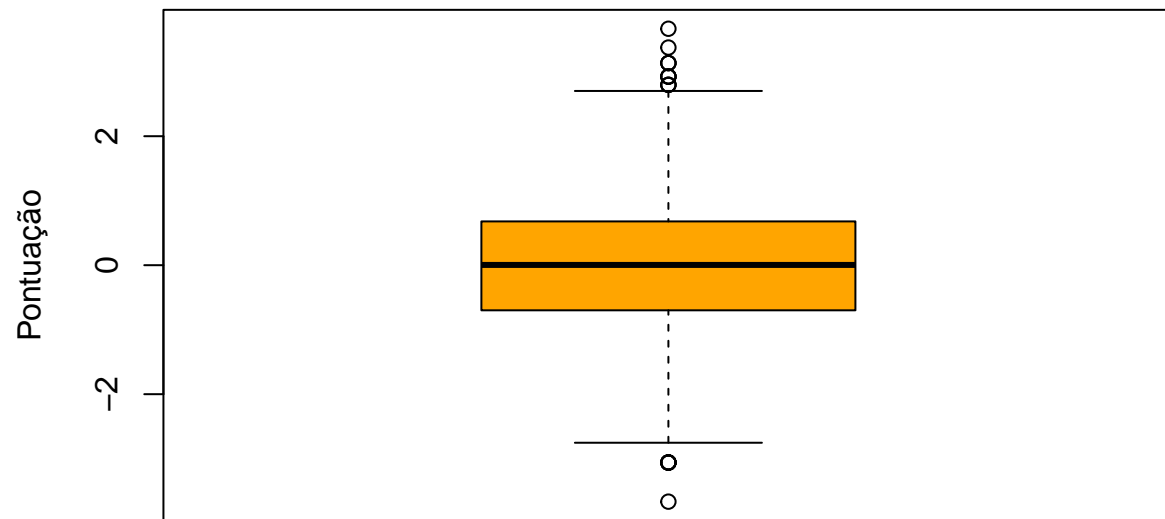


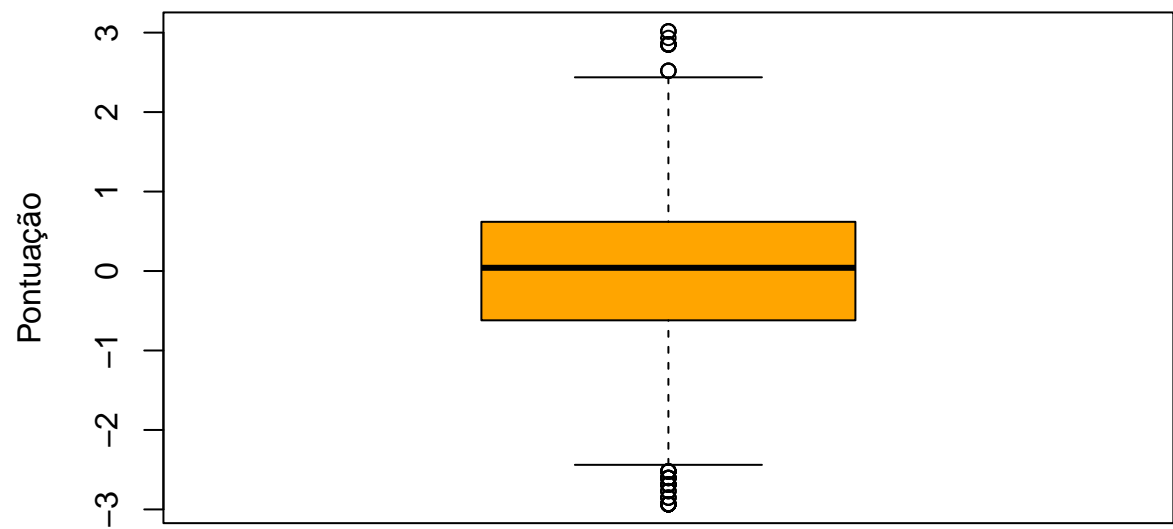
Estudo de caso escolar na região de Londres

Começaremos com uma exploração de dados para conhecimento geral, para as variáveis numéricas achei interessante apresentar boxplots, assim podemos visualizarmos possíveis outliers.

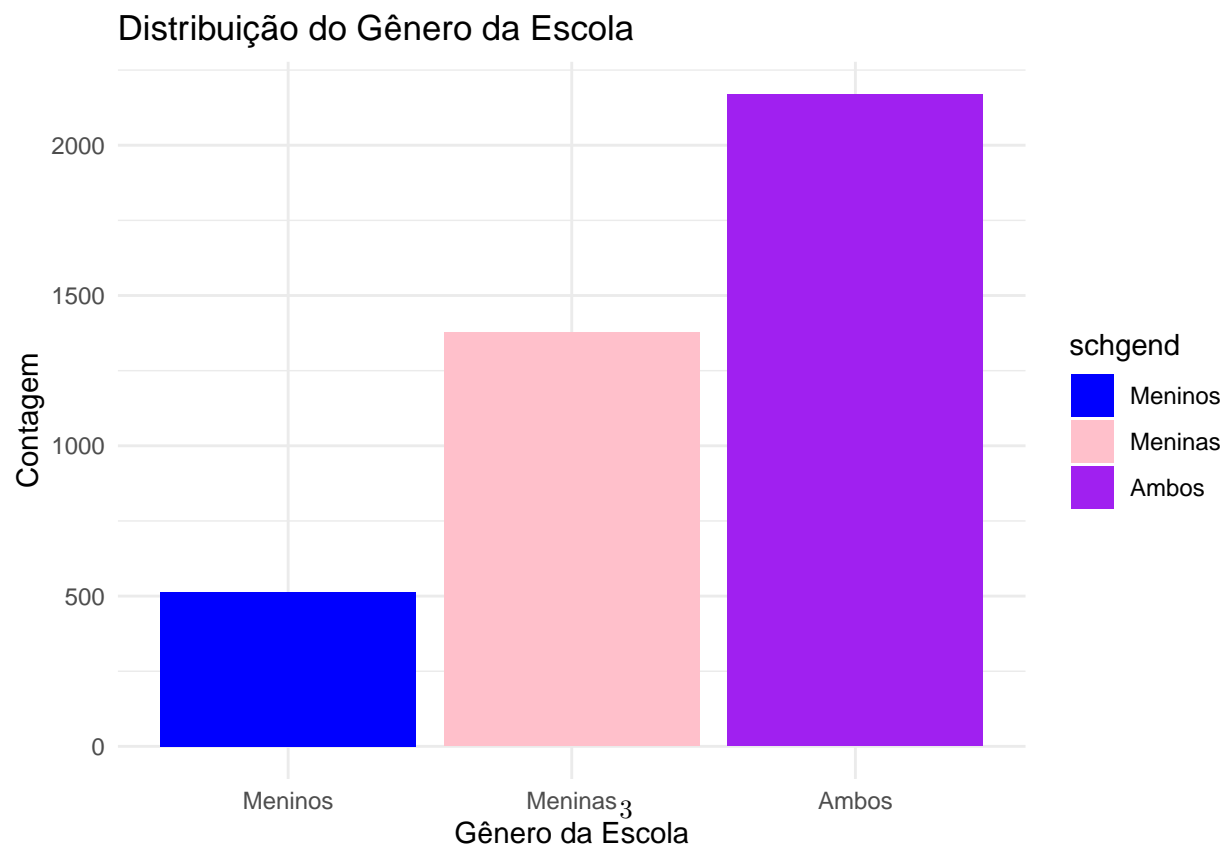
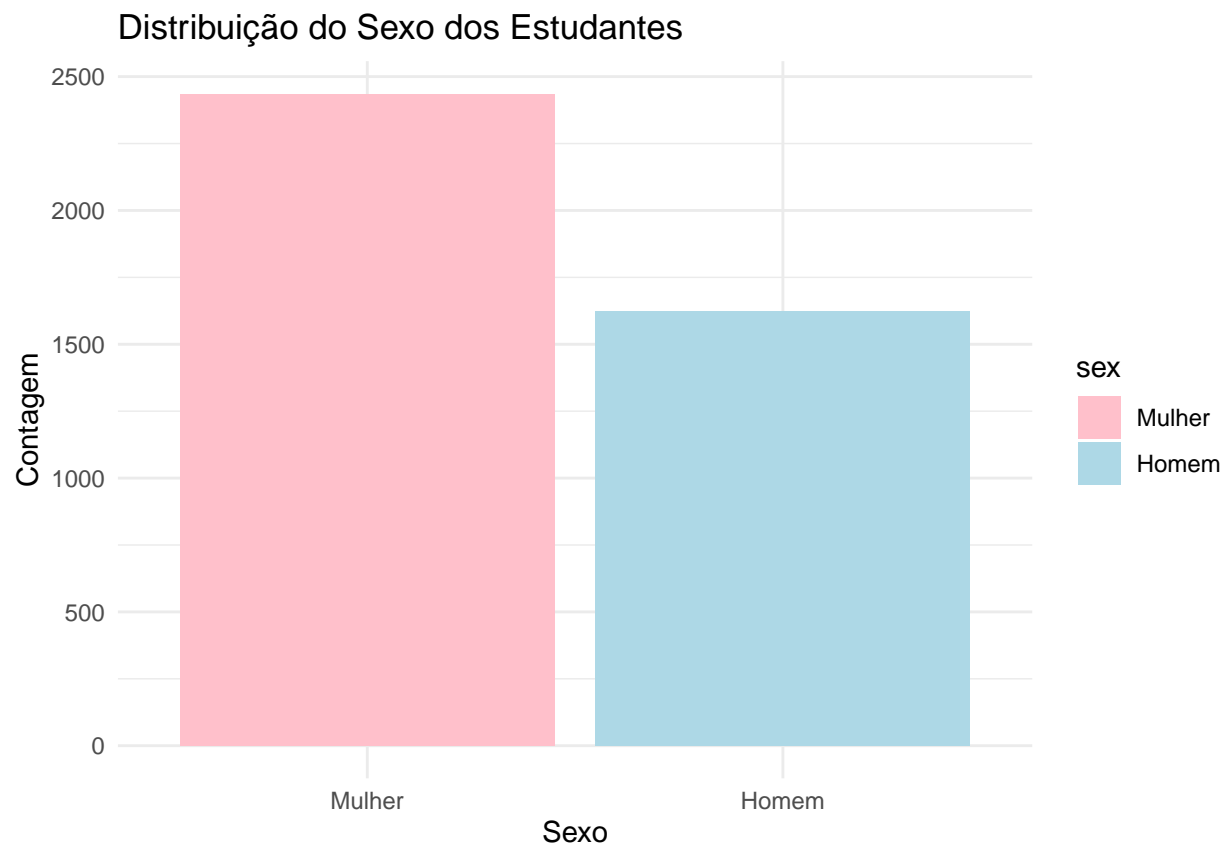
Boxplot da Pontuação Normalizada



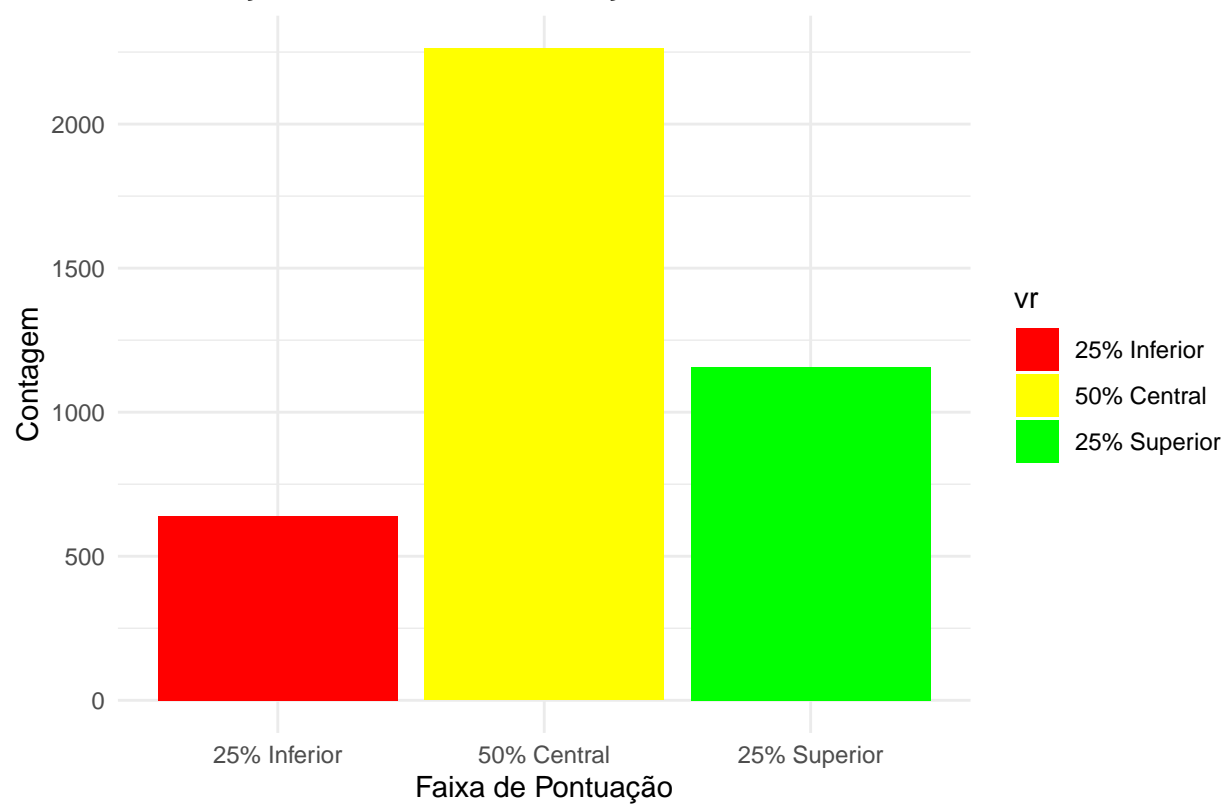
Boxplot da Pontuação do Teste RL

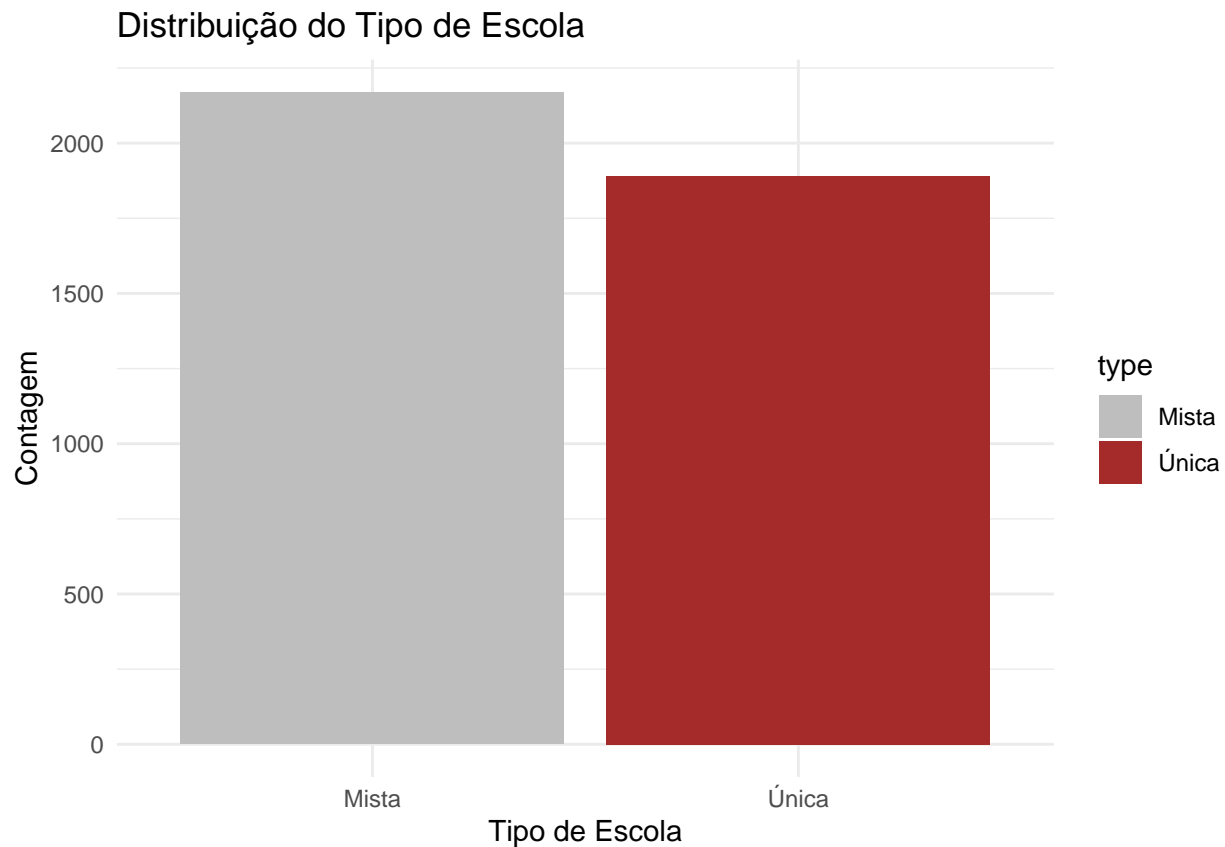


Partindo para as variáveis categóricas separei em alguns gráficos de barras para termos uma visão geral dos dados apresentados no banco, é interessante para visualizarmos a proporção de cada variável também.



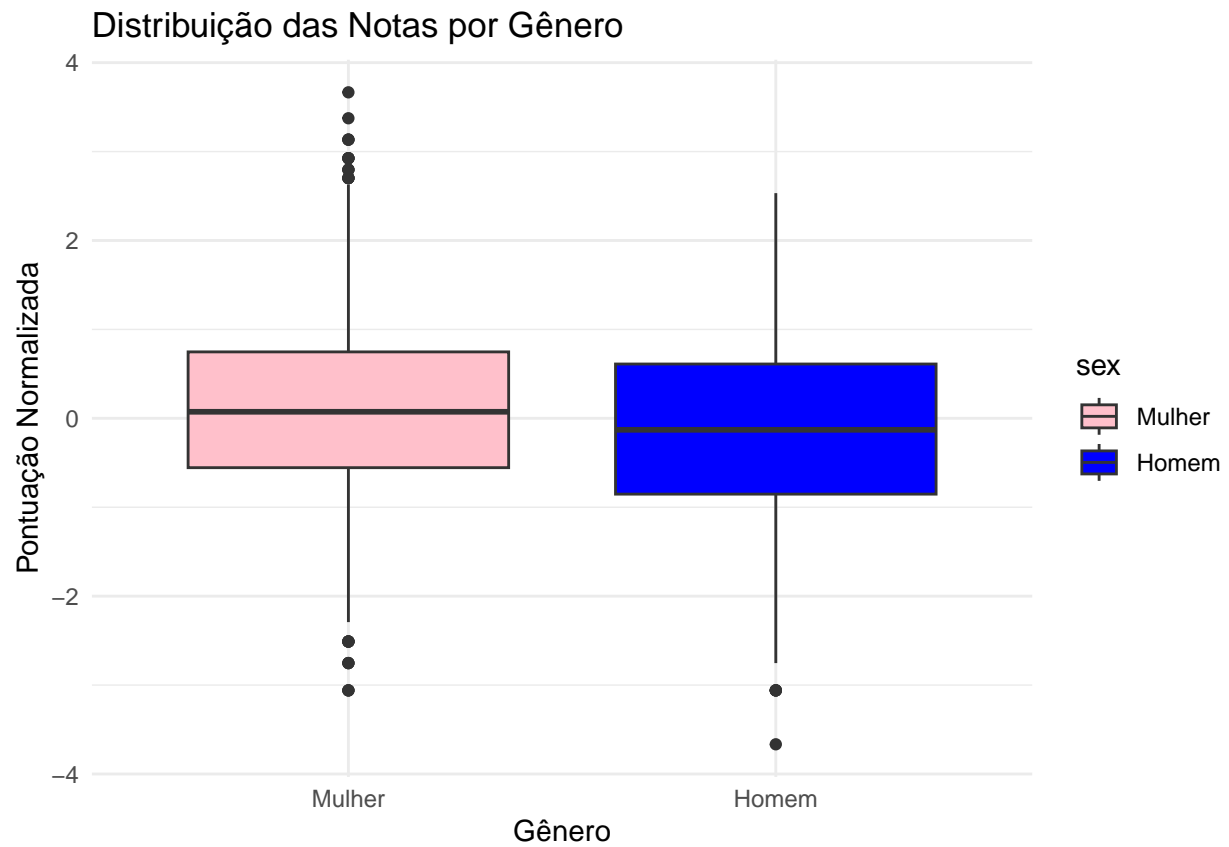
Distribuição da Faixa de Pontuação de Raciocínio Verbal



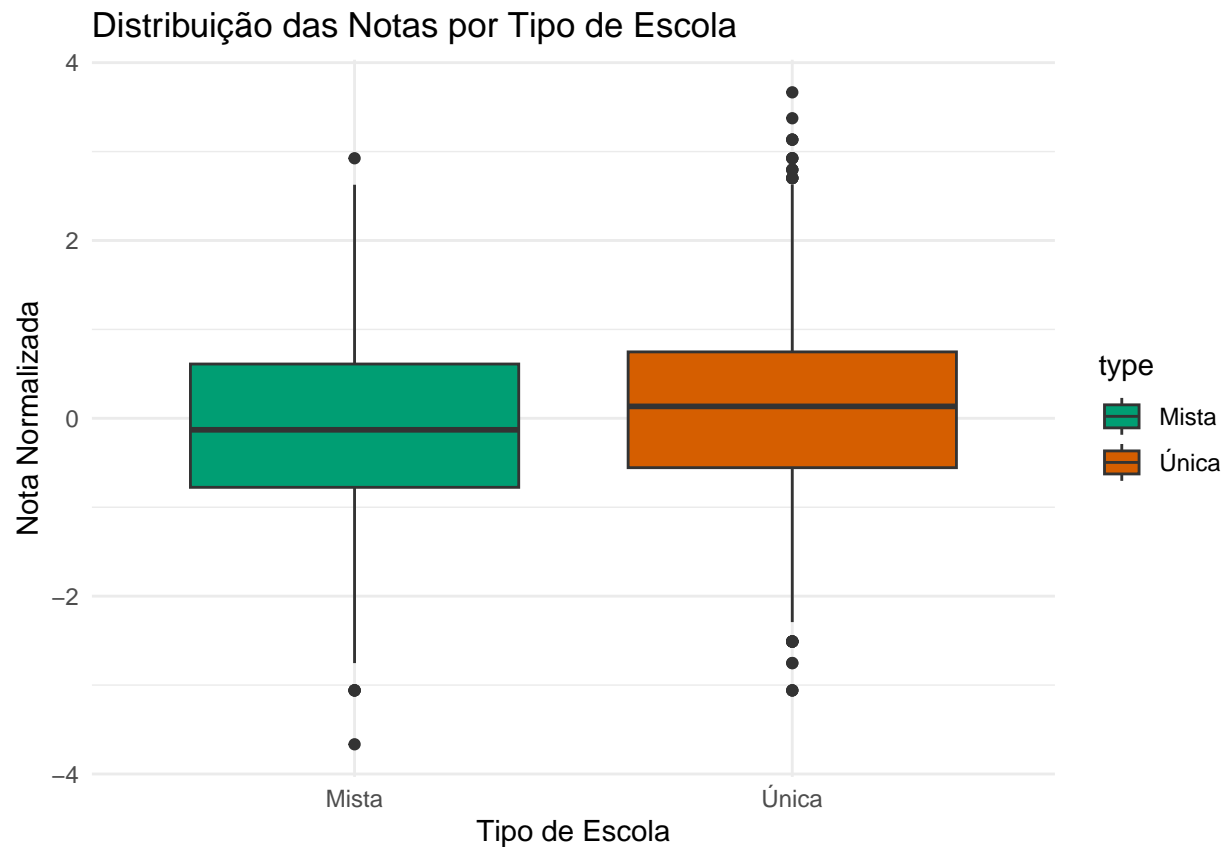


Agora que finalizamos a análise exploratória inicial, partirei para algumas análises mais específicas e objetivas, para tentarmos encontrar padrões entre algumas variáveis.

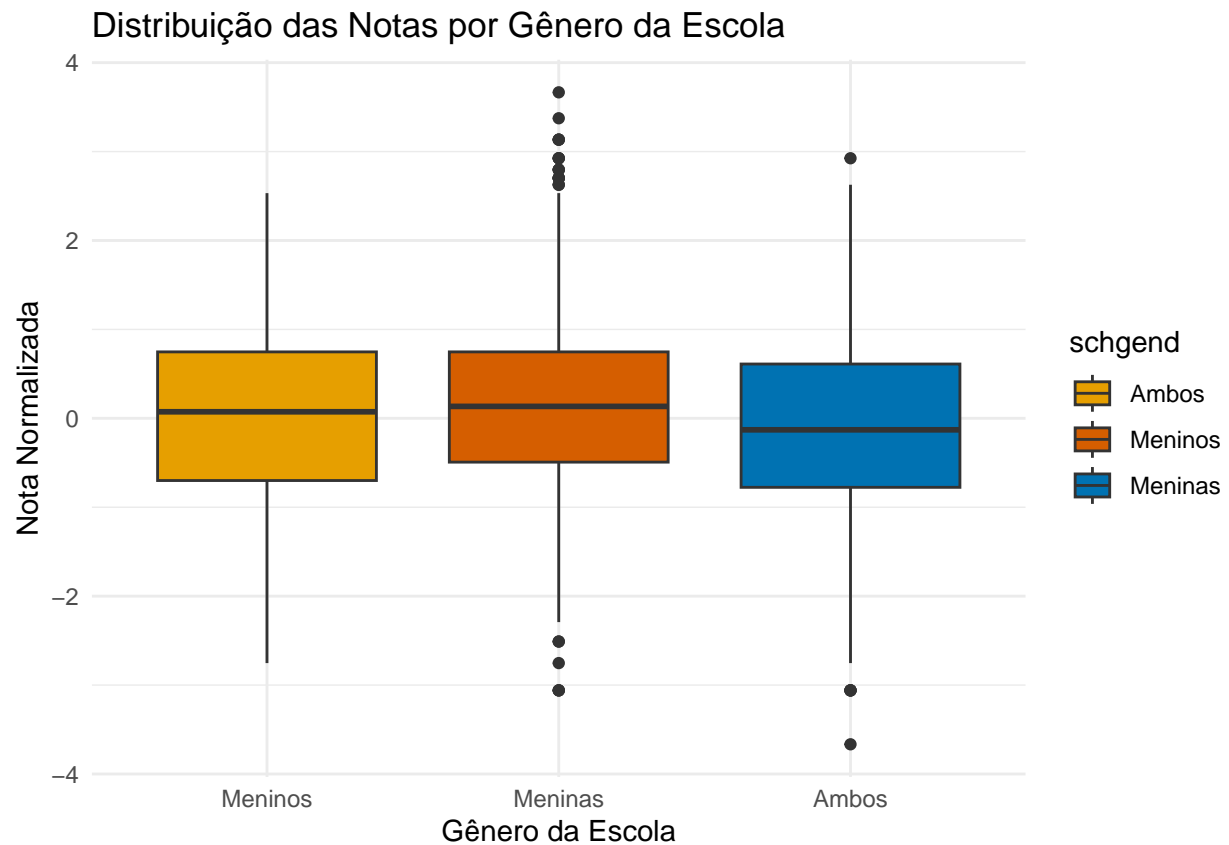
Começarei relacionando diversas variáveis com a variável “normexam”, para visualizarmos quais tem relação com a nota dos alunos serem mais altas/baixas.



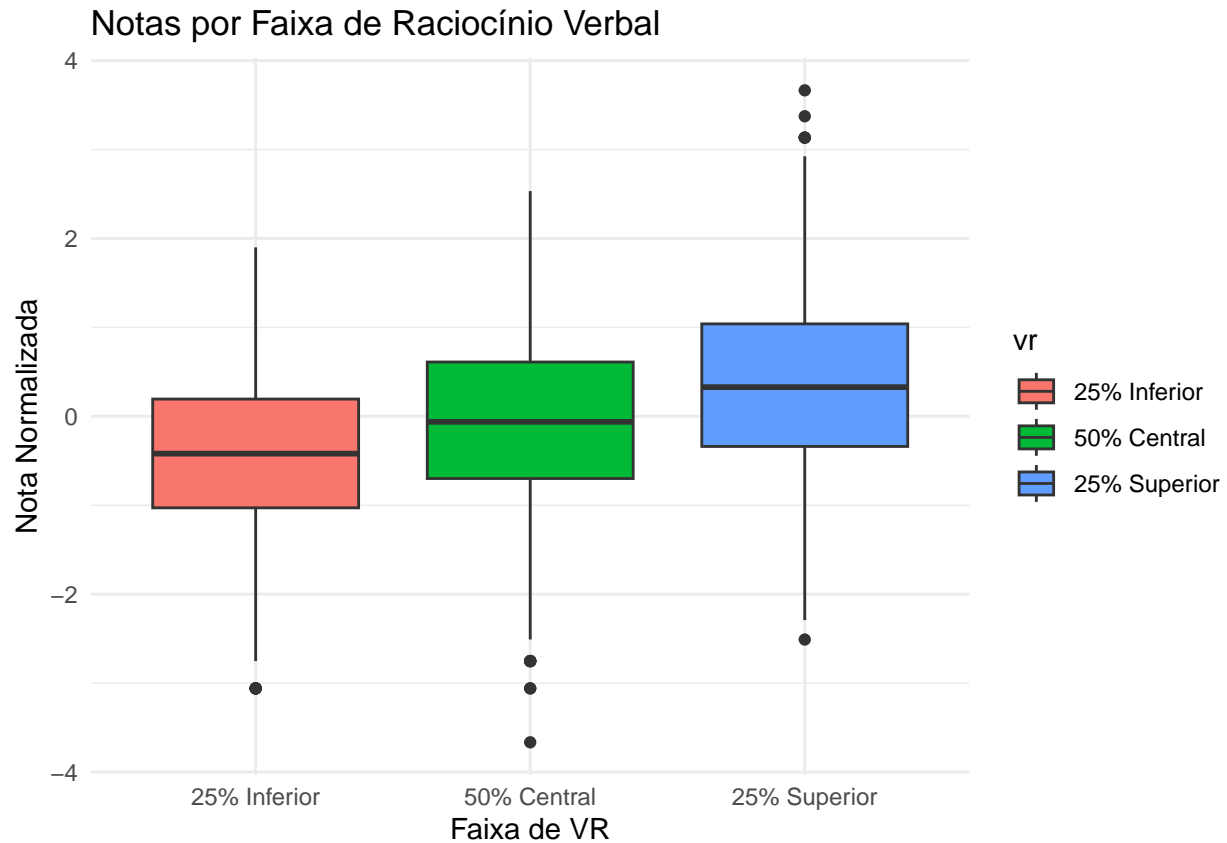
Aqui, notamos que as meninas tem a média de notas um pouco maior que os meninos, além de alguns outliers superiores e inferiores. os meninos apresentaram apenas dois outliers inferiores, e nenhum superior.



Aqui percebemos que o tipo de escola de gênero único tem a média da nota um pouco superior, isso pode se dar por algum fator de ensino mais controlado ou de maior qualidade acerca desse tipo de escola. Também vemos alguns outliers presentes tanto nas escolas de gênero misto quanto único.



Aqui vemos uma semelhança muito grande entre as escolas de meninos e meninas, mas percebe-se que tem uma média maior quando comparado com escolas que misturam os gêneros, como pudemos perceber também no gráfico anterior. Escolas de meninas tem vários outliers superior/inferior.

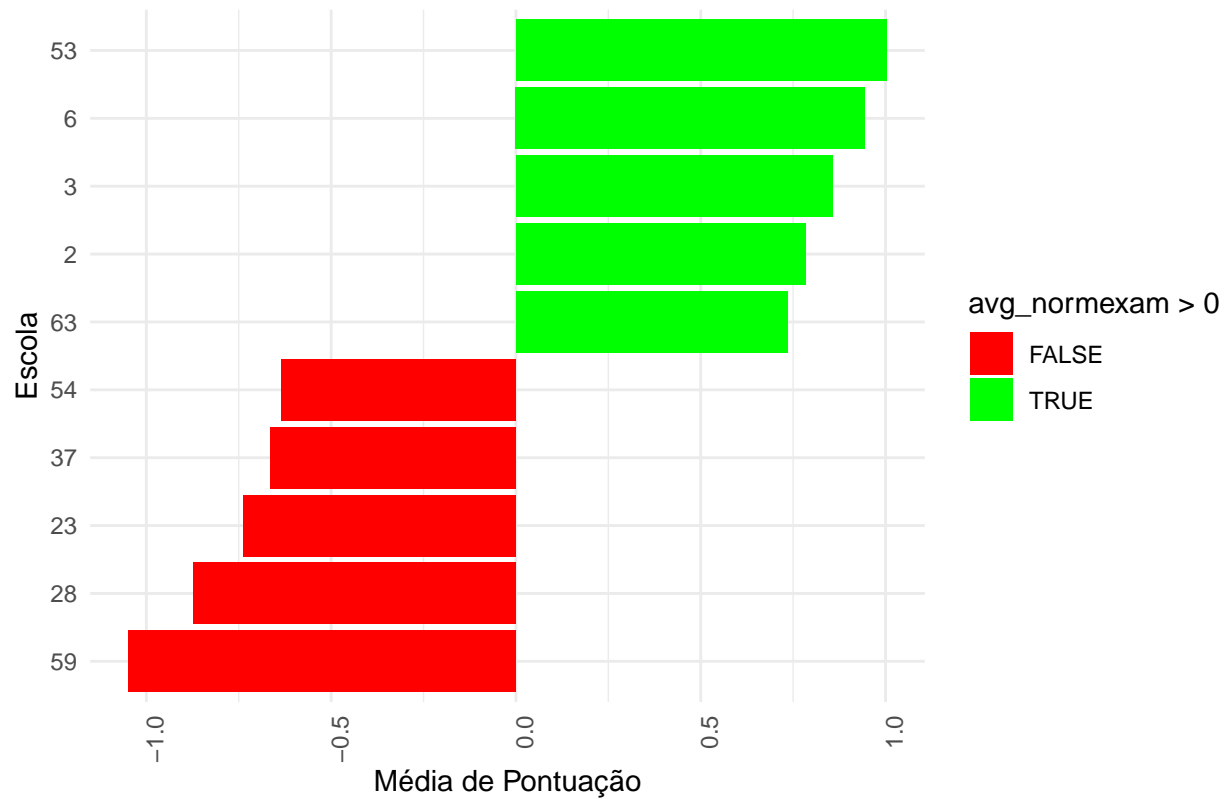


```
##
## Pearson's product-moment correlation
##
## data: data$normexam and data$vr_numeric
## t = 16.579, df = 4057, p-value < 2.2e-16
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.2228624 0.2804922
## sample estimates:
##      cor
## 0.2519006
```

Aqui percebemos uma relação interessante entre o teste VR e a nota média dos alunos. É notável que quanto maior a nota no teste VR, a nota média do aluno também é maior. Isso se deve pelo fato do teste VR medir o raciocínio verbal dos alunos, algo que está bem relacionado com a possível nota média deles.

Essa relação me criou a curiosidade de fazer um teste de correlação entre as duas variáveis, o qual me retornou uma correlação de ~0,25 na escala de Pearson.

5 Escolas com Maior e Menor Média de Notas (Normexam)



```
## # A tibble: 10 x 3
##   school avg_normexam schgend
##   <int>     <dbl> <chr>
## 1     59     -1.05 mixed
## 2     28    -0.873 mixed
## 3     23    -0.738 mixed
## 4     37    -0.665 boys
## 5     54    -0.634 mixed
## 6     63     0.736 mixed
## 7      2     0.783 girls
## 8      3     0.855 mixed
## 9      6     0.945 girls
## 10    53     1.00 girls
```

Esse gráfico mostra as 5 escolas com maior e menor média de notas e logo em seguida uma tabela que mostra essas 10 escolas e seu gênero. As 5 de baixo são as maiores notas e as 5 de cima as menores.

Conclusão

Com base nas 9 variáveis analisadas neste estudo de caso, podemos tirar algumas conclusões importantes sobre o desempenho dos estudantes nas escolas da região central de Londres:

1. **Desempenho Acadêmico (normexam):** A variável “normexam” mostrou uma distribuição com variação considerável nas pontuações entre os estudantes, refletindo a diversidade no desempenho acadêmico das escolas. As escolas com maior média de pontuação tendem a ter uma estrutura diferenciada, enquanto as de menor média de pontuação precisam de intervenções para melhorar a qualidade educacional.
2. **Gênero da Escola (schgend):** A análise revelou que a variável “schgend” (gênero da escola) impacta diretamente no desempenho dos estudantes. As escolas mistas (“mixed”) podem apresentar uma maior diversidade de abordagens pedagógicas, enquanto as escolas exclusivas para meninos ou meninas (“boys” e “girls”) podem ter um foco mais especializado no gênero, o que pode afetar o desempenho dos estudantes de forma distinta.
3. **Faixa de Pontuação de Raciocínio Verbal (vr):** A relação entre a “vr” e o desempenho no exame (“normexam”) mostrou que os estudantes nas faixas superiores de raciocínio verbal tendem a ter melhores desempenhos no exame. Esse achado sugere que intervenções específicas no desenvolvimento do raciocínio verbal podem melhorar o desempenho acadêmico geral.
4. **Tipo de Escola (type):** O tipo de escola (mista ou única) não apresenta uma correlação muito forte com o desempenho acadêmico geral, mas pode influenciar o tipo de ambiente de aprendizado e os recursos disponíveis para os estudantes. A comparação entre escolas mistas e únicas sugere que o contexto escolar pode influenciar, mas não é um fator determinante isolado.
5. **Sexo dos Estudantes (sex):** A diferença de desempenho entre os sexos foi analisada, e embora a média de desempenho entre meninos e meninas tenha apresentado variações, o desempenho acadêmico foi mais influenciado por outros fatores, como o tipo de escola e a faixa de raciocínio verbal do estudante.
6. **Média de Pontuação de Entrada (schavg):** A média de pontuação de entrada das escolas (schavg) foi um indicativo importante do nível geral de desempenho. Escolas com maior média de entrada tendem a ter uma base de alunos com melhores resultados nos exames, o que reflete a seleção do aluno no momento da entrada.

Após esse estudo podemos concluir que o desempenho acadêmico dos estudantes dependem de diversos fatores como o tipo de escola, gênero da escola, raciocínio verbal e outros. É algo muito difícil de relacionar diretamente com apenas uma variável e a que mais se relacionou foi a do teste vr.