Multiple Testing & Cross Validation

```
data = read.csv("C:/Users/rohit/Downloads/semiconductor.csv")
```

```
head(data)
```

	F… <int>	SIG1 <dbl>	SIG2 <dbl>	SIG3 <dbl>	SIG4 <dbl>	SIG5 <dbl>	SIG6 <dbl>	SIG7 <dbl>
1	0	2.2237790	-0.3490181	-1.985797	-2.6009934	0.4510919	0.8566255	0.01999866
2	1	-0.1303779	-0.3813179	-1.285172	-0.7663017	-3.0983019	1.2578523	0.79225764
3	0	-0.8711378	-0.8253003	-1.615008	-0.4488032	-2.5707819	-1.3457985	3.56110704
4	0	-0.8773409	-1.9055982	-2.306114	-1.2107612	-4.0374708	-2.2236083	2.40412779
5	0	-2.1134556	-2.3138448	-2.917153	-3.3334670	-4.4393457	3.0296318	-0.17712967
6	0	1.8048809	-4.5603800	-2.506963	-2.6146047	-1.0183995	-1.1815782	2.91109363

6 rows | 1-10 of 202 columns

```
sum(is.na(data))
```

```
## [1] 0
```

```
summary(data)
```

```
##       FAIL             SIG1               SIG2               SIG3         
##  Min.   :0.0000   Min.   :-101.9871   Min.   :-33.7081   Min.   :-120.5202  
##  1st Qu.:0.0000   1st Qu.:  -0.4486   1st Qu.: -0.5770   1st Qu.:  -1.1517  
##  Median :0.0000   Median :   0.4808   Median :  0.6458   Median :   0.1727  
##  Mean   :0.0677   Mean   :   0.0000   Mean   :  0.0000   Mean   :   0.0000  
##  3rd Qu.:0.0000   3rd Qu.:   1.1676   3rd Qu.:  1.6644   3rd Qu.:   1.2218  
##  Max.   :1.0000   Max.   :   3.8389   Max.   :  5.5605   Max.   :  13.1233  
##       SIG4               SIG5               SIG6               SIG7         
##  Min.   :-17.0922   Min.   :-22.5265   Min.   :-29.0353   Min.   :-29.5962  
##  1st Qu.: -2.2960   1st Qu.: -1.9105   1st Qu.: -1.3330   1st Qu.: -1.7233  
##  Median : -0.1184   Median :  0.1515   Median :  0.1135   Median : -0.0172  
##  Mean   :  0.0000   Mean   :  0.0000   Mean   :  0.0000   Mean   :  0.0000  
##  3rd Qu.:  2.0168   3rd Qu.:  2.0154   3rd Qu.:  1.5496   3rd Qu.:  1.5746  
##  Max.   : 54.1353   Max.   : 54.7029   Max.   : 52.2178   Max.   : 40.9311  
##       SIG8               SIG9              SIG10               SIG11         
##  Min.   :-21.4101   Min.   :-21.5492   Min.   :-12.19927   Min.   :-37.3446  
##  1st Qu.: -1.6366   1st Qu.: -1.6936   1st Qu.: -1.02288   1st Qu.: -0.9082  
##  Median :  0.2617   Median : -0.2023   Median : -0.01815   Median :  0.1906  
##  Mean   :  0.0000   Mean   :  0.0000   Mean   :  0.00000   Mean   :  0.0000  
##  3rd Qu.:  1.9111   3rd Qu.:  1.8282   3rd Qu.:  1.02005   3rd Qu.:  1.2590  
##  Max.   :  9.6902   Max.   : 13.2895   Max.   : 60.32429   Max.   :  6.5503  
##      SIG12              SIG13              SIG14              SIG15         
##  Min.   :-17.8863   Min.   :-18.736   Min.   :-11.0771   Min.   :-12.1857  
##  1st Qu.: -0.6504   1st Qu.: -1.246   1st Qu.: -1.4796   1st Qu.: -1.4519  
##  Median :  0.3806   Median :  0.162   Median :  0.1417   Median : -0.0147  
##  Mean   :  0.0000   Mean   :  0.000   Mean   :  0.0000   Mean   :  0.0000  
##  3rd Qu.:  1.1972   3rd Qu.:  1.451   3rd Qu.:  1.5661   3rd Qu.:  1.4925  
##  Max.   : 19.7004   Max.   :  6.723   Max.   :  8.2992   Max.   : 12.1163  
##      SIG16              SIG17              SIG18              SIG19         
##  Min.   :-22.38819   Min.   :-9.64298   Min.   :-5.60876   Min.   :-22.7792  
##  1st Qu.: -1.34727   1st Qu.:-1.33235   1st Qu.:-1.19694   1st Qu.: -1.0635  
##  Median : -0.02447   Median :-0.03234   Median :-0.08389   Median :  0.1557  
##  Mean   :  0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   :  0.0000  
##  3rd Qu.:  1.36541   3rd Qu.: 1.29055   3rd Qu.: 0.98628   3rd Qu.:  1.2341  
##  Max.   : 11.51787   Max.   :13.05048   Max.   :27.79395   Max.   :  9.9419  
##      SIG20              SIG21              SIG22              SIG23         
##  Min.   :-19.84091   Min.   :-13.0182   Min.   :-12.80742   Min.   :-13.33159  
##  1st Qu.: -0.73161   1st Qu.: -1.2607   1st Qu.: -1.19684   1st Qu.: -1.08998  
##  Median :  0.05093   Median : -0.1542   Median : -0.02671   Median : -0.05439  
##  Mean   :  0.00000   Mean   :  0.0000   Mean   :  0.00000   Mean   :  0.00000  
##  3rd Qu.:  0.87791   3rd Qu.:  1.1401   3rd Qu.:  1.03493   3rd Qu.:  1.11679  
##  Max.   : 50.22213   Max.   : 11.1117   Max.   : 27.92576   Max.   :  9.96918  
##      SIG24              SIG25              SIG26              SIG27         
##  Min.   :-10.94792   Min.   :-20.1552   Min.   :-15.2488   Min.   :-21.226  
##  1st Qu.: -1.27199   1st Qu.: -0.9459   1st Qu.: -0.9715   1st Qu.: -1.006  
##  Median :  0.06046   Median :  0.1307   Median :  0.1170   Median :  0.096  
##  Mean   :  0.00000   Mean   :  0.0000   Mean   :  0.0000   Mean   :  0.000  
##  3rd Qu.:  1.28015   3rd Qu.:  1.1017   3rd Qu.:  1.0872   3rd Qu.:  1.212  
##  Max.   :  8.83394   Max.   :  9.9681   Max.   : 12.4476   Max.   :  8.842  
##      SIG28              SIG29              SIG30              SIG31         
##  Min.   :-9.50186   Min.   :-11.0019   Min.   :-8.83968   Min.   :-9.85075  
##  1st Qu.:-1.14795   1st Qu.: -1.0859   1st Qu.:-1.04928   1st Qu.:-1.09889  
##  Median :-0.01395   Median :  0.1007   Median :-0.07677   Median :-0.01263  
##  Mean   : 0.00000   Mean   :  0.0000   Mean   : 0.00000   Mean   : 0.00000  
##  3rd Qu.: 1.19803   3rd Qu.:  1.1021   3rd Qu.: 0.91281   3rd Qu.: 1.11980
```

```
##  Max.   : 7.75489   Max.   : 13.6621   Max.   :20.81753   Max.   : 9.43231
##      SIG32              SIG33              SIG34               SIG35
##  Min.   :-13.94302   Min.   :-8.13543   Min.   :-20.822911   Min.   :-16.9039
##  1st Qu.: -1.10051   1st Qu.:-0.98473   1st Qu.: -1.006064   1st Qu.: -0.9239
##  Median :  0.03582   Median : 0.04532   Median : -0.009319   Median :  0.1264
##  Mean   :  0.00000   Mean   : 0.00000   Mean   :  0.000000   Mean   :  0.0000
##  3rd Qu.:  1.08339   3rd Qu.: 1.10554   3rd Qu.:  0.967755   3rd Qu.:  1.0063
##  Max.   : 13.90922   Max.   : 9.31362   Max.   :  7.767196   Max.   : 13.5106
##      SIG36              SIG37              SIG38               SIG39
##  Min.   :-14.11753   Min.   :-7.35351   Min.   :-10.86076   Min.   :-14.14282
##  1st Qu.: -1.12415   1st Qu.:-1.08621   1st Qu.: -1.05106   1st Qu.: -1.02701
##  Median : -0.06756   Median :-0.04525   Median :  0.01502   Median : -0.04097
##  Mean   :  0.00000   Mean   : 0.00000   Mean   :  0.00000   Mean   :  0.00000
##  3rd Qu.:  1.09721   3rd Qu.: 1.10117   3rd Qu.:  1.06274   3rd Qu.:  1.01537
##  Max.   :  6.82505   Max.   : 6.22578   Max.   :  8.25255   Max.   : 11.27878
##      SIG40             SIG41              SIG42               SIG43
##  Min.   :-9.3412   Min.   :-6.269998   Min.   :-9.171595   Min.   :-8.25963
##  1st Qu.:-1.0877   1st Qu.:-1.036709   1st Qu.:-1.030773   1st Qu.:-0.83090
##  Median : 0.1056   Median :-0.003993   Median :-0.007489   Median : 0.07565
##  Mean   : 0.0000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 1.1257   3rd Qu.: 1.002898   3rd Qu.: 1.014953   3rd Qu.: 0.99546
##  Max.   : 5.8261   Max.   :20.634600   Max.   : 5.740488   Max.   : 6.98362
##      SIG44              SIG45              SIG46               SIG47
##  Min.   :-15.03661   Min.   :-7.0674   Min.   :-8.765497   Min.   :-11.24369
##  1st Qu.: -0.84051   1st Qu.:-1.0157   1st Qu.:-0.939142   1st Qu.: -0.81610
##  Median :  0.03629   Median :-0.0347   Median : 0.000336   Median :  0.05304
##  Mean   :  0.00000   Mean   : 0.0000   Mean   : 0.000000   Mean   :  0.00000
##  3rd Qu.:  0.92879   3rd Qu.: 1.0387   3rd Qu.: 0.991525   3rd Qu.:  0.93659
##  Max.   :  7.95802   Max.   : 6.7862   Max.   :11.620280   Max.   :  7.18540
##      SIG48              SIG49              SIG50               SIG51
##  Min.   :-5.78420   Min.   :-5.07692   Min.   :-9.58459   Min.   :-10.54550
##  1st Qu.:-1.03537   1st Qu.:-1.03773   1st Qu.:-0.96622   1st Qu.: -0.95093
##  Median :-0.01418   Median :-0.02534   Median :-0.01458   Median : -0.08296
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   :  0.00000
##  3rd Qu.: 1.02709   3rd Qu.: 0.96804   3rd Qu.: 0.97651   3rd Qu.:  0.85119
##  Max.   : 6.39814   Max.   : 9.59684   Max.   : 7.15806   Max.   :  8.08939
##      SIG52              SIG53              SIG54               SIG55
##  Min.   :-5.89025   Min.   :-8.99970   Min.   :-5.94708   Min.   :-7.24906
##  1st Qu.:-0.98536   1st Qu.:-0.92555   1st Qu.:-0.88479   1st Qu.:-0.91480
##  Median :-0.07347   Median : 0.04963   Median :-0.02201   Median :-0.04218
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.92435   3rd Qu.: 0.91180   3rd Qu.: 0.79743   3rd Qu.: 0.87248
##  Max.   : 5.28650   Max.   : 7.05085   Max.   :15.51312   Max.   :11.51112
##      SIG56              SIG57              SIG58               SIG59
##  Min.   :-6.89741   Min.   :-8.30504   Min.   :-4.80839   Min.   :-7.42766
##  1st Qu.:-0.91746   1st Qu.:-0.84952   1st Qu.:-0.90415   1st Qu.:-0.86140
##  Median : 0.04171   Median : 0.02319   Median :-0.01386   Median : 0.02136
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.97637   3rd Qu.: 0.87341   3rd Qu.: 0.89166   3rd Qu.: 0.91770
##  Max.   : 6.32762   Max.   : 8.97124   Max.   : 7.53049   Max.   : 5.46670
##      SIG60              SIG61              SIG62               SIG63
##  Min.   :-7.04094   Min.   :-5.16292   Min.   :-13.61877   Min.   :-6.59393
##  1st Qu.:-0.89744   1st Qu.:-0.94218   1st Qu.: -0.73640   1st Qu.:-0.85018
##  Median :-0.04198   Median : 0.02035   Median : -0.01674   Median :-0.04689
##  Mean   : 0.00000   Mean   : 0.00000   Mean   :  0.00000   Mean   : 0.00000
##  3rd Qu.: 0.91878   3rd Qu.: 0.93130   3rd Qu.:  0.80603   3rd Qu.: 0.78999
```

```
##  Max.   : 6.77590   Max.   : 5.95179   Max.   :  8.66847   Max.   : 6.43623
##      SIG64              SIG65              SIG66              SIG67
##  Min.   :-7.074858   Min.   :-5.391176   Min.   :-4.91779   Min.   :-5.16157
##  1st Qu.:-0.815889   1st Qu.:-0.859629   1st Qu.:-0.83545   1st Qu.:-0.80763
##  Median : 0.007719   Median :-0.007756   Median :-0.01574   Median :-0.03437
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.867190   3rd Qu.: 0.832067   3rd Qu.: 0.86074   3rd Qu.: 0.80592
##  Max.   : 5.613453   Max.   : 5.434328   Max.   : 4.62565   Max.   :11.42766
##      SIG68              SIG69              SIG70              SIG71
##  Min.   :-5.96883   Min.   :-4.27904   Min.   :-5.12622   Min.   :-5.35708
##  1st Qu.:-0.82003   1st Qu.:-0.84790   1st Qu.:-0.73062   1st Qu.:-0.76798
##  Median :-0.03903   Median :-0.01783   Median : 0.02967   Median : 0.02115
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.81946   3rd Qu.: 0.80745   3rd Qu.: 0.76570   3rd Qu.: 0.73659
##  Max.   : 4.59696   Max.   : 7.69613   Max.   :10.55877   Max.   : 7.47180
##      SIG72              SIG73              SIG74              SIG75
##  Min.   :-7.25623   Min.   :-5.98037   Min.   :-7.78315   Min.   :-4.303276
##  1st Qu.:-0.78992   1st Qu.:-0.74707   1st Qu.:-0.76359   1st Qu.:-0.790406
##  Median : 0.01999   Median : 0.02266   Median : 0.03115   Median :-0.002385
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.81480   3rd Qu.: 0.75729   3rd Qu.: 0.76746   3rd Qu.: 0.760592
##  Max.   : 3.44487   Max.   : 9.01094   Max.   : 5.26531   Max.   : 5.135144
##      SIG76              SIG77              SIG78              SIG79
##  Min.   :-8.176320   Min.   :-6.13441   Min.   :-3.62776   Min.   :-4.561328
##  1st Qu.:-0.700350   1st Qu.:-0.67967   1st Qu.:-0.71637   1st Qu.:-0.691591
##  Median : 0.008296   Median :-0.01598   Median :-0.01907   Median : 0.002108
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.747855   3rd Qu.: 0.69081   3rd Qu.: 0.68203   3rd Qu.: 0.679854
##  Max.   : 4.478561   Max.   :12.81258   Max.   : 5.16012   Max.   :10.930587
##      SIG80              SIG81              SIG82              SIG83
##  Min.   :-3.255948   Min.   :-8.67979   Min.   :-3.578811   Min.   :-4.32434
##  1st Qu.:-0.722641   1st Qu.:-0.68349   1st Qu.:-0.673601   1st Qu.:-0.69173
##  Median :-0.001182   Median :-0.03447   Median : 0.007961   Median :-0.03149
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.694637   3rd Qu.: 0.67808   3rd Qu.: 0.699824   3rd Qu.: 0.66787
##  Max.   : 4.484155   Max.   : 4.42965   Max.   : 4.216525   Max.   : 3.44339
##      SIG84              SIG85              SIG86              SIG87
##  Min.   :-4.49480   Min.   :-5.87885   Min.   :-5.820985   Min.   :-3.80180
##  1st Qu.:-0.67301   1st Qu.:-0.71693   1st Qu.:-0.655358   1st Qu.:-0.65368
##  Median : 0.01004   Median : 0.01275   Median :-0.006101   Median :-0.01377
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.65652   3rd Qu.: 0.68401   3rd Qu.: 0.717114   3rd Qu.: 0.67991
##  Max.   : 6.13005   Max.   : 6.41334   Max.   : 3.638504   Max.   : 4.40402
##      SIG88              SIG89              SIG90              SIG91
##  Min.   :-6.842007   Min.   :-3.92223   Min.   :-6.29531   Min.   :-3.674766
##  1st Qu.:-0.636826   1st Qu.:-0.64860   1st Qu.:-0.64690   1st Qu.:-0.647105
##  Median :-0.009184   Median : 0.02623   Median :-0.01269   Median : 0.001969
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.663308   3rd Qu.: 0.64064   3rd Qu.: 0.65213   3rd Qu.: 0.667219
##  Max.   : 6.629338   Max.   : 4.41320   Max.   : 3.43355   Max.   : 6.148528
##      SIG92              SIG93              SIG94              SIG95
##  Min.   :-2.82532   Min.   :-3.618097   Min.   :-3.590730   Min.   :-3.533637
##  1st Qu.:-0.63628   1st Qu.:-0.591245   1st Qu.:-0.651229   1st Qu.:-0.639081
##  Median : 0.01779   Median :-0.000029   Median :-0.002419   Median : 0.006503
##  Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.63875   3rd Qu.: 0.624283   3rd Qu.: 0.643957   3rd Qu.: 0.619388
```

```
##  Max.   : 4.54840   Max.   : 4.404144   Max.   : 6.575099   Max.   : 3.389839
##      SIG96              SIG97              SIG98              SIG99
##  Min.   :-4.08914   Min.   :-8.562327   Min.   :-4.99488   Min.   :-4.26849
##  1st Qu.:-0.64829   1st Qu.:-0.604842   1st Qu.:-0.60111   1st Qu.:-0.66174
##  Median : 0.02979   Median :-0.006103   Median :-0.01491   Median :-0.01941
##  Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.62250   3rd Qu.: 0.588882   3rd Qu.: 0.61523   3rd Qu.: 0.64033
##  Max.   : 3.30692   Max.   : 4.026245   Max.   : 3.24742   Max.   : 4.10811
##      SIG100             SIG101             SIG102             SIG103
##  Min.   :-3.724956   Min.   :-4.07592   Min.   :-5.98295   Min.   :-3.92634
##  1st Qu.:-0.610324   1st Qu.:-0.58297   1st Qu.:-0.59055   1st Qu.:-0.56798
##  Median :-0.006575   Median : 0.01684   Median : 0.01445   Median :-0.03009
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.606321   3rd Qu.: 0.60866   3rd Qu.: 0.58356   3rd Qu.: 0.61837
##  Max.   : 3.263935   Max.   : 3.21367   Max.   : 4.11534   Max.   : 3.44913
##      SIG104             SIG105             SIG106             SIG107
##  Min.   :-2.522288   Min.   :-6.190204   Min.   :-3.57915   Min.   :-4.40136
##  1st Qu.:-0.580911   1st Qu.:-0.625148   1st Qu.:-0.57287   1st Qu.:-0.56873
##  Median :-0.002154   Median : 0.007258   Median : 0.02847   Median :-0.01599
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.581832   3rd Qu.: 0.618050   3rd Qu.: 0.55472   3rd Qu.: 0.57065
##  Max.   : 3.994206   Max.   : 2.719232   Max.   : 3.53131   Max.   : 4.96022
##      SIG108             SIG109             SIG110             SIG111
##  Min.   :-4.60239   Min.   :-4.49481   Min.   :-4.14125   Min.   :-3.70955
##  1st Qu.:-0.55804   1st Qu.:-0.57173   1st Qu.:-0.56249   1st Qu.:-0.54644
##  Median :-0.02313   Median :-0.02831   Median : 0.02217   Median :-0.01207
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.55852   3rd Qu.: 0.54563   3rd Qu.: 0.54807   3rd Qu.: 0.56933
##  Max.   : 3.24992   Max.   : 5.86196   Max.   : 3.56868   Max.   : 2.91407
##      SIG112             SIG113             SIG114
##  Min.   :-3.090543   Min.   :-3.880808   Min.   :-2.871864
##  1st Qu.:-0.550602   1st Qu.:-0.547696   1st Qu.:-0.563736
##  Median :-0.002473   Median : 0.006092   Median : 0.003679
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.523325   3rd Qu.: 0.569969   3rd Qu.: 0.520873
##  Max.   : 2.525851   Max.   : 2.988446   Max.   : 2.979917
##      SIG115             SIG116             SIG117             SIG118
##  Min.   :-4.703784   Min.   :-3.60979   Min.   :-3.23877   Min.   :-6.64385
##  1st Qu.:-0.500679   1st Qu.:-0.50021   1st Qu.:-0.52588   1st Qu.:-0.47829
##  Median :-0.000438   Median :-0.01715   Median :-0.02887   Median :-0.04138
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.501102   3rd Qu.: 0.50523   3rd Qu.: 0.50948   3rd Qu.: 0.45626
##  Max.   : 4.774178   Max.   : 2.86306   Max.   : 3.59635   Max.   : 7.83149
##      SIG119             SIG120             SIG121             SIG122
##  Min.   :-2.850314   Min.   :-7.206683   Min.   :-5.895363   Min.   :-7.75879
##  1st Qu.:-0.528997   1st Qu.:-0.498230   1st Qu.:-0.519440   1st Qu.:-0.45029
##  Median :-0.005412   Median : 0.002604   Median :-0.000604   Median : 0.01433
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.494769   3rd Qu.: 0.503689   3rd Qu.: 0.489914   3rd Qu.: 0.49091
##  Max.   : 4.393550   Max.   : 3.182156   Max.   : 4.429347   Max.   : 6.52183
##      SIG123             SIG124             SIG125
##  Min.   :-4.354327   Min.   :-4.191655   Min.   :-2.470692
##  1st Qu.:-0.496143   1st Qu.:-0.514498   1st Qu.:-0.534924
##  Median :-0.000755   Median : 0.000294   Median : 0.000737
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.511315   3rd Qu.: 0.502951   3rd Qu.: 0.503479
```

```
##  Max.   : 2.914657   Max.   : 2.502614   Max.   : 3.630760
##      SIG126              SIG127              SIG128              SIG129
##  Min.   :-2.535163   Min.   :-3.267873   Min.   :-3.44750   Min.   :-3.85773
##  1st Qu.:-0.511448   1st Qu.:-0.468936   1st Qu.:-0.45908   1st Qu.:-0.45684
##  Median :-0.009111   Median : 0.009787   Median : 0.03757   Median :-0.02526
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.481334   3rd Qu.: 0.452768   3rd Qu.: 0.47801   3rd Qu.: 0.48666
##  Max.   : 3.606985   Max.   : 3.257084   Max.   : 3.12571   Max.   : 3.43947
##      SIG130              SIG131              SIG132              SIG133
##  Min.   :-2.544242   Min.   :-4.429391   Min.   :-2.43988   Min.   :-3.60047
##  1st Qu.:-0.468480   1st Qu.:-0.449746   1st Qu.:-0.45728   1st Qu.:-0.41255
##  Median : 0.006637   Median : 0.005748   Median :-0.02273   Median : 0.02499
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.463451   3rd Qu.: 0.443822   3rd Qu.: 0.46902   3rd Qu.: 0.46374
##  Max.   : 4.758497   Max.   : 3.395100   Max.   : 2.51004   Max.   : 2.77262
##      SIG134              SIG135              SIG136              SIG137
##  Min.   :-2.73904   Min.   :-2.33344   Min.   :-3.62048   Min.   :-2.9084434
##  1st Qu.:-0.45712   1st Qu.:-0.44568   1st Qu.:-0.44901   1st Qu.:-0.4016432
##  Median : 0.01026   Median :-0.02563   Median : 0.02238   Median : 0.0003719
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.0000000
##  3rd Qu.: 0.43393   3rd Qu.: 0.47043   3rd Qu.: 0.46284   3rd Qu.: 0.4157009
##  Max.   : 3.01331   Max.   : 3.98639   Max.   : 2.78841   Max.   : 2.8361914
##      SIG138              SIG139              SIG140              SIG141
##  Min.   :-3.361862   Min.   :-3.081770   Min.   :-3.121260   Min.   :-2.65085
##  1st Qu.:-0.412055   1st Qu.:-0.418268   1st Qu.:-0.406682   1st Qu.:-0.43294
##  Median : 0.009899   Median : 0.009435   Median :-0.004945   Median :-0.01291
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.447420   3rd Qu.: 0.447397   3rd Qu.: 0.419751   3rd Qu.: 0.42131
##  Max.   : 3.591805   Max.   : 2.725717   Max.   : 4.053833   Max.   : 2.28415
##      SIG142              SIG143              SIG144              SIG145
##  Min.   :-1.935480   Min.   :-3.93812   Min.   :-3.43894   Min.   :-6.22517
##  1st Qu.:-0.422757   1st Qu.:-0.37052   1st Qu.:-0.34331   1st Qu.:-0.35072
##  Median :-0.005777   Median :-0.02555   Median : 0.01001   Median :-0.01431
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.397312   3rd Qu.: 0.36987   3rd Qu.: 0.37622   3rd Qu.: 0.33144
##  Max.   : 3.829648   Max.   : 4.18459   Max.   : 2.97530   Max.   : 4.22344
##      SIG146              SIG147              SIG148              SIG149
##  Min.   :-1.70417   Min.   :-2.42690   Min.   :-2.076808   Min.   :-2.10344
##  1st Qu.:-0.41601   1st Qu.:-0.38138   1st Qu.:-0.358942   1st Qu.:-0.35091
##  Median :-0.01903   Median :-0.01903   Median : 0.002168   Median : 0.00853
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.36210   3rd Qu.: 0.37709   3rd Qu.: 0.374784   3rd Qu.: 0.37519
##  Max.   : 2.67147   Max.   : 2.08890   Max.   : 3.136832   Max.   : 2.61507
##      SIG150              SIG151              SIG152              SIG153
##  Min.   :-4.079790   Min.   :-4.13845   Min.   :-1.83219   Min.   :-2.37002
##  1st Qu.:-0.323292   1st Qu.:-0.35398   1st Qu.:-0.35619   1st Qu.:-0.30925
##  Median :-0.007923   Median :-0.00366   Median :-0.02447   Median :-0.01199
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.334002   3rd Qu.: 0.34440   3rd Qu.: 0.29346   3rd Qu.: 0.31476
##  Max.   : 4.783899   Max.   : 2.71800   Max.   : 4.63668   Max.   : 3.13480
##      SIG154              SIG155              SIG156              SIG157
##  Min.   :-2.79400   Min.   :-2.7712928   Min.   :-3.70068   Min.   :-4.543349
##  1st Qu.:-0.31379   1st Qu.:-0.3329795   1st Qu.:-0.27869   1st Qu.:-0.258250
##  Median : 0.00735   Median :-0.0009139   Median : 0.01068   Median :-0.002628
##  Mean   : 0.00000   Mean   : 0.0000000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.33345   3rd Qu.: 0.3090514   3rd Qu.: 0.31008   3rd Qu.: 0.294544
```

```
##  Max.   : 2.88042   Max.   : 3.0132257   Max.   : 4.31547   Max.   : 2.166903
##      SIG158             SIG159             SIG160             SIG161
##  Min.   :-2.85736   Min.   :-2.78304   Min.   :-1.763135   Min.   :-2.07186
##  1st Qu.:-0.29934   1st Qu.:-0.28459   1st Qu.:-0.268236   1st Qu.:-0.28778
##  Median : 0.00646   Median : 0.02186   Median :-0.003707   Median :-0.01081
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.30047   3rd Qu.: 0.30791   3rd Qu.: 0.267469   3rd Qu.: 0.28401
##  Max.   : 2.75423   Max.   : 1.63265   Max.   : 5.397322   Max.   : 2.61547
##      SIG162             SIG163             SIG164
##  Min.   :-4.712742   Min.   :-2.629493   Min.   :-2.256023
##  1st Qu.:-0.267136   1st Qu.:-0.277695   1st Qu.:-0.282460
##  Median :-0.003656   Median : 0.009856   Median :-0.008185
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.270475   3rd Qu.: 0.301140   3rd Qu.: 0.298398
##  Max.   : 3.961747   Max.   : 2.028397   Max.   : 2.069854
##      SIG165             SIG166             SIG167             SIG168
##  Min.   :-2.133417   Min.   :-2.95100   Min.   :-1.880815   Min.   :-2.31946
##  1st Qu.:-0.270329   1st Qu.:-0.26444   1st Qu.:-0.265226   1st Qu.:-0.24363
##  Median :-0.003285   Median : 0.01844   Median : 0.005672   Median : 0.01578
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.265240   3rd Qu.: 0.27538   3rd Qu.: 0.242554   3rd Qu.: 0.25128
##  Max.   : 1.659123   Max.   : 1.94107   Max.   : 4.331708   Max.   : 3.01509
##      SIG169             SIG170             SIG171             SIG172
##  Min.   :-1.804001   Min.   :-2.00212   Min.   :-1.873660   Min.   :-1.547631
##  1st Qu.:-0.274198   1st Qu.:-0.24101   1st Qu.:-0.248720   1st Qu.:-0.243518
##  Median : 0.009434   Median : 0.01233   Median :-0.000699   Median : 0.000536
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.273351   3rd Qu.: 0.26105   3rd Qu.: 0.239420   3rd Qu.: 0.246963
##  Max.   : 1.272902   Max.   : 2.45445   Max.   : 2.103239   Max.   : 1.767786
##      SIG173             SIG174             SIG175
##  Min.   :-4.828800   Min.   :-1.994681   Min.   :-1.670022
##  1st Qu.:-0.212614   1st Qu.:-0.228886   1st Qu.:-0.228237
##  Median :-0.004214   Median :-0.007521   Median : 0.003024
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.217075   3rd Qu.: 0.229773   3rd Qu.: 0.218791
##  Max.   : 2.833481   Max.   : 2.054785   Max.   : 1.767400
##      SIG176             SIG177             SIG178             SIG179
##  Min.   :-1.548575   Min.   :-1.006833   Min.   :-1.84071   Min.   :-2.067035
##  1st Qu.:-0.227514   1st Qu.:-0.226432   1st Qu.:-0.21784   1st Qu.:-0.194252
##  Median : 0.005132   Median : 0.001554   Median :-0.00576   Median :-0.001161
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.000000
##  3rd Qu.: 0.229386   3rd Qu.: 0.214566   3rd Qu.: 0.21228   3rd Qu.: 0.196395
##  Max.   : 1.475972   Max.   : 1.786938   Max.   : 2.35934   Max.   : 1.651295
##      SIG180             SIG181             SIG182             SIG183
##  Min.   :-2.016286   Min.   :-1.39252   Min.   :-1.476954   Min.   :-2.140330
##  1st Qu.:-0.198807   1st Qu.:-0.20904   1st Qu.:-0.190893   1st Qu.:-0.189786
##  Median :-0.001669   Median : 0.00601   Median : 0.004069   Median : 0.007624
##  Mean   : 0.000000   Mean   : 0.00000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.199388   3rd Qu.: 0.19893   3rd Qu.: 0.192318   3rd Qu.: 0.193168
##  Max.   : 2.279228   Max.   : 1.73042   Max.   : 1.505848   Max.   : 1.628072
##      SIG184             SIG185             SIG186
##  Min.   :-1.181386   Min.   :-1.760612   Min.   :-1.728785
##  1st Qu.:-0.191347   1st Qu.:-0.175641   1st Qu.:-0.182662
##  Median : 0.005529   Median :-0.005518   Median : 0.006732
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.200309   3rd Qu.: 0.185656   3rd Qu.: 0.188885
```

```
##  Max.   : 1.080420   Max.   : 1.339821   Max.   : 1.270147
##      SIG187              SIG188              SIG189
##  Min.   :-2.391159   Min.   :-1.216079   Min.   :-1.359307
##  1st Qu.:-0.172797   1st Qu.:-0.182385   1st Qu.:-0.168814
##  Median :-0.003753   Median :-0.008081   Median : 0.002325
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.176160   3rd Qu.: 0.166455   3rd Qu.: 0.170720
##  Max.   : 1.977947   Max.   : 1.693073   Max.   : 1.373629
##      SIG190              SIG191              SIG192
##  Min.   :-1.451934   Min.   :-1.505320   Min.   :-0.924336
##  1st Qu.:-0.169803   1st Qu.:-0.162505   1st Qu.:-0.171876
##  Median :-0.004635   Median : 0.008415   Median :-0.007541
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.169735   3rd Qu.: 0.156153   3rd Qu.: 0.160978
##  Max.   : 1.001747   Max.   : 1.239244   Max.   : 1.098903
##      SIG193              SIG194              SIG195
##  Min.   :-1.667591   Min.   :-1.032334   Min.   :-1.214819
##  1st Qu.:-0.147506   1st Qu.:-0.163981   1st Qu.:-0.120257
##  Median :-0.003293   Median :-0.004189   Median : 0.007221
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000
##  3rd Qu.: 0.153791   3rd Qu.: 0.155053   3rd Qu.: 0.128714
##  Max.   : 1.792427   Max.   : 1.347462   Max.   : 2.795549
##      SIG196              SIG197              SIG198              SIG199
##  Min.   :-1.282285   Min.   :-1.039773   Min.   :-1.382971   Min.   :-1.19827
##  1st Qu.:-0.154911   1st Qu.:-0.145165   1st Qu.:-0.137257   1st Qu.:-0.14223
##  Median :-0.002498   Median :-0.001646   Median : 0.004751   Median :-0.01329
##  Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.000000   Mean   : 0.00000
##  3rd Qu.: 0.144669   3rd Qu.: 0.152125   3rd Qu.: 0.149507   3rd Qu.: 0.14130
##  Max.   : 0.955378   Max.   : 1.533994   Max.   : 0.951079   Max.   : 1.11086
##      SIG200
##  Min.   :-2.003092
##  1st Qu.:-0.133725
##  Median : 0.004427
##  Mean   : 0.000000
##  3rd Qu.: 0.136356
##  Max.   : 0.682943
```

CHECK BALANCE

```
table(data$FAIL)
```

```
##
##    0    1
## 1377  100
```

```
duplicate_rows <- data[duplicated(data), ]

if (nrow(duplicate_rows) > 0) {
  cat("Duplicate rows found:\n")
  print(duplicate_rows)
} else {
  cat("No duplicate rows found.\n")
}
```

```
## No duplicate rows found.
```

```
model = glm(FAIL ~ ., data= data, family=binomial)
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
summary(model)
```

```
##
## Call:
## glm(formula = FAIL ~ ., family = binomial, data = data)
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -8.608560   0.995718  -8.646  < 2e-16 ***
## SIG1         0.018091   0.152885   0.118 0.905804
## SIG2        -0.287535   0.057554  -4.996 5.86e-07 ***
## SIG3        -0.168501   0.098584  -1.709 0.087413 .
## SIG4        -0.320748   0.201749  -1.590 0.111871
## SIG5        -0.056681   0.227509  -0.249 0.803253
## SIG6         0.416706   0.282500   1.475 0.140194
## SIG7         0.130161   0.370669   0.351 0.725474
## SIG8         0.250574   0.447558   0.560 0.575570
## SIG9         0.345944   0.199912   1.730 0.083545 .
## SIG10       -0.928454   1.444365  -0.643 0.520346
## SIG11        0.457731   0.769680   0.595 0.552042
## SIG12       -0.045185   0.484454  -0.093 0.925689
## SIG13        0.446287   0.299601   1.490 0.136328
## SIG14        0.281358   0.270999   1.038 0.299166
## SIG15       -0.248939   0.375557  -0.663 0.507425
## SIG16        1.108456   0.473701   2.340 0.019284 *
## SIG17       -0.929254   0.370226  -2.510 0.012074 *
## SIG18       -1.480863   0.647080  -2.289 0.022107 *
## SIG19        1.613104   0.572020   2.820 0.004802 **
## SIG20       -3.425560   1.290658  -2.654 0.007952 **
## SIG21       -1.066873   0.542257  -1.967 0.049129 *
## SIG22       -1.326239   0.400798  -3.309 0.000936 ***
## SIG23        1.236199   0.586537   2.108 0.035064 *
## SIG24        1.153090   0.373703   3.086 0.002032 **
## SIG25        1.085550   0.482553   2.250 0.024475 *
## SIG26        0.629159   0.389761   1.614 0.106480
## SIG27       -0.701359   0.350524  -2.001 0.045405 *
## SIG28       -0.684805   0.345566  -1.982 0.047514 *
## SIG29       -0.535295   0.369203  -1.450 0.147096
## SIG30       -0.833740   0.458474  -1.819 0.068986 .
## SIG31       -0.621157   0.371532  -1.672 0.094548 .
## SIG32        1.126793   0.522159   2.158 0.030932 *
## SIG33       -0.024915   0.296371  -0.084 0.933002
## SIG34        0.222017   0.270560   0.821 0.411883
## SIG35       -0.407722   0.322257  -1.265 0.205796
## SIG36       -0.579793   0.271972  -2.132 0.033022 *
## SIG37        0.563736   0.232728   2.422 0.015423 *
## SIG38       -0.407880   0.419935  -0.971 0.331402
## SIG39       -0.502882   0.332200  -1.514 0.130079
## SIG40       -0.185501   0.292956  -0.633 0.526599
## SIG41       -0.152845   0.257237  -0.594 0.552392
## SIG42        0.502501   0.280858   1.789 0.073589 .
## SIG43        0.761260   0.364657   2.088 0.036834 *
## SIG44       -0.061670   0.450678  -0.137 0.891159
## SIG45        0.050258   0.230701   0.218 0.827545
## SIG46        1.335949   0.486397   2.747 0.006021 **
## SIG47       -1.257049   0.418010  -3.007 0.002637 **
## SIG48       -0.117457   0.223788  -0.525 0.599682
```

```
## SIG49       -1.324478   0.424810  -3.118 0.001822 **
## SIG50        0.445356   0.167530   2.658 0.007852 **
## SIG51       -0.595912   0.448405  -1.329 0.183861
## SIG52       -0.912942   0.313674  -2.910 0.003609 **
## SIG53       -0.400170   0.258319  -1.549 0.121350
## SIG54       -0.337916   0.313391  -1.078 0.280920
## SIG55        0.974337   0.480444   2.028 0.042561 *
## SIG56        0.348016   0.267432   1.301 0.193146
## SIG57        0.139583   0.361918   0.386 0.699737
## SIG58       -0.624836   0.302166  -2.068 0.038654 *
## SIG59        1.504603   0.390603   3.852 0.000117 ***
## SIG60        0.140161   0.281269   0.498 0.618262
## SIG61       -0.999613   0.258161  -3.872 0.000108 ***
## SIG62       -0.197386   0.329263  -0.599 0.548855
## SIG63        0.757650   0.339461   2.232 0.025620 *
## SIG64        0.077071   0.167940   0.459 0.646291
## SIG65       -0.073014   0.208804  -0.350 0.726581
## SIG66        0.241388   0.186575   1.294 0.195741
## SIG67       -0.001067   0.253721  -0.004 0.996645
## SIG68        0.715743   0.372559   1.921 0.054712 .
## SIG69       -0.700273   0.250055  -2.800 0.005103 **
## SIG70       -0.262690   0.303469  -0.866 0.386697
## SIG71       -0.267533   0.215208  -1.243 0.213818
## SIG72       -0.124849   0.227676  -0.548 0.583444
## SIG73       -0.200355   0.257248  -0.779 0.436074
## SIG74        0.457583   0.304281   1.504 0.132629
## SIG75        0.375987   0.240490   1.563 0.117953
## SIG76        0.402953   0.281888   1.429 0.152866
## SIG77       -0.481708   0.361678  -1.332 0.182904
## SIG78        0.124982   0.220077   0.568 0.570104
## SIG79        0.310623   0.355107   0.875 0.381720
## SIG80       -0.243898   0.271481  -0.898 0.368974
## SIG81       -0.169787   0.266519  -0.637 0.524089
## SIG82        0.107366   0.250181   0.429 0.667813
## SIG83       -0.743529   0.296562  -2.507 0.012170 *
## SIG84        0.092639   0.301594   0.307 0.758717
## SIG85       -0.123393   0.422712  -0.292 0.770357
## SIG86       -0.690334   0.279156  -2.473 0.013401 *
## SIG87       -0.072285   0.334877  -0.216 0.829101
## SIG88       -0.250466   0.249287  -1.005 0.315027
## SIG89        0.174486   0.279202   0.625 0.532007
## SIG90        0.014297   0.278342   0.051 0.959034
## SIG91       -0.153724   0.266942  -0.576 0.564704
## SIG92        0.567767   0.323691   1.754 0.079424 .
## SIG93        0.090442   0.356631   0.254 0.799804
## SIG94        0.136603   0.313910   0.435 0.663443
## SIG95       -0.626576   0.374648  -1.672 0.094438 .
## SIG96        0.347690   0.316874   1.097 0.272531
## SIG97       -0.874136   0.614881  -1.422 0.155133
## SIG98        0.871363   0.362304   2.405 0.016170 *
## SIG99        0.641844   0.359433   1.786 0.074146 .
## SIG100      -0.459570   0.315203  -1.458 0.144837
## SIG101      -0.124556   0.372728  -0.334 0.738247
## SIG102       1.063320   0.335961   3.165 0.001551 **
## SIG103       0.620041   0.345621   1.794 0.072815 .
## SIG104       0.556560   0.271995   2.046 0.040735 *
```

```
## SIG105       0.304614   0.466220   0.653 0.513518
## SIG106      -0.358075   0.333100  -1.075 0.282385
## SIG107      -0.385778   0.414251  -0.931 0.351717
## SIG108       0.314609   0.310745   1.012 0.311330
## SIG109       0.831316   0.381950   2.177 0.029518 *
## SIG110      -0.066306   0.338042  -0.196 0.844494
## SIG111      -0.903874   0.310959  -2.907 0.003652 **
## SIG112      -0.005793   0.330141  -0.018 0.986001
## SIG113       0.136990   0.396667   0.345 0.729830
## SIG114      -0.098571   0.315740  -0.312 0.754896
## SIG115      -0.288403   0.401230  -0.719 0.472265
## SIG116       0.659738   0.330281   1.998 0.045770 *
## SIG117      -0.901469   0.515430  -1.749 0.080297 .
## SIG118      -0.136131   0.449586  -0.303 0.762048
## SIG119      -0.112843   0.378791  -0.298 0.765777
## SIG120      -0.344103   0.416093  -0.827 0.408245
## SIG121      -0.645386   0.493129  -1.309 0.190616
## SIG122      -0.097475   0.367137  -0.265 0.790624
## SIG123      -0.013593   0.304340  -0.045 0.964374
## SIG124       0.248213   0.324737   0.764 0.444657
## SIG125      -0.260911   0.376276  -0.693 0.488057
## SIG126       0.061645   0.347610   0.177 0.859243
## SIG127      -0.294943   0.446828  -0.660 0.509202
## SIG128      -0.729096   0.483879  -1.507 0.131869
## SIG129      -0.296278   0.547901  -0.541 0.588680
## SIG130      -0.916643   0.503296  -1.821 0.068564 .
## SIG131      -0.345161   0.519893  -0.664 0.506750
## SIG132      -0.283651   0.428247  -0.662 0.507745
## SIG133      -0.389617   0.439413  -0.887 0.375253
## SIG134      -0.248313   0.480927  -0.516 0.605630
## SIG135       0.159223   0.415424   0.383 0.701513
## SIG136      -1.669057   0.555248  -3.006 0.002647 **
## SIG137       1.104639   0.523326   2.111 0.034789 *
## SIG138       0.230340   0.467320   0.493 0.622086
## SIG139      -0.204627   0.369808  -0.553 0.580035
## SIG140       0.759637   0.494012   1.538 0.124125
## SIG141      -0.675337   0.430668  -1.568 0.116854
## SIG142      -1.511301   0.536923  -2.815 0.004882 **
## SIG143      -0.629039   0.491423  -1.280 0.200532
## SIG144       1.451946   1.000010   1.452 0.146521
## SIG145      -1.516151   0.649946  -2.333 0.019662 *
## SIG146      -1.131616   0.602121  -1.879 0.060192 .
## SIG147      -0.254586   0.426528  -0.597 0.550588
## SIG148      -0.954601   1.000473  -0.954 0.340008
## SIG149      -0.808151   0.465865  -1.735 0.082789 .
## SIG150       1.524165   0.967686   1.575 0.115242
## SIG151       0.007612   0.655771   0.012 0.990738
## SIG152      -1.134141   0.653992  -1.734 0.082886 .
## SIG153      -1.033605   0.990214  -1.044 0.296569
## SIG154      -1.705046   0.589757  -2.891 0.003839 **
## SIG155      -1.108875   0.511208  -2.169 0.030073 *
## SIG156       1.781601   1.128424   1.579 0.114373
## SIG157      -0.284096   0.494972  -0.574 0.565992
## SIG158      -0.377243   0.477127  -0.791 0.429145
## SIG159       0.952180   0.867428   1.098 0.272334
## SIG160       0.715932   0.640298   1.118 0.263514
```

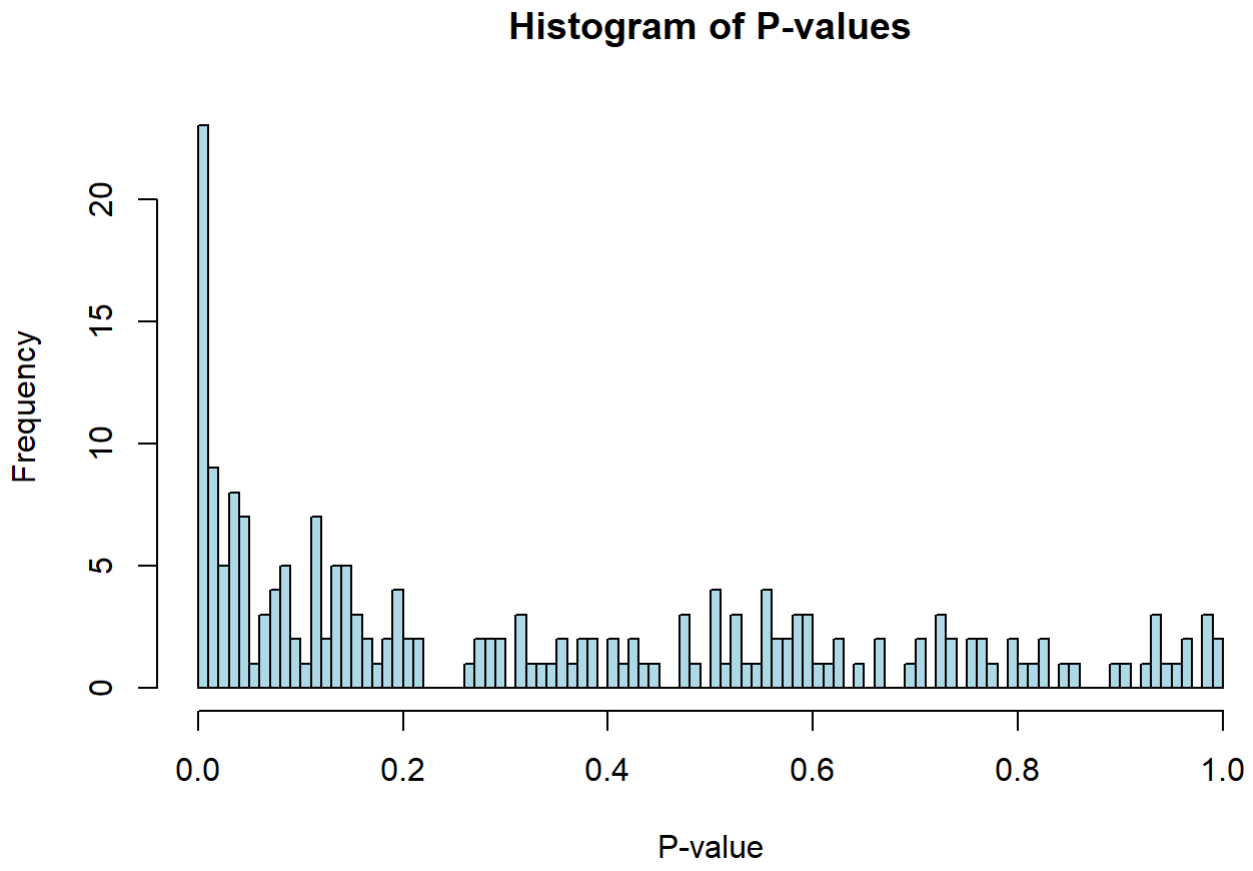
```
## SIG161       0.517449   0.721444   0.717 0.473226
## SIG162      -1.325549   0.570065  -2.325 0.020058 *
## SIG163      -0.604140   0.613569  -0.985 0.324805
## SIG164       1.144243   0.544224   2.103 0.035508 *
## SIG165       1.092038   0.783739   1.393 0.163508
## SIG166      -1.671841   0.562394  -2.973 0.002952 **
## SIG167       1.298798   0.555656   2.337 0.019418 *
## SIG168      -0.146341   0.621524  -0.235 0.813855
## SIG169      -0.185953   0.558242  -0.333 0.739055
## SIG170       0.266786   0.543619   0.491 0.623596
## SIG171      -0.346118   0.650877  -0.532 0.594884
## SIG172       1.006540   0.641979   1.568 0.116911
## SIG173       1.404088   0.707503   1.985 0.047193 *
## SIG174       0.027022   0.674428   0.040 0.968040
## SIG175       0.537696   0.641219   0.839 0.401720
## SIG176       0.010119   0.619808   0.016 0.986974
## SIG177      -0.055417   0.649738  -0.085 0.932030
## SIG178       2.412234   0.731984   3.295 0.000983 ***
## SIG179      -0.553312   0.621018  -0.891 0.372942
## SIG180       1.067855   0.829990   1.287 0.198238
## SIG181       1.384182   0.918782   1.507 0.131928
## SIG182       1.511775   1.025409   1.474 0.140397
## SIG183       0.787824   0.851665   0.925 0.354945
## SIG184       0.070447   0.929197   0.076 0.939567
## SIG185      -0.531064   0.893370  -0.594 0.552211
## SIG186       3.166738   1.236314   2.561 0.010424 *
## SIG187      -1.380091   0.875327  -1.577 0.114874
## SIG188       1.234792   0.872857   1.415 0.157169
## SIG189       3.661363   0.935639   3.913 9.11e-05 ***
## SIG190      -1.152331   0.839602  -1.372 0.169916
## SIG191       0.676733   0.942276   0.718 0.472640
## SIG192       0.511373   0.949583   0.539 0.590216
## SIG193      -1.216210   1.220792  -0.996 0.319131
## SIG194      -1.132121   0.911947  -1.241 0.214446
## SIG195       0.644187   1.702129   0.378 0.705089
## SIG196      -0.062079   0.916134  -0.068 0.945975
## SIG197      -0.021123   0.945818  -0.022 0.982183
## SIG198      -0.933732   1.177299  -0.793 0.427711
## SIG199      -1.661653   1.234706  -1.346 0.178371
## SIG200      -2.531454   0.879994  -2.877 0.004019 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 731.59  on 1476  degrees of freedom
## Residual deviance: 320.33  on 1276  degrees of freedom
## AIC: 722.33
##
## Number of Fisher Scoring iterations: 10
```

```
deviance <- function(y, pred, family=c("gaussian","binomial")){
  # Compute the deviance as a measure of prediction error
  # analogous to the SSR in a regression model.
  # y = vector of observed outcomes (binary)
  # pred = vector of predicted outcomes (probabilities between 0 and 1)
  family <- match.arg(family)
  if(family=="gaussian"){
    return( sum( (y-pred)^2 ) )
    }else{
  if(is.factor(y)) y <- as.numeric(y)>1
    return( -2*sum( y*log(pred) + (1-y)*log(1-pred) ) )
  }
}
R2 <- function(y, pred, family=c("gaussian","binomial")){
  # Compute the R2 as a measure of goodness of fit for
  # glm-estimated models.
  # y = vector of observed outcomes (binary)
  # pred = vector of predicted outcomes (probabilities between 0 and 1)
  fam <- match.arg(family)
  if(fam=="binomial"){
    if(is.factor(y)){ y <- as.numeric(y)>1 }
  }
  dev <- deviance(y, pred, family=fam)
  dev0 <- deviance(y, mean(y), family=fam)
  return(1-dev/dev0)
}
```

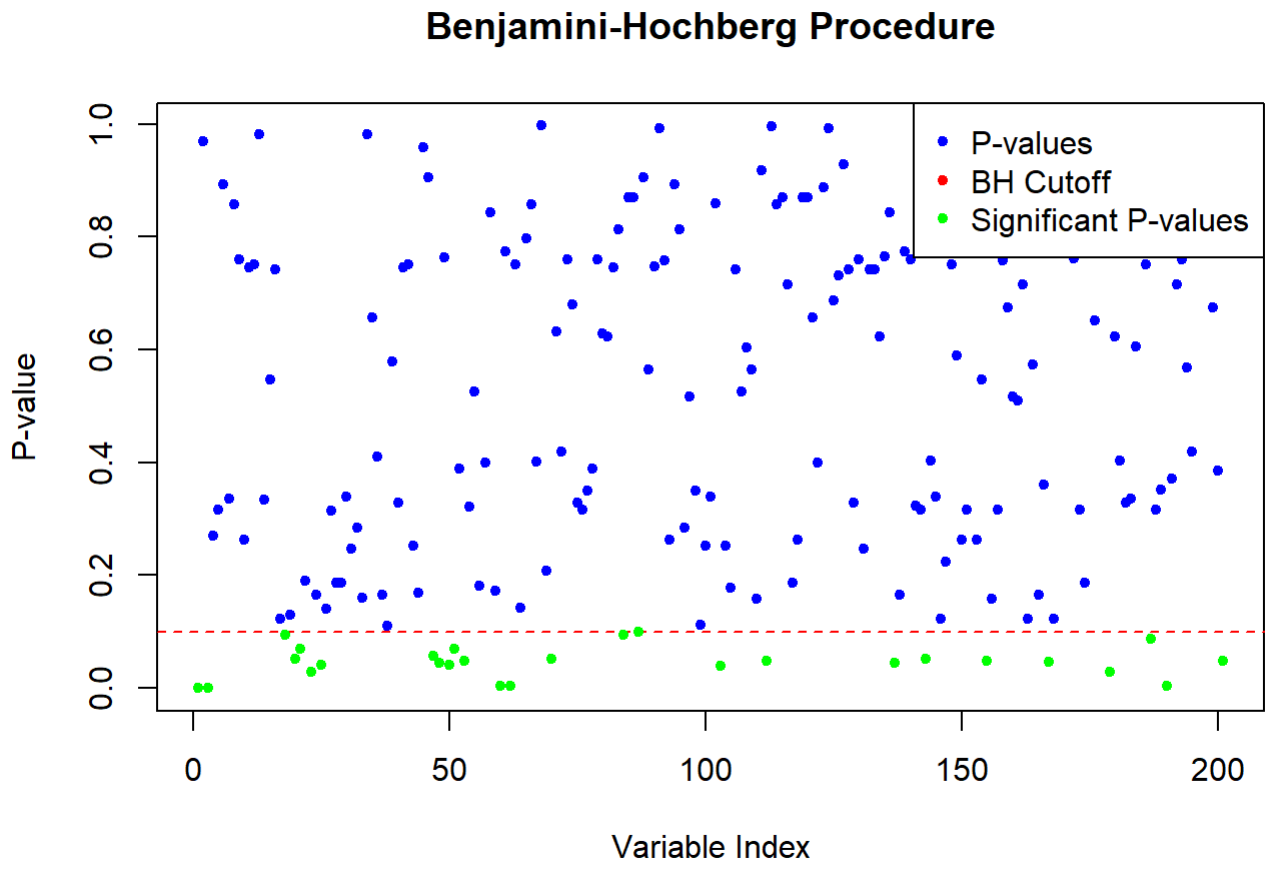
Graph for all p values and BH line cut off showing the ret dots

```
p_values <- summary(model)$coefficients[, "Pr(>|z|)"]
```

```
hist(p_values, breaks = 100, col = "lightblue", main = "Histogram of P-values", xlab = "P-val
ue")
```




```
# Calculate adjusted p-values using BH correction
adjusted_p_values <- p.adjust(p_values, method = "BH")

# Determine the BH cutoff
bh_cutoff <- 0.1  # Define the desired FDR threshold

# Create a plot of the p-values
plot(1:length(p_values), adjusted_p_values, pch = 20, col = "blue",
     xlab = "Variable Index", ylab = "P-value",
     main = "Benjamini-Hochberg Procedure")

# Add a line for the BH cutoff
abline(h = bh_cutoff, col = "red", lty = 2)

# Highlight p-values below the BH cutoff
significant_variables <- which(adjusted_p_values <= bh_cutoff)
points(significant_variables, adjusted_p_values[significant_variables], pch = 20, col = "gree
n")

# Add a legend
legend("topright", legend = c("P-values", "BH Cutoff", "Significant P-values"),
       col = c("blue", "red", "green"), pch = 20)
```




```
p_values[significant_variables]
```

```
##  (Intercept)         SIG2        SIG17        SIG19        SIG20        SIG22
## 5.353317e-18 5.856618e-07 1.207432e-02 4.802170e-03 7.951597e-03 9.363145e-04
##        SIG24        SIG46        SIG47        SIG49        SIG50        SIG52
## 2.031553e-03 6.021210e-03 2.636504e-03 1.822010e-03 7.852088e-03 3.608682e-03
##        SIG59        SIG61        SIG69        SIG83        SIG86       SIG102
## 1.171567e-04 1.079202e-04 5.102665e-03 1.217043e-02 1.340081e-02 1.550787e-03
##       SIG111       SIG136       SIG142       SIG154       SIG166       SIG178
## 3.652313e-03 2.647394e-03 4.881594e-03 3.838949e-03 2.951738e-03 9.825513e-04
##       SIG186       SIG189       SIG200
## 1.042408e-02 9.107346e-05 4.018916e-03
```

```
data_restricted <- data[, significant_variables]

head(data_restricted)
```

	F...	SIG2	SIG17	SIG19	SIG20	SIG22	SIG24	SIG46
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	0	-0.3490181	2.0476967	-1.7756151	-1.8028029	-3.0232213	1.8220742	0.6606199
2	1	-0.3813179	-0.3341238	0.5464913	0.9069240	-0.9950427	5.6063616	0.1634569
3	0	-0.8253003	0.8510338	0.1111051	0.9455983	-1.8249360	2.4579196	-0.4909513
4	0	-1.9055982	5.8602136	-6.2693813	-3.1342653	-2.8211406	2.3396976	0.1468513

	F…	SIG2	SIG17	SIG19	SIG20	SIG22	SIG24	SIG46
	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
5	0	-2.3138448	-2.1279288	-1.6461828	-0.5619549	1.4076090	-0.7329345	0.9815192
6	0	-4.5603800	2.5799909	-0.1919715	0.4303201	0.1563130	2.2196977	1.7460495

```
table(data_restricted$FAIL)
```

```
##
##    0    1
## 1377  100
```

```
dim(data_restricted)
```

```
## [1] 1477   27
```

```
model_restricted <- glm(FAIL ~ ., data = data_restricted, family = "binomial")
```

```
summary(model_restricted)
```

```
##
## Call:
## glm(formula = FAIL ~ ., family = "binomial", data = data_restricted)
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -3.31275    0.16375 -20.230  < 2e-16 ***
## SIG2        -0.08414    0.01750  -4.808 1.52e-06 ***
## SIG17       -0.09557    0.05688  -1.680 0.092926 .
## SIG19        0.07265    0.05795   1.254 0.209927
## SIG20       -0.21954    0.07357  -2.984 0.002845 **
## SIG22       -0.15203    0.05762  -2.639 0.008327 **
## SIG24        0.18985    0.05882   3.228 0.001248 **
## SIG46        0.10152    0.07592   1.337 0.181183
## SIG47       -0.04679    0.07513  -0.623 0.533476
## SIG49       -0.07668    0.07707  -0.995 0.319760
## SIG50        0.23063    0.07028   3.281 0.001033 **
## SIG52       -0.08487    0.07845  -1.082 0.279356
## SIG59        0.16108    0.08135   1.980 0.047678 *
## SIG61       -0.26494    0.08152  -3.250 0.001155 **
## SIG69       -0.17523    0.08930  -1.962 0.049722 *
## SIG83       -0.12113    0.10381  -1.167 0.243277
## SIG86       -0.19059    0.11073  -1.721 0.085216 .
## SIG102       0.25057    0.12404   2.020 0.043375 *
## SIG111      -0.30290    0.13056  -2.320 0.020342 *
## SIG136      -0.44934    0.16098  -2.791 0.005249 **
## SIG142      -0.34157    0.16570  -2.061 0.039264 *
## SIG154      -0.61988    0.20568  -3.014 0.002580 **
## SIG166      -0.90456    0.25463  -3.552 0.000382 ***
## SIG178       0.79811    0.32816   2.432 0.015012 *
## SIG186       0.62163    0.37920   1.639 0.101147
## SIG189       2.05099    0.43694   4.694 2.68e-06 ***
## SIG200      -0.89506    0.48874  -1.831 0.067045 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 731.59  on 1476  degrees of freedom
## Residual deviance: 596.08  on 1450  degrees of freedom
## AIC: 650.08
##
## Number of Fisher Scoring iterations: 6
```

```
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'caret'
```

```
## The following object is masked _by_ '.GlobalEnv':
##
##     R2
```

```
ctrl <- trainControl(method = "cv", number = 5)
```

```
dataf = data
dataf$FAIL <- factor(dataf$FAIL, levels = c("0", "1"))
class(dataf$FAIL)
```

```
## [1] "factor"
```

```
data_restrictedf = data_restricted
data_restrictedf$FAIL <- factor(data_restrictedf$FAIL, levels = c("0", "1"))
class(data_restrictedf$FAIL)
```

```
## [1] "factor"
```

```
#r2_full <- R2(as.numeric(predict(model_full_cv, data), data$FAIL))

#r2_restricted <- R2(predict(model_restricted_cv, data_restricted), data_restricted$FAIL)
```

```
set.seed(123)  # for reproducibility

# Number of folds for cross-validation
num_folds <- 5

# Indices for cross-validation folds
fold_indices <- sample(rep(1:num_folds, length.out = nrow(data)))

# Initialize vectors to store predicted values and actual outcomes
pred_full <- rep(NA, nrow(data))
```

```
# Perform cross-validation
for (fold in 1:num_folds) {
  # Split the data into training and testing sets for the current fold
  train_data <- data[fold_indices != fold, ]
  test_data <- data[fold_indices == fold, ]

  # Train the full model
  model_full_cv <- glm(FAIL ~ ., data = train_data, family = "binomial")

  # Predict on the test set
  pred_full[fold_indices == fold] <- predict(model_full_cv, newdata = test_data, type = "resp
onse")

}
```

```
## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
## Warning: glm.fit: algorithm did not converge
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
set.seed(123)  # for reproducibility

# Number of folds for cross-validation
num_folds <- 5

# Indices for cross-validation folds
fold_indices <- sample(rep(1:num_folds, length.out = nrow(data_restricted)))

# Initialize vectors to store predicted values and actual outcomes
pred_restricted <- rep(NA, nrow(data_restricted))
```

```
# Perform cross-validation
for (fold in 1:num_folds) {
  # Split the data into training and testing sets for the current fold
  train_data <- data_restricted[fold_indices != fold, ]
  test_data <- data_restricted[fold_indices == fold, ]

  # Train the full model
  model_res_cv <- glm(FAIL ~ ., data = train_data, family = "binomial")

  # Predict on the test set
  pred_restricted[fold_indices == fold] <- predict(model_res_cv, newdata = test_data, type =
"response")

}
```

```
deviance <- function(y, pred, family = c("gaussian", "binomial")) {
  # Match the family argument
  family <- match.arg(family)

  if (family == "gaussian") {
    # For Gaussian family, compute the sum of squared differences
    return(sum((y - pred)^2))
  } else {
    # For binomial family (e.g., logistic regression), calculate the binomial deviance
    if (is.factor(y)) y <- as.numeric(y) > 1
    return(-2 * sum(y * log(pred) + (1 - y) * log(1 - pred)))
  }
}
```

```
R2 <- function(y, pred, family = c("gaussian", "binomial")) {
  # Match the family argument
  fam <- match.arg(family)

  if (fam == "binomial") {
    # For binomial family, ensure y is numeric (convert a factor to numeric)
    if (is.factor(y)) y <- as.numeric(y) > 1
  }

  # Calculate deviance for the model and deviance for the null model (mean model)
  dev <- deviance(y, pred, family = fam)
  dev0 <- deviance(y, mean(y), family = fam)

  # Calculate and return R-squared
  return(1 - dev / dev0)
}
```

```
# Calculate R-squared for the full model
r2_full <- R2(data$FAIL, pred_full)

# Calculate R-squared for the restricted model
r2_restricted <- R2(data_restricted$FAIL, pred_restricted)
```

```
# Print results
cat("R2 for the Full Model:", r2_full, "\n")
```

```
## R2 for the Full Model: -1.163335
```

```
cat("R2 for the Restricted Model:", r2_restricted, "\n")
```

```
## R2 for the Restricted Model: 0.05674817
```

Multiple testing - Increased likelihood of finding some significant results by chance even if there’s no true effect