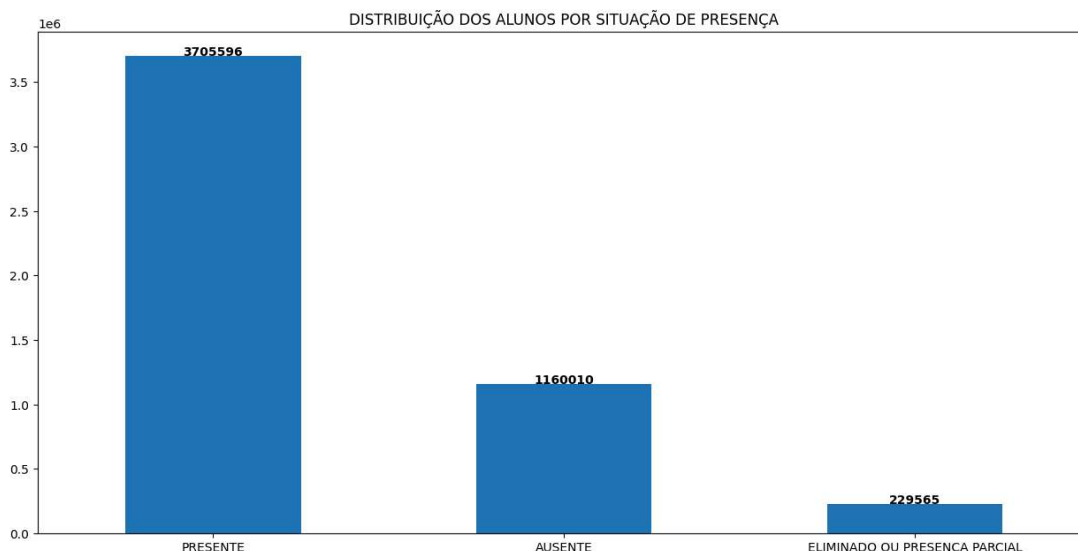


ANÁLISE EXPLORATÓRIA DOS DADOS

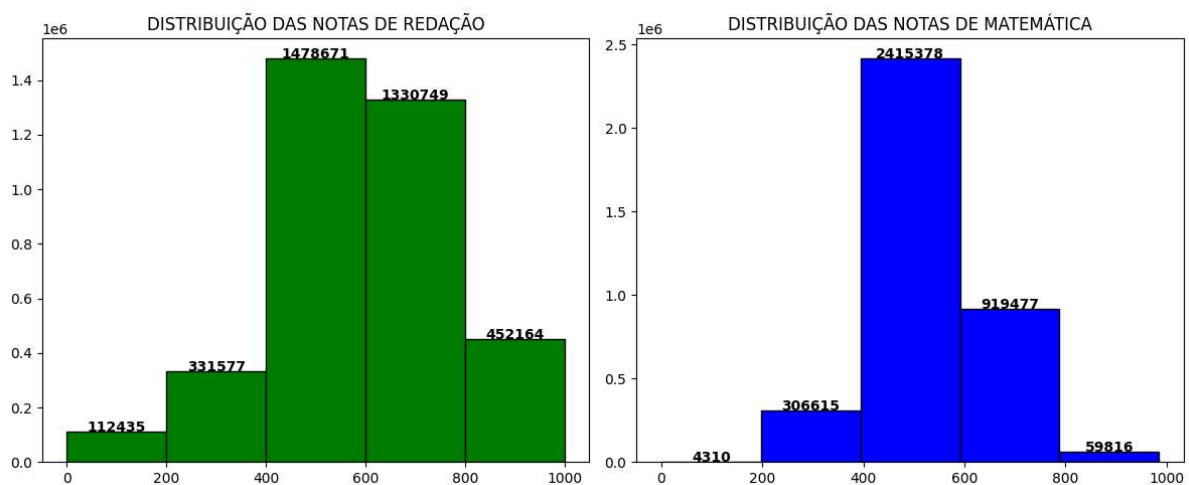
1. DISTRIBUIÇÃO DOS ALUNOS POR SITUAÇÃO DE PRESENÇA



Descrição: O gráfico apresenta a quantidade de alunos presentes, ausentes, eliminados ou com presença parcial.

Interpretação: Entre mais de cinco milhões de inscritos, a maioria esteve presente em todas as provas. Destaca-se, contudo, o elevado número de ausentes (mais de um milhão), além dos casos de eliminação e presença parcial, que merecem análises para compreender os fatores que levaram a esses resultados.

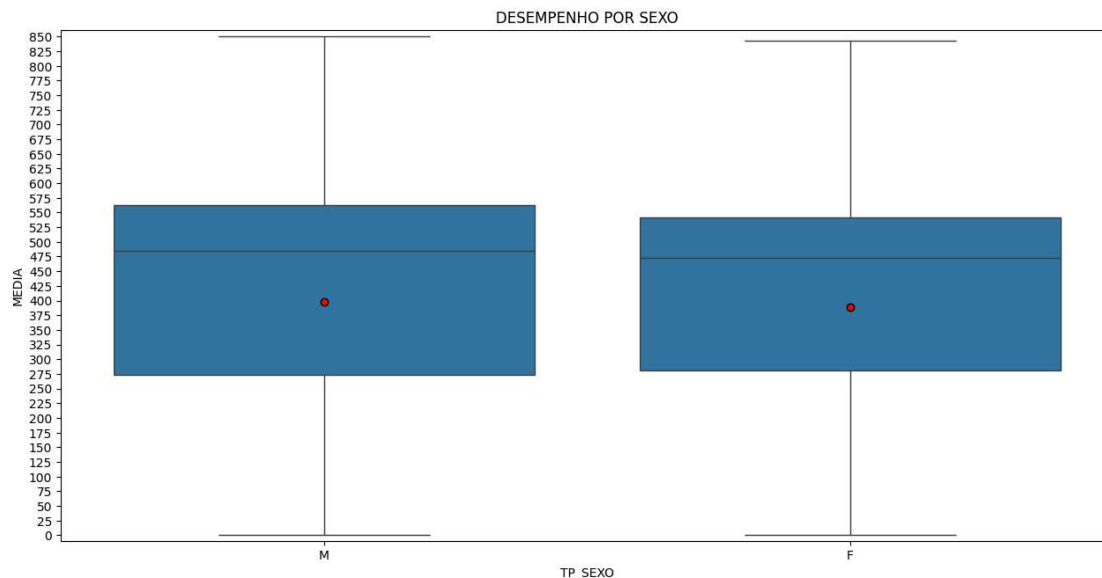
2. DISTRIBUIÇÃO DAS NOTAS DE REDAÇÃO E MATEMÁTICA



Descrição: O gráfico apresenta a distribuição das notas dos alunos presentes, considerando a frequência de desempenho em cada disciplina.

Interpretação: Observa-se que, em redação, pouco mais de 400 mil candidatos obtiveram notas altas, enquanto mais de dois milhões ficaram entre 400 e 800 pontos. Em matemática, apenas cerca de 50 mil atingiram a pontuação máxima, sendo que a maioria, também acima de dois milhões, concentrou-se na faixa de 400 a 800 pontos. Esses resultados evidenciam a necessidade de investigar os fatores que favorecem ou dificultam o bom desempenho, de modo a orientar políticas educacionais mais eficazes.

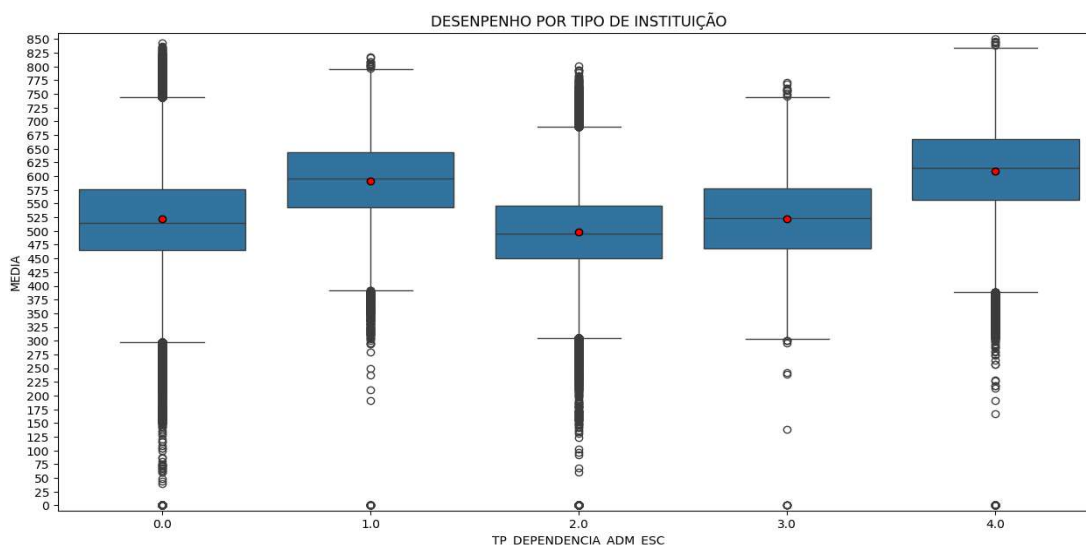
3. DESEMPENHO POR SEXO



Descrição: O gráfico apresenta a distribuição das notas dos alunos presentes, considerando por sexo a média.

Interpretação: o desempenho central e a variabilidade dos resultados são comparáveis entre homens e mulheres, com ambos os grupos atingindo resultados máximos semelhantes.

4. DESEMPENHO POR TIPO DE INSTITUIÇÃO

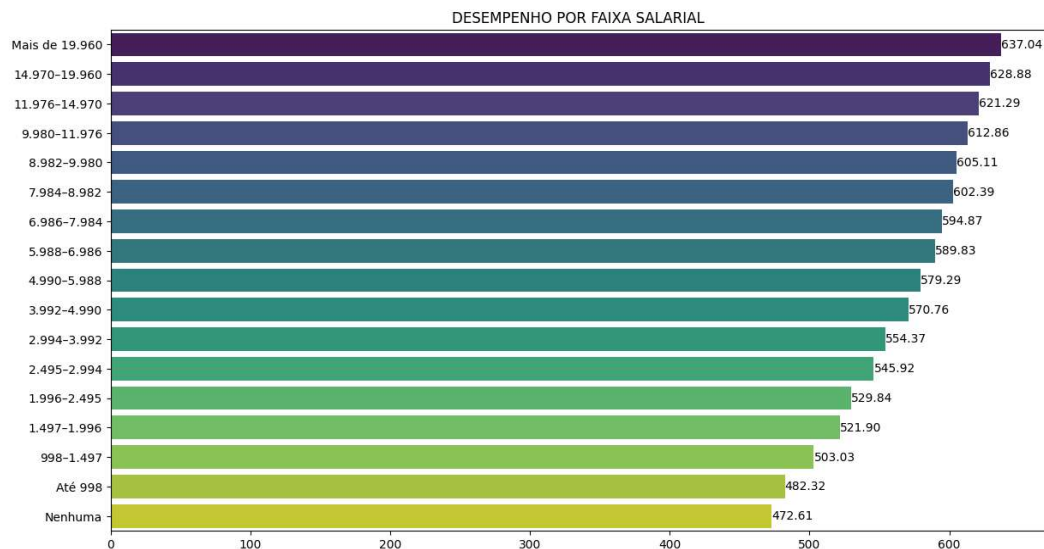


Legenda: 0 = Não declarado, 1 = Federal, 2 = Estadual, 3 = Municipal, 4 = Particular

Descrição: O gráfico apresenta a distribuição da média dos alunos presentes, considerando o tipo de instituição: não declarado, federal, estadual, municipal ou particular.

Interpretação: Os melhores resultados concentram-se em instituições privadas e federais. No entanto, em todos os tipos de instituição há alunos com boas performances. Os outliers superiores e os “bigodes” próximos ao valor máximo (850) indicam que o alto desempenho individual é possível independentemente do tipo de instituição.

5. DESEMPENHO POR FAIXA SALARIAL



Descrição: O gráfico apresenta a média das notas dos alunos presentes, considerando o tipo de renda familiar.

Interpretação: Observa-se que alunos de famílias com renda mais elevada tendem a obter desempenhos superiores em relação aos demais.

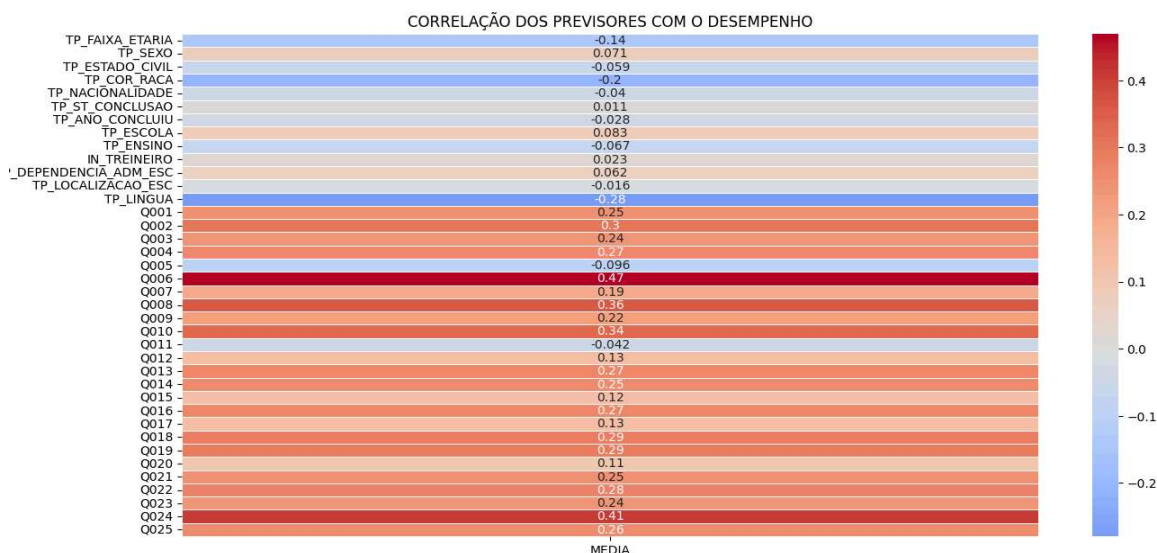
6. ANALISANDO A EXISTÊNCIA DE CORRELAÇÃO

Variáveis analisadas: Média e faixa etária.

Coefficiente de correlação: -0,14

Interpretação: Correlação fraca, praticamente inexistente, indicando que o desempenho no ENEM não tem nenhuma correlação com a idade, nesse caso faixa etária.

7. ANALISANDO AS CORRELAÇÕES COM O DESEMPENHO (MÉDIA)



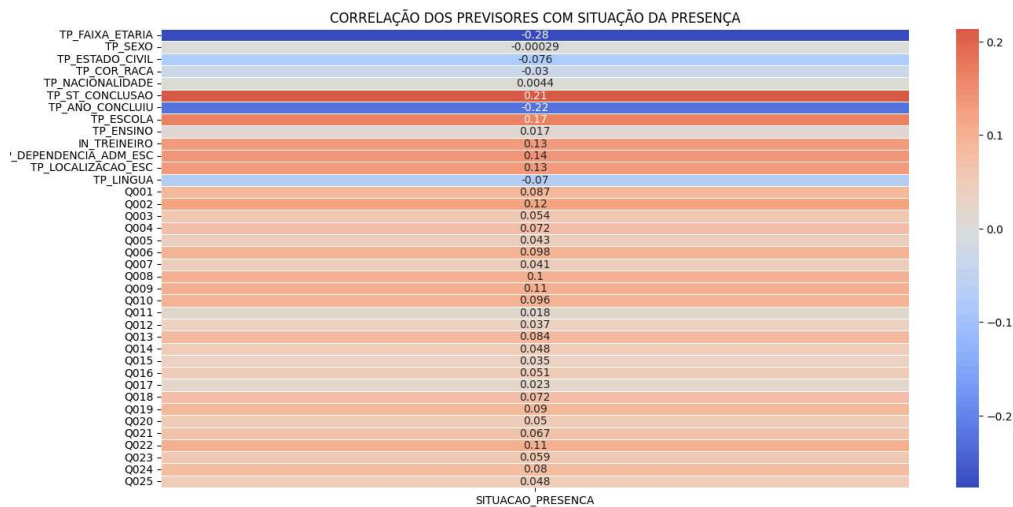
Descrição: O gráfico apresenta como as variáveis predictoras se relacionam com a média (desempenho).

Interpretação: Observa-se que existem bastantes variáveis que possuem uma correlação positiva de **média**. As variáveis que apresentam maior correlação são as seguintes:

- **Q6:** Variável que contém a renda familiar ou pessoal do participante, com valor de aproximadamente 0,47, indica que a renda está **associada** ao desempenho do candidato, pois pode proporcionar maiores oportunidades para o desenvolvimento estudantil.

- **Q24:** Variável que contém a quantidade de computadores na residência, com valor de aproximadamente 0,41, indica que ter computadores em casa está **associado** a melhores resultados na busca de conhecimento, facilitando o estudo.

8. ANALISANDO AS CORRELAÇÕES COM TIPO DE PRESENÇA



Descrição: O gráfico apresenta como as variáveis previsoras se relacionam com a situação da presença.

Interpretação: Observa-se que existem algumas variáveis que possuem maior correlação com a situação de presença codificada numericamente. As variáveis que apresentam maior correlação são as seguintes:

- **TP_ANO_CONCLUIU:** Variável que contém o ano de conclusão do ensino médio do participante, com valor de aproximadamente 0,21, indica que há alguma associação estatística entre o ano de conclusão e a situação de presença do candidato. Isso significa que o ano de conclusão pode ter relação com a codificação atribuída à situação, mas não necessariamente indica causalidade ou efeito direto.
- **TP_ESCOLA:** Variável que contém o tipo de escola no ensino médio, com valor de aproximadamente 0,17, indica que há alguma associação estatística entre o tipo de escola e a situação de presença do candidato. Dessa forma, mesmo que exista uma correlação, não se pode concluir diretamente que o tipo de escola determina a presença, ausência, eliminação ou participação parcial do candidato; outras variáveis devem ser consideradas para entender o comportamento