

# MVC Architecture Code Smell data exploration

In [36]:

```
# Installing required Libraries
install.packages("correlation")
install.packages("bayestestR")
install.packages("see")
install.packages("ggplot2")
install.packages("tidyR")
install.packages("poorman")
```

There is a binary version available but the source version is later:

binary	source	needs_compilation
correlation	0.6.1	0.8.1
		FALSE

installing the source package 'correlation'

There is a binary version available but the source version is later:

binary	source	needs_compilation
bayestestR	0.9.0	0.12.1
		FALSE

installing the source package 'bayestestR'

There is a binary version available but the source version is later:

binary	source	needs_compilation
see	0.6.3	0.7.0
		FALSE

installing the source package 'see'

There is a binary version available but the source version is later:

binary	source	needs_compilation
ggplot2	3.3.3	3.3.6
		FALSE

installing the source package 'ggplot2'

There is a binary version available but the source version is later:

binary	source	needs_compilation
tidyR	1.1.3	1.2.0
		TRUE

Binaries will be installed  
package 'tidyR' successfully unpacked and MD5 sums checked

Warning message:

```
"cannot remove prior installation of package 'tidyR'"Warning message in file.copy(savedcopy, lib, recursive = TRUE):
"problem copying C:\Users\ranja\anaconda3\envs\R-Studio\Lib\R\library\00LOCK\tidyR\libs\x64\tidyR.dll to C:\Users\ranja\anaconda3\envs\R-Studio\Lib\R\library\tidyR\libs\x64\tidyR.dll: Permission denied"Warning message:
"restored 'tidyR'"
```

The downloaded binary packages are in  
 C:\Users\ranja\AppData\Local\Temp\Rtmp69TAt8\downloaded\_packages  
 package 'poorman' successfully unpacked and MD5 sums checked

The downloaded binary packages are in  
 C:\Users\ranja\AppData\Local\Temp\Rtmp69TAt8\downloaded\_packages

In [37]:

```
#Loading the installed libraries
library(readxl)      # reading in data
library(ggplot2)      # visualizing data
library(gridExtra)    # combining multiple plots
library(corrgram)     # visualizing data
library(corrplot)     # visualizing data
library(Hmisc)        # produces correlation matrices with p-values
library(ppcor)        # assesses partial correlations
library(tidyverse)    #general data modeling, wrangling, and visualization tasks
```

Error in library(corrgram): there is no package called 'corrgram'  
 Traceback:

1. library(corrgram)

In [38]:

```
#reading MVC final input file
file_mvc <- read.csv(file = "MVC_input.csv", header=TRUE, stringsAsFactors = F)
```

In [39]:

```
#printing the top rows of file
head(file_mvc)
```

ID	Magic.Number	Long.Statement	Missing.Defaults	Complex.Conditional	Unutilized.abstraction	Cyclic.Dependency	Insufficient.Modularization	
1	153	43	1	0	16	0	4	
2	1	16	0	1	12	3	2	
3	53	32	3	0	94	8	1	
4	3	3	0	0	5	2	2	
5	3	1	1	0	5	0	0	

ID	Magic.Number	Long.Statement	Missing.Defaults	Complex.Conditional	Unutilized.abstraction	Cyclic.Dependency	Insufficient.Modularization	
6	640	165	31	16	219	40	10	

In [40]:

```
str(file_mvc) # shows the structure of the data frame
summary(file_mvc) #provides summary statistics on the columns of the data frame
```

```
'data.frame': 70 obs. of 11 variables:
 $ ID                  : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Magic.Number         : int 153 1 53 3 3 640 80 2 92 78 ...
 $ Long.Statement       : int 43 16 32 3 1 165 49 49 15 11 ...
 $ Missing.Defaults    : int 1 0 3 0 1 31 0 1 3 1 ...
 $ Complex.Conditional : int 0 1 0 0 0 16 14 8 5 10 ...
 $ Unutilized.abstraction : int 16 12 94 5 5 219 35 25 53 25 ...
 $ Cyclic.Dependency   : int 0 3 8 2 0 40 12 5 0 7 ...
 $ Insufficient.Modularization: int 4 2 1 2 0 10 4 0 5 6 ...
 $ Deficient.Encapsulation : int 2 5 15 5 3 204 27 23 19 14 ...
 $ Total.Design.Smells  : int 22 23 140 7 9 491 81 60 91 51 ...
 $ Total.Code.Smells   : int 213 19 93 6 5 947 190 76 131 113 ...
   ID      Magic.Number     Long.Statement  Missing.Defaults
Min. : 1.00  Min. : 0.00  Min. : 0.00  Min. : 0.000
1st Qu.:18.25  1st Qu.: 13.5  1st Qu.: 13.50  1st Qu.: 0.000
Median :35.50  Median : 54.5  Median : 27.50  Median : 1.000
Mean   :35.50  Mean   : 234.7  Mean   : 52.60  Mean   : 4.057
3rd Qu.:52.75  3rd Qu.: 176.0  3rd Qu.: 52.25  3rd Qu.: 2.000
Max.  :70.00  Max.  :4855.0  Max.  :540.00  Max.  :142.000
Complex.Conditional Unutilized.abstraction Cyclic.Dependency
Min. : 0.00  Min. : 1.00  Min. : 0.000
1st Qu.: 0.00  1st Qu.: 16.50  1st Qu.: 0.000
Median : 1.00  Median : 30.00  Median : 0.000
Mean   : 5.90  Mean   : 57.01  Mean   : 4.414
3rd Qu.: 5.75  3rd Qu.: 53.00  3rd Qu.: 5.000
Max.  :120.00  Max.  :442.00  Max.  :50.000
Insufficient.Modularization Deficient.Encapsulation Total.Design.Smells
Min. : 0.000  Min. : 0.00  Min. : 5.00
1st Qu.: 0.000  1st Qu.: 2.00  1st Qu.: 28.25
Median : 1.000  Median : 4.00  Median : 57.00
Mean   : 4.371  Mean   : 14.19  Mean   :106.53
3rd Qu.: 4.750  3rd Qu.: 10.00  3rd Qu.: 90.75
Max.  :52.000  Max.  :204.00  Max.  :838.00
Total.Code.Smells
Min. : 2.00
```

```
1st Qu.: 59.75
Median : 110.50
Mean   : 329.14
3rd Qu.: 251.00
Max.   :5172.00
```

## checking outliers with box plot

In [41]:

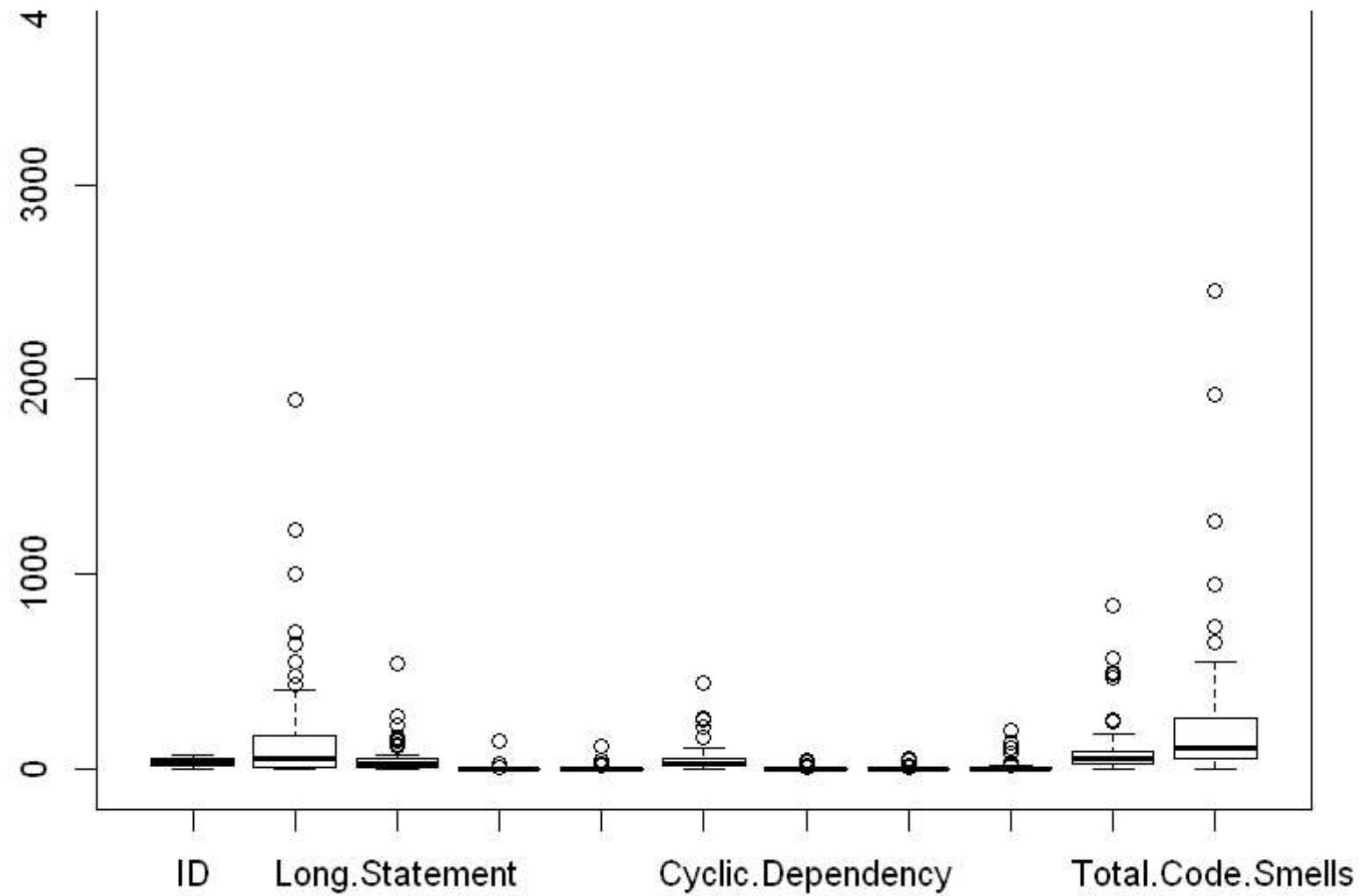
```
# Create a boxplot of the dataset, outliers are shown as two distinct points
boxplot(file_mvc)$out
```

```
1.640
2.481
3.706
4.554
5.1000
6.439
7.1897
8.4855
9.1224
10.165
11.156
12.157
13.224
14.122
15.125
16.270
17.540
18.156
19.145
20.31
21.9
22.7
23.7
```

24. 8  
25. 15  
26. 7  
27. 142  
28. 16  
29. 29  
30. 45  
31. 32  
32. 120  
33. 219  
34. 257  
35. 264  
36. 160  
37. 442  
38. 257  
39. 252  
40. 40  
41. 13  
42. 19  
43. 14  
44. 50  
45. 17  
46. 19  
47. 14  
48. 17  
49. 52  
50. 46  
51. 14  
52. 17  
53. 15  
54. 204  
55. 27  
56. 23

57.41  
58.114  
59.24  
60.140  
61.82  
62.491  
63.570  
64.471  
65.241  
66.838  
67.258  
68.571  
69.502  
70.947  
71.655  
72.733  
73.1270  
74.1269  
75.1925  
76.5172  
77.2458





There are outliers present in all the columns.

In [ ]:

```
#making a copy of original file before changing the column name
df_mvc <- file_mvC
```

Renaming the column name for our ease in furthur calculations

```
In [42]: #LOAD LIBRARY Library(tidyverse)
#rename the column name
names(file_mvc)[names(file_mvc) == "Magic.Number"] <- "MN"
names(file_mvc)[names(file_mvc) == "Long.Statement"] <- "LS"
names(file_mvc)[names(file_mvc) == "Missing.Defaults"] <- "MD"
names(file_mvc)[names(file_mvc) == "Complex.Conditional"] <- "CC"
names(file_mvc)[names(file_mvc) == "Unutilized.abstraction"] <- "Ua"
names(file_mvc)[names(file_mvc) == "Cyclic.Dependency"] <- "CD"
names(file_mvc)[names(file_mvc) == "Insufficient.Modularization"] <- "IM"
names(file_mvc)[names(file_mvc) == "Deficient.Encapsulation"] <- "DE"
names(file_mvc)[names(file_mvc) == "Total.Design.Smells"] <- "TDS"
names(file_mvc)[names(file_mvc) == "Total.Code.Smells"] <- "TCS"
```

```
In [44]: #making a copy of dataframe
df <- file_mvc
```

We can see there are lot of outliers present in our dataset.

We are targeting to remove outliers from some columns for better correlation analysis.

```
In [46]: # finding outliers in our dataset - Long.Statement
Q1 <- quantile(file_mvc$MN, probs=c(.25, .75), na.rm = FALSE)
Q2 <- quantile(file_mvc$LS, probs=c(.25, .75), na.rm = FALSE)
Q3 <- quantile(file_mvc$MD, probs=c(.25, .75), na.rm = FALSE)
Q4 <- quantile(file_mvc$CC, probs=c(.25, .75), na.rm = FALSE)
Q5 <- quantile(file_mvc$Ua, probs=c(.25, .75), na.rm = FALSE)
Q6 <- quantile(file_mvc$CD, probs=c(.25, .75), na.rm = FALSE)
Q7 <- quantile(file_mvc$IM, probs=c(.25, .75), na.rm = FALSE)
Q8 <- quantile(file_mvc$DE, probs=c(.25, .75), na.rm = FALSE)
Q9 <- quantile(file_mvc$TDS, probs=c(.25, .75), na.rm = FALSE)
Q10 <- quantile(file_mvc$TCS, probs=c(.25, .75), na.rm = FALSE)

# how to find outliers in r - calculate Interquartile Range
iqr1 <- IQR(file_mvc$MN)
iqr2 <- IQR(file_mvc$LS)
iqr3 <- IQR(file_mvc$MD)
iqr4 <- IQR(file_mvc$CC)
iqr5 <- IQR(file_mvc$Ua)
iqr6 <- IQR(file_mvc$CD)
iqr7 <- IQR(file_mvc$IM)
iqr8 <- IQR(file_mvc$DE)
iqr9 <- IQR(file_mvc$TDS)
```

```
iqr10 <- IQR(file_mvc$TCS)

# how to find outliers in r - upper and Lower range
#up <- Q[2]+1.5*iqr # Upper Range
#Low<- Q[1]-1.5*iqr # Lower Range

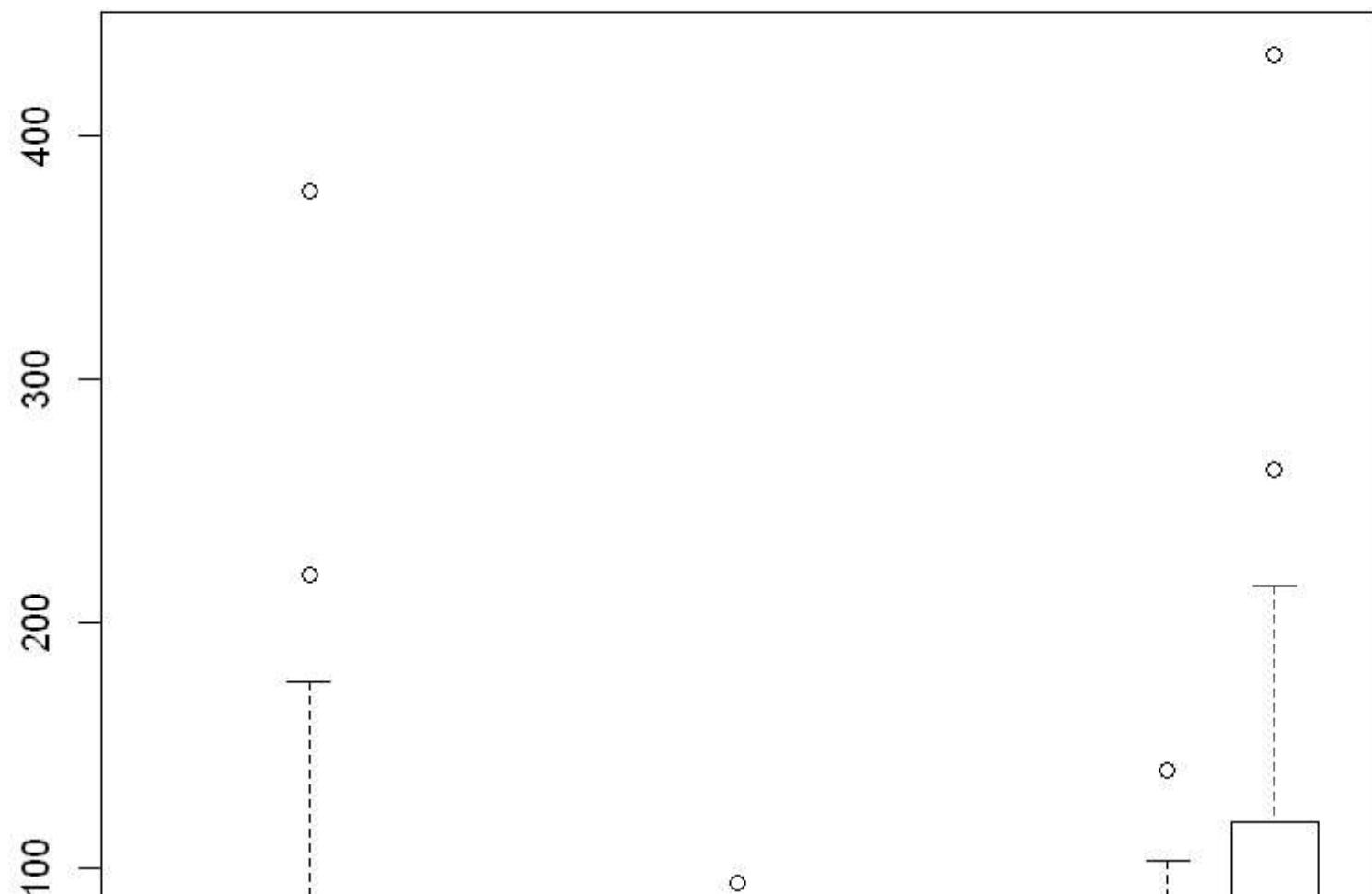
# how to remove outliers in (the removal)
file_mvc<- subset(file_mvc, file_mvc$MN > (Q1[1] - 1.5*iqr1) & file_mvc$MN <(Q1[2]+1.5*iqr1))
file_mvc<- subset(file_mvc, file_mvc$LS > (Q2[1] - 1.5*iqr2) & file_mvc$LS< (Q2[2]+1.5*iqr2))
file_mvc<- subset(file_mvc, file_mvc$MD > (Q3[1] - 1.5*iqr3) & file_mvc$MD< (Q3[2]+1.5*iqr3))
file_mvc<- subset(file_mvc, file_mvc$CC > (Q4[1] - 1.5*iqr4) & file_mvc$CC< (Q4[2]+1.5*iqr4))
file_mvc<- subset(file_mvc, file_mvc$Ua > (Q5[1] - 1.5*iqr5) & file_mvc$Ua< (Q5[2]+1.5*iqr5))
file_mvc<- subset(file_mvc, file_mvc$CD > (Q6[1] - 1.5*iqr6) & file_mvc$CD< (Q6[2]+1.5*iqr6))
file_mvc<- subset(file_mvc, file_mvc$IM > (Q7[1] - 1.5*iqr7) & file_mvc$IM< (Q7[2]+1.5*iqr7))
file_mvc<- subset(file_mvc, file_mvc$DE > (Q8[1] - 1.5*iqr8) & file_mvc$DE< (Q8[2]+1.5*iqr8))
file_mvc<- subset(file_mvc, file_mvc$TDS >(Q9[1] - 1.5*iqr9) & file_mvc$TDS<(Q9[2]+1.5*iqr9))
file_mvc<- subset(file_mvc, file_mvc$TCS >(Q10[1] - 1.5*iqr10) & file_mvc$TCS<(Q10[2]+1.5*iqr10))
```

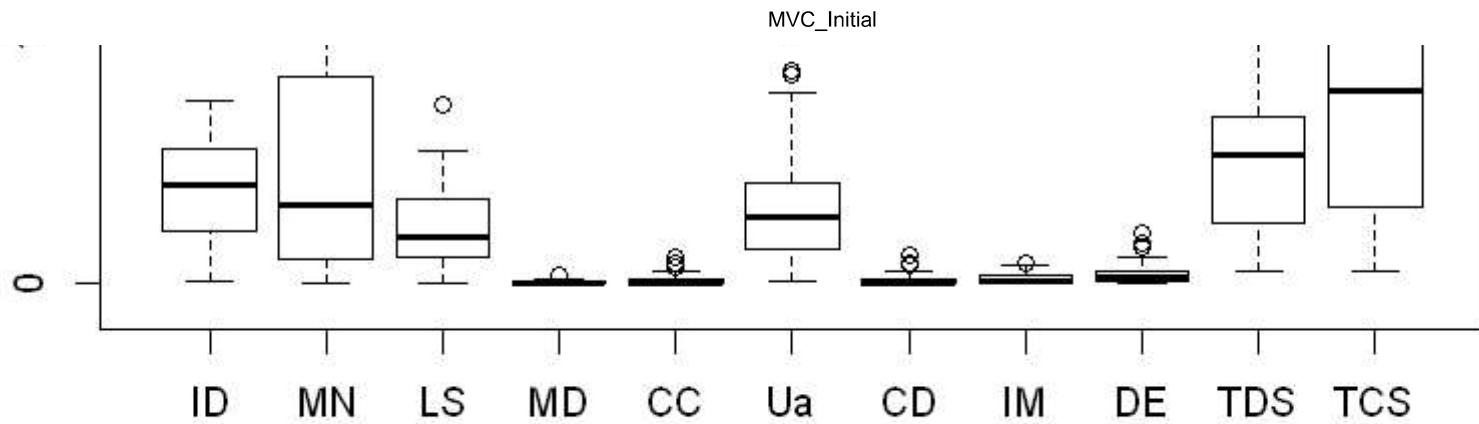
In [47]:

```
# Create a boxplot of the dataset, outliers are shown as two distinct points
boxplot(file_mvc)$out
```

1.220  
2.377  
3.67  
4.3  
5.3  
6.10  
7.6  
8.8  
9.94  
10.79  
11.80  
12.8  
13.7  
14.11  
15.11  
16.7

17.8  
18.15  
19.19  
20.14  
21.140  
22.263  
23.433





We can see above that the outliers has been decreased in a certain amount and we can go for finding correlation between columns.

In [49]:

```
# Correlation test between multiple variables
round(cor(file_mvc, use = "pairwise.complete.obs"), 2)
```

	ID	MN	LS	MD	CC	Ua	CD	IM	DE	TDS	TCS
ID	1.00	-0.02	-0.04	-0.25	-0.04	-0.17	0.07	-0.28	-0.28	-0.19	-0.03
MN	-0.02	1.00	0.24	0.08	0.20	0.04	0.08	0.11	0.01	-0.05	0.97
LS	-0.04	0.24	1.00	0.09	-0.14	0.24	0.06	0.09	-0.03	0.38	0.44
MD	-0.25	0.08	0.09	1.00	0.13	0.24	-0.02	0.06	0.67	0.24	0.12
CC	-0.04	0.20	-0.14	0.13	1.00	0.19	0.19	0.38	0.41	0.21	0.21
Ua	-0.17	0.04	0.24	0.24	0.19	1.00	0.11	0.09	0.25	0.88	0.10
CD	0.07	0.08	0.06	-0.02	0.19	0.11	1.00	0.20	0.40	0.17	0.09
IM	-0.28	0.11	0.09	0.06	0.38	0.09	0.20	1.00	0.27	0.21	0.18
DE	-0.28	0.01	-0.03	0.67	0.41	0.25	0.40	0.27	1.00	0.32	0.04
TDS	-0.19	-0.05	0.38	0.24	0.21	0.88	0.17	0.21	0.32	1.00	0.06
TCS	-0.03	0.97	0.44	0.12	0.21	0.10	0.09	0.18	0.04	0.06	1.00

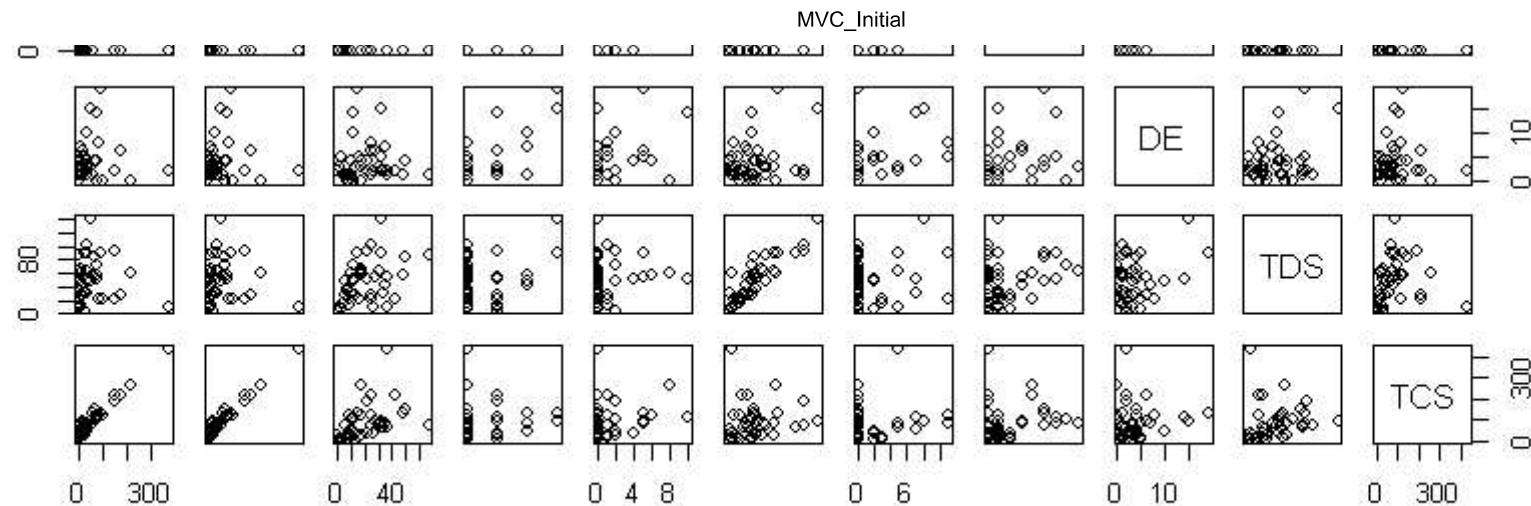
We are plotting three correlation graphs to find our relationship existance between each column present in our dataset.

- 1. Scatter plot matrices to check relationship within each columns of dataset

In [50]:

```
#Scatter plot matrices to check relationship within each columns of dataset
pairs(file_mvc[, c(2, 2:11)])
```





- 1. pairs plot to check correlation pattern between multiple variables

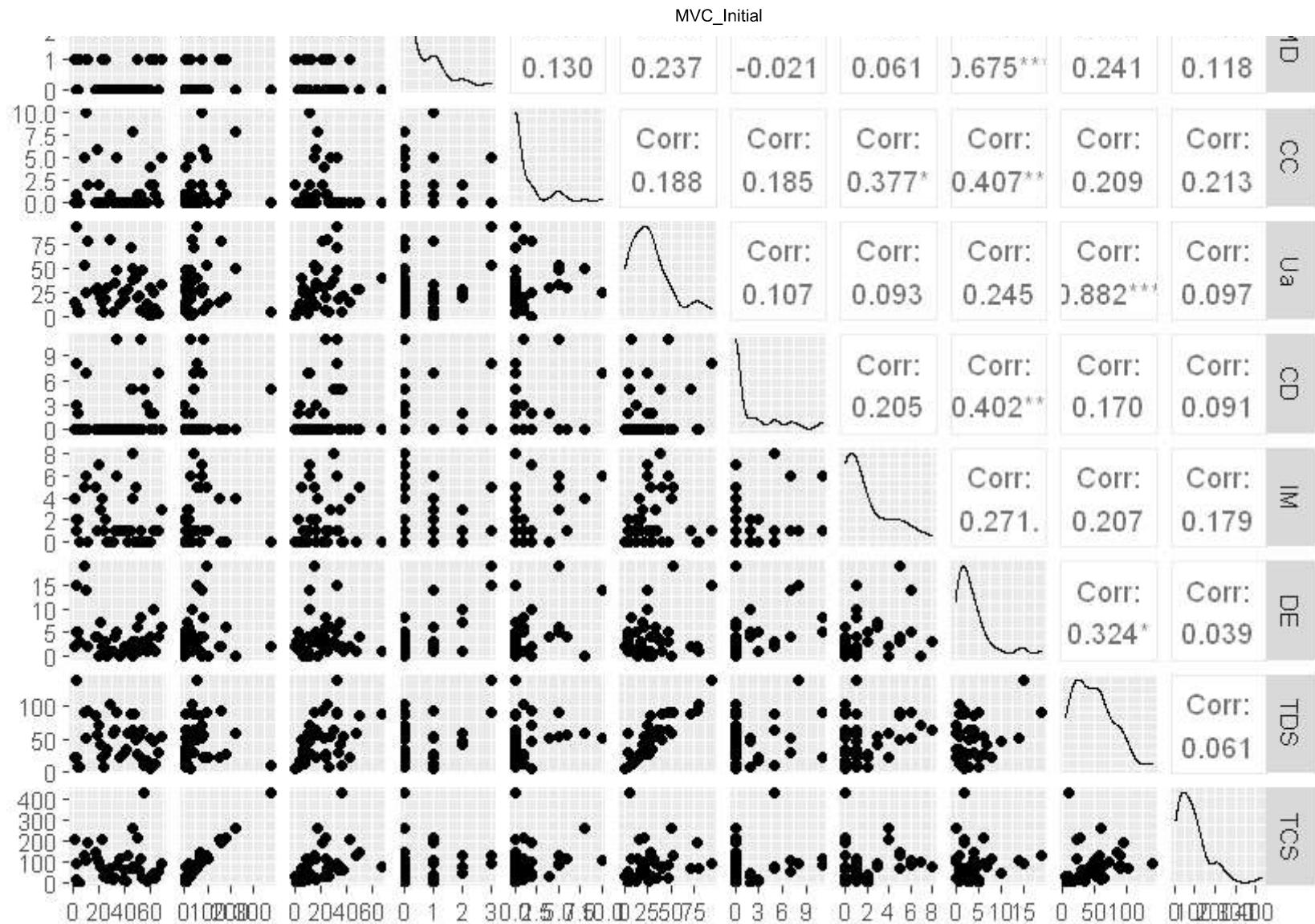
In [52]:

```
#create pairs plot
library(GGally)
ggpairs(file_mvc)
```

Registered S3 method overwritten by 'GGally':

method from  
+.gg ggplot2

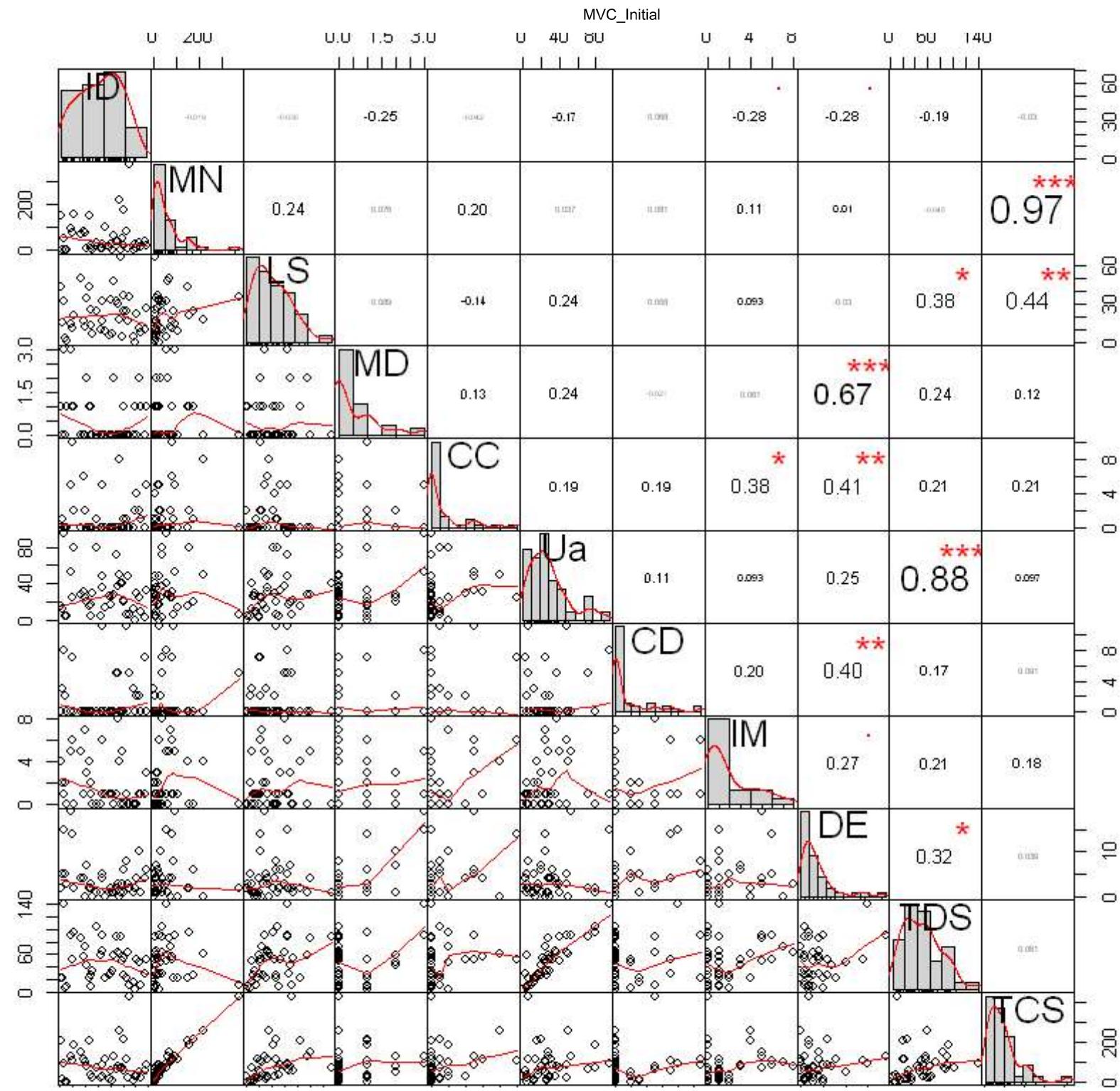




- 1. displaying PerformanceAnalytics chart of correlation matrix between variables

In [23]:

```
#install.packages("PerformanceAnalytics")
#library(PerformanceAnalytics)
chart.Correlation(file_mvc, histogram=TRUE, pch=19)
```





In the above plot:

- The distribution of each variable is shown on the diagonal.
- On the bottom of the diagonal : the bivariate scatter plots with a fitted line are displayed
- On the top of the diagonal : the value of the correlation plus the significance level as stars
- Each significance level is associated to a symbol : p-values(0, 0.001, 0.01, 0.05, 0.1, 1) <=> symbols("\*\*\*", "\*\*", "\*", ".", " ")

From the above two plot we can see few positive correlation , which are :

- MD and DE
- CC and DE
- CD and DE
- TDS and TCS are showing high correlation as its a summation column values for the rest.

Hence we will further do the pearson, kendall, spearman on these columns.

- Pearson correlation evaluates the linear relationship between two continuous variables.
- Kendall and spearman correlation evaluates the monotonic relationship between variables.

In [24]:

```
#pearson1 - MD and DE
p1 <- cor(file_mvc$MD, file_mvc$DE, method = "pearson")
p1
cor.test(file_mvc$MD, file_mvc$DE, use = 'complete.obs')
```

0.674958003507884

Pearson's product-moment correlation

```
data: file_mvc$MD and file_mvc$DE
t = 5.7854, df = 40, p-value = 9.466e-07
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
0.4667829 0.8122613
sample estimates:
cor
0.674958
```

Observation: Our p-value is significant and the 95% confidence interval confirms this as the range does not contain zero.

This suggests the correlation between MD(Missing.Defaults) and DE(Deficient.Encapsulation) per architecture is  $r = 0.67$  with 95% confidence of being between 0.47 and 0.81.

In [25]:

```
#pearson2 - CC and DE
p2 <- cor(file_mvc$CC, file_mvc$DE, method = "pearson")
p2
cor.test(file_mvc$CC, file_mvc$DE, use = 'complete.obs')
```

0.406892758992953

Pearson's product-moment correlation

```
data: file_mvc$CC and file_mvc$DE
t = 2.8172, df = 40, p-value = 0.007491
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1174912 0.6325930
sample estimates:
cor
0.4068928
```

Observation: Our p-value is significant and the 95% confidence interval confirms this as the range does not contain zero.

This suggests the correlation between CC(Complex.Conditional) and DE(Deficient.Encapsulation) per architecture is  $r = 0.41$  with 95% confidence of being between 0.12 and 0.63.

In [26]:

```
#pearson3 - CD and DE
p3 <- cor(file_mvc$CD, file_mvc$DE, method = "pearson")
p3
cor.test(file_mvc$CD, file_mvc$DE, use = 'complete.obs')
```

0.401972246653177

Pearson's product-moment correlation

```
data: file_mvc$CD and file_mvc$DE
t = 2.7765, df = 40, p-value = 0.00832
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.1116858 0.6290512
sample estimates:
```

```
cor
0.4019722
```

**Observation:** Our p-value is significant and the 95% confidence interval confirms this as the range does not contain zero.

This suggests the correlation between CD(Cyclic.Dependency) and DE(Deficient.Encapsulation) per architecture is  $r = 0.41$  with 95% confidence of being between 0.11 and 0.63.

In [27]:

```
#spearman1 - MD and DE
s1 <- cor(file_mvc$MD, file_mvc$DE, method = "spearman")
s1
cor.test(file_mvc$MD, file_mvc$DE, use = 'complete.obs', method = "spearman")
```

0.393180330525326

Warning message in cor.test.default(file\_mvc\$MD, file\_mvc\$DE, use = "complete.obs", :  
"Cannot compute exact p-value with ties"

Spearman's rank correlation rho

data: file\_mvc\$MD and file\_mvc\$DE  
 $S = 7488.8$ , p-value = 0.009999  
alternative hypothesis: true rho is not equal to 0  
sample estimates:  
rho  
0.3931803

**Observation: Our p-value is significant .**

This suggests the Spearman rank correlation rho between MN(Missing.Default) and DE(Deficient.Encapsulation) per architecture is  $r = 0.39$  with P-vale less than 0.05 we can reject the null hypothesis and say there is partial relatiopn between two variables.

In [28]:

```
#spearman2 - CC and DE
s2 <- cor(file_mvc$CC, file_mvc$DE, method = "spearman")
s2
cor.test(file_mvc$CC, file_mvc$DE, use = 'complete.obs', method = "spearman")
```

0.295442221024155

Warning message in cor.test.default(file\_mvc\$CC, file\_mvc\$DE, use = "complete.obs", :  
"Cannot compute exact p-value with ties"

Spearman's rank correlation rho

data: file\_mvc\$CC and file\_mvc\$DE

```
S = 8694.9, p-value = 0.05749
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.2954422
```

**Observation:** Our p-value is not significant and the 95% confidence interval confirms this as the range does not contain zero.

Cant reject the null hypothesis, no relationship present between variables.

In [29]:

```
#spearman3 - CD and DE
s3 <- cor(file_mvc$CD, file_mvc$DE, method = "spearman")
s3
cor.test(file_mvc$CD, file_mvc$DE, use = 'complete.obs', method = "spearman")
```

0.458287615524412

```
Warning message in cor.test.default(file_mvc$CD, file_mvc$DE, use = "complete.obs", :
"Cannot compute exact p-value with ties"
Spearman's rank correlation rho
```

```
data: file_mvc$CD and file_mvc$DE
S = 6685.3, p-value = 0.002272
alternative hypothesis: true rho is not equal to 0
sample estimates:
rho
0.4582876
```

**Observation:** Our p-value is significant .

This suggests the Spearman rank correlation rho between CD(Cyclic.Dependency) and DE(Deficient.Encapsulation) per architecture is r = 0.46 with P-vale less than 0.05 we can reject the null hypothesis and say there is some relatiopn between two variables.

In [30]:

```
#kendall1 - MD and DE
k1 <- cor(file_mvc$MD, file_mvc$DE, method = "kendall")
k1
cor.test(file_mvc$MD, file_mvc$DE, use = 'complete.obs', method = "kendall")
```

0.336342579822996

```
Warning message in cor.test.default(file_mvc$MD, file_mvc$DE, use = "complete.obs", :
"Cannot compute exact p-value with ties"
Kendall's rank correlation tau
```

```
data: file_mvc$MD and file_mvc$DE
z = 2.5923, p-value = 0.009535
alternative hypothesis: true tau is not equal to 0
sample estimates:
tau
0.3363426
```

**Observation:** Our p-value is significant and the 95% confidence interval confirms this as the range does not contain zero.

This suggests the kendall rank correlation tau between MD(Missing.Defaults) and DE(Deficient.Encapsulation) per architecture is r = 0.34 with P-vale less than 0.05 we can reject the null hypothesis and say there is partial relatiopn between two variables.

In [31]:

```
#kendall2 - CC and DE
k2 <- cor(file_mvc$CC, file_mvc$DE, method = "kendall")
k2
cor.test(file_mvc$CC, file_mvc$DE, use = 'complete.obs', method = "kendall")
```

0.23979161560332

Warning message in cor.test.default(file\_mvc\$CC, file\_mvc\$DE, use = "complete.obs", :  
 "Cannot compute exact p-value with ties"  
 Kendall's rank correlation tau

```
data: file_mvc$CC and file_mvc$DE
z = 1.9046, p-value = 0.05683
alternative hypothesis: true tau is not equal to 0
sample estimates:
tau
0.2397916
```

**Observation:** Our p-value is not significant and the 95% confidence interval confirms this as the range does not contain zero.

Cant reject the null hypothesis, no relationship present between variables.

In [32]:

```
#kendall3 - CD and DE
k3 <- cor(file_mvc$CD, file_mvc$DE, method = "kendall")
k3
cor.test(file_mvc$CD, file_mvc$DE, use = 'complete.obs', method = "kendall")
```

0.379566489413746

Warning message in cor.test.default(file\_mvc\$CD, file\_mvc\$DE, use = "complete.obs", :  
 "Cannot compute exact p-value with ties"

## Kendall's rank correlation tau

```
data: file_mvc$CD and file_mvc$DE
z = 2.9831, p-value = 0.002854
alternative hypothesis: true tau is not equal to 0
sample estimates:
tau
0.3795665
```

**Observation:** Our p-value is significant and the 95% confidence interval confirms this as the range does not contain zero.

This suggests the kendall rank correlation tau between CD(Cyclic.Dependency) and DE(Deficient.Encapsulation) per architecture is  $r = 0.38$  with P-vale less than 0.05 we can reject the null hypothesis and say there is partial relatiopn between two variables.

## Summarization of Correlation Processes

In [33]:

```
corr_df <- data.frame(
  Sl_No = c(1,2,3,4,5,6,7,8,9),
  Corr_Method = c("Pearson", "spearman", "kendall", "Pearson", "spearman", "kendall", "Pearson", "spearman", "kendall"),
  Column_related = c("MD and DE", "MD and DE", "MD and DE", "CC and DE", "CD and DE",
                     "CC and DE", "CD and DE", "CD and DE", "CD and DE"),
  corr_value = c(0.67, 0.41, 0.40, 0.39, 0.30, 0.46, 0.34, 0.24, 0.38)
)
```

In [34]:

```
corr_df
```

Sl_No	Corr_Method	Column_related	corr_value
1	Pearson	MD and DE	0.67
2	spearman	MD and DE	0.41
3	kendall	MD and DE	0.40
4	Pearson	CC and DE	0.39
5	spearman	CD and DE	0.30
6	kendall	CC and DE	0.46
7	Pearson	CD and DE	0.34

SI_No	Corr_Method	Column_related	corr_value
8	spearman	CD and DE	0.24
9	kendall	CD and DE	0.38

8	spearman	CD and DE	0.24
9	kendall	CD and DE	0.38

In [35]:

```
knitr::kable(corr_df, format = "markdown")
```

S1_No	Corr_Method	Column_related	corr_value
1	Pearson	MD and DE	0.67
2	spearman	MD and DE	0.41
3	kendall	MD and DE	0.40
4	Pearson	CC and DE	0.39
5	spearman	CD and DE	0.30
6	kendall	CC and DE	0.46
7	Pearson	CD and DE	0.34
8	spearman	CD and DE	0.24
9	kendall	CD and DE	0.38

- From the above correlation tests we can conclude that there is high correlation present within MVC architecture between Missing.Defaults & Deficient.Encapsulation .

In [ ]: