**PARTE 2**

**SITUAÇÃO PROBLEMA:**

Uma empresa de software começou recentemente a realizar estudos para criar um software que prevê a média do aluno em questão utilizando algumas características dos alunos com o objetivo de criar um tipo de software utilizando dados recolhidos do site Kaggle (https://www.kaggle.com/spscientist/students-performance-in-exams). O objetivo é criar um modelo que pudesse indicar a média do aluno em matemática.

**ANÁLISE:**

Para avaliar a média do aluno em matemáticas, decidiu-se criar um modelo de regressão que preveja a nota média do aluno em matemática, para que o usuário final tenha noção de sua média.

**ANÁLISE DESCRITIVA DAS VARIAVEIS:**

• gender – gênero do aluno;

- Qualitativa nominal.

• race – raça ou etnia do aluno;

-Qualitativa nominal.

• parental.level.of.education – nível escolar dos pais;

-Qualitativa ordinal.

• test.preparation.course – se o aluno fez algum tipo de curso preparatório;

-Qualitativa nominal.

• math.score – nota média de matemática;

-Quantitativa discreta.

• reading.score – nota média de escrita;

-Quantitativa discreta.

• writing.score – nota média de leitura;

-Quantitativa discreta.

**gender race parental.level.of.education test.preparation.course**

female:518 group A: 89 associate's degree:222 completed:358

male :482 group B:190 bachelor's degree :118 none :642

group C:319 high school :196

group D:262 master's degree : 59

group E:140 some college :226

some high school :179

**math.score reading.score writing.score**

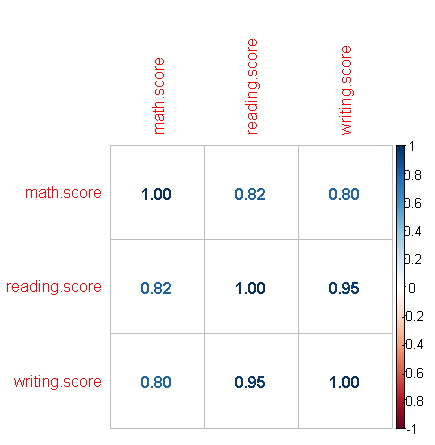
Min. : 0.00 Min. : 17.00 Min. : 10.00

Median : 66.00 Median : 70.00 Median : 69.00

Mean : 66.09 Mean : 69.17 Mean : 68.05

Max. :100.00 Max. :100.00 Max. :100.00

**MATRIZ DE CORRELAÇÃO COM AS VARIÁVEIS QUANTITATIVAS:**



Para entendermos melhor a nossa matriz de correlação , temos que entender que do vermelho -1 até o azul 1, quanto maior o valor maior a correlação entre eles.

Podemos ver na nossa matriz de correlação que todas as variáveis quantitativas têm correlação entre elas.

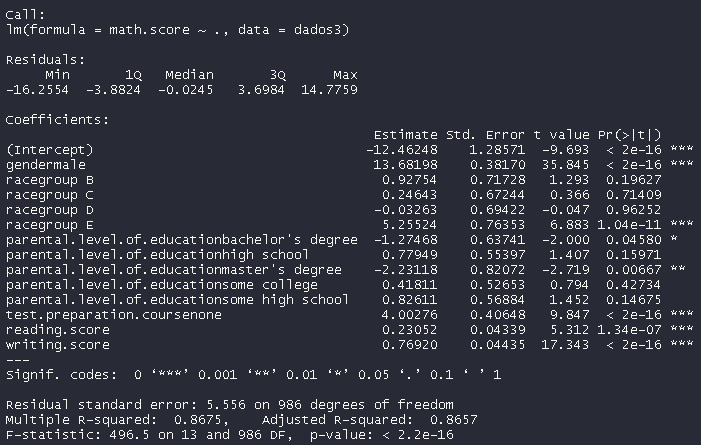
**ANÁLISE DOS DADOS COM GRAFICOS:**

Como podemos ver nos gráficos abaixo , quanto maior a nota em leitura ou escrita do aluno maior é a sua nota em matemática, normalmente não é uma exceção seguindo um padrão crescente.

Gráfico, Gráfico de dispersão

Descrição gerada automaticamenteGráfico, Gráfico de dispersão

Descrição gerada automaticamente

**MODELO DE REGRESSÃO:**

**PREDIÇÃO DA NOTA MÉDIA DE MATEMATICA PARA UM ALUNO COM DETERMINADAS CARACTERISTICAS:**

****

Podemos ver que média de matemática de um aluno com o gênero feminino, com a raça do grupo C, com o ensino dos pais em some college, e com curso preparatório, nota 90 em leitura, nota 88 em escrita, teria uma média de 76 em matemática.