

Um Estudo de gênero, raça e financiamento sobre as Instituições de Ensino Superior do Brasil

Alberto Santos, Luiz Felipe, Pedro Moura

*Universidade Federal de Pernambuco – Centro de Informática
Caixa Postal 7.851 – 50.732-970 – Recife, PE – Brasil
ass6@cin.ufpe.br, lfvg@cin.ufpe.br, pvbm@cin.ufpe.br*

Abstract

Este documento foi elaborado com o objetivo de agregar uma série de informações referentes à nossa pesquisa, mais especificamente sobre informações raciais, de gênero e de financiamento das Instituições de Ensino Superior Brasileiras. Com o auxílio de diversas ferramentas disponíveis para ciência dos dados, buscamos compreender como acontece a distribuição de gênero nos cursos da área de ciências exatas, distribuição racial nos cursos mais populares em relação aos estados brasileiros, e como o financiamento nas instituições de ensino superior são realizadas, assim como suas implicações, como uma forma de melhor compreender e delinear uma visualização do cenário do ensino superior brasileiro.

1. Introdução

O Brasil possui uma alta miscigenação racial, onde as mais predominantes segundo estatísticas do IBGE são de raça branca e parda, onde 45,5% da população se declara branca, 45% se declara de cor parda, 8,6% se declaram de
5 raça negra, enquanto que o resto da população se declara como indígena ou Amarelo. Há também uma diferença entre as regiões do país, onde na Região Sul 76% se declara branca, nas Regiões Norte e Nordeste temos 69,3% e 61,9% considerados como pardos. Com tantas diferenças por região e suas raças mais predominantes, fizemos uma relação para ver como é a distribuição racial no
10 ensino superior brasileiro. Também focamos em alguns outros ramos, como a

avaliação de gêneros nas instituições de ensino superior brasileiras, assim como, questões financeiras dentre elas.

2. Método

Objetivamos conduzir análises no website do INEP. Para a concretização de
15 nosso estudo, utilizamos a linguagem de programação Python 3 ao decorrer de
diversas etapas — desde a coleta até a análise. A partir desta, desenvolvemos
scripts para realizar a coleta dos dados. Em seguida, realizamos etapas de pré-
processamento para que pudéssemos transformar os dados de forma que estes
se tornassem compatíveis com as análises que pretendíamos realizar.

20 Isso feito, o próximo passo fora utilizar de diversas bibliotecas da linguagem
de programação que possibilitassem a realização de nossas análises, exemplo a
biblioteca *Holoviews*. Através da elaboração de visualizações, buscamos com-
preender os critérios que influenciam os diferentes problemas referentes as insti-
tuições de ensino superior brasileiras.

25 3. Coleta

Inicialmente, fizemos o *download* dos dados contidos no site do INEP, mais
especificamente na seção dos microdados ([http://download.inep.gov.br/microdados/
microdados_censo_superior_2016.zip](http://download.inep.gov.br/microdados/microdados_censo_superior_2016.zip)) referentes ao censo da educação supe-
rior do ano 2016. Para isso, utilizamos as bibliotecas *os*, de Python, que foram
30 essenciais para este processo de download e extração dos dados. Modelamos
o dados coletados no formato de um *DataFrame*, conforme utilizado pela bib-
lioteca *Pandas*, e salvamos os conteúdos extraídos em arquivos no formato CSV.

4. Pré-processamento

Adiante, realizamos a criação de um *notebook* a partir da ferramenta *Jupyter*
35 *Notebook*, utilizada para a criação de um documento que possibilitasse o acom-
panhamento de cada etapa realizada pela equipe. Inicialmente, realizamos a

leitura dos arquivos CSVs gerados durante a coleta e partimos para a etapa de pré-processamento dos dados. Fizemos uma pré-seleção dos dados a serem utilizados e removemos aqueles que foram considerados desnecessários, converte-
40 mos cada dado para seu tipo de atributo correspondente e, por conseguinte, usamos os valores absolutos para a execução da análise. Como extraímos um grande conjunto de dados, a alternativa que escolhemos para o tratamento dos dados ausentes fora simplesmente ignorá-los durante a análise, visto que isso não impactaria tanto nossas observações. Por fim, realizamos a integração dos
45 dados em vários dataframes a serem utilizados para a análise.

5. Análise exploratória

Na parte da análise exploratória dos dados, fizemos as perguntas: “qual a distribuição racial nos cursos mais populares da rede pública e particular de ensino superior?” “Qual a distribuição de determinada raça em uma região
50 específica do país?”.

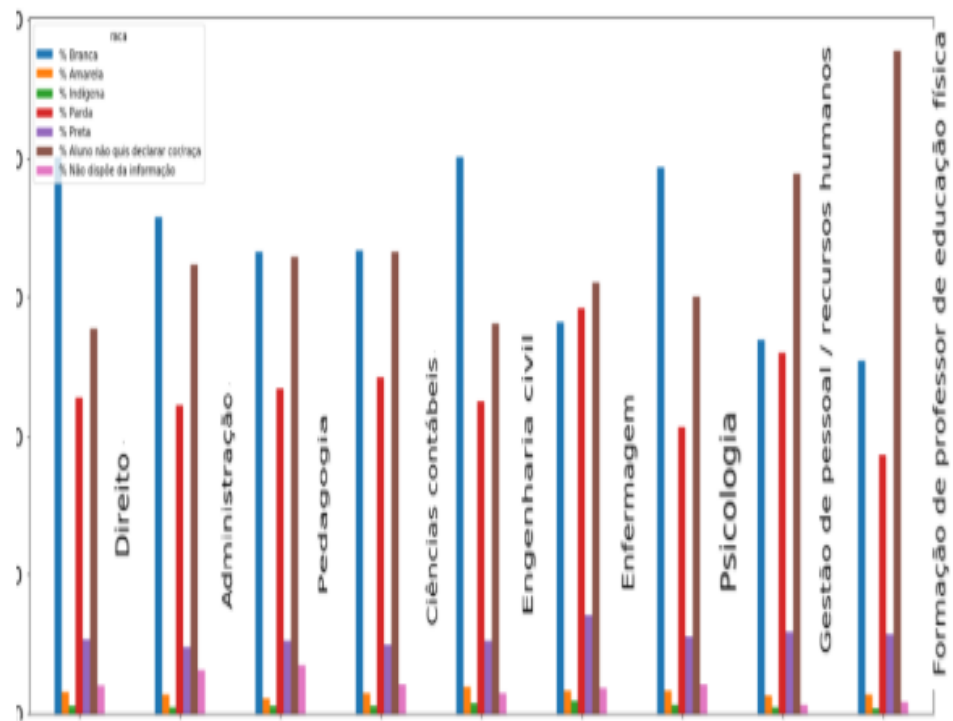


Figure 1: Distribuição racial dos principais cursos da rede Particular de Ensino.

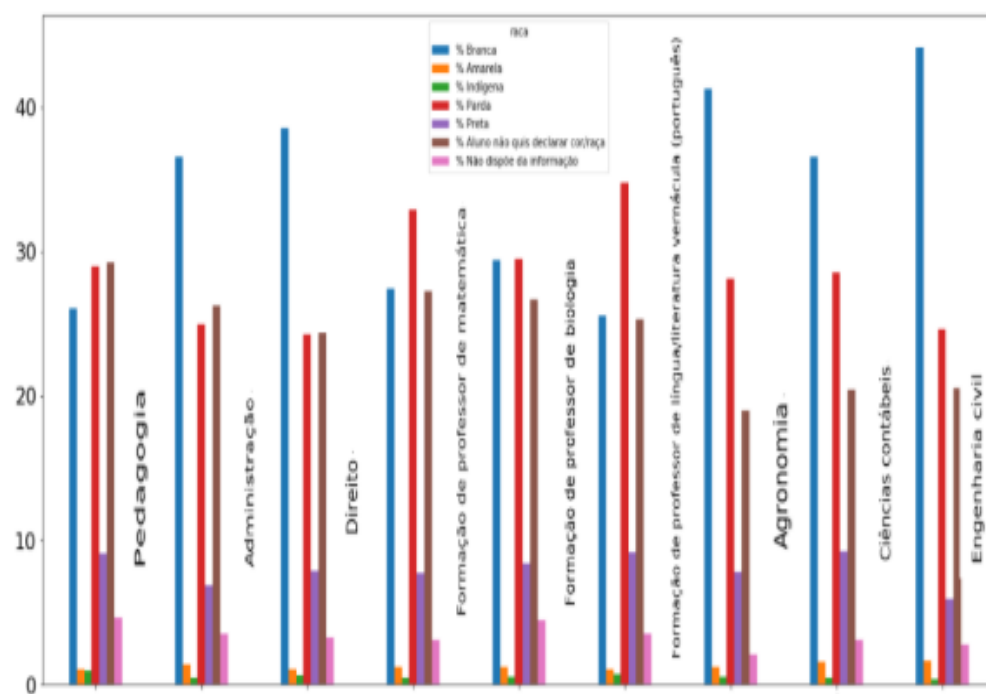


Figure 2: Distribuição racial dos principais cursos da rede Pública de Ensino.

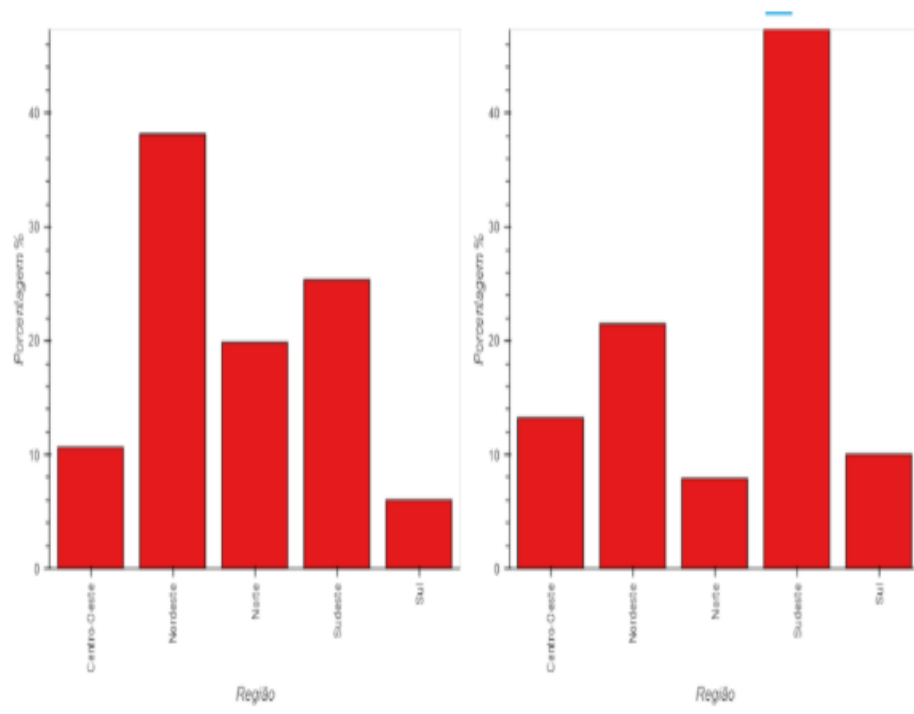


Figure 3: Distribuição de Pardos no âmbito Público(esquerda) e Particular(direita).

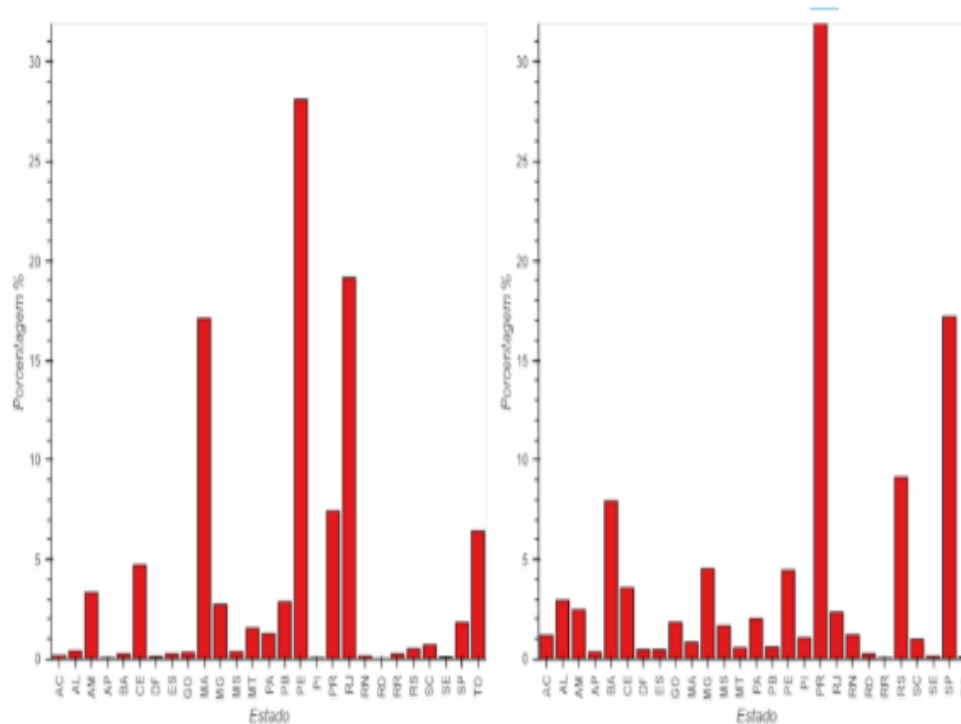


Figure 4: Relação de dados faltantes por estado.

Esses foram apenas alguns exemplos de capturas de informações, mas no geral foi possível perceber:

Que há uma predominância de brancos no ensino superior, tanto no âmbito público como no privado.

55 Os cursos mudaram quando se diz respeito ao paralelo entre Público e Privado, mostrando que as pessoas que podem pagar pela graduação, têm preferências diferentes em relação aos alunos da rede pública.

Há dados faltantes e dados onde os alunos escolheram não disponibilizar, sendo separados principalmente em estados e regiões com maior predominância
60 branca.

Quando se está analisando outras raças além da branca, comumente aparece uma distribuição mais homogênea na rede pública, como por exemplo: os pardos

que é uma raça predominante na população do Nordeste, sendo a região com mais pardos também na rede pública de ensino, mas quando se observa a rede particular esse cenário muda, tendo maior predominância no Sudeste.

Quanto à análise com relação ao gênero, buscamos responder se “existe uma região ou estado do país aonde a presença de alunas nas ciências se destaca das demais?” ; “Como se compara a proporção de alunas nas ciências exatas com a sua presença geral no ensino superior?”; “Existe alguma relação entre a presença de docentes do sexo feminino com a presença de alunas na área de ciências exatas?”; “Quais as características das IESs com mais e menos presença feminina nos cursos de exatas?”.

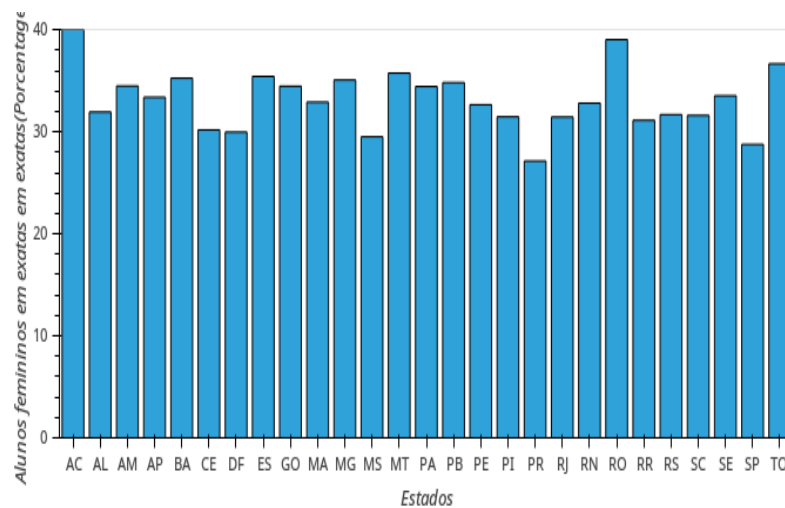


Figure 5: Distribuição da quantidade de bolsistas por curso.

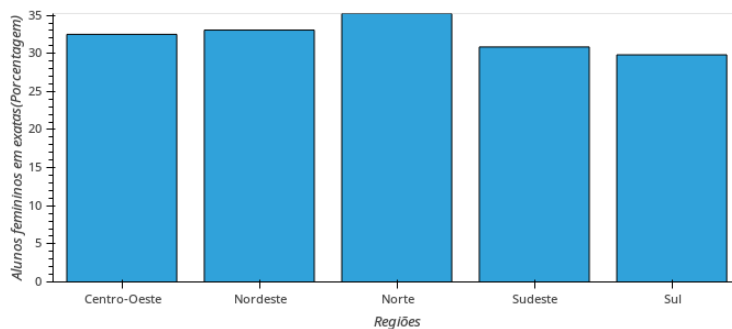


Figure 6: Distribuição da quantidade de bolsistas por curso.

No geral, é possível identificar uma tendência, com valores muito próximos a participação de mulheres nos cursos da área de ciências exatas em todo o país e sem uma grande variação. Uma tendência que se acentua ainda mais quando analisamos os dados por região, com valores ainda mais próximos.

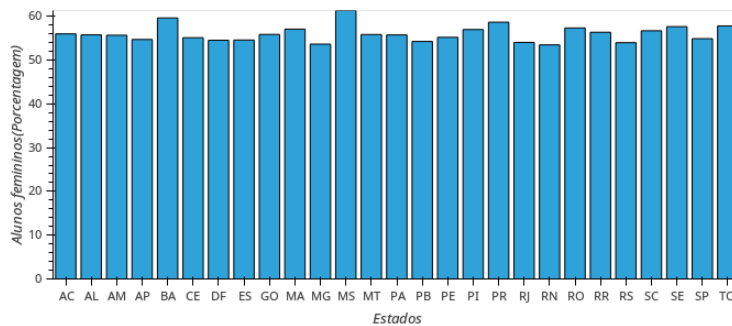


Figure 7: Distribuição da quantidade de bolsistas por curso.

Em comparação com a presença de mulheres no ensino superior, no geral, é notada uma presença muito mais baixa nos cursos da área de ciências exatas. Com uma participação que varia de 53% à 61% no ensino superior e uma penetração de 27% à 40% nas ciências exatas.

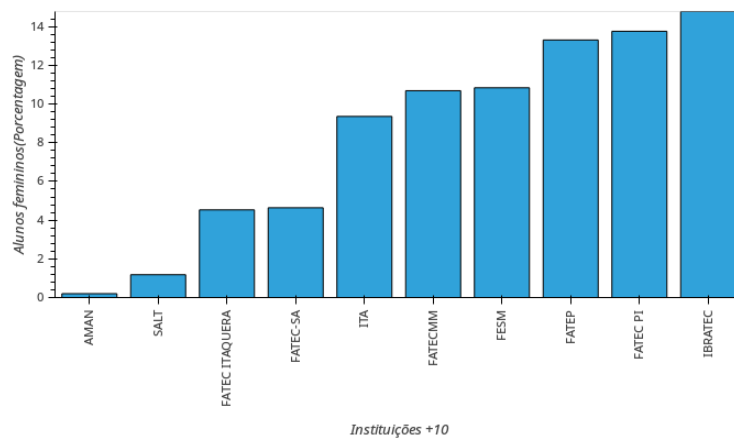


Figure 8: Distribuição da quantidade de bolsistas por curso.

Nas instituições com maior presença de mulheres na área de exatas não se distingue nenhuma característica específica, no entanto nas IESs com menor participação feminina se destaca a presença de instituições militares assim como de faculdades de tecnologia.

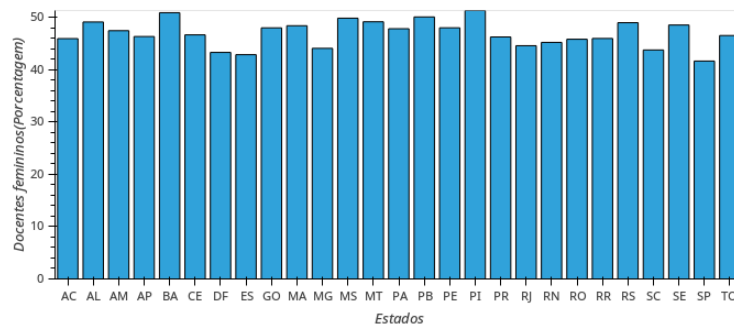


Figure 9: Distribuição da quantidade de bolsistas por curso.

85 Por fim não foi identificada nenhuma relação clara entre a presença de docentes femininas com a maior penetração de mulheres nos cursos de ciências exatas ou no ensino superior no geral, além da porcentagem de docentes ser ser tão estável quanto as outras métricas por todo o país(mas ainda assim ficando abaixo da presença masculina).

90 Entrando na parte financeira do problema, podemos inferir algumas questões
como: "Quais faculdades investem mais em pesquisa, e são públicas ou pri-
vadas?" ou ainda "Quais estados possuem mais bolsistas?".

Em ambos os gráficos abaixo, conseguimos inferir onde há uma maior con-
centração de bolsistas, tanto comparando os valores absolutos nos 10 cursos
95 com os maiores números, tanto nos estados, onde claramente há uma maior
concentração no estado de São Paulo.

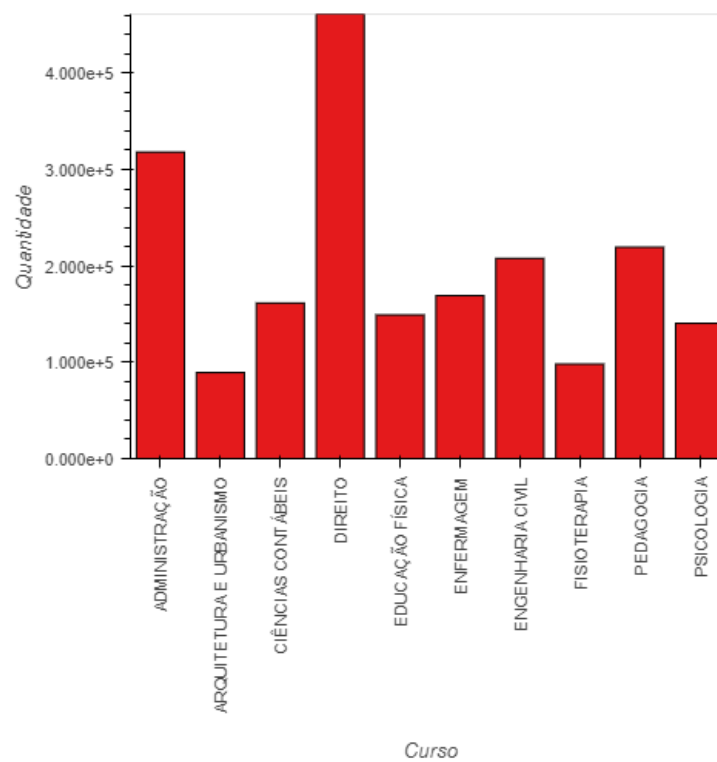


Figure 10: Distribuição da quantidade de bolsistas por curso.

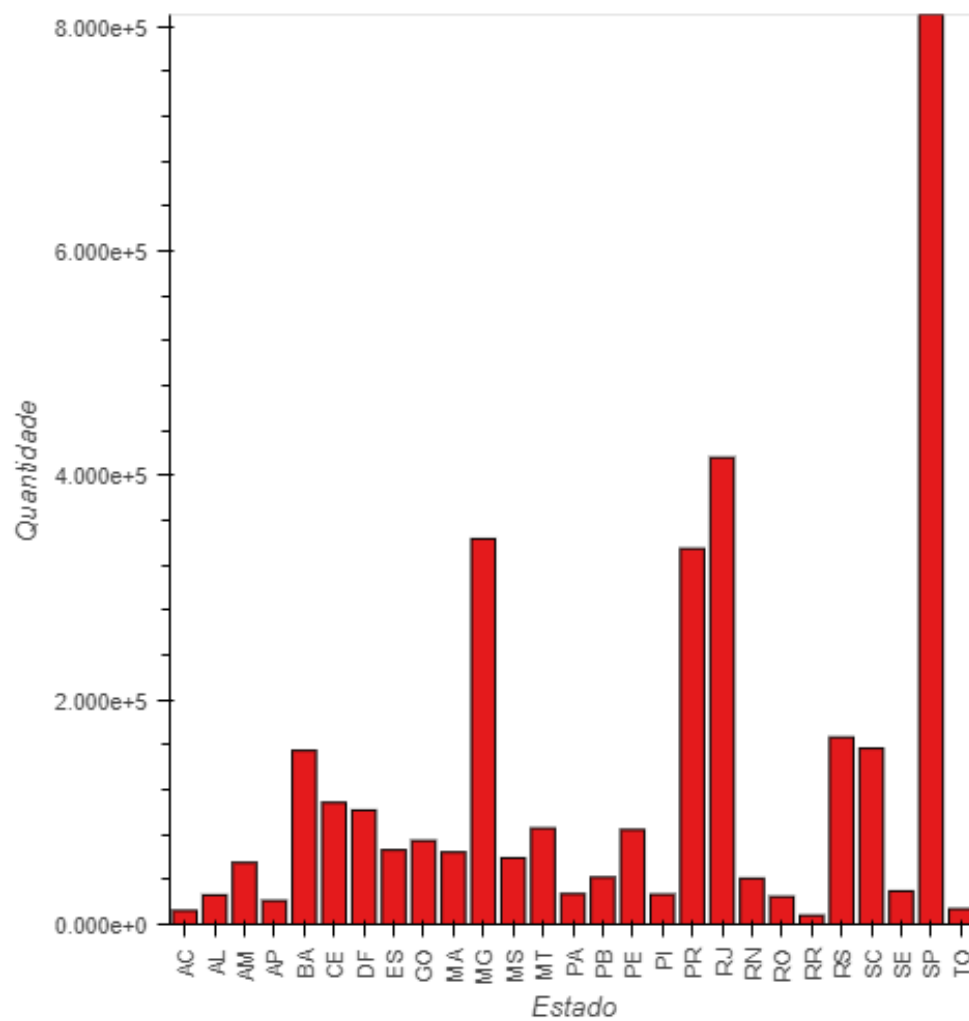


Figure 11: Distribuição da quantidade de bolsistas por estado.

	Instituicao	Porcentagem	Rede
0	FACULDADE ORTODOXA	42.524677	Privada
1	INSTITUTO TEOLÓGICO FRANCISCANO	35.883252	Privada
2	FACULDADE DE AGUDOS	35.036037	Privada
3	FACULDADE PROF. WLADEMIR DOS SANTOS	34.132792	Privada
4	FACULDADE DE TECNOLOGIA E CIÊNCIAS DA BAHIA	27.124667	Privada
5	CENTRO UNIVERSITÁRIO CAMPOS DE ANDRADE	26.419986	Privada
6	FACULDADE DE GETÚLIO VARGAS	26.289966	Privada
7	FACULDADE MAURÍCIO DE NASSAU DE FEIRA DE SANTANA	24.165400	Privada
8	FACULDADE REINALDO RAMOS	23.493733	Privada
9	FACULDADE DE SANTA CRUZ DA BAHIA	22.023140	Privada
10	FACULDADE DE EDUCAÇÃO ACRIANA EUCLIDES DA CUNHA	20.000000	Privada
11	INSTITUTO SUPERIOR DE EDUCAÇÃO CAMPO LIMPO PAU...	19.360856	Privada
12	FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE B...	18.943132	Privada
13	FACULDADE ALFREDO NASSER	18.910117	Privada
14	INSTITUTO DE EDUCAÇÃO SUPERIOR UNYAHNA LUIS ED...	18.443889	Privada
15	FACULDADE DE TECNOLOGIA DE AMPÉRE	18.254703	Privada
16	ESCOLA DE ECONOMIA DE SÃO PAULO	18.208217	Privada
17	UNIVERSIDADE ESTADUAL DE MONTES CLAROS	17.943344	Publica
18	FACULDADE TECNOLOGIA EDUVALE - AVARÉ	17.927918	Privada
19	Faculdade Cosmopolita	17.449533	Privada
20	FACULDADE DE CIÊNCIAS SOCIAIS DE GUARANTÃ DO N...	17.163888	Privada
21	FACULDADE FUCAPE	17.151796	Privada
22	FACULDADE CATÓLICA DE RONDONIA	16.852682	Privada
23	FACULDADE MERCÚRIO	16.666667	Privada
24	INSTITUTO DE EDUCAÇÃO SUPERIOR PRESIDENTE KENN...	16.666667	Publica
25	Faculdade da Região Sisaleira	16.571772	Privada
26	FACULDADES INTEGRADAS DE CASSILÂNDIA	15.109465	Privada
27	FACULDADE CAMPO LIMPO PAULISTA	14.402037	Privada
28	FACULDADES NETWORK - CAMPUS SUMARÉ	14.400110	Privada
29	FACULDADE NETWORK	14.400110	Privada

Table 1: Porcentagem de despesas com pesquisa por IES.

Na tabela, concluímos que as universidades que investem mais com pesquisa

de acordo com os dados, são instituições privadas.

6. Considerações finais

100 A possibilidade da realização de nossas análises nos permitiu uma melhor compreensão do sistema nas IES. Por sua vez, a elaboração de *notebooks* e utilização de diversas bibliotecas que conduzissem nosso estudo demonstraram a utilidade e a importância do domínio destas ferramentas no campo de ciência dos dados, bem como a do conhecimento da linguagem de programação Python.

105 Por conseguinte, as diversas visualizações mostradas ao decorrer deste relatório corroboram com o processo de delineamento da influência de diversos fatores nos escores. A partir das análises realizadas, foi possível identificar que essas avaliações, apresentaram certos padrões em diversas ocasiões.

Uma vez que conseguimos realizar um mapeamento geral dos fornecidos
110 pelo site do INEP, há a possibilidade de condução de estudos posteriores que busquem outros tipos de problemas similares dentro do contexto, como partindo do critério religioso.