

Classification

Abhirup Sen

11/05/2021

```
# tinytex::install_tinytex()

library(ISLR)
library(tinytex)
names(Smarket)

## [1] "Year"      "Lag1"       "Lag2"       "Lag3"       "Lag4"       "Lag5"
## [7] "Volume"    "Today"      "Direction"

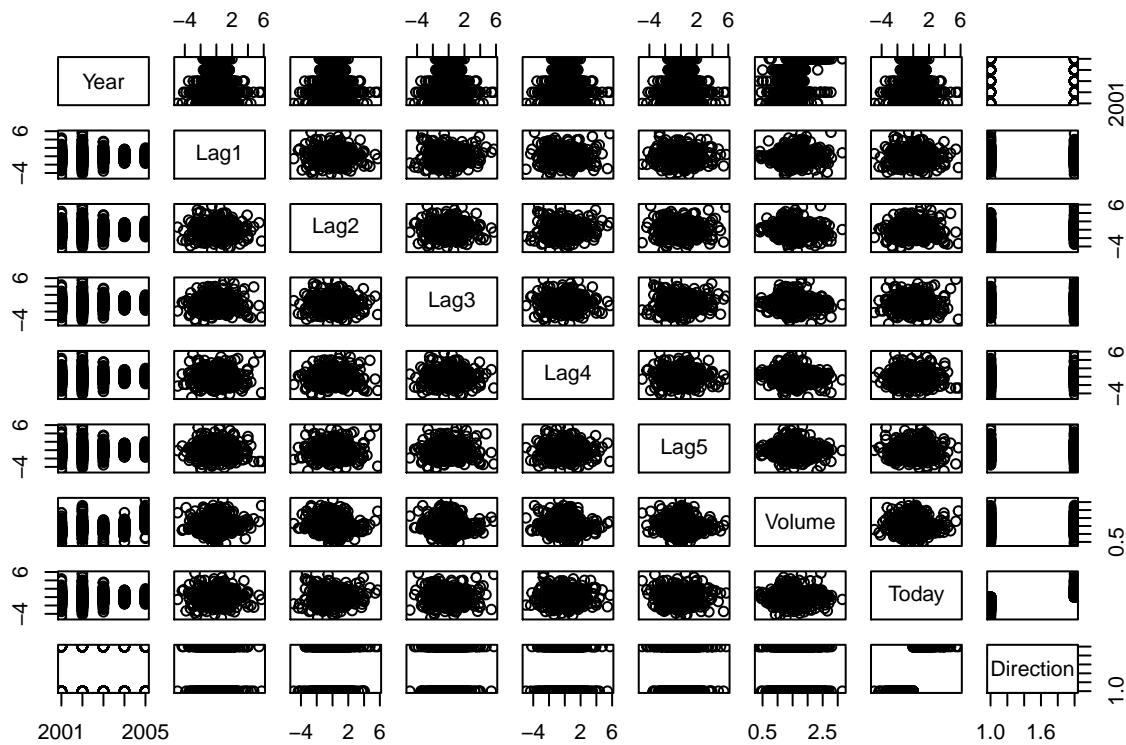
dim(Smarket)

## [1] 1250     9

summary(Smarket)

##      Year          Lag1          Lag2          Lag3          Lag4          Lag5
##  Min.   :2001   Min.   :-4.922000   Min.   :-4.922000   Min.   :-4.922000
##  1st Qu.:2002  1st Qu.:-0.639500  1st Qu.:-0.639500  1st Qu.:-0.640000
##  Median :2003  Median : 0.039000  Median : 0.039000  Median : 0.038500
##  Mean   :2003  Mean   : 0.003834  Mean   : 0.003919  Mean   : 0.001716
##  3rd Qu.:2004 3rd Qu.: 0.596750  3rd Qu.: 0.596750  3rd Qu.: 0.596750
##  Max.   :2005  Max.   : 5.733000  Max.   : 5.733000  Max.   : 5.733000
##      Lag4          Lag5          Volume        Today
##  Min.   :-4.922000   Min.   :-4.92200   Min.   :0.3561   Min.   :-4.922000
##  1st Qu.:-0.640000  1st Qu.:-0.64000  1st Qu.:1.2574  1st Qu.:-0.639500
##  Median : 0.038500  Median : 0.03850  Median :1.4229  Median : 0.038500
##  Mean   : 0.001636  Mean   : 0.00561  Mean   :1.4783  Mean   : 0.003138
##  3rd Qu.: 0.596750  3rd Qu.: 0.59700  3rd Qu.:1.6417  3rd Qu.: 0.596750
##  Max.   : 5.733000  Max.   : 5.73300  Max.   :3.1525  Max.   : 5.733000
##      Direction
##  Down:602
##  Up :648
## 
## 
## 
## 

pairs(Smarket)
```



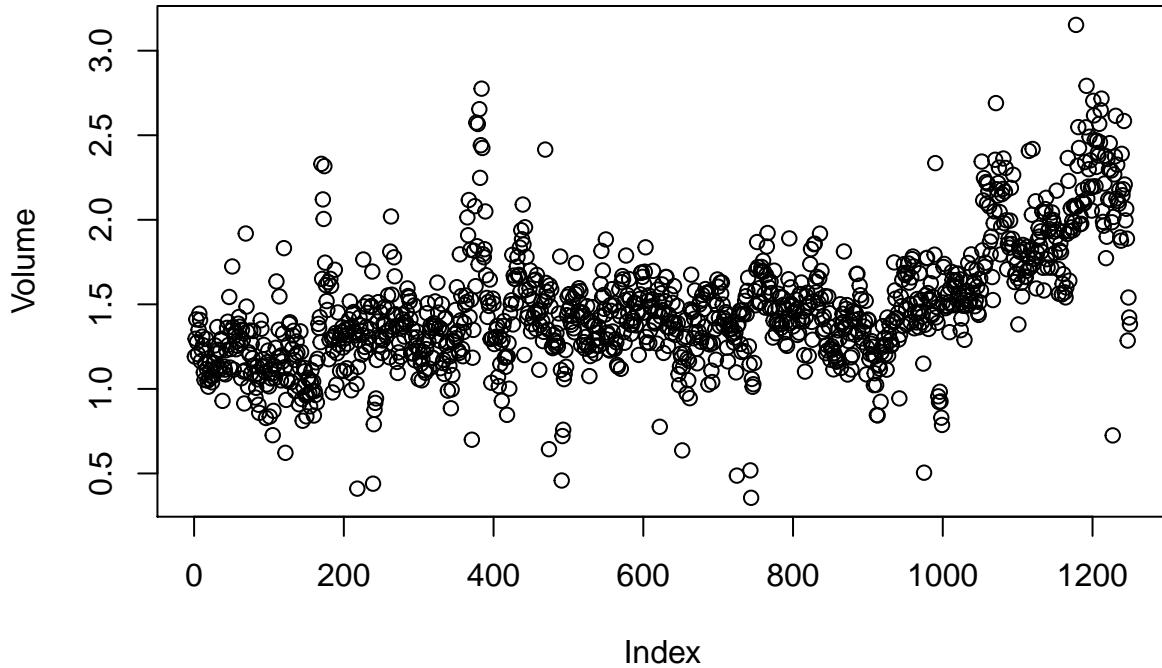
```
# cor(Smarket)
```

removing the Direction col

```
cor(Smarket[,-9])
```

```
##          Year      Lag1      Lag2      Lag3      Lag4
## Year  1.00000000  0.029699649  0.030596422  0.033194581  0.035688718
## Lag1  0.02969965  1.000000000 -0.026294328 -0.010803402 -0.002985911
## Lag2  0.03059642 -0.026294328  1.000000000 -0.025896670 -0.010853533
## Lag3  0.03319458 -0.010803402 -0.025896670  1.000000000 -0.024051036
## Lag4  0.03568872 -0.002985911 -0.010853533 -0.024051036  1.000000000
## Lag5  0.02978799 -0.005674606 -0.003557949 -0.018808338 -0.027083641
## Volume 0.53900647  0.040909908 -0.043383215 -0.041823686 -0.048414246
## Today  0.03009523 -0.026155045 -0.010250033 -0.002447647 -0.006899527
##          Lag5      Volume     Today
## Year    0.029787995  0.53900647  0.030095229
## Lag1   -0.005674606  0.04090991 -0.026155045
## Lag2   -0.003557949 -0.04338321 -0.010250033
## Lag3   -0.018808338 -0.04182369 -0.002447647
## Lag4   -0.027083641 -0.04841425 -0.006899527
## Lag5    1.000000000 -0.02200231 -0.034860083
## Volume -0.022002315  1.000000000  0.014591823
## Today  -0.034860083  0.01459182  1.000000000
```

```
attach(Smarket)
plot(Volume)
```



4.6.2 Logistic Regression

** `glm()` ** : generated Linear model, similar to `lm()` except that we must pass `family = binomial` in order to tell R that its logistic regression

```
glm.fits = glm(Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 + Volume, data = Smarket, family = binomial)
summary(glm.fits)
```

```
##
## Call:
## glm(formula = Direction ~ Lag1 + Lag2 + Lag3 + Lag4 + Lag5 +
##       Volume, family = binomial, data = Smarket)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max 
## -1.446  -1.203   1.065   1.145   1.326 
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)    
## (Intercept) -0.126000  0.240736 -0.523   0.601    
## Lag1        -0.073074  0.050167 -1.457   0.145    
## Lag2        -0.042301  0.050086 -0.845   0.398    
## Lag3         0.011085  0.049939  0.222   0.824
```

```

## Lag4          0.009359  0.049974  0.187    0.851
## Lag5          0.010313  0.049511  0.208    0.835
## Volume        0.135441  0.158360  0.855    0.392
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 1731.2  on 1249  degrees of freedom
## Residual deviance: 1727.6  on 1243  degrees of freedom
## AIC: 1741.6
##
## Number of Fisher Scoring iterations: 3

```

```
coef(glm.fits)
```

```

## (Intercept)      Lag1       Lag2       Lag3       Lag4       Lag5
## -0.126000257 -0.073073746 -0.042301344  0.011085108  0.009358938  0.010313068
##           Volume
##  0.135440659

```

```
summary(glm.fits)$coef
```

	Estimate	Std. Error	z value	Pr(> z)
## (Intercept)	-0.126000257	0.24073574	-0.5233966	0.6006983
## Lag1	-0.073073746	0.05016739	-1.4565986	0.1452272
## Lag2	-0.042301344	0.05008605	-0.8445733	0.3983491
## Lag3	0.011085108	0.04993854	0.2219750	0.8243333
## Lag4	0.009358938	0.04997413	0.1872757	0.8514445
## Lag5	0.010313068	0.04951146	0.2082966	0.8349974
## Volume	0.135440659	0.15835970	0.8552723	0.3924004

```
summary(glm.fits)$coef[,4]
```

```

## (Intercept)      Lag1       Lag2       Lag3       Lag4       Lag5
##  0.6006983   0.1452272   0.3983491   0.8243333   0.8514445   0.8349974
##           Volume
##  0.3924004

```

predict() to predict the probability

type = "response" is of importance ; it tells R to output probabilities as opposed to other information such as logit. If no data is supplied to predict()function, then the probabilities are computed for the training data that was used to fit the logistic regression model.

```

glm.probs = predict(glm.fits, type = "response")
glm.probs

```

```

##      1      2      3      4      5      6      7      8
## 0.5070841 0.4814679 0.4811388 0.5152224 0.5107812 0.5069565 0.4926509 0.5092292
##      9     10     11     12     13     14     15     16
## 0.5176135 0.4888378 0.4965211 0.5197834 0.5183031 0.4963852 0.4864892 0.5153660
##     17     18     19     20     21     22     23     24

```

```

## 0.5053976 0.5319322 0.5167163 0.4983272 0.5229797 0.5221677 0.5340639 0.4920971
##      25      26      27      28      29      30      31      32
## 0.5008751 0.5191211 0.4919679 0.5373796 0.5571306 0.5574363 0.5307063 0.5083437
##      33      34      35      36      37      38      39      40
## 0.4714015 0.4949516 0.5471346 0.5260910 0.5203319 0.4902306 0.4760375 0.4850667
##      41      42      43      44      45      46      47      48
## 0.4979462 0.5535387 0.6176442 0.5284518 0.5301251 0.5153681 0.5355161 0.4945795
##      49      50      51      52      53      54      55      56
## 0.5249470 0.5735593 0.5455080 0.4765992 0.4531691 0.4544318 0.5379025 0.5591742
##      57      58      59      60      61      62      63      64
## 0.4991918 0.5219328 0.5861722 0.5560757 0.4289095 0.4896402 0.5133387 0.4601776
##      65      66      67      68      69      70      71      72
## 0.4962909 0.4847361 0.4972136 0.5013727 0.4537779 0.4608736 0.5284874 0.5541953
##      73      74      75      76      77      78      79      80
## 0.5580780 0.4900787 0.4795402 0.4708035 0.5026603 0.4960765 0.5020786 0.5402471
##      81      82      83      84      85      86      87      88
## 0.4971418 0.4894053 0.5092301 0.5158414 0.5121983 0.5110322 0.4992040 0.4979377
##      89      90      91      92      93      94      95      96
## 0.4626107 0.4781316 0.5076532 0.4834586 0.5076192 0.5437706 0.5200800 0.5139898
##      97      98      99      100     101     102     103     104
## 0.5265218 0.5379726 0.5111662 0.4800603 0.4794640 0.4756823 0.5139061 0.5123646
##     105     106     107     108     109     110     111     112
## 0.5059202 0.5255465 0.5096551 0.5209010 0.5503284 0.5454886 0.5126250 0.4992420
##     113     114     115     116     117     118     119     120
## 0.4877559 0.4895564 0.5159692 0.5301499 0.5199546 0.5171795 0.4914312 0.5169617
##     121     122     123     124     125     126     127     128
## 0.4918327 0.4835776 0.5302389 0.5613952 0.5149406 0.5204423 0.5258673 0.4652345
##     129     130     131     132     133     134     135     136
## 0.4681630 0.5195629 0.5106937 0.5171932 0.5106220 0.5060068 0.5383249 0.5549366
##     137     138     139     140     141     142     143     144
## 0.4959159 0.4643444 0.4838769 0.5011806 0.5049573 0.5038605 0.5003239 0.5078093
##     145     146     147     148     149     150     151     152
## 0.5258066 0.5093433 0.5313691 0.5212320 0.4837498 0.4858783 0.5042435 0.5235971
##     153     154     155     156     157     158     159     160
## 0.5069808 0.5259487 0.4998119 0.5114614 0.5053207 0.4938316 0.4720827 0.4828294
##     161     162     163     164     165     166     167     168
## 0.5409544 0.5398367 0.5501315 0.5024233 0.4939689 0.5121701 0.5527710 0.5745127
##     169     170     171     172     173     174     175     176
## 0.5134565 0.6178172 0.5779814 0.5586072 0.5973840 0.5948993 0.4627514 0.4493427
##     177     178     179     180     181     182     183     184
## 0.5184195 0.5092985 0.4822288 0.4935678 0.4995467 0.4824233 0.5149404 0.5199915
##     185     186     187     188     189     190     191     192
## 0.5222882 0.5336493 0.4741573 0.4711892 0.5100753 0.5197549 0.5064399 0.5464963
##     193     194     195     196     197     198     199     200
## 0.5454220 0.5084035 0.4684574 0.5001289 0.5212669 0.4926755 0.4915752 0.5476582
##     201     202     203     204     205     206     207     208
## 0.5729136 0.5303325 0.4619934 0.4668795 0.4840479 0.4788247 0.5164763 0.5263822
##     209     210     211     212     213     214     215     216
## 0.5064011 0.5074269 0.4831733 0.4948615 0.5192656 0.5229545 0.5019649 0.5151649
##     217     218     219     220     221     222     223     224
## 0.5225812 0.4659844 0.4828791 0.5183050 0.5601603 0.5180282 0.4991824 0.5220576
##     225     226     227     228     229     230     231     232
## 0.4954441 0.4737651 0.5018794 0.5345163 0.5543575 0.5395897 0.5131541 0.5413004
##     233     234     235     236     237     238     239     240

```

```

## 0.5184810 0.4843930 0.4872959 0.4997228 0.5333760 0.5326903 0.4801647 0.4887269
## 241 242 243 244 245 246 247 248
## 0.4802926 0.4885029 0.5200361 0.5130254 0.4924633 0.4985257 0.5191093 0.5298686
## 249 250 251 252 253 254 255 256
## 0.5322041 0.5144131 0.5216854 0.5317916 0.5060478 0.5374491 0.5120715 0.5166303
## 257 258 259 260 261 262 263 264
## 0.5374606 0.5072765 0.5045054 0.5064177 0.5089125 0.5846415 0.5466275 0.4738281
## 265 266 267 268 269 270 271 272
## 0.5084511 0.5689293 0.5664230 0.5353604 0.5205079 0.4820923 0.4626234 0.4995321
## 273 274 275 276 277 278 279 280
## 0.5023880 0.5104609 0.5419878 0.5554827 0.5115796 0.5210413 0.5141462 0.4673050
## 281 282 283 284 285 286 287 288
## 0.4960389 0.5178064 0.5266471 0.4842465 0.4623225 0.5182483 0.5111942 0.5212888
## 289 290 291 292 293 294 295 296
## 0.5178603 0.4975261 0.5175677 0.5358673 0.5231700 0.4968623 0.4938653 0.5045076
## 297 298 299 300 301 302 303 304
## 0.5393732 0.5316516 0.5131948 0.5332988 0.5098507 0.4878379 0.4943066 0.5019081
## 305 306 307 308 309 310 311 312
## 0.5281208 0.5387422 0.5190878 0.5061119 0.5005126 0.5168113 0.5039684 0.5484832
## 313 314 315 316 317 318 319 320
## 0.5269962 0.5076957 0.4784089 0.4874184 0.5256741 0.5119433 0.5410865 0.5423497
## 321 322 323 324 325 326 327 328
## 0.5299964 0.5249800 0.5345663 0.5423733 0.5086142 0.4836342 0.5055525 0.5349018
## 329 330 331 332 333 334 335 336
## 0.5568112 0.5392905 0.4460445 0.4864598 0.5586683 0.4931917 0.4598562 0.5020327
## 337 338 339 340 341 342 343 344
## 0.5109182 0.4989720 0.5241709 0.5453403 0.5079910 0.4800945 0.5053793 0.5319222
## 345 346 347 348 349 350 351 352
## 0.5263558 0.5213901 0.5035706 0.5515162 0.5425999 0.4892929 0.5441519 0.5336913
## 353 354 355 356 357 358 359 360
## 0.5025770 0.5338046 0.5302534 0.5240041 0.5348673 0.4548259 0.4754708 0.5475335
## 361 362 363 364 365 366 367 368
## 0.5636227 0.5671526 0.5251079 0.5342514 0.5526009 0.4956588 0.5193831 0.5564217
## 369 370 371 372 373 374 375 376
## 0.5937056 0.5296956 0.4084461 0.4832595 0.5777391 0.6224823 0.5606276 0.5084422
## 377 378 379 380 381 382 383 384
## 0.5568850 0.5596453 0.5636721 0.5626635 0.6486435 0.6312180 0.6175849 0.4613402
## 385 386 387 388 389 390 391 392
## 0.4752021 0.5055134 0.4177204 0.4835240 0.5329494 0.5861040 0.6098456 0.5977531
## 393 394 395 396 397 398 399 400
## 0.4905924 0.4287278 0.4379590 0.4750538 0.5310447 0.5709875 0.4779914 0.4498097
## 401 402 403 404 405 406 407 408
## 0.5066746 0.4780597 0.5259072 0.5149705 0.4771218 0.5376513 0.5164203 0.5303509
## 409 410 411 412 413 414 415 416
## 0.5552679 0.5229186 0.4971848 0.5814800 0.5216554 0.5145818 0.4894369 0.4596034
## 417 418 419 420 421 422 423 424
## 0.4900315 0.4922908 0.5626621 0.5359886 0.4912912 0.5469931 0.5433529 0.5750779
## 425 426 427 428 429 430 431 432
## 0.5508110 0.5234160 0.5633526 0.4862726 0.4575218 0.5582288 0.5929242 0.4728985
## 433 434 435 436 437 438 439 440
## 0.5188711 0.5688275 0.5814117 0.5815766 0.5086407 0.5508001 0.4986800 0.4152333
## 441 442 443 444 445 446 447 448
## 0.4621799 0.4526092 0.5362726 0.5383721 0.4886116 0.4923419 0.5226693 0.5333180
## 449 450 451 452 453 454 455 456

```

```

## 0.5492493 0.5024581 0.5068784 0.5484071 0.5063423 0.5239311 0.4905745 0.4894922
##      457      458      459      460      461      462      463      464
## 0.4974757 0.5051423 0.5584688 0.5637579 0.5509476 0.5175236 0.4970779 0.4705108
##      465      466      467      468      469      470      471      472
## 0.4749748 0.5330979 0.5393989 0.4935729 0.4886010 0.5089374 0.5302971 0.5651396
##      473      474      475      476      477      478      479      480
## 0.4906811 0.4596468 0.5328854 0.5480893 0.5500662 0.5311520 0.5068863 0.5390923
##      481      482      483      484      485      486      487      488
## 0.5080554 0.4884507 0.5175151 0.5396414 0.4854819 0.4964361 0.5523240 0.5430983
##      489      490      491      492      493      494      495      496
## 0.5137434 0.4817761 0.4905121 0.5060156 0.5290269 0.5112511 0.4930500 0.4458654
##      497      498      499      500      501      502      503      504
## 0.4696628 0.4870720 0.5167604 0.5651517 0.5041312 0.4986487 0.5187366 0.5069508
##      505      506      507      508      509      510      511      512
## 0.5416984 0.5440907 0.5411760 0.5540412 0.5483272 0.5112815 0.5540233 0.5733639
##      513      514      515      516      517      518      519      520
## 0.5027563 0.4876537 0.5463778 0.5228772 0.4859849 0.5376500 0.5410476 0.5351724
##      521      522      523      524      525      526      527      528
## 0.5335242 0.5005738 0.5138356 0.5409118 0.5321515 0.4751630 0.4472618 0.4995613
##      529      530      531      532      533      534      535      536
## 0.5436897 0.5098577 0.5304137 0.5246311 0.5269624 0.5056177 0.4874281 0.5203492
##      537      538      539      540      541      542      543      544
## 0.5475823 0.5144239 0.5146099 0.5067363 0.5450126 0.5616834 0.5166626 0.4563592
##      545      546      547      548      549      550      551      552
## 0.4738574 0.4680552 0.4855794 0.5170218 0.5144186 0.5008520 0.5557676 0.5376840
##      553      554      555      556      557      558      559      560
## 0.5064371 0.5200785 0.5145790 0.5587840 0.5111725 0.4552042 0.4931822 0.5163615
##      561      562      563      564      565      566      567      568
## 0.5217705 0.5187314 0.5394454 0.5151600 0.5031941 0.4732094 0.4826784 0.5440299
##      569      570      571      572      573      574      575      576
## 0.5068339 0.4962702 0.4890202 0.4866715 0.5400445 0.5544985 0.4989469 0.4917022
##      577      578      579      580      581      582      583      584
## 0.5269621 0.5196191 0.4994595 0.5089309 0.5166641 0.5230148 0.5442023 0.4975397
##      585      586      587      588      589      590      591      592
## 0.4754143 0.5089796 0.5291386 0.5143780 0.5173626 0.5635708 0.5480624 0.5062685
##      593      594      595      596      597      598      599      600
## 0.4898031 0.4912634 0.4862750 0.5008255 0.5385601 0.5082821 0.5072216 0.5088997
##      601      602      603      604      605      606      607      608
## 0.4941632 0.5085562 0.5370514 0.5429199 0.5118280 0.4840160 0.5047466 0.5311339
##      609      610      611      612      613      614      615      616
## 0.4893106 0.4940756 0.5249119 0.5525286 0.5457100 0.5361444 0.5233631 0.5234480
##      617      618      619      620      621      622      623      624
## 0.5020146 0.5123299 0.5372136 0.5028409 0.4902092 0.4965594 0.4954123 0.4977675
##      625      626      627      628      629      630      631      632
## 0.5362072 0.5517770 0.5105123 0.4928262 0.5181697 0.5402265 0.5567637 0.5052164
##      633      634      635      636      637      638      639      640
## 0.5193665 0.5140651 0.4991592 0.5358988 0.4901286 0.4996379 0.5383760 0.5282392
##      641      642      643      644      645      646      647      648
## 0.5215891 0.5285143 0.5175201 0.5409217 0.5336321 0.4935632 0.4881271 0.4915967
##      649      650      651      652      653      654      655      656
## 0.4885411 0.5141199 0.5077863 0.4841429 0.4926770 0.4980247 0.5147534 0.5157893
##      657      658      659      660      661      662      663      664
## 0.5306417 0.5119942 0.4995151 0.4993400 0.4951714 0.4855550 0.4899775 0.5057790
##      665      666      667      668      669      670      671      672

```

```

## 0.5168753 0.5337336 0.5059373 0.5204697 0.5541198 0.5148738 0.4980080 0.5086283
##       673      674      675      676      677      678      679      680
## 0.4926429 0.5054973 0.5020893 0.5131346 0.5453881 0.5166303 0.5485826 0.5492208
##       681      682      683      684      685      686      687      688
## 0.5291884 0.4991276 0.5231577 0.4915181 0.4819566 0.5086150 0.4883487 0.5073765
##       689      690      691      692      693      694      695      696
## 0.5197602 0.5235144 0.5029203 0.4929137 0.4965112 0.5235698 0.5160337 0.5316181
##       697      698      699      700      701      702      703      704
## 0.5108934 0.5086419 0.5499579 0.5314997 0.5191487 0.5137153 0.4892423 0.5033963
##       705      706      707      708      709      710      711      712
## 0.5278697 0.5183127 0.5008094 0.5191424 0.5269509 0.5096122 0.5197273 0.5259132
##       713      714      715      716      717      718      719      720
## 0.5146990 0.4927122 0.5007648 0.5299717 0.5372249 0.5384462 0.5047714 0.5143588
##       721      722      723      724      725      726      727      728
## 0.5161020 0.4785049 0.4938868 0.4986330 0.4855421 0.5005281 0.5107896 0.5281587
##       729      730      731      732      733      734      735      736
## 0.5130082 0.5220670 0.5045090 0.5244253 0.5297834 0.4948311 0.4923716 0.5281752
##       737      738      739      740      741      742      743      744
## 0.5179504 0.5100178 0.5003240 0.5134494 0.5094866 0.5009249 0.4903862 0.4809698
##       745      746      747      748      749      750      751      752
## 0.4810915 0.4901907 0.5027633 0.5143963 0.5063505 0.5032933 0.5237583 0.5226434
##       753      754      755      756      757      758      759      760
## 0.5418162 0.5225328 0.5264919 0.5108012 0.5120090 0.5150793 0.5213615 0.5191492
##       761      762      763      764      765      766      767      768
## 0.5251725 0.5322374 0.4994091 0.5308814 0.5677182 0.5383561 0.5204351 0.5144302
##       769      770      771      772      773      774      775      776
## 0.5103340 0.5400230 0.5273724 0.4924219 0.5027941 0.5113269 0.5039247 0.5197308
##       777      778      779      780      781      782      783      784
## 0.5320850 0.5061810 0.5132429 0.5338037 0.5272695 0.5232549 0.5238552 0.5065726
##       785      786      787      788      789      790      791      792
## 0.5073700 0.5194708 0.5023770 0.5204310 0.5198428 0.5040965 0.5107259 0.5241776
##       793      794      795      796      797      798      799      800
## 0.5402452 0.5558591 0.5721521 0.5012907 0.5265812 0.5203313 0.4865639 0.5061675
##       801      802      803      804      805      806      807      808
## 0.5403799 0.5567640 0.5335781 0.5194513 0.4846808 0.4934042 0.4970572 0.4954182
##       809      810      811      812      813      814      815      816
## 0.5259132 0.5163142 0.5066267 0.4957005 0.5149443 0.5376630 0.5214907 0.4971675
##       817      818      819      820      821      822      823      824
## 0.5339295 0.5369032 0.5188348 0.5067181 0.4979545 0.5479893 0.5356904 0.4962524
##       825      826      827      828      829      830      831      832
## 0.4979004 0.5208287 0.5252665 0.5564187 0.5593540 0.5381805 0.5064232 0.5047162
##       833      834      835      836      837      838      839      840
## 0.5122662 0.5311056 0.5596874 0.5658732 0.5126638 0.5069630 0.5120350 0.5153048
##       841      842      843      844      845      846      847      848
## 0.5390508 0.5131147 0.5151237 0.5106017 0.5014467 0.5040224 0.4901981 0.4962639
##       849      850      851      852      853      854      855      856
## 0.5112092 0.5078467 0.5162694 0.5058412 0.5216645 0.5054250 0.4737006 0.4897039
##       857      858      859      860      861      862      863      864
## 0.5311568 0.5150608 0.5234349 0.5130017 0.4951320 0.5139064 0.5151655 0.5128727
##       865      866      867      868      869      870      871      872
## 0.5138191 0.4971222 0.5128839 0.5450081 0.5239514 0.5114615 0.5059881 0.5329044
##       873      874      875      876      877      878      879      880
## 0.5234786 0.5288519 0.5163672 0.5232837 0.5090872 0.4962471 0.5058528 0.5222876
##       881      882      883      884      885      886      887      888

```

```

## 0.5288821 0.5304776 0.5172761 0.5018410 0.5398361 0.5351295 0.5264592 0.5296120
##     889      890      891      892      893      894      895      896
## 0.5014806 0.5075264 0.5111801 0.5074540 0.5062914 0.5220634 0.5258972 0.5461220
##     897      898      899      900      901      902      903      904
## 0.5645934 0.5131409 0.4774042 0.5005335 0.5403863 0.5204280 0.4822326 0.4898656
##     905      906      907      908      909      910      911      912
## 0.4879958 0.5084123 0.5085081 0.5031259 0.5113733 0.4936805 0.4957225 0.4942599
##     913      914      915      916      917      918      919      920
## 0.5107151 0.5095718 0.4979602 0.4843877 0.4950698 0.5061066 0.5134665 0.5200389
##     921      922      923      924      925      926      927      928
## 0.4994528 0.5054387 0.5040797 0.5236454 0.5108634 0.5045200 0.5142824 0.5078334
##     929      930      931      932      933      934      935      936
## 0.5341520 0.5368553 0.5092255 0.5175319 0.5069883 0.4994501 0.5241793 0.4955929
##     937      938      939      940      941      942      943      944
## 0.5009416 0.5194662 0.5093266 0.5331448 0.5390057 0.5033860 0.5125153 0.5327662
##     945      946      947      948      949      950      951      952
## 0.5414487 0.5238549 0.4958327 0.5365916 0.5351533 0.5197389 0.5324592 0.5253747
##     953      954      955      956      957      958      959      960
## 0.4972998 0.4863660 0.5081191 0.5197463 0.5200221 0.5287070 0.5089951 0.4882255
##     961      962      963      964      965      966      967      968
## 0.5059058 0.5195172 0.5277737 0.5267627 0.5006757 0.4933777 0.5107235 0.5322961
##     969      970      971      972      973      974      975      976
## 0.5272588 0.5097961 0.5388631 0.5165224 0.5097323 0.4995745 0.4783265 0.5232677
##     977      978      979      980      981      982      983      984
## 0.5331663 0.5057540 0.5126442 0.5234650 0.5172269 0.5451995 0.5226483 0.5053486
##     985      986      987      988      989      990      991      992
## 0.5120922 0.5015982 0.5063676 0.5218671 0.5338596 0.5668168 0.5247708 0.4997596
##     993      994      995      996      997      998      999      1000
## 0.4976013 0.4970796 0.5102657 0.4965390 0.4922132 0.4976166 0.4980502 0.5376846
##    1001     1002     1003     1004     1005     1006     1007     1008
## 0.5560722 0.5436973 0.5135738 0.5116102 0.5113422 0.5259038 0.5220768 0.5299009
##    1009     1010     1011     1012     1013     1014     1015     1016
## 0.5124372 0.4955240 0.5270000 0.5517132 0.5451431 0.5304140 0.5134074 0.5063355
##    1017     1018     1019     1020     1021     1022     1023     1024
## 0.5154540 0.5299581 0.5150724 0.5045083 0.5100974 0.5259957 0.5117403 0.5061873
##    1025     1026     1027     1028     1029     1030     1031     1032
## 0.5199482 0.5363334 0.5230987 0.5017409 0.5028611 0.5141279 0.5180293 0.5390666
##    1033     1034     1035     1036     1037     1038     1039     1040
## 0.5290903 0.5518620 0.5227948 0.4936795 0.4930808 0.5308783 0.5286152 0.5185462
##    1041     1042     1043     1044     1045     1046     1047     1048
## 0.5253744 0.5056068 0.5054550 0.5287203 0.5527815 0.5319080 0.5261735 0.5118123
##    1049     1050     1051     1052     1053     1054     1055     1056
## 0.5232334 0.5473123 0.5245777 0.5441706 0.5362158 0.5618332 0.5528768 0.5236738
##    1057     1058     1059     1060     1061     1062     1063     1064
## 0.5206989 0.5521634 0.5230082 0.5283617 0.5569550 0.5417814 0.5223468 0.5195234
##    1065     1066     1067     1068     1069     1070     1071     1072
## 0.5197425 0.5359940 0.5322446 0.5250376 0.5529107 0.5777024 0.5974807 0.5503854
##    1073     1074     1075     1076     1077     1078     1079     1080
## 0.5170978 0.5555977 0.5226235 0.5276745 0.5244775 0.5410871 0.5492334 0.5565967
##    1081     1082     1083     1084     1085     1086     1087     1088
## 0.5399225 0.5100433 0.5401469 0.5258198 0.5314995 0.5353962 0.5228285 0.5474604
##    1089     1090     1091     1092     1093     1094     1095     1096
## 0.5350994 0.5480502 0.5615687 0.5132205 0.5064815 0.5187024 0.5126273 0.5285573
##    1097     1098     1099     1100     1101     1102     1103     1104

```

```

## 0.5253966 0.5242548 0.5353673 0.5169853 0.5068378 0.5418127 0.5206434 0.5174404
## 1105 1106 1107 1108 1109 1110 1111 1112
## 0.5356032 0.5268589 0.5310284 0.5299015 0.5214264 0.5232587 0.5239843 0.5189234
## 1113 1114 1115 1116 1117 1118 1119 1120
## 0.5254496 0.5202817 0.5388529 0.5247842 0.5339999 0.5338470 0.5570775 0.5746293
## 1121 1122 1123 1124 1125 1126 1127 1128
## 0.5332632 0.5080069 0.5165660 0.5545847 0.5267333 0.5108068 0.5367991 0.5400948
## 1129 1130 1131 1132 1133 1134 1135 1136
## 0.5093598 0.5082497 0.5248010 0.5309203 0.5371867 0.5240954 0.5324175 0.5322879
## 1137 1138 1139 1140 1141 1142 1143 1144
## 0.5219523 0.5483767 0.5268796 0.5288086 0.5359359 0.5227080 0.5221915 0.5378127
## 1145 1146 1147 1148 1149 1150 1151 1152
## 0.5360140 0.5240494 0.5272447 0.5483266 0.5574668 0.5424229 0.5196106 0.5337884
## 1153 1154 1155 1156 1157 1158 1159 1160
## 0.5225350 0.5301958 0.5258271 0.5480381 0.5436486 0.5265972 0.5194657 0.5166158
## 1161 1162 1163 1164 1165 1166 1167 1168
## 0.5298550 0.5496361 0.5240216 0.5269980 0.5163351 0.5300618 0.5347421 0.5306196
## 1169 1170 1171 1172 1173 1174 1175 1176
## 0.5317611 0.5157188 0.5227221 0.5420646 0.5282723 0.5297885 0.5552796 0.5502475
## 1177 1178 1179 1180 1181 1182 1183 1184
## 0.5412309 0.5565057 0.5375656 0.5684072 0.5795859 0.5520448 0.5246195 0.5326840
## 1185 1186 1187 1188 1189 1190 1191 1192
## 0.5336880 0.5390441 0.5251032 0.5287862 0.5442739 0.5699425 0.5933956 0.5825836
## 1193 1194 1195 1196 1197 1198 1199 1200
## 0.5311230 0.5447316 0.5537287 0.5636082 0.5543141 0.5242266 0.5217321 0.5584029
## 1201 1202 1203 1204 1205 1206 1207 1208
## 0.5457889 0.5683503 0.5674201 0.5074017 0.5340958 0.5632313 0.5763697 0.5324472
## 1209 1210 1211 1212 1213 1214 1215 1216
## 0.5202811 0.5514015 0.5466790 0.5469814 0.5368595 0.5341844 0.5425346 0.5456582
## 1217 1218 1219 1220 1221 1222 1223 1224
## 0.5313672 0.5142526 0.5328266 0.5592255 0.5437027 0.5267398 0.5329257 0.5279180
## 1225 1226 1227 1228 1229 1230 1231 1232
## 0.5350592 0.5288983 0.4893357 0.5534329 0.5567250 0.5592505 0.5400175 0.5230599
## 1233 1234 1235 1236 1237 1238 1239 1240
## 0.5529143 0.5413269 0.5495887 0.5495068 0.5271413 0.5263293 0.5375299 0.5281697
## 1241 1242 1243 1244 1245 1246 1247 1248
## 0.5428337 0.5650049 0.5587216 0.5425798 0.5313870 0.5198924 0.5059256 0.5392683
## 1249 1250
## 0.5261183 0.5179166

```

```
contrasts(Direction)
```

```

##      Up
## Down 0
## Up   1

```

There for in order to make a prediction as whether the market will go up or down on a particular day, we must covert these predicted probabilities into class labels up or down. Done based on the value .5

```

glm.pred=rep("Down",1250)
glm.pred[glm.probs>.5]="Up"

```

```

table(glm.pred,Direction)

##          Direction
## glm.pred Down Up
##      Down 145 141
##      Up   457 507

mean(glm.pred == Direction)

## [1] 0.5216

train =(Year<2005)
Smarket.2005 = Smarket[!train,]
dim(Smarket.2005)

## [1] 252   9

Direction.2005 =Direction[!train]

glm.fits = glm(Direction~Lag1 + Lag2+ Lag3 + Lag4 + Lag5 + Volume, data = Smarket, family =binomial, sub
glm.probs=predict(glm.fits,Smarket.2005, type = "response")

glm.pred=rep("Down", 252)
glm.pred[glm.probs>.5]="Up"
table(glm.pred,Direction.2005)

##          Direction.2005
## glm.pred Down Up
##      Down 77 97
##      Up   34 44

mean(glm.pred!=Direction.2005)

## [1] 0.5198413

mean(glm.pred==Direction.2005)

## [1] 0.4801587

glm.fits = glm(Direction ~ Lag1 + Lag2 , data =Smarket, family =binomial, subset = train)
glm.probs = predict(glm.fits, Smarket.2005, type ="response")
glm.pred = rep("Down",252)
glm.pred[glm.probs > .5] = "Up"
table(glm.pred, Direction.2005)

##          Direction.2005
## glm.pred Down Up
##      Down 35 35
##      Up   76 106

```

```

mean(glm.pred==Direction.2005)

## [1] 0.5595238

predict(glm.fits, newdata = data.frame(Lag1 = c(1.2, 1.5), Lag2=c(1.1, -0.8)), type="response")

##           1          2
## 0.4791462 0.4960939

```

4.6.3 Linear Discriminant Analysis

```

library(MASS)
lda.fit = lda(Direction ~ Lag1 + Lag2, data = Smarket, subset = train)
lda.fit

```

```

## Call:
## lda(Direction ~ Lag1 + Lag2, data = Smarket, subset = train)
##
## Prior probabilities of groups:
##     Down      Up
## 0.491984 0.508016
##
## Group means:
##             Lag1      Lag2
## Down  0.04279022 0.03389409
## Up    -0.03954635 -0.03132544
##
## Coefficients of linear discriminants:
##             LD1
## Lag1 -0.6420190
## Lag2 -0.5135293

```

```

lda.pred = predict(lda.fit, Smarket.2005)
names(lda.pred)

```

```

## [1] "class"      "posterior"   "x"

```

```

lda.class = lda.pred$class
table(lda.class, Direction.2005)

```

```

##           Direction.2005
## lda.class Down Up
##     Down   35 35
##     Up     76 106

```

```

mean(lda.class==Direction.2005)

```

```

## [1] 0.5595238

```

Applying a 50% threshold to the posterior probabilities allows us to recreate the predictions.

```
sum(lda.pred$posterior[,1]>=.5)

## [1] 70

sum(lda.pred$posterior[,1]<.5)

## [1] 182

lda.pred$posterior[1:20,1]

##      999     1000     1001     1002     1003     1004     1005     1006
## 0.4901792 0.4792185 0.4668185 0.4740011 0.4927877 0.4938562 0.4951016 0.4872861
##      1007     1008     1009     1010     1011     1012     1013     1014
## 0.4907013 0.4844026 0.4906963 0.5119988 0.4895152 0.4706761 0.4744593 0.4799583
##      1015     1016     1017     1018
## 0.4935775 0.5030894 0.4978806 0.4886331

lda.class[1:20]

## [1] Up   Down Up   Up   Up
## [16] Up   Up   Down Up   Up
## Levels: Down Up

sum(lda.pred$posterior[,1]>.9)

## [1] 0
```

4.6.4 Quadratic Discriminant Analysis

```
qda.fit = qda(Direction ~ Lag1 + Lag2, data = Smarket, subset=train)
qda.fit

## Call:
## qda(Direction ~ Lag1 + Lag2, data = Smarket, subset = train)
##
## Prior probabilities of groups:
##      Down      Up
## 0.491984 0.508016
##
## Group means:
##           Lag1      Lag2
## Down  0.04279022 0.03389409
## Up   -0.03954635 -0.03132544
```

```
qda.class=predict(qda.fit, Smarket.2005)$class  
table(qda.class, Direction.2005)
```

```
##          Direction.2005  
## qda.class Down Up  
##       Down   30  20  
##       Up    81 121  
  
mean(qda.class == Direction.2005)  
  
## [1] 0.5992063
```

4.6.5 K - Nearest Neighbors

```
library(class)  
train.X=cbind(Lag1,Lag2)[train,]  
test.X=cbind(Lag1,Lag2)[!train,]  
train.Direction = Direction[train]  
  
dim(train.X)  
  
## [1] 998 2  
  
dim(test.X)  
  
## [1] 252 2  
  
set.seed(1)  
knn.pred = knn(train.X,test.X,train.Direction, k = 1)  
table(knn.pred , Direction.2005)  
  
##          Direction.2005  
## knn.pred Down Up  
##       Down   43 58  
##       Up    68 83  
  
mean(knn.pred == Direction.2005)  
  
## [1] 0.5
```

An application to Caravan Insurance Data

```
dim(Caravan)
```

```
## [1] 5822 86
```

```
attach(Caravan)
summary(Caravan)
```

##	MOSTYPE	MAANTHUI	MGEMOMV	MGEMLEEF	
##	Min. : 1.00	Min. : 1.000	Min. : 1.000	Min. : 1.000	
##	1st Qu.: 10.00	1st Qu.: 1.000	1st Qu.: 2.000	1st Qu.: 2.000	
##	Median : 30.00	Median : 1.000	Median : 3.000	Median : 3.000	
##	Mean : 24.25	Mean : 1.111	Mean : 2.679	Mean : 2.991	
##	3rd Qu.: 35.00	3rd Qu.: 1.000	3rd Qu.: 3.000	3rd Qu.: 3.000	
##	Max. : 41.00	Max. : 10.000	Max. : 5.000	Max. : 6.000	
##	MOSHOOFD	MGODRK	MGODPR	MGODOV	
##	Min. : 1.000	Min. : 0.0000	Min. : 0.000	Min. : 0.00	
##	1st Qu.: 3.000	1st Qu.: 0.0000	1st Qu.: 4.000	1st Qu.: 0.00	
##	Median : 7.000	Median : 0.0000	Median : 5.000	Median : 1.00	
##	Mean : 5.774	Mean : 0.6965	Mean : 4.627	Mean : 1.07	
##	3rd Qu.: 8.000	3rd Qu.: 1.0000	3rd Qu.: 6.000	3rd Qu.: 2.00	
##	Max. : 10.000	Max. : 9.0000	Max. : 9.000	Max. : 5.00	
##	MGODGE	MRELGE	MRELSA	MRELOV	
##	Min. : 0.000	Min. : 0.000	Min. : 0.0000	Min. : 0.00	
##	1st Qu.: 2.000	1st Qu.: 5.000	1st Qu.: 0.0000	1st Qu.: 1.00	
##	Median : 3.000	Median : 6.000	Median : 1.0000	Median : 2.00	
##	Mean : 3.259	Mean : 6.183	Mean : 0.8835	Mean : 2.29	
##	3rd Qu.: 4.000	3rd Qu.: 7.000	3rd Qu.: 1.0000	3rd Qu.: 3.00	
##	Max. : 9.000	Max. : 9.000	Max. : 7.0000	Max. : 9.00	
##	MFALLEEN	MFGEKIND	MFWEKIND	MOPLHOOG	MOPLMIDD
##	Min. : 0.000	Min. : 0.00	Min. : 0.0	Min. : 0.000	Min. : 0.000
##	1st Qu.: 0.000	1st Qu.: 2.00	1st Qu.: 3.0	1st Qu.: 0.000	1st Qu.: 2.000
##	Median : 2.000	Median : 3.00	Median : 4.0	Median : 1.000	Median : 3.000
##	Mean : 1.888	Mean : 3.23	Mean : 4.3	Mean : 1.461	Mean : 3.351
##	3rd Qu.: 3.000	3rd Qu.: 4.00	3rd Qu.: 6.0	3rd Qu.: 2.000	3rd Qu.: 4.000
##	Max. : 9.000	Max. : 9.00	Max. : 9.0	Max. : 9.000	Max. : 9.000
##	MOPLLAAG	MBERHOOG	MBERZELF	MBERBOER	
##	Min. : 0.000	Min. : 0.000	Min. : 0.000	Min. : 0.0000	
##	1st Qu.: 3.000	1st Qu.: 0.000	1st Qu.: 0.000	1st Qu.: 0.0000	
##	Median : 5.000	Median : 2.000	Median : 0.000	Median : 0.0000	
##	Mean : 4.572	Mean : 1.895	Mean : 0.398	Mean : 0.5223	
##	3rd Qu.: 6.000	3rd Qu.: 3.000	3rd Qu.: 1.000	3rd Qu.: 1.0000	
##	Max. : 9.000	Max. : 9.000	Max. : 5.000	Max. : 9.0000	
##	MBERMIDD	MBERARBG	MBERARBO	MSKA	MSKB1
##	Min. : 0.000	Min. : 0.00	Min. : 0.000	Min. : 0.000	Min. : 0.000
##	1st Qu.: 2.000	1st Qu.: 1.00	1st Qu.: 1.000	1st Qu.: 0.000	1st Qu.: 1.000
##	Median : 3.000	Median : 2.00	Median : 2.000	Median : 1.000	Median : 2.000
##	Mean : 2.899	Mean : 2.22	Mean : 2.306	Mean : 1.621	Mean : 1.607
##	3rd Qu.: 4.000	3rd Qu.: 3.00	3rd Qu.: 3.000	3rd Qu.: 2.000	3rd Qu.: 2.000
##	Max. : 9.000	Max. : 9.00	Max. : 9.000	Max. : 9.000	Max. : 9.000
##	MSKB2	MSKC	MSKD	MHHUUR	
##	Min. : 0.000	Min. : 0.000	Min. : 0.000	Min. : 0.000	
##	1st Qu.: 1.000	1st Qu.: 2.000	1st Qu.: 0.000	1st Qu.: 2.000	
##	Median : 2.000	Median : 4.000	Median : 1.000	Median : 4.000	
##	Mean : 2.203	Mean : 3.759	Mean : 1.067	Mean : 4.237	
##	3rd Qu.: 3.000	3rd Qu.: 5.000	3rd Qu.: 2.000	3rd Qu.: 7.000	
##	Max. : 9.000	Max. : 9.000	Max. : 9.000	Max. : 9.000	
##	MHKOOP	MAUT1	MAUT2	MAUTO	MZFONDS

```

## Min. :0.000 Min. :0.00 Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.:2.000 1st Qu.:5.00 1st Qu.:0.000 1st Qu.:1.000 1st Qu.:5.000
## Median :5.000 Median :6.00 Median :1.000 Median :2.000 Median :7.000
## Mean :4.772 Mean :6.04 Mean :1.316 Mean :1.959 Mean :6.277
## 3rd Qu.:7.000 3rd Qu.:7.00 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:8.000
## Max. :9.000 Max. :9.00 Max. :7.000 Max. :9.000 Max. :9.000
## MZPART      MINKM30      MINK3045      MINK4575
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:2.000 1st Qu.:1.000
## Median :2.000 Median :2.000 Median :4.000 Median :3.000
## Mean :2.729 Mean :2.574 Mean :3.536 Mean :2.731
## 3rd Qu.:4.000 3rd Qu.:4.000 3rd Qu.:5.000 3rd Qu.:4.000
## Max. :9.000 Max. :9.000 Max. :9.000 Max. :9.000
## MINK7512     MINK123M     MINKGEM      MKOOPKLA
## Min. :0.0000 Min. :0.0000 Min. :0.000 Min. :1.000
## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:3.000 1st Qu.:3.000
## Median :0.0000 Median :0.0000 Median :4.000 Median :4.000
## Mean :0.7961 Mean :0.2027 Mean :3.784 Mean :4.236
## 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:4.000 3rd Qu.:6.000
## Max. :9.0000 Max. :9.0000 Max. :9.000 Max. :8.000
## PWAPART      PWABEDR      PWALAND     PPERSAUT
## Min. :0.0000 Min. :0.00000 Min. :0.00000 Min. :0.00
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00
## Median :0.0000 Median :0.00000 Median :0.00000 Median :5.00
## Mean :0.7712 Mean :0.04002 Mean :0.07162 Mean :2.97
## 3rd Qu.:2.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:6.00
## Max. :3.0000 Max. :6.00000 Max. :4.00000 Max. :8.00
## PBESAUT     PMOTSCO      PVRAAUT     PAANHANG
## Min. :0.00000 Min. :0.0000 Min. :0.000000 Min. :0.00000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.000000 1st Qu.:0.00000
## Median :0.00000 Median :0.0000 Median :0.000000 Median :0.00000
## Mean :0.04827 Mean :0.1754 Mean :0.009447 Mean :0.02096
## 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.000000 3rd Qu.:0.00000
## Max. :7.00000 Max. :7.0000 Max. :9.000000 Max. :5.00000
## PTRACTOR     PWERKT      PBROM       PLEVEN
## Min. :0.00000 Min. :0.0000 Min. :0.000 Min. :0.0000
## 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000
## Median :0.00000 Median :0.0000 Median :0.000 Median :0.0000
## Mean :0.09258 Mean :0.01305 Mean :0.215 Mean :0.1948
## 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:0.000 3rd Qu.:0.0000
## Max. :6.00000 Max. :6.00000 Max. :6.000 Max. :9.0000
## PPERSONG     PGEZONG      PWAOREG     PBRAND
## Min. :0.00000 Min. :0.00000 Min. :0.00000 Min. :0.000
## 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000
## Median :0.00000 Median :0.00000 Median :0.00000 Median :2.000
## Mean :0.01374 Mean :0.01529 Mean :0.02353 Mean :1.828
## 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:4.000
## Max. :6.00000 Max. :3.00000 Max. :7.00000 Max. :8.000
## PZEILPL      PPLEZIER     PFIETS      PINBOED
## Min. :0.0000000 Min. :0.00000 Min. :0.00000 Min. :0.000000
## 1st Qu.:0.0000000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.000000
## Median :0.0000000 Median :0.00000 Median :0.00000 Median :0.000000
## Mean :0.0008588 Mean :0.01889 Mean :0.02525 Mean :0.01563
## 3rd Qu.:0.0000000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.000000

```

```

##  Max.   :3.0000000  Max.   :6.00000  Max.   :1.00000  Max.   :6.00000
##  PBYSTAND      AWAPART      AWABEDR      AWALAND
##  Min.   :0.00000  Min.   :0.000  Min.   :0.00000  Min.   :0.00000
##  1st Qu.:0.00000 1st Qu.:0.000 1st Qu.:0.00000 1st Qu.:0.00000
##  Median :0.00000  Median :0.000  Median :0.00000  Median :0.00000
##  Mean   :0.04758  Mean   :0.403  Mean   :0.01477  Mean   :0.02061
##  3rd Qu.:0.00000 3rd Qu.:1.000 3rd Qu.:0.00000 3rd Qu.:0.00000
##  Max.   :5.00000  Max.   :2.000  Max.   :5.00000  Max.   :1.00000
##  APERSAUT      ABESAUT      AMOTSCO      AVRAAUT
##  Min.   :0.0000  Min.   :0.00000  Min.   :0.00000  Min.   :0.0000000
##  1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.0000000
##  Median :1.0000  Median :0.00000  Median :0.00000  Median :0.0000000
##  Mean   :0.5622  Mean   :0.01048  Mean   :0.04105  Mean   :0.002233
##  3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000000
##  Max.   :7.0000  Max.   :4.00000  Max.   :8.00000  Max.   :3.00000
##  AAANHANG      ATRACTOR     AWERKT      ABROM
##  Min.   :0.0000  Min.   :0.00000  Min.   :0.00000  Min.   :0.00000
##  1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.00000
##  Median :0.0000  Median :0.00000  Median :0.00000  Median :0.00000
##  Mean   :0.01254 Mean   :0.03367 Mean   :0.006183 Mean   :0.07042
##  3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.00000
##  Max.   :3.0000  Max.   :4.00000  Max.   :6.00000  Max.   :2.00000
##  ALEVEN        APERSONG     AGEZONG     AWAOREG
##  Min.   :0.0000  Min.   :0.00000  Min.   :0.00000  Min.   :0.0000000
##  1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:0.00000 1st Qu.:0.0000000
##  Median :0.0000  Median :0.00000  Median :0.00000  Median :0.0000000
##  Mean   :0.07661 Mean   :0.005325 Mean   :0.006527 Mean   :0.004638
##  3rd Qu.:0.0000 3rd Qu.:0.00000 3rd Qu.:0.00000 3rd Qu.:0.0000000
##  Max.   :8.0000  Max.   :1.00000  Max.   :1.00000  Max.   :2.00000
##  ABRAND        AZEILPL      APLEZIER     AFIETS
##  Min.   :0.0000  Min.   :0.0000000  Min.   :0.00000  Min.   :0.00000
##  1st Qu.:0.0000 1st Qu.:0.0000000 1st Qu.:0.00000 1st Qu.:0.00000
##  Median :1.0000  Median :0.0000000  Median :0.00000  Median :0.00000
##  Mean   :0.5701  Mean   :0.0005153 Mean   :0.006012 Mean   :0.03178
##  3rd Qu.:1.0000 3rd Qu.:0.0000000 3rd Qu.:0.00000 3rd Qu.:0.00000
##  Max.   :7.0000  Max.   :1.0000000 Max.   :2.00000  Max.   :3.00000
##  AINBOED      ABYSTAND     Purchase
##  Min.   :0.00000  Min.   :0.00000  No :5474
##  1st Qu.:0.00000 1st Qu.:0.00000  Yes: 348
##  Median :0.00000  Median :0.00000
##  Mean   :0.007901 Mean   :0.01426
##  3rd Qu.:0.00000 3rd Qu.:0.00000
##  Max.   :2.00000  Max.   :2.00000

```

```

standardized.X= scale(Caravan[,-86])
var(Caravan[,1])

```

```

## [1] 165.0378

```

```

var(Caravan[,2])

```

```

## [1] 0.1647078

```

```
var(standardized.X[,1])  
  
## [1] 1  
  
var(standardized.X[,2])  
  
## [1] 1  
  
test = 1:1000  
train.X=standardized.X[-test,]  
test.X = standardized.X[test,]  
train.Y = Purchase[-test]  
test.Y = Purchase[test]  
set.seed(1)  
knn.pred = knn(train.X,test.X,train.Y,k=1)  
mean(test.Y != knn.pred)  
  
## [1] 0.118  
  
mean(test.Y != "No")  
  
## [1] 0.059
```