Table of Content

Link to the GitHub Repository

https://github.com/CasedUgr/PanCancerClassification

Links to the Specific Cancer Datasets

Bladder Dataset

Brain Dataset

Breast Dataset

Colon Dataset

Corpus Uteri Dataset

Esophagus Dataset

HeadNeck Dataset

Kidney Dataset

Larynx Dataset

Liver Dataset

Lung Dataset

Prostate Dataset

Stomach Dataset

Thyroid Dataset

ANOVA (Without Interactions)

Analysis of Variance for F1Score - Type III Sums of Squares

Analysis of Variance for F1Score - Type III Sums of Squares

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-Value*
MAIN EFFECTS					
A:Class	0,000177968	1	0,000177968	0,04	0,8387
B:FS	1,39563	1	1,39563	325,39	0,0000
C:LFC	0,128913	2	0,0644563	15,03	0,0000
D:COV	0,141607	6	0,0236011	5,50	0,0000
E:nGenes	7,39956	6	1,23326	287,53	0,0000
RESIDUAL	2,44908	571	0,00428911		
TOTAL (CORRECTED)	11,515	587			

All F-ratios are based on the residual mean square error.

The StatAdvisor

The ANOVA table decomposes the variability of F1Score into contributions due to various factors. Since

Type III sums of squares (the default) have been chosen, the contribution of each factor is measured having removed the effects of all other factors. The P-values test the statistical significance of each of the factors. Since 4 P-values are less than 0,05, these factors have a statistically significant effect on F1Score at the 95,0% confidence level.

Table of Least Squares Means for F1Score with 95,0% Confidence Intervals

Intervals

Level	*Count*	*Mean*	*Stnd. Error*	*Lower Limit*	*Upper Limit*
GRAND MEAN	588	0,815923			
Class					
KNN	294	0,816473	0,00381953	0,808971	0,823975
SVM	294	0,815373	0,00381953	0,807871	0,822875
FS					
mrmr	294	0,864642	0,00381953	0,85714	0,872144
nofs	294	0,767204	0,00381953	0,759702	0,774706
LFC					
1	196	0,794988	0,00467795	0,7858	0,804176
1.5	196	0,826772	0,00467795	0,817584	0,83596
2	196	0,826009	0,00467795	0,816821	0,835197
COV					
35	84	0,792312	0,00714569	0,778277	0,806347
40	84	0,797923	0,00714569	0,783888	0,811958
45	84	0,805984	0,00714569	0,791949	0,820019
50	84	0,825288	0,00714569	0,811253	0,839323
55	84	0,829339	0,00714569	0,815304	0,843374
60	84	0,833777	0,00714569	0,819742	0,847812
65	84	0,826835	0,00714569	0,8128	0,84087
nGenes					
3	84	0,568757	0,00714569	0,554722	0,582792
6	84	0,757906	0,00714569	0,743871	0,771941
9	84	0,821996	0,00714569	0,807961	0,836031
12	84	0,864608	0,00714569	0,850572	0,878643
15	84	0,887229	0,00714569	0,873194	0,901264
18	84	0,898411	0,00714569	0,884376	0,912446
21	84	0,912553	0,00714569	0,898518	0,926588

The StatAdvisor

This table shows the mean F1Score for each level of the factors. It also shows the standard error of each mean, which is a measure of its sampling variability. The rightmost two columns show 95,0% confidence intervals for each of the means. You can display these means and intervals by selecting Means Plot from the list of Graphical Options.

Multiple Range Tests for F1Score by Class

Method: 95,0 percent LSD

Multiple Range Tests for F1Score by Class

Method: 95,0 percent LSD

Class	*Count*	*LS Mean*	*LS Sigma*	*Homogeneous Groups*
SVM	294	0,815373	0,00381953	X
KNN	294	0,816473	0,00381953	X

Contrast	*Sig.*	*Difference*	*+/- Limits*

Multiple Range Tests for F1Score by FS

Multiple Range Tests for F1Score by FS

Method: 95,0 percent LSD

FS	Count	LS Mean	LS Sigma	Homogeneous Groups
nofs	294	0,767204	0,00381953	X
mrmr	294	0,864642	0,00381953	X

Contrast	Sig.	Difference	+/- Limits
mrmr - nofs	*	0,0974375	0,0106095

* denotes a statistically significant difference.

The StatAdvisor

This table applies a multiple comparison procedure to determine which means are significantly different from which others. The bottom half of the output shows the estimated difference between each pair of means. An asterisk has been placed next to 1 pair, indicating that this pair shows a statistically significant difference at the 95,0% confidence level. At the top of the page, 2 homogenous groups are identified using columns of X's. Within each column, the levels containing X's form a group of means within which there are no statistically significant differences. The method currently being used to discriminate among the means is Fisher's least significant difference (LSD) procedure. With this method, there is a 5,0% risk of calling each pair of means significantly different when the actual difference equals 0.

Multiple Range Tests for F1Score by LFC

Multiple Range Tests for F1Score by LFC

Method: 95,0 percent LSD

LFC	Count	LS Mean	LS Sigma	Homogeneous Groups
1	196	0,794988	0,00467795	X
2	196	0,826009	0,00467795	X
1.5	196	0,826772	0,00467795	X

Contrast	Sig.	Difference	+/- Limits
1 - 1.5	*	-0,0317846	0,0129939
1 - 2	*	-0,031021	0,0129939
1.5 - 2		0,000763603	0,0129939

* denotes a statistically significant difference.

Multiple Range Tests for F1Score by COV

Multiple Range Tests for F1Score by COV

Method: 95,0 percent LSD

COV	Count	LS Mean	LS Sigma	Homogeneous Groups
35	84	0,792312	0,00714569	X
40	84	0,797923	0,00714569	X
45	84	0,805984	0,00714569	XX
50	84	0,825288	0,00714569	XX

65	84	0,826835	0,00714569	X
55	84	0,829339	0,00714569	X
60	84	0,833777	0,00714569	X

Contrast	*Sig.*	*Difference*	*+/- Limits*
35 - 40		-0,00561104	0,0198486
35 - 45		-0,0136715	0,0198486
35 - 50	*	-0,0329759	0,0198486
35 - 55	*	-0,037027	0,0198486
35 - 60	*	-0,041465	0,0198486
35 - 65	*	-0,0345229	0,0198486
40 - 45		-0,00806048	0,0198486
40 - 50	*	-0,0273648	0,0198486
40 - 55	*	-0,0314159	0,0198486
40 - 60	*	-0,035854	0,0198486
40 - 65	*	-0,0289119	0,0198486
45 - 50		-0,0193044	0,0198486
45 - 55	*	-0,0233554	0,0198486
45 - 60	*	-0,0277935	0,0198486
45 - 65	*	-0,0208514	0,0198486
50 - 55		-0,00405109	0,0198486
50 - 60		-0,00848914	0,0198486
50 - 65		-0,00154701	0,0198486
55 - 60		-0,00443806	0,0198486
55 - 65		0,00250407	0,0198486
60 - 65		0,00694213	0,0198486

* denotes a statistically significant difference.

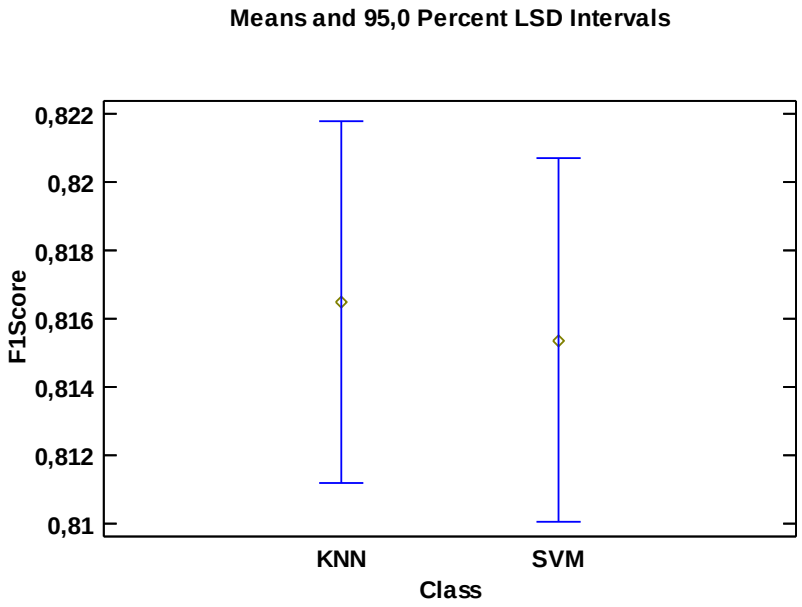
Multiple Range Tests for F1Score by nGenes

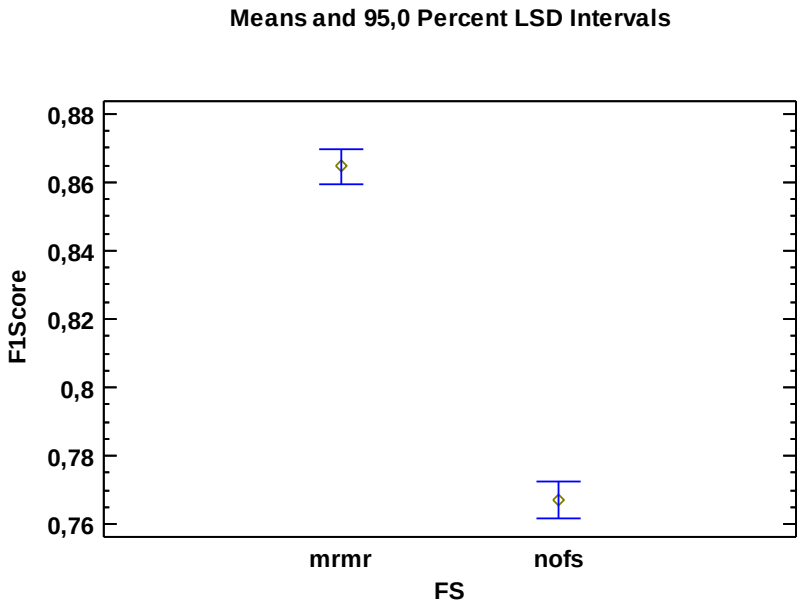
Multiple Range Tests for F1Score by nGenes

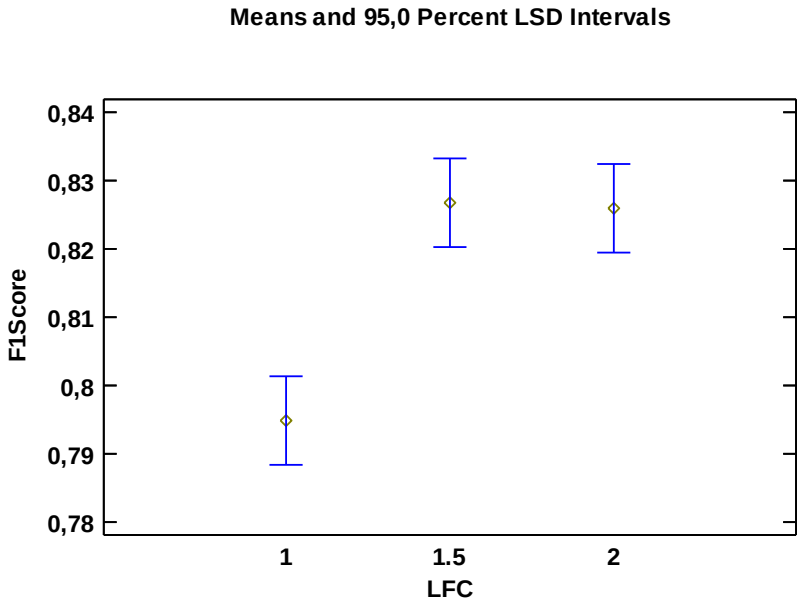
Method: 95,0 percent LSD

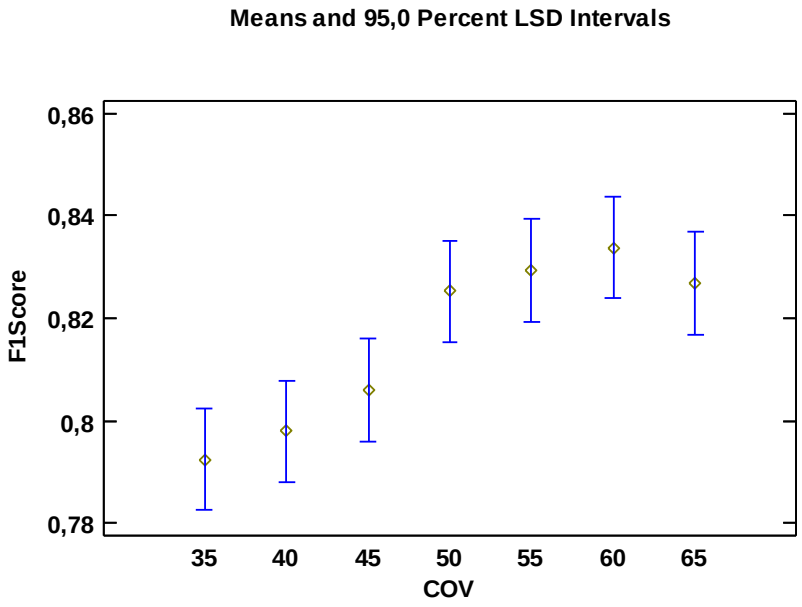
nGenes	*Count*	*LS Mean*	*LS Sigma*	*Homogeneous Groups*
3	84	0,568757	0,00714569	X
6	84	0,757906	0,00714569	X
9	84	0,821996	0,00714569	X
12	84	0,864608	0,00714569	X
15	84	0,887229	0,00714569	X
18	84	0,898411	0,00714569	XX
21	84	0,912553	0,00714569	X

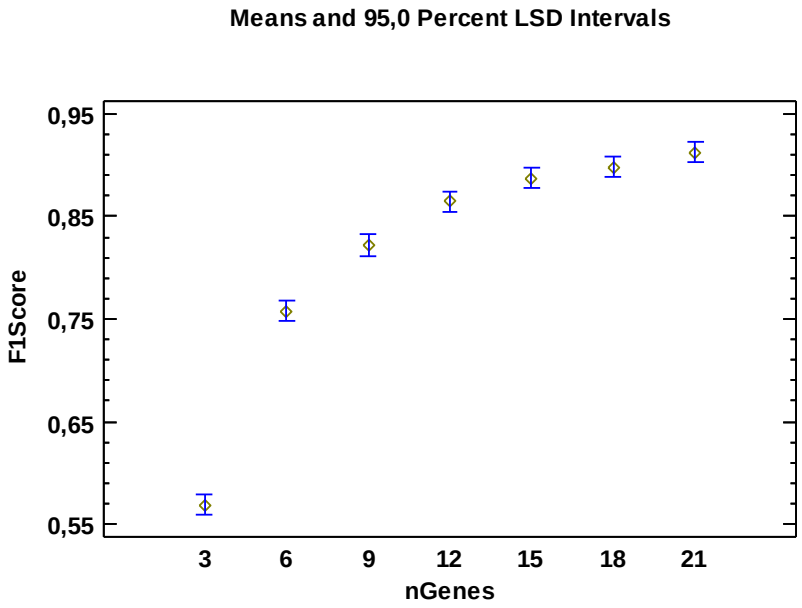
Graphic Results




Means and 95,0 Percent LSD Intervals
0,88
0,86
0,84
0,82
0,8
0,78
0,76
F1Score
mrmr
nofs
FS


Means and 95,0 Percent LSD Intervals
0,84
0,83
0,82
0,81
0,8
0,79
0,78
F1Score
1
1.5
2
LFC


Means and 95,0 Percent LSD Intervals
F1Score
0,86
0,84
0,82
0,8
0,78
35
40
45
50
55
60
65
COV


Means and 95,0 Percent LSD Intervals
F1Score
0,95
0,85
0,75
0,65
0,55
3
6
9
12
15
18
21
nGenes

ANOVA (First Level Interactions).

Analysis of Variance for F1Score - Type III Sums of Squares

Analysis of Variance for F1Score - Type III Sums of Squares

Source	*Sum of Squares*	*Df*	*Mean Square*	*F-Ratio*	*P-Value*
MAIN EFFECTS					
A:Class	0,000177968	1	0,000177968	0,10	0,7466
B:FS	1,39563	1	1,39563	819,58	0,0000
C:LFC	0,128913	2	0,0644563	37,85	0,0000
D:COV	0,141607	6	0,0236011	13,86	0,0000
E:nGenes	7,39956	6	1,23326	724,23	0,0000
INTERACTIONS					
AB	0,00660914	1	0,00660914	3,88	0,0494
AC	0,0242377	2	0,0121188	7,12	0,0009
AD	0,0225239	6	0,00375399	2,20	0,0414
AE	0,00303191	6	0,000505318	0,30	0,9384
BC	0,222048	2	0,111024	65,20	0,0000
BD	0,255062	6	0,0425104	24,96	0,0000
BE	0,671531	6	0,111922	65,73	0,0000
CD	0,309054	12	0,0257545	15,12	0,0000
CE	0,0425097	12	0,00354248	2,08	0,0169
DE	0,0716982	36	0,00199162	1,17	0,2344
RESIDUAL	0,820776	482	0,00170285		
TOTAL (CORRECTED)	11,515	587			

All F-ratios are based on the residual mean square error.

The StatAdvisor

The ANOVA table decomposes the variability of F1Score into contributions due to various factors. Since Type III sums of squares (the default) have been chosen, the contribution of each factor is measured having removed the effects of all other factors. The P-values test the statistical significance of each of the factors. Since 12 P-values are less than 0,05, these factors have a statistically significant effect on F1Score at the 95,0% confidence level.

Table of Least Squares Means for F1Score with 95,0% Confidence Intervals

Table of Least Squares Means for F1Score with 95,0% Confidence Intervals

Level	*Count*	*Mean*	*Stnd. Error*	*Lower Limit*	*Upper Limit*
GRAND MEAN	588	0,815923			
Class					
KNN	294	0,816473	0,00240666	0,811744	0,821202
SVM	294	0,815373	0,00240666	0,810644	0,820102
FS					
mrmr	294	0,864642	0,00240666	0,859913	0,86937
nofs	294	0,767204	0,00240666	0,762475	0,771933
LFC					
1	196	0,794988	0,00294755	0,789196	0,800779
1.5	196	0,826772	0,00294755	0,820981	0,832564
2	196	0,826009	0,00294755	0,820217	0,8318
COV					
35	84	0,792312	0,00450245	0,783466	0,801159
40	84	0,797923	0,00450245	0,789077	0,80677
45	84	0,805984	0,00450245	0,797137	0,814831
50	84	0,825288	0,00450245	0,816441	0,834135
55	84	0,829339	0,00450245	0,820492	0,838186
60	84	0,833777	0,00450245	0,824931	0,842624
65	84	0,826835	0,00450245	0,817988	0,835682
nGenes					
3	84	0,568757	0,00450245	0,55991	0,577604
6	84	0,757906	0,00450245	0,74906	0,766753
9	84	0,821996	0,00450245	0,813149	0,830843
12	84	0,864608	0,00450245	0,855761	0,873454
15	84	0,887229	0,00450245	0,878382	0,896076
18	84	0,898411	0,00450245	0,889564	0,907258
21	84	0,912553	0,00450245	0,903706	0,9214
Class by FS					
KNN;mrmr	147	0,861839	0,00340353	0,855152	0,868527
KNN;nofs	147	0,771107	0,00340353	0,764419	0,777794
SVM;mrmr	147	0,867444	0,00340353	0,860756	0,874132
SVM;nofs	147	0,763301	0,00340353	0,756614	0,769989
Class by LFC					
KNN;1	98	0,804606	0,00416846	0,796415	0,812796
KNN;1.5	98	0,823191	0,00416846	0,815	0,831382
KNN;2	98	0,821623	0,00416846	0,813432	0,829813
SVM;1	98	0,78537	0,00416846	0,777179	0,79356
SVM;1.5	98	0,830354	0,00416846	0,822163	0,838544
SVM;2	98	0,830395	0,00416846	0,822204	0,838585
Class by COV					
KNN;35	42	0,801911	0,00636743	0,789399	0,814422
KNN;40	42	0,804936	0,00636743	0,792424	0,817447
KNN;45	42	0,811565	0,00636743	0,799053	0,824076
KNN;50	42	0,822438	0,00636743	0,809927	0,834949
KNN;55	42	0,826232	0,00636743	0,81372	0,838743
KNN;60	42	0,826287	0,00636743	0,813775	0,838798
KNN;65	42	0,821944	0,00636743	0,809432	0,834455
SVM;35	42	0,782714	0,00636743	0,770203	0,795226
SVM;40	42	0,790911	0,00636743	0,7784	0,803422
SVM;45	42	0,800403	0,00636743	0,787892	0,812915
SVM;50	42	0,828139	0,00636743	0,815627	0,84065
SVM;55	42	0,832447	0,00636743	0,819936	0,844958
SVM;60	42	0,841268	0,00636743	0,828757	0,853779

SVM;65	42	0,831727	0,00636743	0,819215	0,844238
Class by nGenes					
KNN;3	42	0,56502	0,00636743	0,552509	0,577531
KNN;6	42	0,760059	0,00636743	0,747548	0,77257
KNN;9	42	0,823257	0,00636743	0,810746	0,835768
KNN;12	42	0,864539	0,00636743	0,852028	0,877051
KNN;15	42	0,890141	0,00636743	0,87763	0,902653
KNN;18	42	0,901156	0,00636743	0,888645	0,913667
KNN;21	42	0,911139	0,00636743	0,898627	0,92365
SVM;3	42	0,572495	0,00636743	0,559983	0,585006
SVM;6	42	0,755754	0,00636743	0,743242	0,768265
SVM;9	42	0,820735	0,00636743	0,808224	0,833247
SVM;12	42	0,864676	0,00636743	0,852164	0,877187
SVM;15	42	0,884317	0,00636743	0,871805	0,896828
SVM;18	42	0,895666	0,00636743	0,883154	0,908177
SVM;21	42	0,913967	0,00636743	0,901456	0,926478
FS by LFC					
mrmr;1	98	0,870351	0,00416846	0,86216	0,878541
mrmr;1.5	98	0,868	0,00416846	0,85981	0,876191
mrmr;2	98	0,855574	0,00416846	0,847383	0,863764
nofs;1	98	0,719624	0,00416846	0,711434	0,727815
nofs;1.5	98	0,785544	0,00416846	0,777354	0,793735
nofs;2	98	0,796444	0,00416846	0,788253	0,804634
FS by COV					
mrmr;35	42	0,869558	0,00636743	0,857047	0,88207
mrmr;40	42	0,870683	0,00636743	0,858172	0,883194
mrmr;45	42	0,868283	0,00636743	0,855772	0,880795
mrmr;50	42	0,870774	0,00636743	0,858263	0,883286
mrmr;55	42	0,865117	0,00636743	0,852606	0,877629
mrmr;60	42	0,861605	0,00636743	0,849094	0,874116
mrmr;65	42	0,84647	0,00636743	0,833958	0,858981
nofs;35	42	0,715066	0,00636743	0,702555	0,727578
nofs;40	42	0,725164	0,00636743	0,712652	0,737675
nofs;45	42	0,743684	0,00636743	0,731173	0,756196
nofs;50	42	0,779802	0,00636743	0,767291	0,792314
nofs;55	42	0,793561	0,00636743	0,78105	0,806073
nofs;60	42	0,80595	0,00636743	0,793438	0,818461
nofs;65	42	0,807201	0,00636743	0,794689	0,819712
FS by nGenes					
mrmr;3	42	0,684063	0,00636743	0,671552	0,696574
mrmr;6	42	0,836947	0,00636743	0,824436	0,849458
mrmr;9	42	0,87327	0,00636743	0,860759	0,885781
mrmr;12	42	0,897319	0,00636743	0,884808	0,90983
mrmr;15	42	0,91266	0,00636743	0,900148	0,925171
mrmr;18	42	0,920648	0,00636743	0,908137	0,93316
mrmr;21	42	0,927584	0,00636743	0,915073	0,940096
nofs;3	42	0,453452	0,00636743	0,44094	0,465963
nofs;6	42	0,678866	0,00636743	0,666354	0,691377
nofs;9	42	0,770722	0,00636743	0,758211	0,783234
nofs;12	42	0,831896	0,00636743	0,819385	0,844407
nofs;15	42	0,861798	0,00636743	0,849287	0,87431
nofs;18	42	0,876173	0,00636743	0,863662	0,888685
nofs;21	42	0,897521	0,00636743	0,88501	0,910033
LFC by COV					
1;35	28	0,747715	0,00779848	0,732392	0,763038
1;40	28	0,746385	0,00779848	0,731062	0,761708
1;45	28	0,743269	0,00779848	0,727946	0,758592
1;50	28	0,81755	0,00779848	0,802227	0,832873

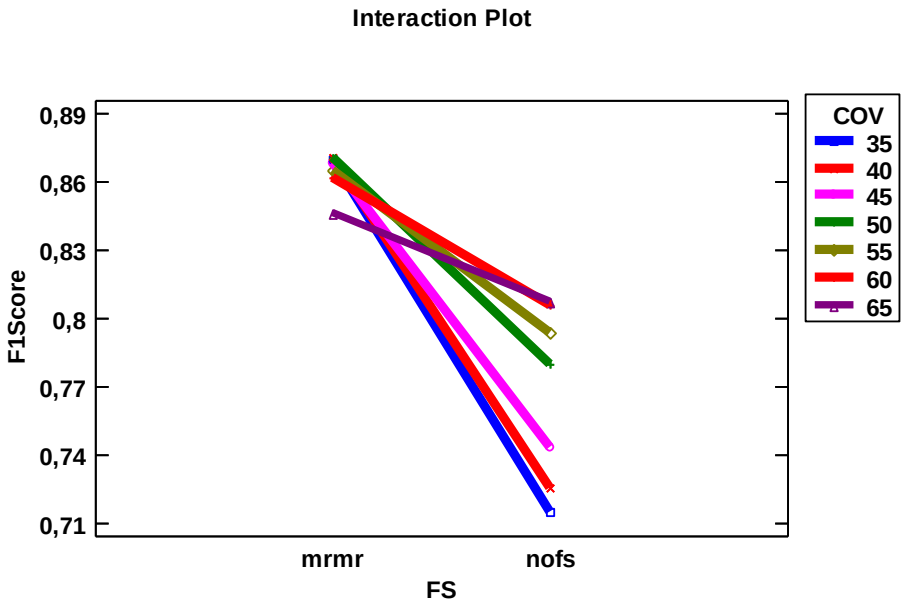
1;55	28	0,832298	0,00779848	0,816975	0,847621
1;60	28	0,831681	0,00779848	0,816358	0,847004
1;65	28	0,846015	0,00779848	0,830692	0,861339
1.5;35	28	0,790848	0,00779848	0,775524	0,806171
1.5;40	28	0,813202	0,00779848	0,797879	0,828525
1.5;45	28	0,842828	0,00779848	0,827504	0,858151
1.5;50	28	0,833201	0,00779848	0,817878	0,848524
1.5;55	28	0,832889	0,00779848	0,817566	0,848213
1.5;60	28	0,840541	0,00779848	0,825218	0,855865
1.5;65	28	0,833897	0,00779848	0,818573	0,84922
2;35	28	0,838375	0,00779848	0,823051	0,853698
2;40	28	0,834183	0,00779848	0,81886	0,849507
2;45	28	0,831855	0,00779848	0,816532	0,847178
2;50	28	0,825114	0,00779848	0,809791	0,840437
2;55	28	0,822831	0,00779848	0,807507	0,838154
2;60	28	0,82911	0,00779848	0,813786	0,844433
2;65	28	0,800594	0,00779848	0,785271	0,815917
LFC by nGenes					
1;3	28	0,536306	0,00779848	0,520983	0,551629
1;6	28	0,723662	0,00779848	0,708339	0,738985
1;9	28	0,788902	0,00779848	0,773579	0,804225
1;12	28	0,846567	0,00779848	0,831244	0,861891
1;15	28	0,875661	0,00779848	0,860338	0,890984
1;18	28	0,884908	0,00779848	0,869585	0,900231
1;21	28	0,908907	0,00779848	0,893584	0,924231
1.5;3	28	0,584809	0,00779848	0,569486	0,600132
1.5;6	28	0,767954	0,00779848	0,752631	0,783277
1.5;9	28	0,838533	0,00779848	0,82321	0,853856
1.5;12	28	0,877323	0,00779848	0,862	0,892646
1.5;15	28	0,895122	0,00779848	0,879799	0,910445
1.5;18	28	0,908344	0,00779848	0,893021	0,923667
1.5;21	28	0,915321	0,00779848	0,899998	0,930644
2;3	28	0,585157	0,00779848	0,569833	0,60048
2;6	28	0,782104	0,00779848	0,76678	0,797427
2;9	28	0,838554	0,00779848	0,82323	0,853877
2;12	28	0,869932	0,00779848	0,854609	0,885255
2;15	28	0,890904	0,00779848	0,875581	0,906227
2;18	28	0,901981	0,00779848	0,886657	0,917304
2;21	28	0,91343	0,00779848	0,898107	0,928753
COV by nGenes					
35;3	12	0,558454	0,0119124	0,535048	0,581861
35;6	12	0,706791	0,0119124	0,683384	0,730197
35;9	12	0,774201	0,0119124	0,750795	0,797608
35;12	12	0,843647	0,0119124	0,82024	0,867053
35;15	12	0,873202	0,0119124	0,849795	0,896608
35;18	12	0,881822	0,0119124	0,858415	0,905228
35;21	12	0,90807	0,0119124	0,884664	0,931477
40;3	12	0,525025	0,0119124	0,501618	0,548431
40;6	12	0,734277	0,0119124	0,71087	0,757683
40;9	12	0,803863	0,0119124	0,780456	0,827269
40;12	12	0,853354	0,0119124	0,829947	0,876761
40;15	12	0,875511	0,0119124	0,852105	0,898918
40;18	12	0,888726	0,0119124	0,865319	0,912132
40;21	12	0,904709	0,0119124	0,881302	0,928115
45;3	12	0,549845	0,0119124	0,526438	0,573251
45;6	12	0,739481	0,0119124	0,716074	0,762888
45;9	12	0,800666	0,0119124	0,777259	0,824073
45;12	12	0,860927	0,0119124	0,83752	0,884334

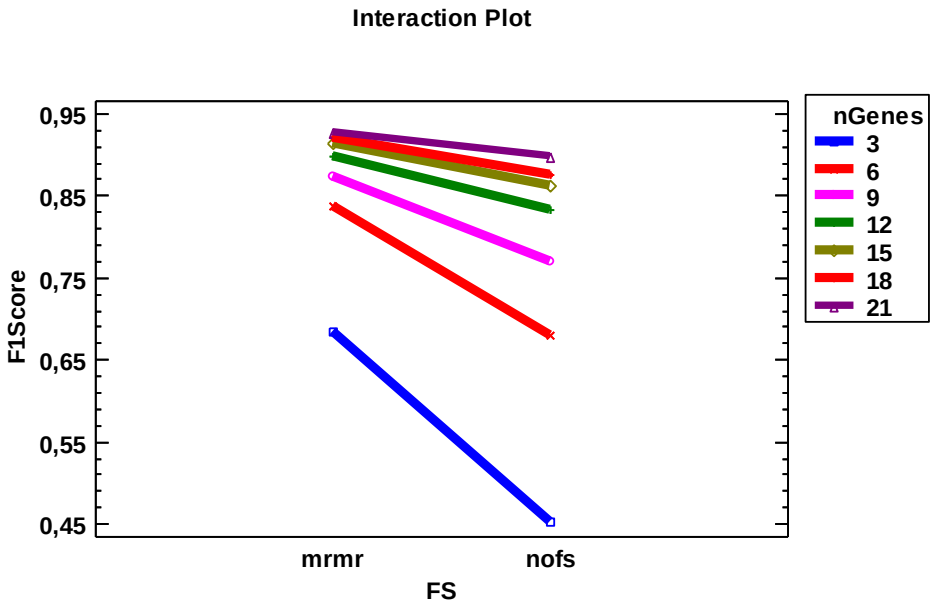
45;15	12	0,89098	0,0119124	0,867574	0,914387
45;18	12	0,887128	0,0119124	0,863722	0,910535
45;21	12	0,91286	0,0119124	0,889453	0,936267
50;3	12	0,582896	0,0119124	0,55949	0,606303
50;6	12	0,771029	0,0119124	0,747623	0,794436
50;9	12	0,836064	0,0119124	0,812657	0,85947
50;12	12	0,873201	0,0119124	0,849794	0,896608
50;15	12	0,893797	0,0119124	0,87039	0,917203
50;18	12	0,903825	0,0119124	0,880419	0,927232
50;21	12	0,916205	0,0119124	0,892799	0,939612
55;3	12	0,571441	0,0119124	0,548035	0,594848
55;6	12	0,781546	0,0119124	0,75814	0,804953
55;9	12	0,848252	0,0119124	0,824845	0,871658
55;12	12	0,872071	0,0119124	0,848664	0,895477
55;15	12	0,896757	0,0119124	0,873351	0,920164
55;18	12	0,914248	0,0119124	0,890841	0,937655
55;21	12	0,92106	0,0119124	0,897654	0,944467
60;3	12	0,600661	0,0119124	0,577254	0,624068
60;6	12	0,784595	0,0119124	0,761188	0,808001
60;9	12	0,849549	0,0119124	0,826142	0,872956
60;12	12	0,880747	0,0119124	0,85734	0,904153
60;15	12	0,895021	0,0119124	0,871614	0,918427
60;18	12	0,908654	0,0119124	0,885247	0,93206
60;21	12	0,917216	0,0119124	0,893809	0,940623
65;3	12	0,592979	0,0119124	0,569572	0,616385
65;6	12	0,787626	0,0119124	0,764219	0,811032
65;9	12	0,841379	0,0119124	0,817972	0,864785
65;12	12	0,868307	0,0119124	0,8449	0,891714
65;15	12	0,885335	0,0119124	0,861928	0,908742
65;18	12	0,904473	0,0119124	0,881066	0,927879
65;21	12	0,907749	0,0119124	0,884343	0,931156

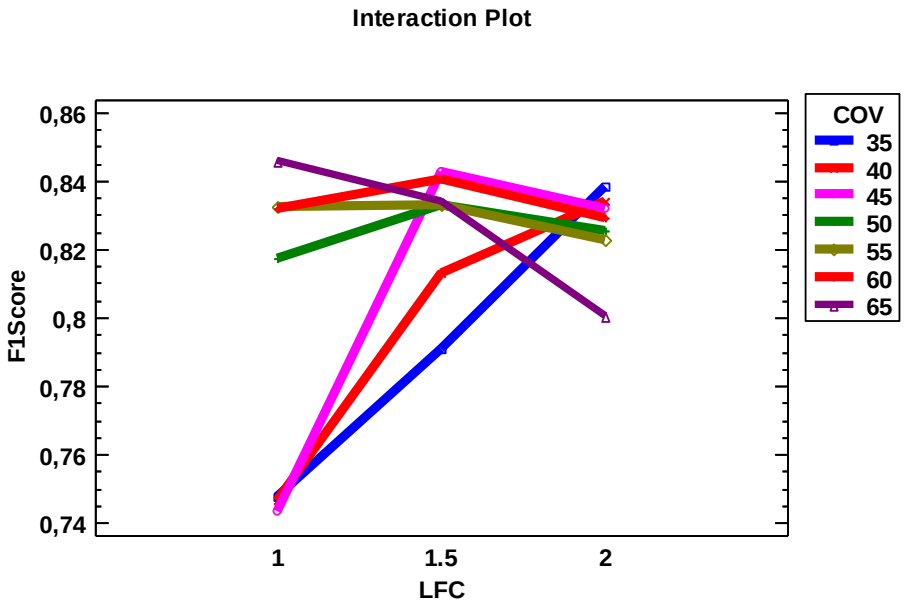
The StatAdvisor
This table shows the mean F1Score for each level of the factors. It also shows the standard error of each mean, which is a measure of its sampling variability. The rightmost two columns show 95,0% confidence intervals for each of the means. You can display these means and intervals by selecting Means Plot from the list of Graphical Options.

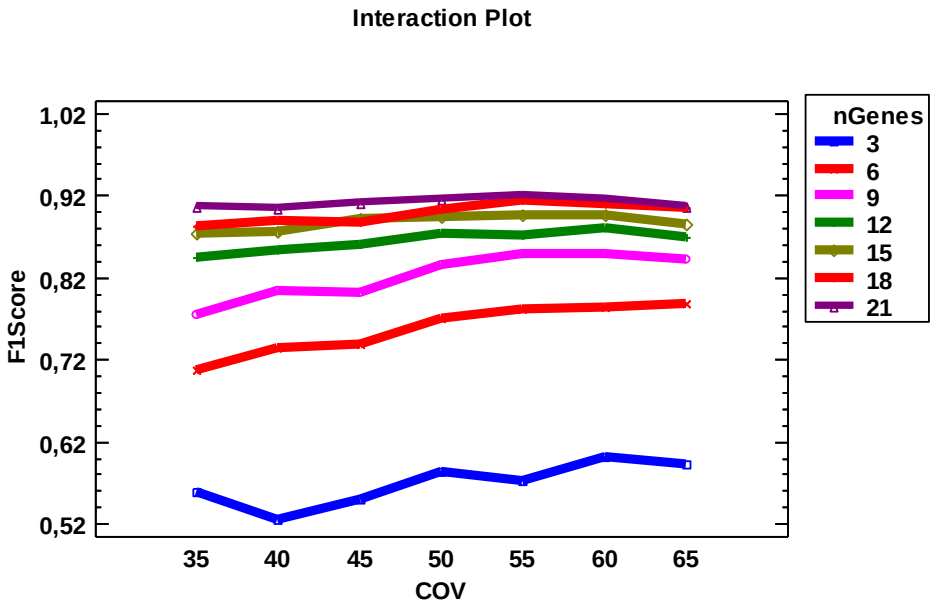
Visual Interaction Analysis:

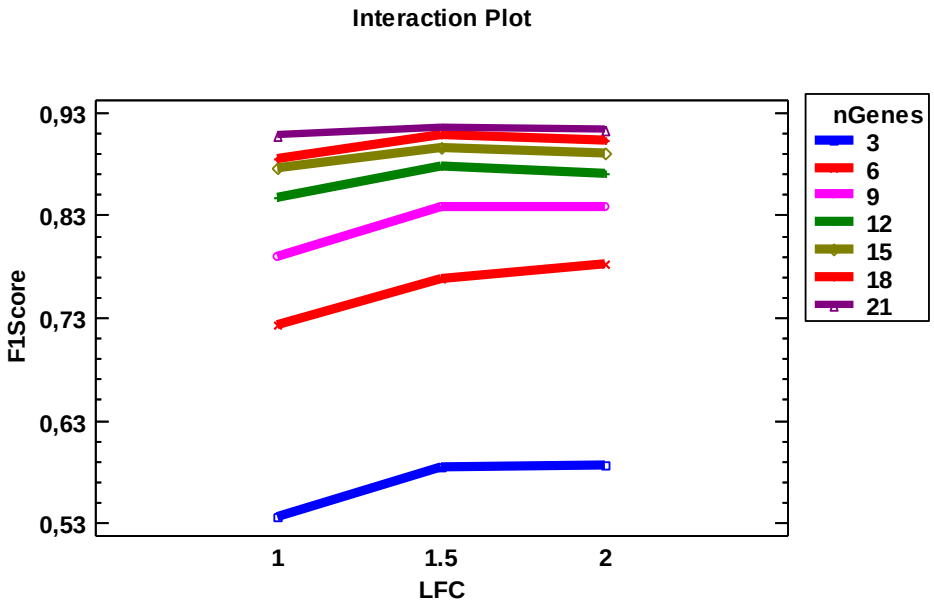

Interaction Plot
0,89
0,86
0,83
0,8
0,77
0,74
0,71
F1Score
mrmr
nofs
FS
LFC
1
1.5
2


Interaction Plot
0,89
0,86
0,83
0,8
0,77
0,74
0,71
F1Score
mrmr
nofs
FS
COV
35
40
45
50
55
60
65


Interaction Plot
0,95
0,85
0,75
0,65
0,55
0,45
F1Score
mrmr
nofs
FS
nGenes
3
6
9
12
15
18
21


Interaction Plot
0,86
0,84
0,82
0,8
0,78
0,76
0,74
F1Score
1
1.5
2
LFC
COV
35
40
45
50
55
60
65


Interaction Plot
1,02
0,92
0,82
0,72
0,62
0,52
F1Score
35
40
45
50
55
60
65
COV
nGenes
3
6
9
12
15
18
21


Interaction Plot
0,93
0,83
0,73
0,63
0,53
F1Score
1
1.5
2
LFC
nGenes
3
6
9
12
15
18
21