

Relatório de Análise dos Dados do Prouni

Introdução

Este relatório apresenta uma análise detalhada dos dados sobre as bolsas concedidas pelo Programa Universidade para Todos (Prouni) - <https://dadosabertos.mec.gov.br/prouni/item/89-bolsas-e-perfil-2018>, focando em diferentes dimensões, como características das bolsas, cursos e perfil dos beneficiários. O conjunto de dados abrange informações sobre o detalhamento quantitativo das bolsas concedidas por ano, distribuídas por região, UF (Unidade da Federação), município, instituição de educação superior, curso, modalidade de ensino (presencial ou EAD), turno e tipo de bolsa. Além disso, o perfil dos beneficiários é segmentado por sexo, raça/cor, faixa etária e deficiência.

O objetivo principal da análise é explorar as características das bolsas e os dados sobre os beneficiários, identificando tendências e padrões ao longo do tempo. A análise também visa fornecer insights sobre desigualdades e potenciais melhorias no sistema de concessão de bolsas do Prouni.

Metodologia

A análise foi conduzida utilizando a linguagem de programação Python e as bibliotecas pandas, matplotlib e seaborn. O processo de análise seguiu as seguintes etapas:

Tratamento de Dados:

- Limpeza de Dados: Identificação e tratamento de valores ausentes (nulos) com a substituição por valores de tendência central, como médias ou modas.
- Conversão de Tipos de Dados: Conversão de colunas para os tipos corretos, como datetime para datas e int/float para valores numéricos.

Exploração Inicial dos Dados:

- Distribuições de Notas: Utilização de histogramas para entender a distribuição das notas dos candidatos e das notas de corte.
- Gráficos de Barras: Identificação das UFs com maior quantidade de bolsas concedidas e análise da quantidade de vagas por tipo de concorrência.
- Boxplots: Comparação das notas mínimas exigidas por tipo de concorrência.

Análise de Performance dos Candidatos:

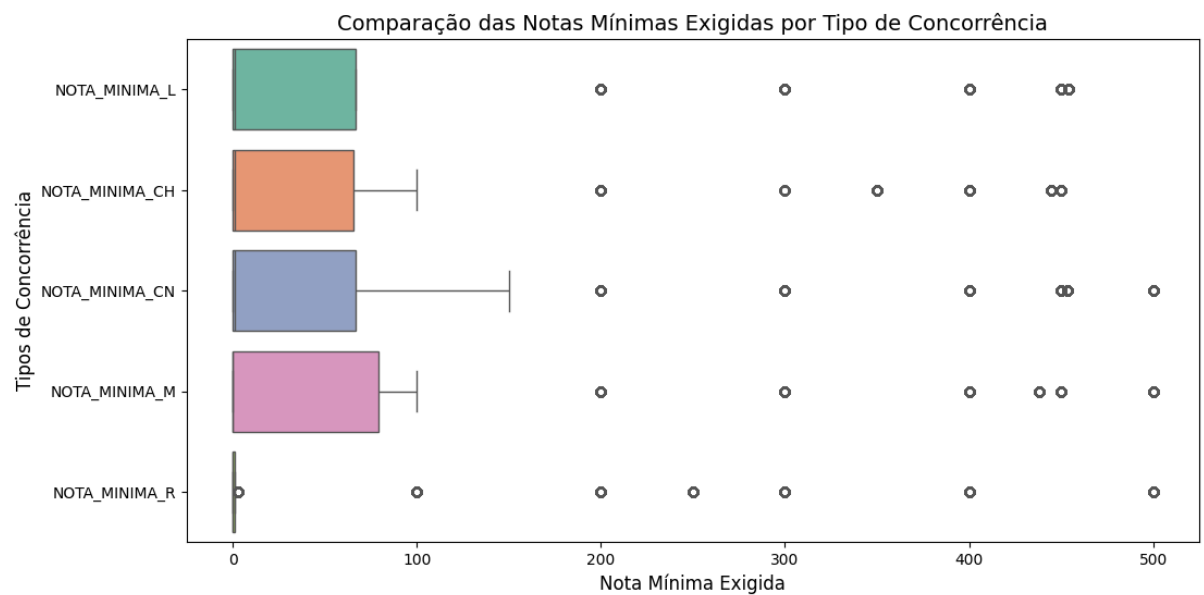
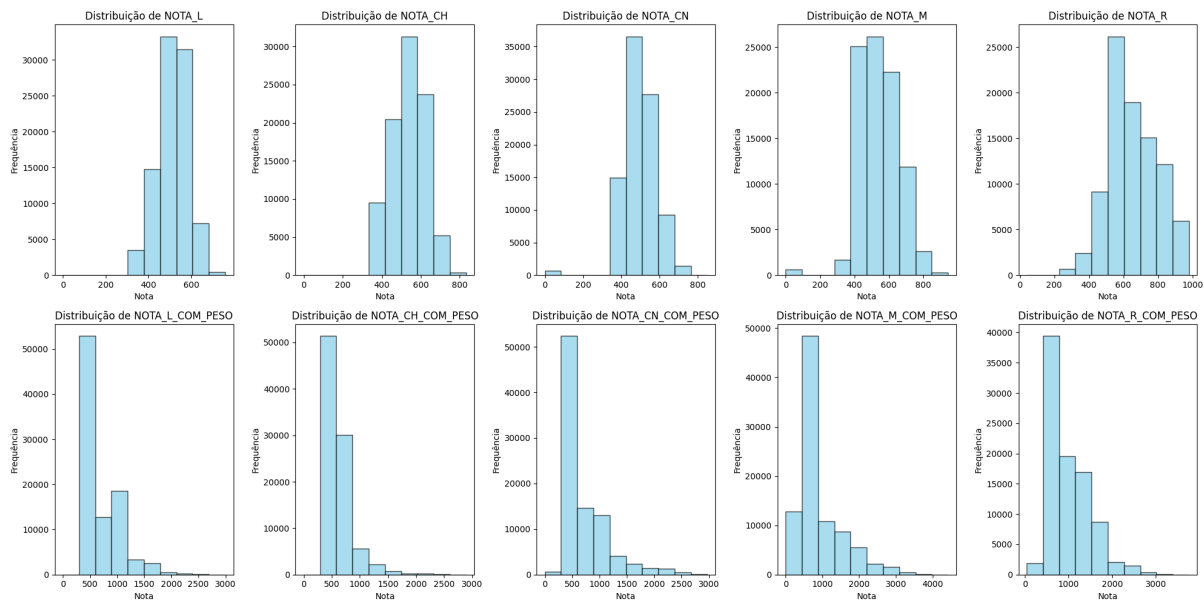
- Gráficos de Barras e Setores: Análise do percentual de aprovação por sexo, para identificar possíveis desigualdades entre os gêneros.
- Análise de Perfil: Análise do perfil dos beneficiários por características demográficas como sexo, raça/cor, idade, etc.

Resultados

Os resultados da análise são apresentados através de gráficos e tabelas que fornecem insights sobre os dados. A seguir, são destacados alguns dos resultados mais relevantes:

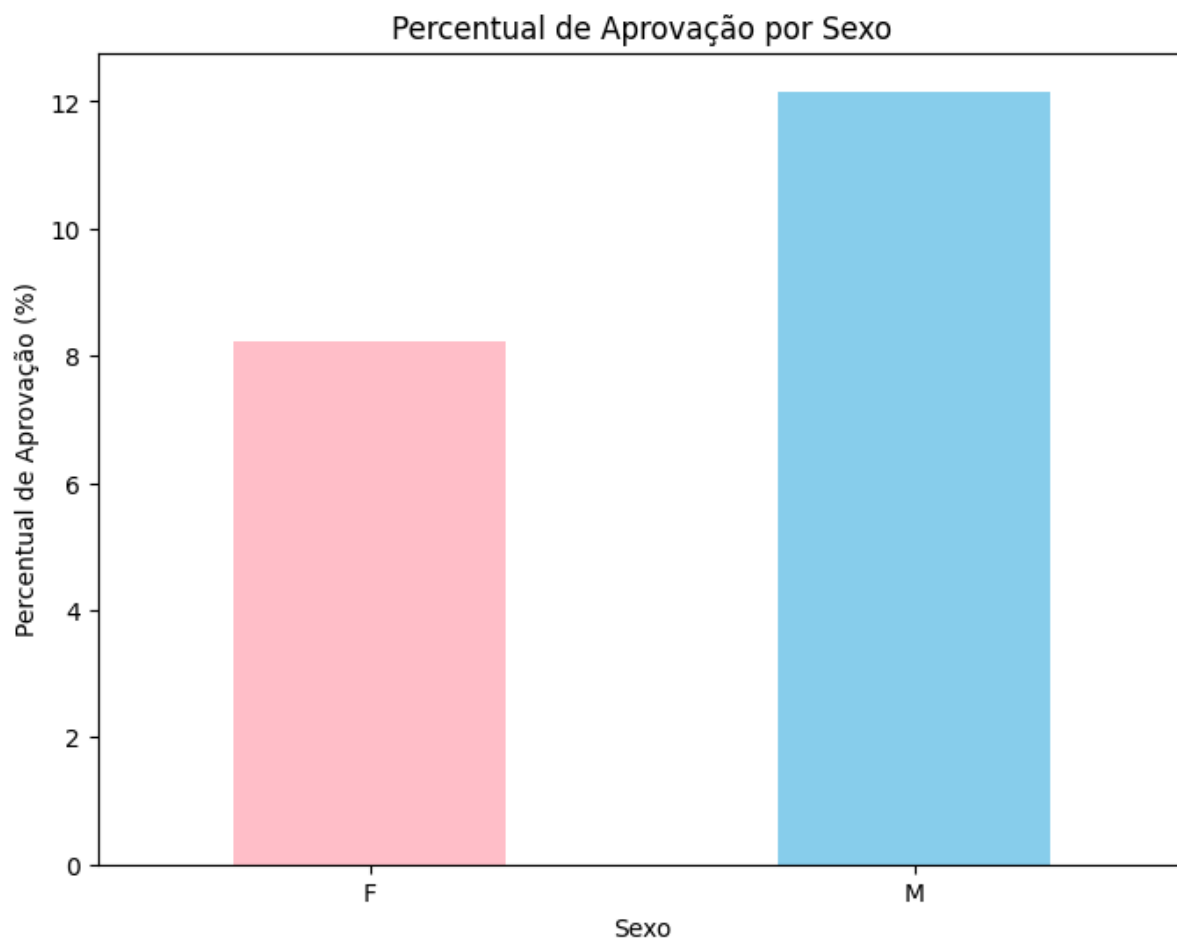
Distribuição das Notas:

A distribuição das notas dos candidatos e das notas ponderadas foi visualizada em histogramas, permitindo observar como as notas variam entre as diferentes disciplinas (L, CH, CN, M, R) e as notas de corte (NOTA_CORTE).



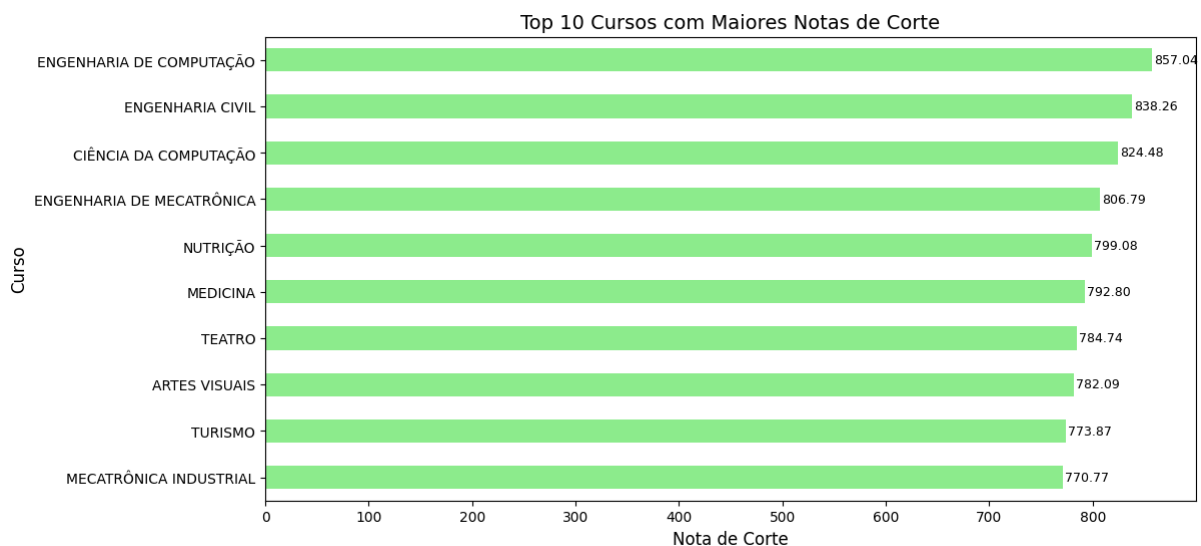
Aprovação por Sexo:

O gráfico de barras ou setores revelou que a distribuição de aprovações entre homens e mulheres apresenta uma leve desigualdade, com uma ligeira predominância de aprovações entre as mulheres.



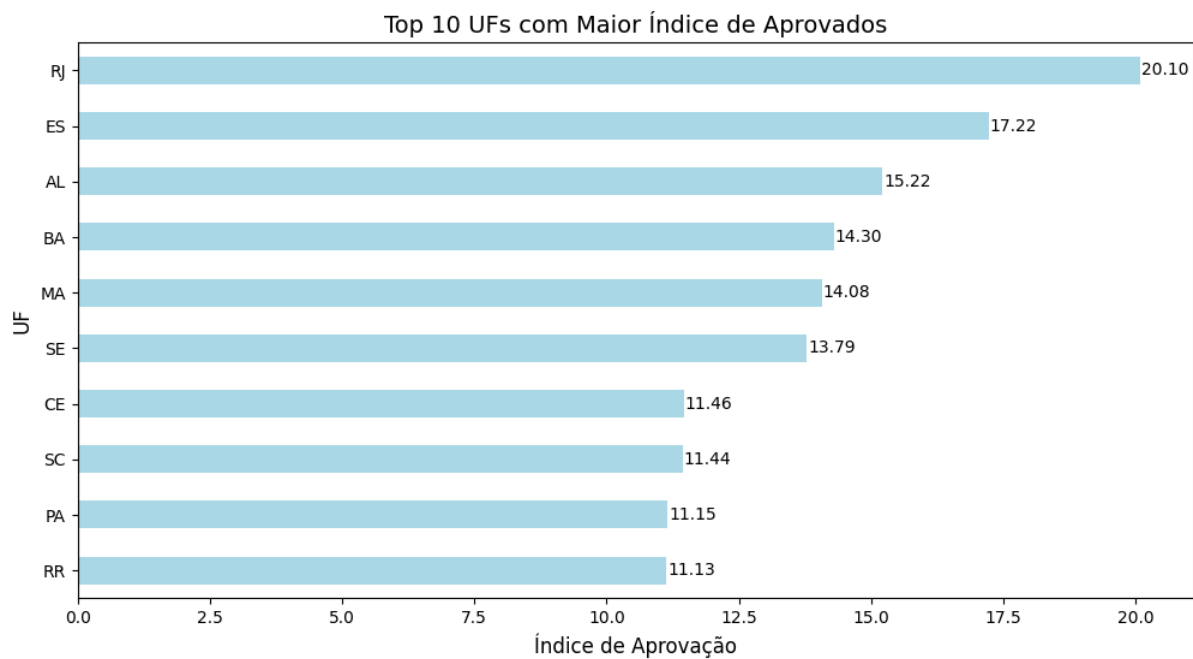
Cursos com Maiores Notas de Corte:

A análise dos cursos com as maiores notas de corte foi realizada com um gráfico de barras horizontais, destacando os cursos de alta demanda no programa.



Distribuição por UF:

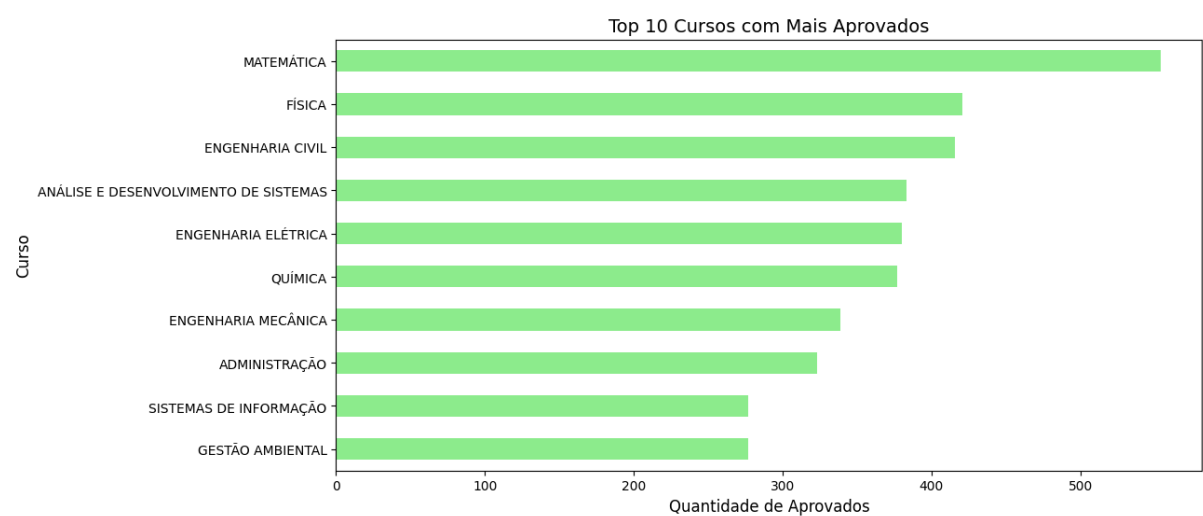
Um gráfico de barras empilhadas mostrou a distribuição dos aprovados por diferentes estados, com as UFs que mais concederam bolsas.



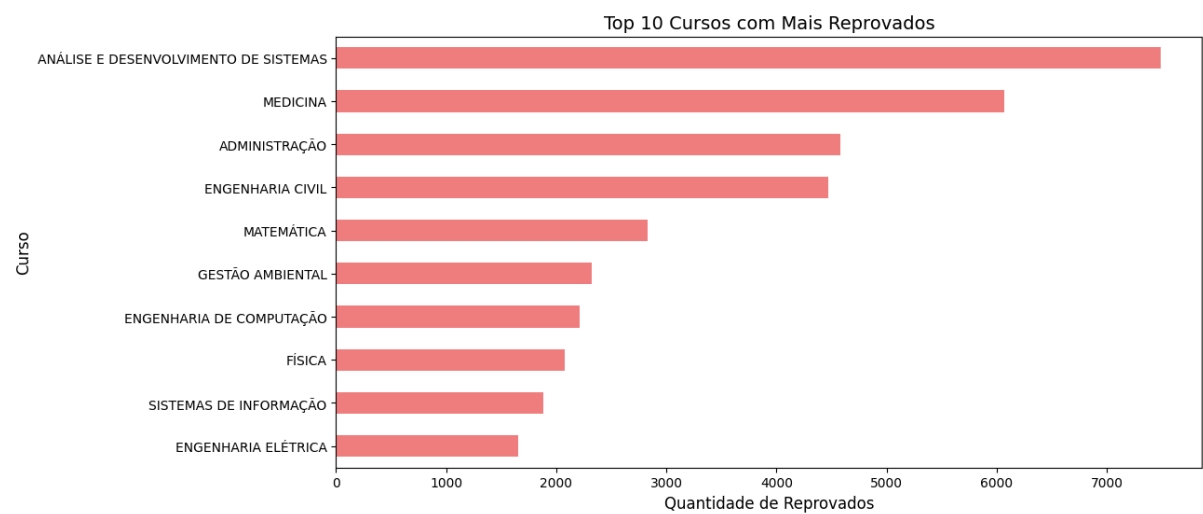
Cursos com Mais Aprovados e Reprovados:

A análise dos cursos com mais aprovados e cursos com mais reprovados foi realizada com gráficos de barras, permitindo observar quais cursos tiveram o maior número de aprovações e reprovações. Esses resultados podem ser úteis para entender a dificuldade e a competitividade dos cursos.

Cursos com Mais Aprovados: Um gráfico de barras horizontais mostra os cursos que têm o maior número de aprovados, com destaque para os cursos com maior taxa de sucesso.



Cursos com Mais Reprovados: Da mesma forma, outro gráfico de barras horizontais foi gerado para exibir os cursos com o maior número de reprovados, ajudando a identificar cursos com maior taxa de reprovação



Conclusão

A análise dos dados do Prouni revela várias informações interessantes que podem ajudar a entender melhor o funcionamento e as tendências do programa. Algumas conclusões importantes incluem:

Desigualdade de Gênero: Existe uma leve desigualdade na taxa de aprovação entre homens e mulheres, com as mulheres apresentando uma taxa de aprovação ligeiramente superior.


Distribuição Geográfica das Bolsas: As bolsas estão distribuídas de forma desigual entre os estados e municípios, com algumas UF's apresentando uma concentração maior de bolsas.

Concorrência e Notas de Corte: Cursos com maiores notas de corte tendem a ser os mais concorridos, com uma alta demanda de candidatos, especialmente em áreas específicas.

Desempenho por Curso: Cursos com maior taxa de aprovação possuem características distintas em relação à concorrência, enquanto cursos com mais reprovação podem sugerir dificuldades maiores ou questões relacionadas à acessibilidade.

Esses achados podem ser usados para otimizar a distribuição das bolsas e para desenvolver políticas mais inclusivas que atendam às necessidades dos candidatos, considerando fatores como sexo, raça/cor e deficiência.

Apêndice

 Atividade P1.ipynb