Bolded → In the model

Yellow highlight → not significant

Green in old model, but not in new

```
> rcorr(as.matrix(newVarModel[woDBMS]), type="pearson")
```

	opCt	NUMFROM	NumAgg	pk_pr	sq_pr	CEPS	num_repeats
opCt	1.00	0.24	0.31	0.39	0.15	**0.59**	**0.45**
NUMFROM	0.24	1.00	-0.02	0.26	0.16	**0.45**	**0.11**
NumAgg	0.31	-0.02	1.00	0.03	0.02	**0.07**	**0.04**
pk_pr	0.39	0.26	0.03	1.00	0.40	**0.26**	**0.15**
sq_pr	0.15	0.16	0.02	0.40	1.00	**0.17**	**-0.05**
CEPS	0.59	0.45	0.07	0.26	0.17	1.00	0.56
num_repeats	0.45	0.11	0.04	0.15	-0.05	0.56	1.00
NumDiscontOperators	0.56	0.36	0.28	0.22	0.21	**0.63**	0.09
log_SUBOPT	0.29	0.31	0.03	0.12	0.18	0.50	0.25
sec_idx_pr	0.40	0.15	0.04	0.41	-0.23	**0.30**	**0.35**

	NumDiscontOperators	log_SUBOPT	sec_idx_pr
opCt	**0.56**	**0.29**	0.40
NUMFROM	**0.36**	0.31	0.15
NumAgg	**0.28**	0.03	0.04
pk_pr	**0.22**	0.12	0.41
sq_pr	**0.21**	0.18	-0.23
CEPS	**0.63**	**0.50**	0.30
num_repeats	0.09	**0.25**	0.35
NumDiscontOperators	1.00	**0.48**	0.15
log_SUBOPT	0.48	1.00	0.07
sec_idx_pr	**0.15**	0.07	1.00

n= 1757

P

	opCt	NUMFROM	NumAgg	pk_pr	sq_pr	CEPS	num_repeats
opCt		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
NUMFROM	0.0000		0.3392	0.0000	0.0000	0.0000	0.0000
NumAgg	0.0000	0.3392		0.1503	0.4338	0.0025	0.1385
pk_pr	0.0000	0.0000	0.1503		0.0000	0.0000	0.0000
sq_pr	0.0000	0.0000	0.4338	0.0000		0.0000	0.0441
CEPS	0.0000	0.0000	0.0025	0.0000	0.0000		0.0000
num_repeats	0.0000	0.0000	0.1385	0.0000	0.0441	0.0000	
NumDiscontOperators	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Log_SUBOPT	0.0000	0.0000	0.2153	0.0070	0.0000	0.0000	0.0000
sec_idx_pr	0.0000	0.0000	0.0752	0.0000	0.0000	0.0000	0.0000

	NumDiscontOperators	log_SUBOPT	sec_idx_pr
opCt	0.0000	0.0000	0.0000
NUMFROM	0.0000	0.0000	0.0000
NumAgg	0.0000	0.2153	0.0752
pk_pr	0.0000	0.0070	0.0000
sq_pr	0.0000	0.0000	0.0000
CEPS	0.0000	0.0000	0.0000
num_repeats	0.0000	0.0000	0.0000
NumDiscontOperators		0.0000	0.0000
SUBOPT	0.0000		0.0028
sec_idx_pr	0.0000	0.0028	

```
> rcorr(as.matrix(newVarModel[inter]), type="pearson")
```

	int_si_op	int_si_nf	int_si_na	int_si_pp	pk_pr	int_si_sp
int_si_op	1.00	0.87	0.70	0.88	0.36	NaN
int_si_nf	0.87	1.00	0.70	0.98	0.40	NaN
int_si_na	0.70	0.70	1.00	0.72	0.29	NaN
int_si_pp	0.88	0.98	0.72	1.00	0.41	NaN
pk_pr	0.36	0.40	0.29	0.41	1.00	NaN
int_si_sp	NaN	NaN	NaN	NaN	NaN	1
CEPS	0.44	0.33	0.22	0.30	0.26	NaN
num_repeats	0.53	0.35	0.26	0.35	0.15	NaN
NumDiscontOperators	0.23	0.17	0.20	0.15	0.22	NaN
sec_idx_pr	0.88	0.98	0.72	1.00	0.41	NaN

	CEPS	num_repeats	NumDiscontOperators	sec_idx_pr
int_si_op	**0.44**	**0.53**	**0.23**	0.88
int_si_nf	**0.33**	**0.35**	**0.17**	0.98
int_si_na	**0.22**	**0.26**	**0.20**	0.72
int_si_pp	**0.30**	**0.35**	**0.15**	1.00
pk_pr	0.26	0.15	0.22	0.41
int_si_sp	NaN	NaN	NaN	NaN
CEPS	1.00	0.56	0.63	0.30
num_repeats	0.56	1.00	0.09	0.35
NumDiscontOperators	0.63	0.09	1.00	0.15
sec_idx_pr	0.30	0.35	0.15	1.00

n= 1757

	int_si_op	int_si_nf	int_si_na	int_si_pp	pk_pr	int_si_sp
int_si_op		0	0	0	0	
int_si_nf	0		0	0	0	
int_si_na	0	0		0	0	
int_si_pp	0	0	0		0	
pk_pr	0	0	0	0		
int_si_sp						
CEPS	0	0	0	0	0	
num_repeats	0	0	0	0	0	
NumDiscontOperators	0	0	0	0	0	
sec_idx_pr	0	0	0	0	0	

	CEPS	num_repeats	NumDiscontOperators	sec_idx_pr
int_si_op	0	0	0	0
int_si_nf	0	0	0	0
int_si_na	0	0	0	0
int_si_pp	0	0	0	0
pk_pr	0	0	0	0
int_si_sp				
CEPS		0	0	0
num_repeats	0		0	0
NumDiscontOperators	0	0		0
sec_idx_pr	0	0	0	

```
Call:
lm(formula = log_SUBOPT ~ NumDiscontOperators + num_repeats +
    CEPS + pk_pr + opCt, data = newVarModel)
# added back opCt & pk_pr
Residuals:
    Min      1Q  Median      3Q     Max
-1.6963 -0.2655 -0.1030  0.2698  1.4076

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          0.2439227  0.0223886  10.895  < 2e-16 ***
NumDiscontOperators  0.0160000  0.0013619  11.748  < 2e-16 ***
num_repeats          0.0017539  0.0004096   4.282 1.95e-05 ***
CEPS                 0.0374127  0.0046781   7.997 2.29e-15 ***
pk_pr                0.0027903  0.0208573   0.134    0.894
opCt                -0.0266091  0.0061343  -4.338 1.52e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3962 on 1751 degrees of freedom
Multiple R-squared:  0.3026,    Adjusted R-squared:  0.3006
F-statistic: 151.9 on 5 and 1751 DF,  p-value: < 2.2e-16


Call:
lm(formula = log_SUBOPT ~ NumDiscontOperators + num_repeats +
    CEPS, data = newVarModel)

Residuals:
   Min     1Q Median     3Q    Max
-1.646 -0.265 -0.106  0.275  1.418
```

```
Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         0.175150   0.015791   11.09  < 2e-16 ***
NumDiscontOperators 0.013399   0.001239   10.82  < 2e-16 ***
num_repeats         0.001108   0.000385    2.88    0.004 **
CEPS                0.035984   0.004678    7.69  2.4e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.4 on 1753 degrees of freedom
Multiple R-squared:  0.294,     Adjusted R-squared:  0.293
F-statistic:  244 on 3 and 1753 DF,  p-value: <2e-16
-----------------------------------
>
> fit_NumRepeats <- lm(num_repeats ~ opCt + NUMFROM + NumAgg + pk_pr + sq_pr +
int_si_op + int_si_nf + int_si_na + int_si_pp + int_si_sp, data=newVarModel)
> fit_CEPS <- lm(CEPS ~ opCt + NUMFROM + NumAgg + pk_pr + sq_pr + int_si_op +
int_si_nf + int_si_na + int_si_pp + int_si_sp, data=newVarModel)
> fit_NumDisconOp <- lm(NumDiscontOperators ~ opCt + NUMFROM + NumAgg + pk_pr +
sq_pr + int_si_op + int_si_nf + int_si_na + int_si_pp + int_si_sp,
data=newVarModel)
>
> summary(fit_NumRepeats) # show results


Call:
lm(formula = num_repeats ~ opCt + NUMFROM + NumAgg + pk_pr +
    sq_pr + int_si_op + int_si_nf + int_si_na + int_si_pp + int_si_sp,
    data = newVarModel)


Residuals:
     Min       1Q   Median       3Q      Max
-103.677   -5.359   -3.026    0.254  176.158
```

```
Coefficients: (1 not defined because of singularities)
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.0304     2.8905  -1.394  0.16338
opCt          1.3898     0.4662   2.981  0.00291 **
NUMFROM       1.4433     0.8514   1.695  0.09021 .
NumAgg       -0.4465     1.5290  -0.292  0.77029
pk_pr        -2.6299     1.7338  -1.517  0.12948
sq_pr         1.9778     2.0614   0.959  0.33747
int_si_op    11.1835     0.6965  16.056  < 2e-16 ***
int_si_nf    -2.9195     2.5813  -1.131  0.25820
int_si_na   -13.8289     3.3594  -4.116 4.03e-05 ***
int_si_pp   -20.8386     9.7150  -2.145  0.03209 *
int_si_sp         NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 26.4 on 1747 degrees of freedom
Multiple R-squared:  0.3496,    Adjusted R-squared:  0.3462
F-statistic: 104.3 on 9 and 1747 DF,  p-value: < 2.2e-16

> summary(fit_CEPS) # show results

Call:
lm(formula = CEPS ~ opCt + NUMFROM + NumAgg + pk_pr + sq_pr +
    int_si_op + int_si_nf + int_si_na + int_si_pp + int_si_sp,
    data = newVarModel)

Residuals:
    Min      1Q  Median      3Q     Max
-7.2173 -1.6064 -0.3485  1.2129 28.7143

Coefficients: (1 not defined because of singularities)
```

```
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -2.24051    0.27038  -8.287 2.29e-16 ***
opCt         0.58097    0.04361  13.322  < 2e-16 ***
NUMFROM      1.25795    0.07964  15.795  < 2e-16 ***
NumAgg      -0.24458    0.14303  -1.710  0.08745 .
pk_pr       -0.66982    0.16218  -4.130 3.80e-05 ***
sq_pr        1.28557    0.19282   6.667 3.49e-11 ***
int_si_op    0.45322    0.06515   6.956 4.93e-12 ***
int_si_nf    1.06967    0.24146   4.430 1.00e-05 ***
int_si_na   -0.81692    0.31424  -2.600  0.00941 **
int_si_pp   -4.49056    0.90876  -4.941 8.50e-07 ***
int_si_sp         NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.47 on 1747 degrees of freedom
Multiple R-squared:  0.4876,    Adjusted R-squared:  0.485
F-statistic: 184.7 on 9 and 1747 DF,  p-value: < 2.2e-16

> summary(fit_NumDisconOp) # show results

Call:
lm(formula = NumDiscontOperators ~ opCt + NUMFROM + NumAgg +
    pk_pr + sq_pr + int_si_op + int_si_nf + int_si_na + int_si_pp +
    int_si_sp, data = newVarModel)

Residuals:
    Min      1Q  Median      3Q     Max
-22.396  -5.581  -0.980   4.528  61.141

Coefficients: (1 not defined because of singularities)
            Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) -11.4396     0.9009 -12.699  < 2e-16 ***
opCt          3.0181     0.1453  20.771  < 2e-16 ***
NUMFROM       2.7553     0.2653  10.384  < 2e-16 ***
NumAgg        1.8184     0.4765   3.816 0.000140 ***
pk_pr        -1.8730     0.5404  -3.466 0.000541 ***
sq_pr         2.6021     0.6425   4.050 5.34e-05 ***
int_si_op    -1.5551     0.2171  -7.164 1.15e-12 ***
int_si_nf     2.7946     0.8045   3.474 0.000526 ***
int_si_na     3.1654     1.0470   3.023 0.002537 **
int_si_pp    -4.6419     3.0278  -1.533 0.125439
int_si_sp         NA         NA      NA       NA
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.228 on 1747 degrees of freedom
Multiple R-squared:  0.4214,    Adjusted R-squared:  0.4184
F-statistic: 141.3 on 9 and 1747 DF,  p-value: < 2.2e-16
```