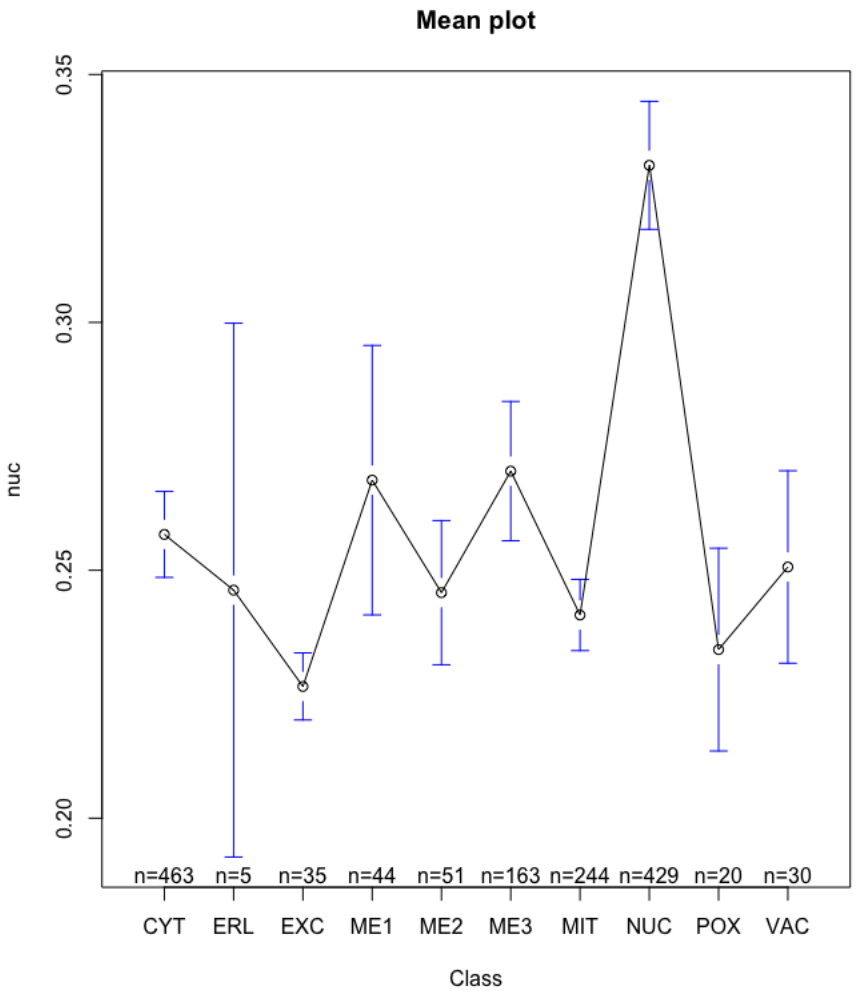
Console ~/work/ACADGILD/

```
> yeast_data=read.table(url("https://archive.ics.uci.edu/ml/machine-learning-databases/yeast/yeast.data"),col.names = col_names)
> col_names <- c("Sequence Name","mcg","gvh","alm","mit","erl","pox","vac","nuc","class")
> summary(yeast_data)
    Sequence.Name        mcg              gvh              alm              mit              erl              pox              vac              nuc              class       
 EF1A_YEAST:   2   Min.   :0.1100   Min.   :0.1300   Min.   :0.21   Min.   :0.0000   Min.   :0.5000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   CYT    :463  
 H3_YEAST  :   2   1st Qu.:0.4100   1st Qu.:0.4200   1st Qu.:0.46   1st Qu.:0.1700   1st Qu.:0.5000   1st Qu.:0.0000   1st Qu.:0.4800   1st Qu.:0.2200   NUC    :429  
 H4_YEAST  :   2   Median :0.4900   Median :0.4900   Median :0.51   Median :0.2200   Median :0.5000   Median :0.0000   Median :0.5100   Median :0.2200   MIT    :244  
 IF4A_YEAST:   2   Mean   :0.5001   Mean   :0.4999   Mean   :0.50   Mean   :0.2612   Mean   :0.5047   Mean   :0.0075   Mean   :0.4999   Mean   :0.2762   ME3    :163  
 MAT2_YEAST:   2   3rd Qu.:0.5800   3rd Qu.:0.5700   3rd Qu.:0.55   3rd Qu.:0.3200   3rd Qu.:0.5000   3rd Qu.:0.0000   3rd Qu.:0.5300   3rd Qu.:0.3000   ME2    : 51  
 MTC_YEAST :   2   Max.   :1.0000   Max.   :1.0000   Max.   :1.00   Max.   :1.0000   Max.   :1.0000   Max.   :0.8300   Max.   :0.7300   Max.   :1.0000   ME1    : 44  
 (Other)   :1472                                                                                                                                  (Other): 90  
> #a.Perform ANOVA test on the discriminant analysis scores of nuclear localization signals of both nuclear
> #and non-nuclear proteins by class variables (Target).
> library(lattice)
> test1 =aov(yeast_data$nuc ~ yeast_data$class , yeast_data)
> summary(test1)
                Df Sum Sq Mean Sq F value Pr(>F)    
yeast_data$class    9  1.993 0.22141   22.01 <2e-16 ***
Residuals        1474 14.825 0.01006                   
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> TukeyHSD(test1)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = yeast_data$nuc ~ yeast_data$class, data = yeast_data)

$`yeast_data$class`
                diff         lwr          upr     p adj
ERL-CYT -0.0112354212 -0.15411077  0.131639929 0.9999999
EXC-CYT -0.0306639926 -0.08636974  0.025041758 0.7698696
ME1-CYT  0.0109463970 -0.03918352  0.061076319 0.9995610
ME2-CYT -0.0117452251 -0.05862827  0.035137825 0.9986751
ME3-CYT  0.0127645788 -0.01617637  0.041705527 0.9281817
MIT-CYT -0.0162518146 -0.04139001  0.008886385 0.5645771
NUC-CYT  0.0744429005  0.05314812  0.095737681 0.0000000
POX-CYT -0.0232354212 -0.09580890  0.049338059 0.9914395
VAC-CYT -0.0065687545 -0.06643503  0.053297518 0.9999988
EXC-ERL -0.0194285714 -0.17135063  0.132493492 0.9999953
ME1-ERL  0.0221818182 -0.12778548  0.172149120 0.9999834
ME2-ERL -0.0005098039 -0.14942321  0.148403604 1.0000000
ME3-ERL  0.0240000000 -0.12027322  0.168273217 0.9999547
MIT-ERL -0.0050163934 -0.14857513  0.138542346 1.0000000
NUC-ERL  0.0856783217 -0.05725750  0.228614148 0.6696330
POX-ERL -0.0120000000 -0.17088390  0.146883897 1.0000000
VAC-ERL  0.0046666667 -0.14882976  0.158163094 1.0000000
ME1-EXC  0.0416103896 -0.03036153  0.113582306 0.7147855
ME2-EXC  0.0189187675 -0.05083056  0.088668090 0.9975341
ME3-EXC  0.0434285714 -0.01577047  0.102627610 0.3735707
MIT-EXC  0.0144121780 -0.04302367  0.071848027 0.9986581
NUC-EXC  0.1051068931  0.04924622  0.160967571 0.0000001
POX-EXC  0.0074285714 -0.08164363  0.096500777 0.9999999
VAC-EXC  0.0240952381 -0.05496751  0.103157987 0.9940385
ME2-ME1 -0.0226916221 -0.08807393  0.042690687 0.9847913
ME3-ME1  0.0018181818 -0.05216704  0.055803406 1.0000000
MIT-ME1 -0.0271982116 -0.07924391  0.024847491 0.8199565
NUC-ME1  0.0634965035  0.01319448  0.113798528 0.0026743
POX-ME1 -0.0341818182 -0.11987742  0.051513783 0.9613426
VAC-ME1 -0.0175151515 -0.09275342  0.057723114 0.9992623
ME3-ME2  0.0245098039 -0.02647466  0.075494270 0.8829855
MIT-ME2 -0.0045065895 -0.05343272  0.044419543 0.9999997
NUC-ME2  0.0861881256  0.03912110  0.133255152 0.0000004
POX-ME2 -0.0114901961 -0.09532782  0.072347429 0.9999914
VAC-ME2  0.0051764706 -0.06793856  0.078291502 1.0000000
MIT-ME3 -0.0290163934 -0.06116176  0.003128972 0.1179577
NUC-ME3  0.0616783217  0.03244028  0.090916364 0.0000000
POX-ME3 -0.0360000000 -0.11128815  0.039288150 0.8863429
VAC-ME3 -0.0193333333 -0.08246309  0.043796422 0.9938197
NUC-MIT  0.0906947151  0.06521504  0.116174388 0.0000000
POX-MIT -0.0069836066 -0.08089339  0.066926177 0.9999997
VAC-MIT  0.0096830601 -0.05179634  0.071162460 0.9999716
POX-NUC -0.0976783217 -0.17037079 -0.024985855 0.0009144
VAC-NUC -0.0810116550 -0.14102211 -0.021001196 0.0008397
VAC-POX  0.0166666667 -0.07506499  0.108398327 0.9999044

> library(gplots)
> plotmeans(nuc ~ class,xlab="Class",ylab="nuc",main="Mean plot",data = yeast_data)
```


Mean plot
nuc
0.35
0.30
0.25
0.20
n=463
n=5
n=35
n=44
n=51
n=163
n=244
n=429
n=20
n=30
CYT
ERL
EXC
ME1
ME2
ME3
MIT
NUC
POX
VAC
Class