Relatorio Tese - Capítulo 3

Denilson Junio Marques Soares

2023-03-09

Análises da Tese - Capítulo 3

Carregamento de Pacotes

```
library(readxl)
library(dplyr)
library(haven)
library(psych)
library(lm.beta)
library(car)
library(rstatix)
library(olsrr)
library(ggplot2)
library(GGally)
library(lmtest)
library(data.table)
library(performance)
library(see)
library(patchwork)
```

Leitura dos dados

```
setwd("C:\\Users\\UFES\\Desktop\\Tese_DenilsonSoares\\Capítulo 3")
dados <- read_excel("Dados_contexto.xlsx")
dados$IRD=as.numeric(dados$IRD)
```

Função para as análises gráficas

```
diag_fun <- function(data, mapping, hist=list(), ...){

  X = eval_data_col(data, mapping$x)
  mn = mean(X)
  s = sd(X)

  ggplot(data, mapping) +
    do.call(function(...) geom_histogram(aes(y =..density..), ...), hist) +
    stat_function(fun = dnorm, args = list(mean = mn, sd = s), ...)
}
```

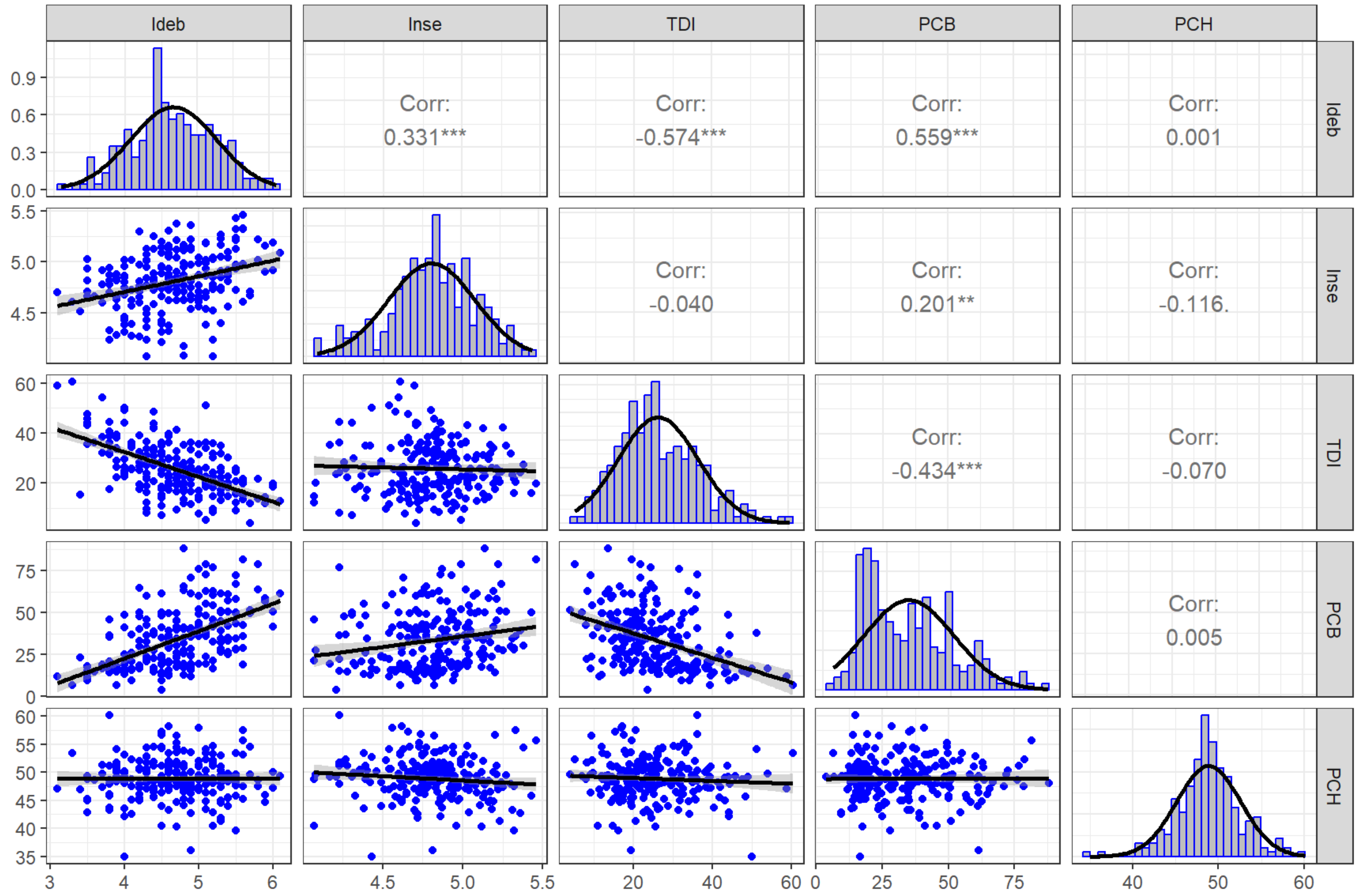
Indicadores de contexto relacionados aos alunos

```
alunos=data.frame(dados$Ideb, dados$Inse, dados$TDI, dados$PCB, dados$PCH)
names(alunos)[1:5] <- c("Ideb", "Inse", "TDI", "PCB", "PCH")
summary(alunos)
```

```
##       Ideb            Inse            TDI             PCB        
##  Min.   :3.100   Min.   :4.070   Min.   : 4.10   Min.   : 4.023  
##  1st Qu.:4.300   1st Qu.:4.660   1st Qu.:18.52   1st Qu.:18.710  
##  Median :4.700   Median :4.820   Median :24.70   Median :30.391  
##  Mean   :4.667   Mean   :4.810   Mean   :25.90   Mean   :33.548  
##  3rd Qu.:5.100   3rd Qu.:5.005   3rd Qu.:32.52   3rd Qu.:45.003  
##  Max.   :6.100   Max.   :5.460   Max.   :60.40   Max.   :88.000  
##       PCH       
##  Min.   :35.00  
##  1st Qu.:46.99  
##  Median :48.78  
##  Mean   :48.85  
##  3rd Qu.:51.07  
##  Max.   :60.09  
```

FIGURA 20

```
ggpairs(alunos,  diag = list(continuous = wrap(diag_fun, hist=list(fill="gray", colour="Blu
e"),
                                               colour="Black", lwd=1)),
        lower = list(continuous = wrap("smooth", color="Blue", se=T)))  +theme_bw()
```



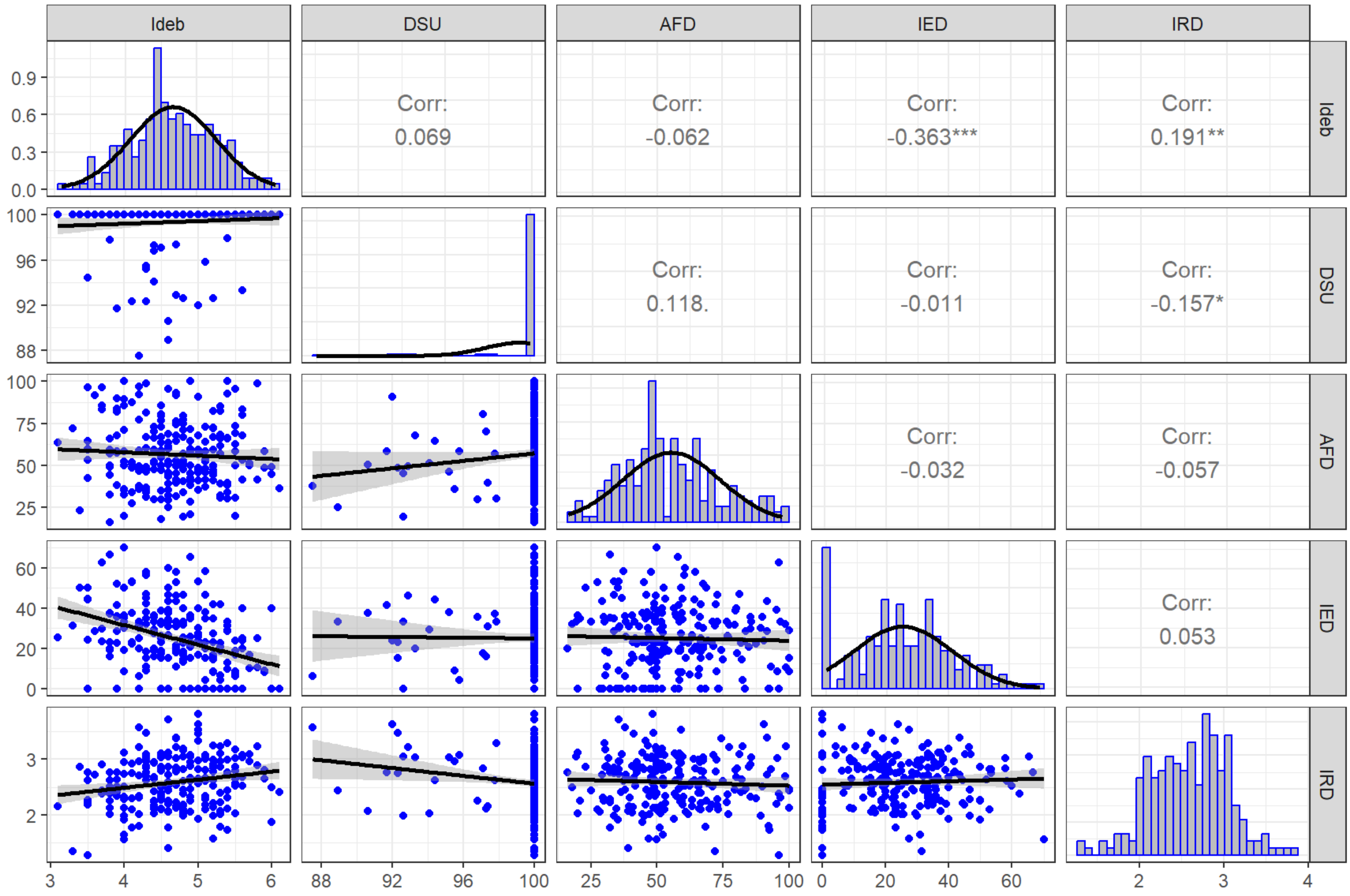
Indicadores de contexto relacionados aos docentes

```
docentes=data.frame(dados$Ideb, dados$DSU, dados$AFD, dados$IED, dados$IRD)
names(docentes)[1:5] <- c("Ideb", "DSU", "AFD", "IED", "IRD")
summary(docentes)
```

```
##       Ideb             DSU              AFD               IED       
##  Min.   :3.100   Min.   : 87.5   Min.   : 16.20   Min.   : 0.00  
##  1st Qu.:4.300   1st Qu.:100.0   1st Qu.: 44.40   1st Qu.:15.25  
##  Median :4.700   Median :100.0   Median : 52.75   Median :24.10  
##  Mean   :4.667   Mean   : 99.4   Mean   : 56.58   Mean   :25.13  
##  3rd Qu.:5.100   3rd Qu.:100.0   3rd Qu.: 67.67   3rd Qu.:35.42  
##  Max.   :6.100   Max.   :100.0   Max.   :100.00   Max.   :70.00  
##                                                                  
##       IRD       
##  Min.   :1.277  
##  1st Qu.:2.251  
##  Median :2.622  
##  Mean   :2.589  
##  3rd Qu.:2.903  
##  Max.   :3.810  
##  NA's   :6
```

FIGURA 21

```
ggpairs(docentes,  diag = list(continuous = wrap(diag_fun, hist=list(fill="gray", colour="Blu
e"),
                                                   colour="Black", lwd=1)),
        lower = list(continuous = wrap("smooth", color="Blue", se=T)))  +theme_bw()
```



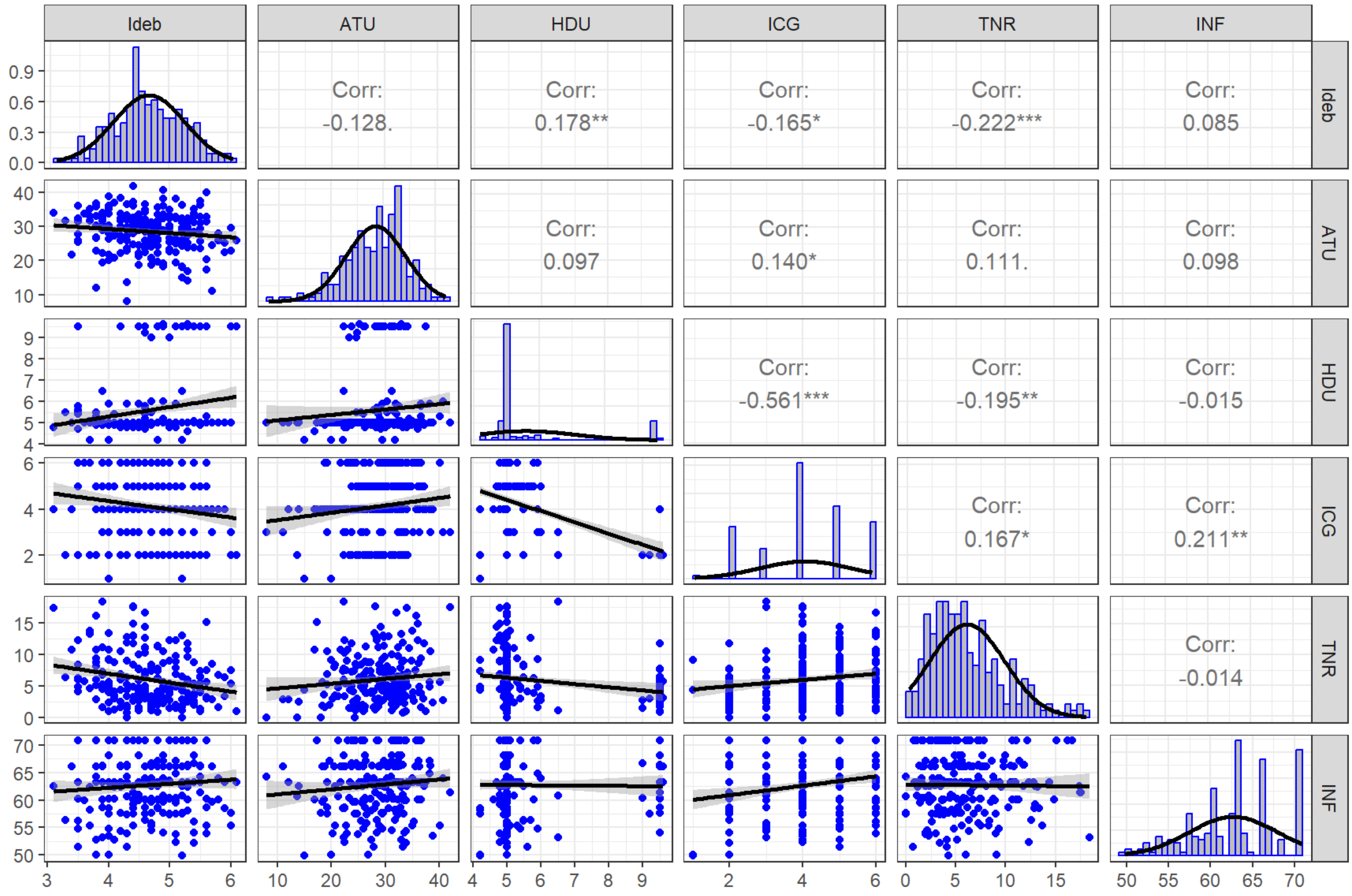
Indicadores de contexto relacionados às escolas

```
escolas=data.frame(dados$Ideb, dados$ATU, dados$HDU, dados$ICG, dados$TNR, dados$INF)
names(escolas)[1:6] <- c("Ideb", "ATU", "HDU", "ICG", "TNR", "INF")
summary(escolas)
```

```
##       Ideb            ATU             HDU             ICG       
##  Min.   :3.100   Min.   : 8.00   Min.   :4.200   Min.   :1.000  
##  1st Qu.:4.300   1st Qu.:25.05   1st Qu.:5.000   1st Qu.:3.000  
##  Median :4.700   Median :29.25   Median :5.000   Median :4.000  
##  Mean   :4.667   Mean   :28.56   Mean   :5.598   Mean   :4.131  
##  3rd Qu.:5.100   3rd Qu.:32.50   3rd Qu.:5.000   3rd Qu.:5.000  
##  Max.   :6.100   Max.   :42.00   Max.   :9.600   Max.   :6.000  
##       TNR              INF       
##  Min.   : 0.000   Min.   :49.80  
##  1st Qu.: 3.125   1st Qu.:59.60  
##  Median : 5.150   Median :63.30  
##  Mean   : 6.029   Mean   :62.75  
##  3rd Qu.: 8.075   3rd Qu.:66.20  
##  Max.   :18.400   Max.   :70.90  
```

FIGURA 22

```
ggpairs(escolas,  diag = list(continuous = wrap(diag_fun, hist=list(fill="gray", colour="Blue"),
                                                              colour="Black", lwd=1)),
          lower = list(continuous = wrap("smooth", color="Blue", se=T)))  +theme_bw()
```



Ajuste do Modelo de Regressão Linear Múltipla

```
fit <- lm(data = dados, Ideb ~ Inse + TDI + PCB + IED + IRD + HDU + ICG + TNR)
step_fit_p <- ols_step_backward_p(model = fit, prem = 0.05, details = TRUE)
```

```
## Backward Elimination Method
## ---------------------------
##
## Candidate Terms:
##
## 1 . Inse
## 2 . TDI
## 3 . PCB
## 4 . IED
## 5 . IRD
## 6 . HDU
## 7 . ICG
## 8 . TNR
##
## We are eliminating variables based on p value...
##
## - ICG
##
## Backward Elimination: Step 1
##
##  Variable ICG Removed
##
##                         Model Summary
## --------------------------------------------------------------
## R                       0.725       RMSE               0.422
## R-Squared               0.526       Coef. Var          9.048
## Adj. R-Squared          0.510       MSE                0.178
## Pred R-Squared          0.487       MAE                0.335
## --------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                ANOVA
## ---------------------------------------------------------------------
##                Sum of
##               Squares         DF    Mean Square      F         Sig.
## ---------------------------------------------------------------------
## Regression     41.068          7          5.867    32.993    0.0000
## Residual       36.987        208          0.178
## Total          78.056        215
## ---------------------------------------------------------------------
##
##                                   Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta      t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)     2.753         0.579                  4.753    0.000     1.611     3.895
##        Inse     0.473         0.112        0.218     4.238    0.000     0.253     0.693
##         TDI    -0.026         0.004       -0.451    -6.888    0.000    -0.033    -0.019
##         PCB     0.011         0.002        0.323     5.694    0.000     0.007     0.015
##         IED    -0.003         0.002       -0.087    -1.407    0.161    -0.008     0.001
##         IRD    -0.005         0.072       -0.004    -0.070    0.945    -0.148     0.137
##         HDU    -0.006         0.027       -0.013    -0.208    0.835    -0.060     0.048
##         TNR     0.012         0.009        0.080     1.403    0.162    -0.005     0.030
## ----------------------------------------------------------------------------------------
##
##
## - IRD
##
## Backward Elimination: Step 2
##
##  Variable IRD Removed
##
##                         Model Summary
## -------------------------------------------------------------
## R                       0.723       RMSE                0.422
## R-Squared               0.523       Coef. Var           9.036
## Adj. R-Squared          0.509       MSE                 0.178
## Pred R-Squared          0.488       MAE                 0.336
## -------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                ANOVA
## --------------------------------------------------------------------
##                Sum of
##               Squares         DF    Mean Square      F         Sig.
## --------------------------------------------------------------------
## Regression     41.851          6          6.975    39.218    0.0000
## Residual       38.239        215          0.178
## Total          80.090        221
## --------------------------------------------------------------------
##
##                                   Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta      t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)     2.683         0.520                  5.161    0.000     1.658     3.708
##        Inse     0.473         0.110        0.220     4.296    0.000     0.256     0.690
##         TDI    -0.025         0.004       -0.434    -7.148    0.000    -0.032    -0.018
##         PCB     0.011         0.002        0.321     5.739    0.000     0.007     0.015
##         IED    -0.004         0.002       -0.108    -1.750    0.082    -0.009     0.001
##         HDU     0.002         0.025        0.005     0.074    0.941    -0.048     0.051
##         TNR     0.013         0.009        0.086     1.528    0.128    -0.004     0.031
## ----------------------------------------------------------------------------------------
##
##
## - HDU
##
## Backward Elimination: Step 3
##
##  Variable HDU Removed
##
##                         Model Summary
## -------------------------------------------------------------
## R                       0.723       RMSE                0.421
## R-Squared               0.523       Coef. Var           9.015
## Adj. R-Squared          0.511       MSE                 0.177
## Pred R-Squared          0.494       MAE                 0.336
## -------------------------------------------------------------
##  RMSE: Root Mean Square Error 
##  MSE: Mean Square Error 
##  MAE: Mean Absolute Error 
## 
##                                 ANOVA                                 
## --------------------------------------------------------------------
##                Sum of                                               
##               Squares         DF    Mean Square      F         Sig. 
## --------------------------------------------------------------------
## Regression     41.850          5          8.370    47.278    0.0000 
## Residual       38.240        216          0.177                     
## Total          80.090        221                                    
## --------------------------------------------------------------------
## 
##                                    Parameter Estimates                                     
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta      t        Sig      lower     upper 
## ----------------------------------------------------------------------------------------
## (Intercept)     2.688         0.515                  5.223    0.000     1.673     3.702 
##        Inse     0.475         0.106        0.221     4.480    0.000     0.266     0.684 
##         TDI    -0.025         0.004       -0.434    -7.166    0.000    -0.032    -0.018 
##         PCB     0.011         0.002        0.320     5.981    0.000     0.007     0.015 
##         IED    -0.004         0.002       -0.111    -2.177    0.031    -0.008     0.000 
##         TNR     0.013         0.009        0.085     1.546    0.124    -0.004     0.030 
## ----------------------------------------------------------------------------------------
## 
## 
## - TNR 
## 
## Backward Elimination: Step 4 
## 
##  Variable TNR Removed 
## 
##                         Model Summary                          
## -------------------------------------------------------------
## R                       0.719       RMSE                0.422 
## R-Squared               0.517       Coef. Var           9.044 
## Adj. R-Squared          0.508       MSE                 0.178 
## Pred R-Squared          0.494       MAE                 0.340 
## -------------------------------------------------------------
##  RMSE: Root Mean Square Error 
##  MSE: Mean Square Error 
##  MAE: Mean Absolute Error 
## 
##                                 ANOVA                                 
## --------------------------------------------------------------------
##                Sum of                                               
##               Squares         DF    Mean Square      F         Sig. 
## --------------------------------------------------------------------
## Regression     41.427          4         10.357    58.128    0.0000 
## Residual       38.663        217          0.178                     
## Total          80.090        221                                    
## --------------------------------------------------------------------
## 
##                                    Parameter Estimates                                     
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta      t        Sig      lower      upper
## ----------------------------------------------------------------------------------------
## (Intercept)     2.638         0.515                    5.119    0.000      1.622      3.653
##        Inse     0.490         0.106        0.227       4.621    0.000      0.281      0.698
##         TDI    -0.023         0.003       -0.392      -7.222    0.000     -0.029     -0.016
##         PCB     0.011         0.002        0.317       5.901    0.000      0.007      0.015
##         IED    -0.004         0.002       -0.110      -2.157    0.032     -0.008      0.000
## ----------------------------------------------------------------------------------------
##
##
##
## No more variables satisfy the condition of p value = 0.05
##
##
## Variables Removed:
##
## - ICG
## - IRD
## - HDU
## - TNR
##
##
## Final Model Output
## ------------------
##
##                         Model Summary
## --------------------------------------------------------------
## R                       0.719       RMSE               0.422
## R-Squared               0.517       Coef. Var          9.044
## Adj. R-Squared          0.508       MSE                0.178
## Pred R-Squared          0.494       MAE                0.340
## --------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                ANOVA
## ---------------------------------------------------------------------
##               Sum of
##              Squares         DF    Mean Square      F         Sig.
## ---------------------------------------------------------------------
## Regression     41.427          4         10.357    58.128    0.0000
## Residual       38.663        217          0.178
## Total          80.090        221
## ---------------------------------------------------------------------
##
##                                   Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta      t        Sig      lower      upper
## ----------------------------------------------------------------------------------------
## (Intercept)     2.638         0.515                    5.119    0.000      1.622      3.653
##        Inse     0.490         0.106        0.227       4.621    0.000      0.281      0.698
##         TDI    -0.023         0.003       -0.392      -7.222    0.000     -0.029     -0.016
##         PCB     0.011         0.002        0.317       5.901    0.000      0.007      0.015
##         IED    -0.004         0.002       -0.110    -2.157    0.032    -0.008     0.000
## -----------------------------------------------------------------------------------------
```

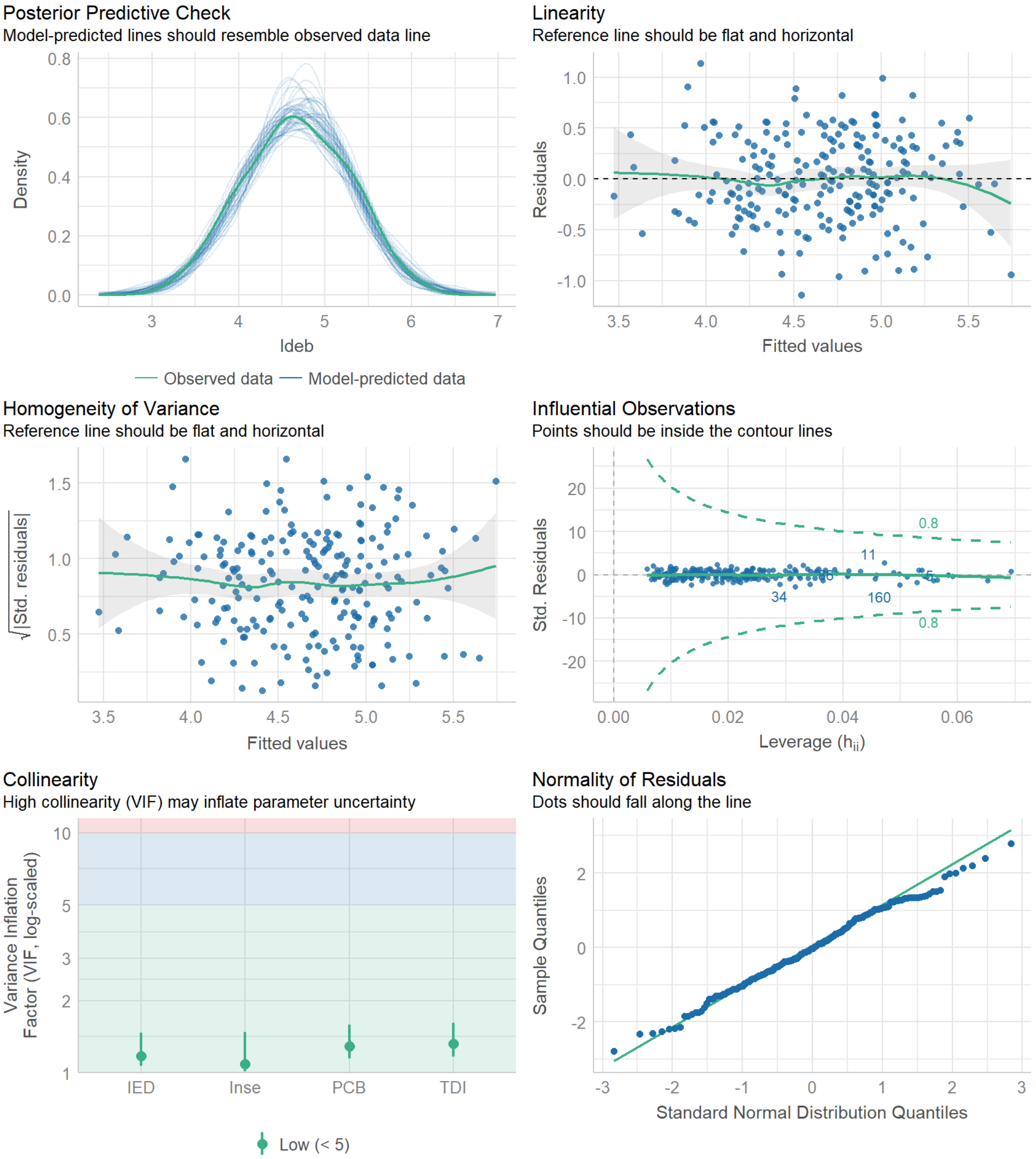
```
fit <- lm(data = dados, Ideb ~ Inse + TDI + PCB + IED)
```

Diagnóstico do Modelo

Análise gráfica:

FIGURA 23

```
check_model(fit)
```



Normalidade dos resíduos:

```
shapiro.test(fit$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit$residuals
## W = 0.99371, p-value = 0.475
```

Independência dos resíduos (Durbin-Watson):

```
durbinWatsonTest(fit)
```

```
##  lag Autocorrelation D-W Statistic p-value
##    1     0.007748487      1.982225   0.844
##  Alternative hypothesis: rho != 0
```

Homocedasticidade (Breusch-Pagan):

```
bptest(fit)
```

```
## 
##  studentized Breusch-Pagan test
## 
## data:  fit
## BP = 2.0247, df = 4, p-value = 0.7312
```

Outliers nos resíduos:

```
summary(rstandard(fit))
```

```
##       Min.    1st Qu.     Median       Mean    3rd Qu.       Max. 
## -2.7500281 -0.6868858 -0.0284504 -0.0001924  0.7912897  2.7499475
```

Multicolinearidade

```
vif(fit)
```

```
##     Inse      TDI      PCB      IED 
## 1.088480 1.324397 1.293899 1.177553
```

Discretização dos dados

```
dados=data.frame(dados$Ideb, dados$Inse, dados$TDI, dados$PCB, dados$IED)
names(dados)[1:5] <- c("Ideb", "Inse", "TDI", "PCB", "IED")

RRRR = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI > median(dados$TDI))

RRRB = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI <= median(dados$TDI))

RRBR = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI > median(dados$TDI))

RRBB = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI <= median(dados$TDI))

RBRR = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI > median(dados$TDI))

RBRB = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI <= median(dados$TDI))

RBBR = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI > median(dados$TDI))

RBBB = subset(dados, dados$Inse <= median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI <= median(dados$TDI))

BRRR = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI > median(dados$TDI))

BRRB = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI <= median(dados$TDI))

BRBR = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI > median(dados$TDI))

BRBB = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED > median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI <= median(dados$TDI))

BBRR = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI > median(dados$TDI))

BBRB = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB <= median(dados$PCB)  & dados$TDI <= median(dados$TDI))

BBBR = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI > median(dados$TDI))

BBBB = subset(dados, dados$Inse > median(dados$Inse)  & dados$IED <= median(dados$IED)
              & dados$PCB > median(dados$PCB) & dados$TDI <= median(dados$TDI))
```

4 indicadores abaixo da mediana (B - baixo) e 0 indicadores acima da mediana (A - alto)

```
dados_4B0A = RRRR
summary(dados_4B0A)
```

```
##       Ideb            Inse            TDI              PCB              IED
##  Min.   :3.100   Min.   :4.17   Min.   :25.60   Min.   : 6.711   Min.   :25.5
##  1st Qu.:3.700   1st Qu.:4.60   1st Qu.:28.70   1st Qu.:13.505   1st Qu.:35.0
##  Median :4.000   Median :4.70   Median :34.30   Median :15.734   Median :38.9
##  Mean   :4.067   Mean   :4.65   Mean   :36.76   Mean   :15.450   Mean   :41.6
##  3rd Qu.:4.400   3rd Qu.:4.77   3rd Qu.:43.30   3rd Qu.:16.844   3rd Qu.:46.7
##  Max.   :5.300   Max.   :4.82   Max.   :60.40   Max.   :28.395   Max.   :70.0
```

3 indicadores abaixo da mediana (B - baixo) e 1 indicadores acima da mediana (A - alto)

```
dados_3B1A = rbind(RRRB, RRBR, RBRR, BRRR)
summary(dados_3B1A)
```

```
##       Ideb            Inse             TDI              PCB
##  Min.   :3.400   Min.   :4.210   Min.   :15.40   Min.   : 4.023
##  1st Qu.:4.100   1st Qu.:4.540   1st Qu.:24.80   1st Qu.:17.722
##  Median :4.400   Median :4.750   Median :31.80   Median :22.162
##  Mean   :4.392   Mean   :4.711   Mean   :31.64   Mean   :24.270
##  3rd Qu.:4.600   3rd Qu.:4.840   3rd Qu.:38.40   3rd Qu.:26.165
##  Max.   :5.400   Max.   :5.300   Max.   :51.20   Max.   :60.588
##       IED
##  Min.   : 0.00
##  1st Qu.:25.60
##  Median :33.30
##  Mean   :32.69
##  3rd Qu.:37.50
##  Max.   :66.70
```

2 indicadores abaixo da mediana (B - baixo) e 2 indicadores acima da mediana (A - alto)

```
dados_2B2A = rbind(RRBB, RBRB, RBBR, BRRB, BRBR, BBRR)
summary(dados_2B2A)
```

```
##       Ideb            Inse            TDI             PCB        
##  Min.   :3.500   Min.   :4.070   Min.   : 8.30   Min.   : 9.807  
##  1st Qu.:4.300   1st Qu.:4.652   1st Qu.:19.43   1st Qu.:19.073  
##  Median :4.700   Median :4.840   Median :25.35   Median :27.302  
##  Mean   :4.606   Mean   :4.800   Mean   :25.03   Mean   :32.660  
##  3rd Qu.:4.900   3rd Qu.:5.013   3rd Qu.:31.40   3rd Qu.:41.279  
##  Max.   :5.600   Max.   :5.380   Max.   :44.40   Max.   :76.744  
##       IED       
##  Min.   : 0.00  
##  1st Qu.:16.10  
##  Median :22.40  
##  Mean   :23.28  
##  3rd Qu.:30.80  
##  Max.   :65.40  
```

1 indicadores abaixo da mediana (B - baixo) e 3 indicadores acima da mediana (A - alto)

```
dados_1B3A = rbind(RBBB, BRBB, BBRB, BBBR)
summary(dados_1B3A)
```

```
##       Ideb            Inse            TDI             PCB       
##  Min.   :3.800   Min.   :4.070   Min.   : 4.10   Min.   :12.76  
##  1st Qu.:4.625   1st Qu.:4.713   1st Qu.:15.10   1st Qu.:34.66  
##  Median :5.050   Median :4.895   Median :19.40   Median :41.55  
##  Mean   :5.009   Mean   :4.880   Mean   :19.66   Mean   :44.26  
##  3rd Qu.:5.300   3rd Qu.:5.048   3rd Qu.:23.02   3rd Qu.:50.54  
##  Max.   :6.000   Max.   :5.460   Max.   :36.20   Max.   :81.52  
##       IED       
##  Min.   : 0.00  
##  1st Qu.: 6.80  
##  Median :18.35  
##  Mean   :18.74  
##  3rd Qu.:25.00  
##  Max.   :58.30  
```

0 indicadores abaixo da mediana (B - baixo) e 4 indicadores acima da mediana (A - alto)

```
dados_0B4A = BBBB
summary(dados_0B4A)
```

```
##       Ideb            Inse            TDI             PCB       
##  Min.   :4.300   Min.   :4.830   Min.   : 8.40   Min.   :31.74  
##  1st Qu.:4.875   1st Qu.:4.907   1st Qu.:13.75   1st Qu.:41.43  
##  Median :5.300   Median :5.025   Median :19.25   Median :49.90  
##  Mean   :5.268   Mean   :5.049   Mean   :17.68   Mean   :50.64  
##  3rd Qu.:5.500   3rd Qu.:5.168   3rd Qu.:21.20   3rd Qu.:58.84  
##  Max.   :6.100   Max.   :5.430   Max.   :24.50   Max.   :88.00  
##       IED        
##  Min.   : 0.000  
##  1st Qu.: 3.150  
##  Median : 8.500  
##  Mean   : 9.304  
##  3rd Qu.:14.300  
##  Max.   :23.800  
```

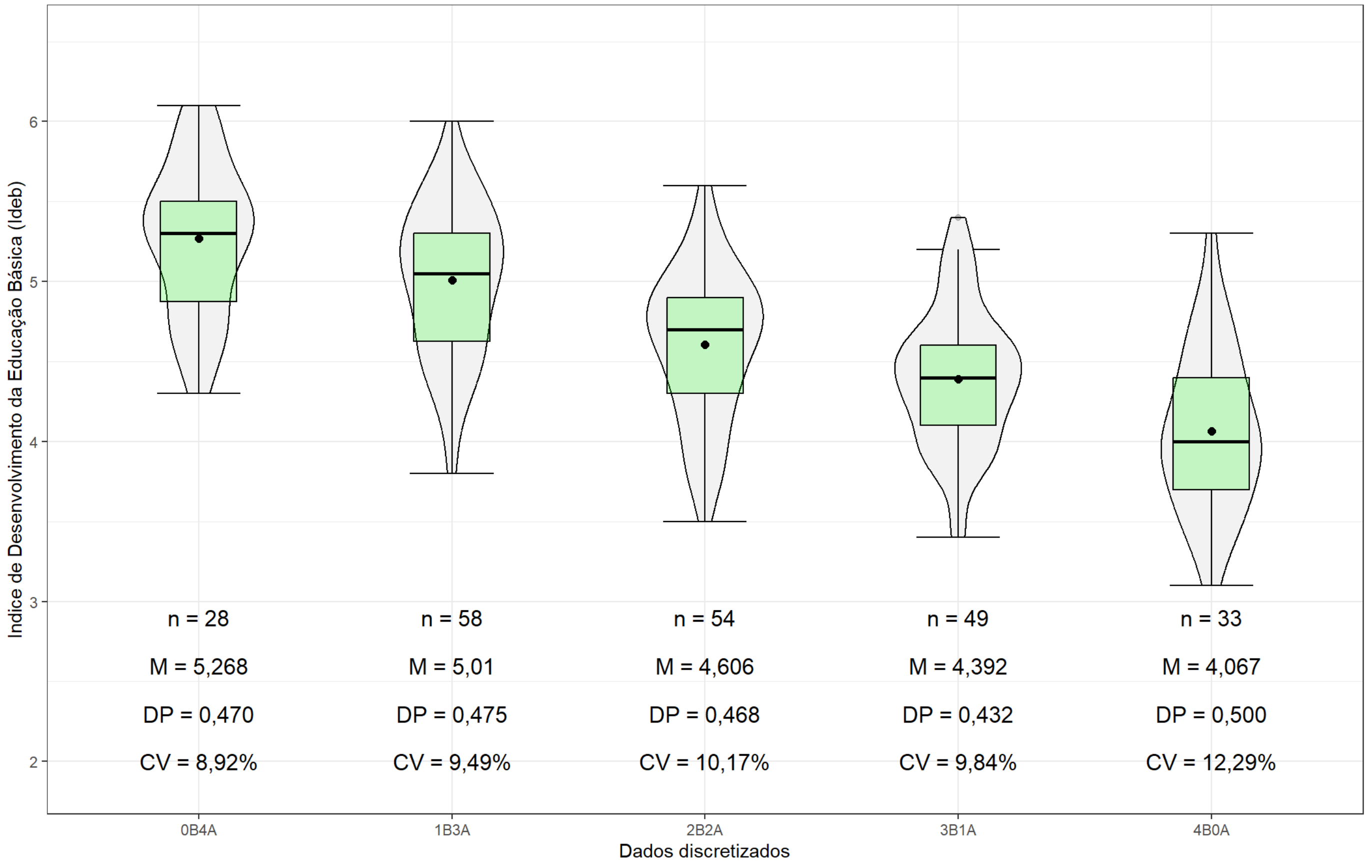
Boxplots para as notas no Ideb, considerando a discretização dos dados

```
name=c( rep('0B4A', 28), rep("1B3A",58), rep("2B2A",54),rep("3B1A",49), rep("4B0A",33))
value=c(dados_0B4A$Ideb, dados_1B3A$Ideb,dados_2B2A$Ideb, dados_3B1A$Ideb,  dados_4B0A$Ideb)
data=data.frame(name,value)

sample_size = data %>% group_by(name) %>% summarize(num=n())

# Plot
a=data %>%
  ggplot( aes(x=name, y=value, fill=name)) +
  stat_boxplot(geom = "errorbar", width = .33) +
  geom_violin(width=0.5, fill = "grey95", colour = "Black") +
  geom_boxplot(width=0.3, fill = "green", color="black", alpha=0.2) +
  stat_summary(aes(shape = "média"),
               geom = "point",
               color="Black",
               fun = mean,
               size = 2) +
  theme_bw() +
  labs(x = "Dados discretizados", y = "Indice de Desenvolvimento da Educação Básica (Ideb)")
+
  ylim(1.9,6.5)+
  theme(legend.position = "none")
```

FIGURA 24


Indice de Desenvolvimento da Educação Básica (Ideb)
6
5
4
3
2
n = 28
M = 5,268
DP = 0,470
CV = 8,92%
n = 58
M = 5,01
DP = 0,475
CV = 9,49%
n = 54
M = 4,606
DP = 0,468
CV = 10,17%
n = 49
M = 4,392
DP = 0,432
CV = 9,84%
n = 33
M = 4,067
DP = 0,500
CV = 12,29%
0B4A
1B3A
2B2A
3B1A
4B0A
Dados discretizados