

Relatorio

Fabio Firanzi, Heitor Dias, Julia Fideles, Matheus Soares, Tiago Braga

2025-11-11

| Presença na prova | Percentual | Frequência |
|-------------------|------------|------------|
| Ausente | 26.74% | 115880 |
| Presente | 73.12% | 316843 |
| Eliminado | 0.13% | 571 |

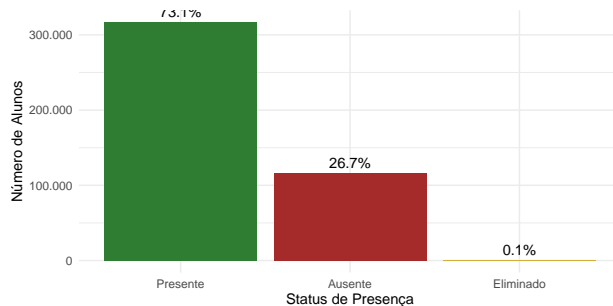


Tabela e gráfico de presença

A tabela “Frequência de presença na prova de LC” e o Gráfico “Presença na prova de LC” exibem a porcentagem de alunos presentes, ausentes e eliminados na prova de linguagens e códigos. Com base nessa tabela, pode-se identificar que a grande maioria, 73,1%, dos alunos estava presente na avaliação, 26,7% ausente e uma minoria de 0,1% foi eliminada antes, durante ou após a prova. Embora a maioria tenha completado a avaliação, o alto índice de ausência sugere a necessidade de estudos futuros para investigar os fatores que contribuem para essa abstenção.

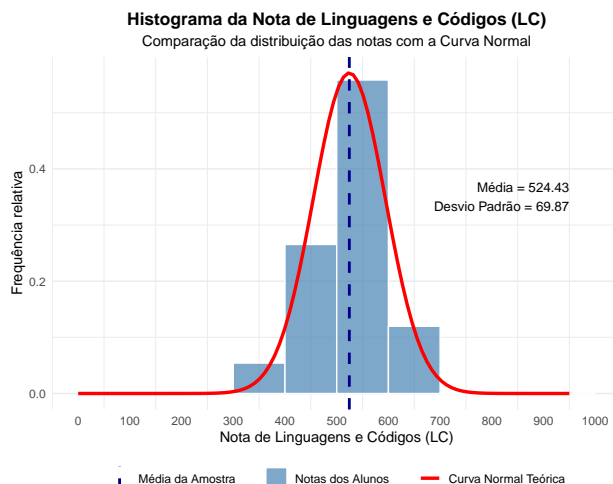


Tabela 1: Tabela Comparativa: LC vs MT (para 298.976 alunos presentes em ambas e com nota > 0)

| Prova | Média | Mediana | Desv. Padrão | Min | Máx |
|-------|--------|---------|--------------|-------|-------|
| LC | 526.70 | 533.3 | 68.03 | 298.8 | 795.8 |
| MT | 527.24 | 499.1 | 113.91 | 342.8 | 961.9 |

Histograma da nota de LC

Para a criação desse histograma foi utilizado a nota dos alunos presentes e não eliminados na avaliação de Linguagens e códigos, com esse dado sendo analisado em sua frequência, média, desvio padrão e distribuição.

O “Histograma da Nota de LC” ilustra a distribuição de frequência das notas de Linguagens e Códigos para todos os alunos presentes. Ele foi construído sobrepondo uma curva normal teórica (linha vermelha), calculada a partir da média e do desvio padrão da amostra, sobre o histograma das notas reais (barras azuis).

A maior concentração de notas é de 500-600 pontos, contendo aproximadamente 55% dos alunos, o desvio padrão das notas é de 69,87.

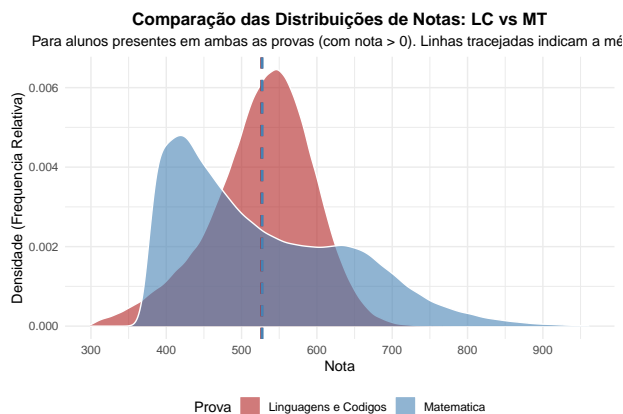
Tabela e gráfico LC vs MT

A tabela “Tabela comparativa: LC vs MT” ilustra as, médias, medianas, desvios-padrão, notas mínimas e notas máximas das provas de linguagens e matemática, considerando apenas os alunos que estiveram presentes em ambas delas.

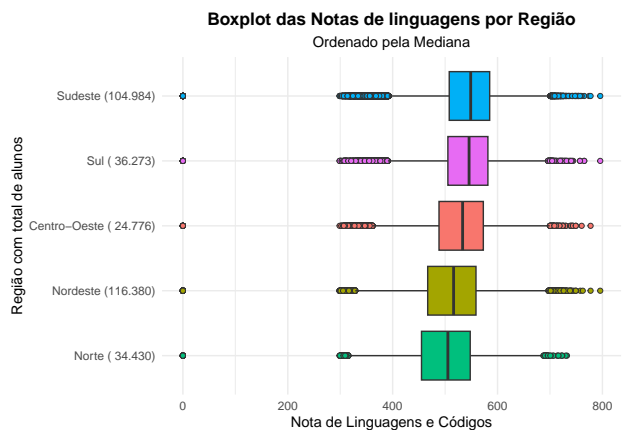
O gráfico “comparação das distribuições de notas:LC vs MT” exhibe os gráficos de distribuição de notas de linguagens e matemática sobrepostos, oferecendo uma comparação visual direta e clara da distribuição de notas das duas matérias.

Embora as médias de linguagens e matemática tenham sido muito próximas, 526,7 e 527,2 respectivamente, a distribuição de notas da prova de matemática é

totalmente diferente da de linguagens (que se aproxima de uma normal perfeita) tendo suas notas muito mais distribuídas no gráfico, evidenciando um desvio padrão muito maior.



O Gráfico de Densidade oferece uma comparação visual direta entre as distribuições das notas de Linguagens (vermelho) e Matemática (azul). Este gráfico foi gerado utilizando apenas os alunos que compareceram e obtiveram nota maior que zero em ambas as provas. Percebe-se que a curva de Matemática é mais achatada e espalhada (refletindo o maior desvio padrão), indicando que há uma variabilidade muito maior nas notas. Em contraste, as notas de Linguagens são mais “pontudas” e concentradas em torno de sua média.



O Boxplot complementa a tabela estatística anterior, visualizando a distribuição das notas de Linguagens e Códigos por região. Ele foi criado agrupando os alunos por região e ordenando o eixo Y pela mediana das notas, da maior para a menor. Este gráfico permite uma visualização clara não apenas da mediana (a linha central na caixa), mas também da dispersão do “meio” dos alunos (o tamanho da caixa, ou Intervalo Interquartil) e dos outliers (pontos). Observa-se que as regiões Sudeste e Sul apresentam as medianas mais elevadas, enquanto as regiões Nordeste e Norte mostram um desempenho mediano inferior e uma

| Região | Média | Mediana | Var | D. Padrão | Min | Máx |
|--------|--------|---------|---------|-----------|-------|-------|
| S | 542.13 | 548.0 | 3935.23 | 62.73 | 298.8 | 795.8 |
| CO | 527.52 | 533.5 | 4542.66 | 67.40 | 298.8 | 777.4 |
| NE | 510.60 | 516.3 | 4889.35 | 69.92 | 298.8 | 795.8 |
| N | 499.47 | 505.4 | 4805.12 | 69.32 | 298.8 | 732.3 |

dispersão de notas (tamanho da caixa) ligeiramente maior, indicando maior variação no desempenho dos alunos dessas regiões.

Tabela e boxplot de notas LC por região

A tabela “Estatísticas das notas de LC por Região” exibe as médias, medianas, variâncias, desvios-padrão, notas mínimas, notas máximas e número aproximado de alunos de cada região do Brasil.

O “Boxplot das notas de linguagens por Região”, ilustra a performance dos alunos de cada uma das regiões do Brasil exibindo a distribuição das notas. As “caixas” representam os 50% centrais dos alunos, com os pontos sendo os “outliers”, ou seja, os fora da média, tanto para baixo quanto para cima e a linha dentro da caixa representa a mediana. Com base no boxplot, a região Sudoeste obteve o melhor resultado, com sua média de 542,74 sendo a mais alta seguida de perto pela região Sul e sua média de 540.36, o boxplot também deixa evidente a desigualdade do país, com as regiões Norte e nordeste ficando significativamente atrás das demais.