1.86859958     -1.41713935         1
## 383  0.54603163      2.31048343         1
## 389  0.82943904     -1.06365788         1
## 392  0.82943904     -1.41713935         1
## 395  0.07368593     -0.26029090         0
## 400  1.01837732     -0.99938852         1
```

Ajustar el modelo de Regresión Logística con el conjunto de entrenamiento

```
classifier = glm(formula = Purchased ~.,
                 data = training_set,
                 family = binomial)
summary(classifier)
```

```
##
## Call:
## glm(formula = Purchased ~ ., family = binomial, data = training_set)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.0753  -0.5235  -0.1161   0.3224   2.3977
##
## Coefficients:
```

```
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -1.1923     0.2018  -5.908 3.47e-09 ***
## Age               2.6324     0.3461   7.606 2.83e-14 ***
## EstimatedSalary   1.3947     0.2326   5.996 2.03e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 390.89  on 299  degrees of freedom
## Residual deviance: 199.78  on 297  degrees of freedom
## AIC: 205.78
##
## Number of Fisher Scoring iterations: 6
```

Predicción de los resultados con el conjunto testing

```
prob_pred = predict(classifier,
                    type = 'response',
                    newdata = testing_set[, -3])
prob_pred
```

```
##           2           4           5           9          12          18
## 0.016239537 0.011714838 0.003784646 0.002452746 0.007333944 0.206157658
##          19          20          22          29          32          34
## 0.266993507 0.385147569 0.544857878 0.010300564 0.299492214 0.008416879
##          35          38          45          46          48          52
## 0.049447195 0.017164148 0.048505130 0.000834306 0.010256162 0.000705535
##          66          69          74          75          82          84
## 0.005844846 0.004453495 0.393346849 0.007106567 0.106858918 0.258008495
##          85          86          87          89         103         104
## 0.030324893 0.330364917 0.005113292 0.026386185 0.131014806 0.764977231
##         107         108         109         117         124         126
## 0.003436779 0.047382710 0.032796511 0.162604929 0.067549405 0.218965851
##         127         131         134         139         148         154
## 0.414256249 0.032433775 0.004345784 0.016353871 0.103059060 0.075109325
##         156         159         162         163         170         175
## 0.004855698 0.002748726 0.030664790 0.046355572 0.012298141 0.116901671
##         176         193         199         200         208         213
## 0.001193661 0.010300564 0.025258942 0.017735391 0.987085981 0.945335997
##         224         226         228         229         230         234
## 0.996945445 0.106443057 0.997939388 0.370509342 0.580752796 0.911776284
##         236         237         239         241         255         264
## 0.781727341 0.231067293 0.803799604 0.968270671 0.668600783 0.145116928
##         265         266         273         274         281         286
## 0.906031141 0.829311241 0.956852035 0.678106429 0.992695540 0.417048639
##         292         299         302         305         307         310
## 0.922009699 0.736349886 0.824773682 0.255813682 0.993200711 0.117805893
##         316         324         326         332         339         341
## 0.344284549 0.395813865 0.305941244 0.972503555 0.143160230 0.984279548
##         343         347         353         363         364         367
## 0.207327301 0.937190970 0.684394006 0.555947912 0.569802886 0.944051224
##         368         369         372         373         380         383
```

```
## 0.8427877409 0.2549836305 0.9928717092 0.3243409327 0.8519685008 0.9697473704
##          389          392          395          400
## 0.3793408625 0.2718336775 0.2040229226 0.5236436275
```

Lo anterior fue calcular un vector de probabilidades de en purchased sea 1. Con esto construimos la prediccion de la variable dependiente

```
y_pred = ifelse(prob_pred > 0.5, 1, 0)
y_pred
```

```
##   2   4   5   9  12  18  19  20  22  29  32  34  35  38  45  46  48  52  66  69
##   0   0   0   0   0   0   0   0   1   0   0   0   0   0   0   0   0   0   0   0
##  74  75  82  84  85  86  87  89 103 104 107 108 109 117 124 126 127 131 134 139
##   0   0   0   0   0   0   0   0   0   1   0   0   0   0   0   0   0   0   0   0
## 148 154 156 159 162 163 170 175 176 193 199 200 208 213 224 226 228 229 230 234
##   0   0   0   0   0   0   0   0   0   0   0   0   1   1   1   0   1   0   1   1
## 236 237 239 241 255 264 265 266 273 274 281 286 292 299 302 305 307 310 316 324
##   1   0   1   1   1   0   1   1   1   1   1   0   1   1   1   0   1   0   0   0
## 326 332 339 341 343 347 353 363 364 367 368 369 372 373 380 383 389 392 395 400
##   0   1   0   1   0   1   1   1   1   1   1   0   1   0   1   1   0   0   0   1
```

Comparar uno a uno los resultados predichos con los esperados no es una buena técnica por lo que se construye la matriz de confusión

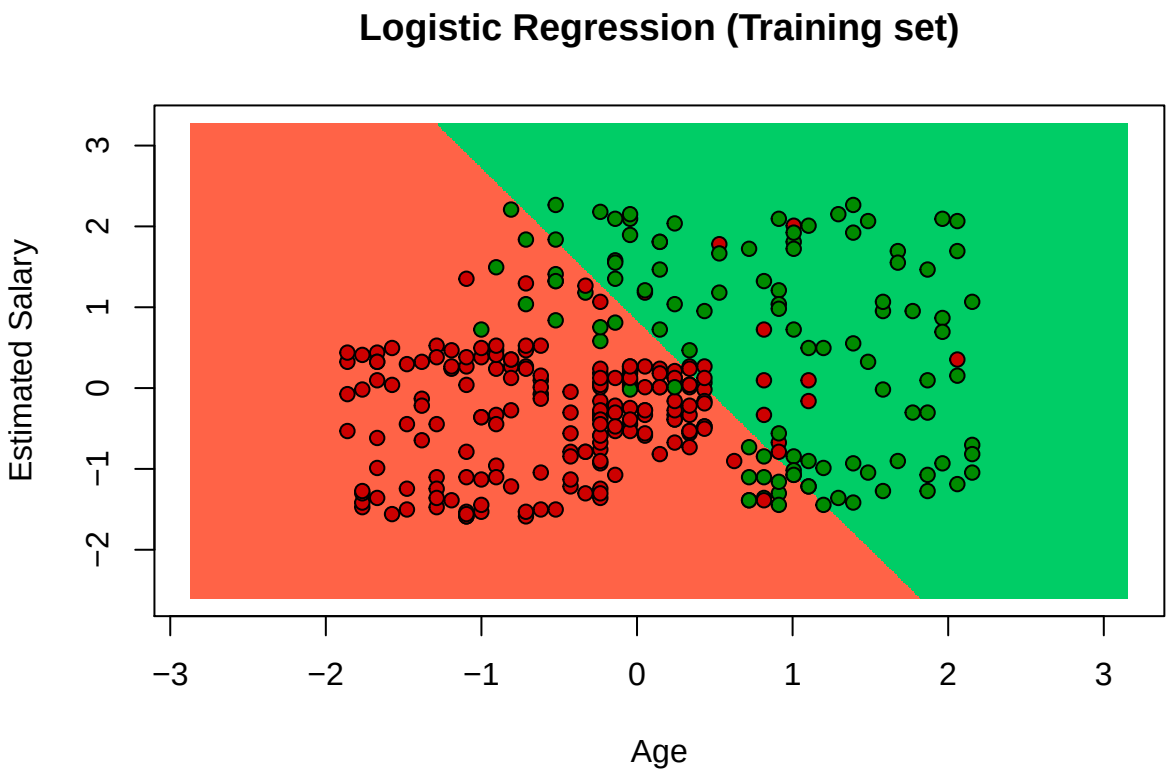
```
cm = table(testing_set[, 3], y_pred)
cm
```

```
##    y_pred
##       0  1
##   0 57  7
##   1 10 26
```

La diagonal principal es la cantidad de datos que son predichos correctamente.La diagonal secundaria son los fallos

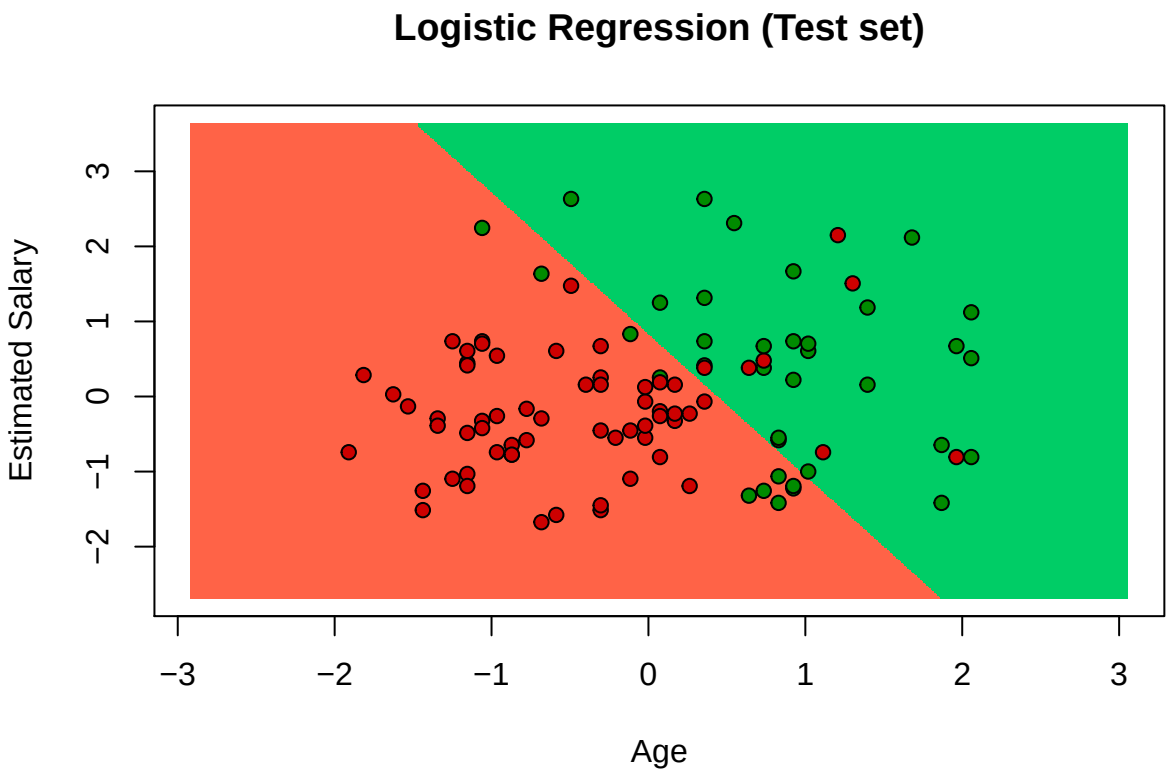
Visualización del conjunto de entranmiento

```
library(ElemStatLearn)
set = training_set
X1 = seq(min(set[, 1]) - 1, max(set[, 1]) + 1, by = 0.01)
X2 = seq(min(set[, 2]) - 1, max(set[, 2]) + 1, by = 0.01)
grid_set = expand.grid(X1, X2)
colnames(grid_set) = c('Age', 'EstimatedSalary')
prob_set = predict(classifier, type = 'response', newdata = grid_set)
y_grid = ifelse(prob_set > 0.5, 1, 0)
plot(set[, -3],
     main = 'Logistic Regression (Training set)',
     xlab = 'Age', ylab = 'Estimated Salary',
     xlim = range(X1), ylim = range(X2))
contour(X1, X2, matrix(as.numeric(y_grid), length(X1), length(X2)), add = TRUE)
points(grid_set, pch = '.', col = ifelse(y_grid == 1, 'springgreen3', 'tomato'))
points(set, pch = 21, bg = ifelse(set[, 3] == 1, 'green4', 'red3'))
```



Visualising the Test set results

```
library(ElemStatLearn)
set = testing_set
X1 = seq(min(set[, 1]) - 1, max(set[, 1]) + 1, by = 0.01)
X2 = seq(min(set[, 2]) - 1, max(set[, 2]) + 1, by = 0.01)
grid_set = expand.grid(X1, X2)
colnames(grid_set) = c('Age', 'EstimatedSalary')
prob_set = predict(classifier, type = 'response', newdata = grid_set)
y_grid = ifelse(prob_set > 0.5, 1, 0)
plot(set[, -3],
     main = 'Logistic Regression (Test set)',
     xlab = 'Age', ylab = 'Estimated Salary',
     xlim = range(X1), ylim = range(X2))
contour(X1, X2, matrix(as.numeric(y_grid), length(X1), length(X2)), add = TRUE)
points(grid_set, pch = '.', col = ifelse(y_grid == 1, 'springgreen3', 'tomato'))
points(set, pch = 21, bg = ifelse(set[, 3] == 1, 'green4', 'red3'))
```


Logistic Regression (Test set)
Estimated Salary
Age
−3
−2
−1
0
1
2
3