



**CATÁLOGO** DE  
PROJETOS  
DA **ENCE**  
2023

# Sumário

## 7 Introdução

---

### PESQUISA

- 10** Avaliação Sistêmica e Multicêntrica do Programa Dinheiro Direto na Escola
- 11** Big Data e Indicadores ODS
- 12** Inteligência Artificial Aplicada às Políticas Públicas
- 14** Produtividade, transição demográfica e condições de vida: um estudo dos efeitos de interação das tendências macroeconômicas com as mudanças na estrutura etária da população brasileira
- 16** Qualidade em Big Data
- 17** Teoria Algébrica dos Grafos e Aplicações
- 18** Uso de Big Data em Institutos de Estatística na América Latina e Caribe
- 19** Custos sociais dos agrotóxicos no Brasil
- 20** Dinâmica populacional e mudança do clima no Brasil: trajetórias e cenários para a adaptação
- 21** Estimadores de mobilidade, instabilidade de renda e transições com amostras longitudinais complexas
- 22** Métodos de otimização aplicados a problemas de estatística: melhorando a eficácia e a eficiência no levantamento e divulgação das estatísticas oficiais
- 23** Panorama da população brasileira a partir dos recortes de Gênero, Orientação Sexual e Identidade de Gênero: análise de vulnerabilidades nos campos da saúde sexual e reprodutiva, arranjos familiares, conjugais e parentais e exposição a violências a partir dos dados populacionais existentes.
- 24** Violência contra mulheres no Brasil: permanências entre discursos e práticas
- 26** Análise de dados amostrais complexos utilizando redes neurais

- 27** Arranjos institucionais de gestão territorial e planejamento em regiões metropolitanas: complexidade local - regional
  - 28** Parâmetros espectrais de grafos e suas aplicações
  - 29** Percepções sobre tatuagem e sua relação com vivências no mercado de trabalho no Brasil
  - 31** Uma linguagem unificada na modelagem de sinais do mundo
  - 32** Uso de big data em estatísticas oficiais na América Latina e Caribe
  - 34** Combinando sub-amostras de cotas e probabilísticas selecionadas por setores para produzir estimativas mais confiáveis
  - 35** Algoritmo de otimização para definição de áreas de ponderação do Censo Demográfico 2020
  - 36** Análise espaço-temporal e circunstancial dos óbitos de policiais militares do Estado do Rio de Janeiro (2015-2019)
  - 37** Informação estatística e políticas públicas no Brasil
  - 39** Preferências reprodutivas, fecundidade não realizada e contradições na América Latina
  - 40** Produção de estatísticas experimentais para indicadores de mercado de trabalho baseados em modelos usando dados da PNAD Contínua
  - 42** Ajuste de modelos multivariados multiníveis sob amostragem informativa
  - 43** Doenças transmitidas pelo *Aedes aegypti* e a Agenda 2030: um desafio para o Brasil
  - 44** Fatores de risco associados à ocorrência de maus-tratos em pessoas idosas
  - 45** Mercado de trabalho e fecundidade
  - 46** Perfil da saúde sexual e reprodutiva das mulheres brasileiras
  - 47** Modelos de espaços latentes para análise de redes sociais
  - 48** Utilização da capacidade produtiva instalada e produto potencial no Brasil: alternativas não convencionais de mensuração e construção de indicadores para o nível da economia nacional
  - 50** Planejamento amostral e amostragem preferencial em Geoestatística
- Programa Institucional de Bolsas de Iniciação Científica - PIBIC**
- 52** Análises de dados de Violências por Sexo, Orientação Sexual e Espaços de Prevalência: Uma Análise a partir da Pesquisa Nacional de Saúde – PNS 2019
  - 53** Proposta de Algoritmo de Agrupamento Fuzzy

- 54** Análise de dados musicais no R
- 55** Estudo sobre qualidade de Big Data em estatísticas públicas
- 56** Inferência para localizações faltantes em Geoestatística associadas à ocorrência de valores extremos
- 57** Medidas de Centralidade em grafos e o autovetor associado ao raio espectral da matriz Aalpa
- 58** Mulheres sob ataque na Baixada Fluminense (2018-2022): Gênero, raça, classe e território como catalisadores da violência
- 60** Projeção da população a partir de cenários: estimativas do nível educacional e da fecundidade
- 61** Uso de big data para o estudo de assentamentos humanos precários: o potencial de imagens de satélite
- 62** A saúde do homem brasileiro por perfil de atividade sexual e orientação sexual: uma análise a partir da Pesquisa Nacional de Saúde - PNS 2019
- 63** Estudo sobre o uso de big data para a gestão do serviço de coleta de resíduos sólidos urbanos
- 65** Identidade de gênero e orientação sexual em estatísticas oficiais
- 66** Música no R: revisão sistemática e aplicação
- 67** O mercado de pinturas no Brasil entre 2000 e 2022: um estudo a partir dos leilões da Bolsa de Arte
- 68** Estudo sobre o uso de big data em estatísticas oficiais
- 70** Similaridade Semântica de Nomes de Produtos e Serviços Utilizando Grafos de Conhecimento
- 72** Tabulação de Dados Amostrais - PNAD-Contínua
- 73** Avaliação da iniciação sexual de estudantes no município do Rio de Janeiro: Uma análise da PeNSE 2015
- 74** Avaliação de classificadores interpretáveis em bases de dados de pesquisas do Instituto Brasileiro de Geografia e Estatística - IBGE
- 75** Percepção da população sobre os impactos da pandemia da COVID-19 nos Direitos Humanos
- 77** Acompanhamento dos indicadores de desenvolvimento sustentável para a saúde no Brasil
- 78** Ensinando Python para alunos do ensino fundamental II através de um jogo 2D
- 80** Estudo dos modelos de regressão assimétricos
- 82** Seleção de ordem em modelos autorregressivos

- 83** Amostras complexas e painéis rotativos: o uso do R na construção de um pacote para a estimação de modelos lineares multivariados longitudinais multiníveis
- 84** Classificação multirrótulo de documentos texto utilizando a relevância binária e o algoritmo naïve bayes
- 85** Modelos dinâmicos para dados censurados assimétricos e com caudas pesadas
- 86** Produtividade, terciarização e emprego no Brasil: desafios para o crescimento econômico inclusivo com melhora sustentável das condições de vida da população
- 88** Distribuição funcional da renda e evolução da wage share no Brasil - determinantes estruturais e seus impactos sobre os rendimentos do trabalho no período 1995-2013
- Programa Institucional de Projeto de Pesquisa para Alunos de Graduação - PPAG**
- 90** Análise de dados musicais no R: uma análise da progressão de acordes
- 91** Aplicação e comparação do método CHAID em amostragem aleatória e complexa: um estudo exploratório
- 92** E depois da formatura? Uma análise do perfil dos egressos da Graduação em Estatística da ENCE
- 94** Análise dos determinantes do tempo de amamentação: Uma abordagem de análise de sobrevivência utilizando o Random Survival Forest
- 96** Investigação das percepções sobre os impactos da COVID-19 nos Direitos Humanos: construção de um banco de perguntas
- 98** Comparação do ajuste de modelos de regressão censurados em diferentes pacotes estatísticos
- 99** Desenvolvimento de um aplicativo *web* com o Shiny para o ajuste de modelos de regressão censurados
- 100** Estudo e aplicações da classe de modelos fatoriais bayesianos: uma análise de indicadores sociais
- 101** Estudo e aplicações da classe de modelos fatoriais bayesianos: uma análise temporal

---

## EXTENSÃO

- 103** Aprendizado colaborativo: a matemática como via de comunicação entre escola e faculdade
- 105** Educação estatística: a estatística como aliada na construção do pensamento crítico
- 106** Ferramentas computacionais no ensino e aprendizado escolar
- 108** Retratos socioeconômicos da favela do Jacarezinho: Etapas de planejamento e campo
- 109** Rio em dados 2023
- 111** Estudo dos impactos da COVID-19 no Jacarezinho
- 113** Rio em Dados 2020-2022
- 114** Aprendizado colaborativo: a matemática como via de comunicação entre escola e faculdade - 2022

---

## OUTROS PROJETOS

- 117** Pesquisa por Amostra de Domicílios de Niterói - PMADNIt
- 118** Pesquisa de prevalência de COVID-19 nas escolas públicas e privadas do Estado da Paraíba
- 119** Pesquisa sobre as condições de trabalho na agricultura brasileira
- 120** Programa de capacitação intermediário em Ciência de Dados no Instituto Brasileiro de Geografia e Estatística - IBGE
- 121** Pesquisa Domiciliar Participativa (PDP) do Projeto Rio Doce
- 122** Métodos de pareamento de dados sobre agropecuária
- 124** Programa de capacitação em Ciência de Dados aplicada a Geociências
- 125** Revisão do livro *Análise de Dados Amostrais com R*

### Livros

- 127** Zona Portuária do Rio de Janeiro: Múltiplos olhares sobre um espaço em mutação
- 128** Linguagem R: Uma visão computacional
- 129** Pandas Python: Data Wrangling para Ciência de Dados
- 131** Lógica de Programação com Pascal: Um curso aplicado

## Introdução

**E**m sua segunda edição, o Catálogo de Projetos da ENCE apresenta 90 projetos coordenados por docentes da Escola e, na sua maioria, com participação discente. O Catálogo traz 69 projetos de pesquisa, dentre os quais figuram 29 projetos do Programa Institucional de Bolsas de Iniciação Científica (PIBIC) e 9 do Programa Institucional de Projeto de Pesquisa para Alunos de Graduação (PPAG). Somam-se a estes 8 projetos de extensão, 12 projetos em colaboração com outras áreas do IBGE ou outras instituições, além de 1 livro em revisão e 4 livros publicados como resultado de estudos e pesquisas realizadas como parte da atividade docente e discente na Escola. Além da participação de docentes, os projetos aqui apresentados oportunizaram o desenvolvimento de estudos e de habilidades em pesquisa a 165 discentes de graduação, 29 de mestrado, 23 de doutorado e 2 de especialização.

Na sua maioria, os projetos estão concentrados nas áreas de probabilidade e estatística e de demografia, mostrando aderência dos projetos aos programas de graduação em Estatística e de pós-graduação na subárea de Demografia na área de Planejamento Urbano e Regional/Demografia. Os demais projetos revelam a ampla gama de áreas de conhecimento de interesse e nas quais docentes e discentes da ENCE contribuem com produção de conhecimento. Serão encontrados neste Catálogo projetos das áreas de ciência da computação,

ciência de dados, computação, computação e estatística, econometria, economia ambiental, estatísticas públicas, geografia, matemática básica, otimização e estatística e teoria espectral dos grafos.

Os temas abordados são variados, transitando do teórico ao aplicado, buscando solução de problemas reais para implementação de pesquisa, capacitação, dentre outros. Além disso, temas de interesse da comunidade externa são abordados por meio dos projetos de extensão cujos impactos esperados passam pela redução das dificuldades apresentadas no aprendizado de matemática por alunos da Escola Municipal Orlando Villas Boas, pela construção de um panorama da qualidade de vida da população da favela do Jacarezinho, pelo aprendizado, o debate e a compreensão da importância da estatística por parte da sociedade alcançada pelas mídias sociais, entre outras temáticas.

O Catálogo traz um resumo de cada projeto, o período de realização, nomes de docentes, discentes e outros pesquisadores envolvidos, além de nome de outra instituição ou área do IBGE envolvida no projeto e de instituição financiadora quando aplicável.

Esperamos que essas informações ajudem a ter uma visão geral da relevância da produção da ENCE e que sua leitura seja proveitosa.



# PESQUISA



## Avaliação Sistêmica e Multicêntrica do Programa Dinheiro Direto na Escola

### Palavras-chave

Políticas Públicas, Avaliação, Estatísticas Públicas

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 -

Trata-se de projeto de avaliação do programa Dinheiro Direto na Escolas (PDDE), um dos principais programas públicos de infraestrutura escolar no país, com quase 30 anos de existência, e que atende mais de 135 mil escolas públicas (urbanas, rurais, indígenas, quilombolas etc). A abordagem empregada diferencia-se de metodologias convencionais de avaliação, ao assumir a perspectiva sistêmica do programa- com suas interações com o outros programas públicos e contexto sociopolítico - e natureza multicêntrica de realização- com a participação de cinco universidades no país, uma em cada região. Busca-se levantar evidências sobre o atendimento do programa em relação a seis critérios avaliativos- relevância do objetivo, coerência das ações, eficácia dos resultados, efetividade dos desdobramentos, sustentabilidade dos efeitos e eficiência dos meios- adotados pela OCDE. Envolve um conjunto amplo de estudos avaliativos, baseados em registros administrativos, integração de bases de dados e levantamento amostral de escolas pelo país e design etnográfico de campo.

### Equipe

#### Coordenação

Paulo de Martino Jannuzzi  
Vicente Rocha Ferreira

#### Demais docente

Bárbara Cobo

#### Discentes

Fernanda Teixeira  
Lais Fortes

#### Outros pesquisadores

Andre Coelho Vaz Henriques  
Raphael Molina Guimarães

### Outra instituição ou área do IBGE envolvida no projeto

Centro de Pesquisas Aplicadas ao Setor Público da UFG

### Financiamento

Fundo Nacional de Desenvolvimento da Educação - FNDE

Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

## Big Data e Indicadores ODS

### Palavras-chave

Big data, Agenda 2030. Objetivos de Desenvolvimento Sustentável - ODS

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 -

A produção de indicadores para o monitoramento da Agenda 2030 depende da disponibilidade de dados oportunos e de qualidade. O uso de fontes tradicionais, como pesquisas e registros administrativos, nem sempre é suficiente, assim considerar o uso de big data tornou-se fundamental. O objetivo da linha é investigar o uso de big data para o cálculo de indicadores dos Objetivos de Desenvolvimento Sustentável. Essa atividade é parte do plano de trabalho do Hub Regional da ONU para Big Data no Brasil.

### Equipe

#### Coordenação

Andrea Diniz da Silva

Marcelo Rodrigues de Albuquerque Maranhão

#### Demais docente

Ian Monteiro Nunes

#### Discentes

Alexandre Mario Oliveira Santos da Paz

André Coelho Vaz Henriques

Bernardo Braga Martins

Bruno Sá Guedes

Cassio Puzanovsky Magalhães

Eduardo Soares Jangutta

Jedielso Sales de Souza

João Paulo Pereira Leonardo

Lívia Médice de Souza Lima

Miguel do Nascimento Faria Conforto

#### Outros pesquisadores

Andre Coelho Vaz Henriques

Raphael Molina Guimarães

#### Outra instituição ou área do IBGE envolvida no projeto

Diretoria de Geociências do IBGE e Centro de Documentação e Disseminação de Informações, do IBGE

#### Financiamento

Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Instituto Brasileiro de Geografia e Estatística - IBGE

## Inteligência Artificial Aplicada às Políticas Públicas

### Palavras-chave

Inteligência Artificial, Políticas Públicas, Modelos de Linguagem

### Área de conhecimento

Administração Pública

### Início Término

2023 -

Este projeto, tem o objetivo de contribuir para formação e consolidação da Rede de Pesquisa em Inteligência Artificial Aplicada em Políticas Públicas, que tem como objetivo geral de fortalecer a capacidade de gestão de Políticas Públicas mediante o uso de evidências, estudos e avaliações. Nesse sentido, o projeto tem três objetivos específicos: 1) desenvolver uma ferramenta com recursos de Inteligência Artificial (ChatPP) aberta à sociedade, gestores, técnicos do setor público e conselheiros de fóruns de participação social para possibilitar acesso conversacional sobre dúvidas, demandas e perguntas típicas relativas ao processo de formulação e gestão de políticas e programas públicos; 2) Manter e atualizar repositório de artigos e material sobre políticas públicas, com curadoria e alimentado por uma rede de instituições universitárias e de pesquisa capilarizada pelo país (GeraPP); 3) Qualificar técnicos municipais e estaduais e conselheiros de participação social para no acesso e uso das informações para o debate público, formulação e gestão de programas públicos em todo o território nacional (QualificaPP).

### Equipe

#### Coordenação

Paulo de Martino Jannuzzi  
Vicente Rocha Ferreira  
Ana Lucia Gonçalves da Silva  
Marcelo Proni

#### Demais docentes

Andréa Diniz da Silva  
Paulo Tostes

#### Outra instituição ou área do IBGE envolvida no projeto

Centro Colaborador Interinstitucional de Inteligência Artificial Aplicada às Políticas Públicas da UFG  
Núcleo de Estudos de Políticas Públicas da UNICAMP  
Centro de Documentação e Disseminação de Informações do IBGE  
Diretoria de Tecnologia da Informação do IBGE

## Financiamento

Secretaria de Gestão do Trabalho e da  
Educação na Saúde (SGTES) do Ministério  
da Saúde

Fundação de Apoio à Pesquisa de Goiás

## Produtividade, transição demográfica e condições de vida: um estudo dos efeitos de interação das tendências macroeconômicas com as mudanças na estrutura etária da população brasileira

### Palavras-chave

produtividade, transição demográfica, poupança, investimento, crescimento econômico, condições de vida

### Área de conhecimento

área interdisciplinar pois abrange as interações entre dinâmica demográfica e dinâmica econômica.

Início Término

2023 2026

O processo de transição demográfica nos países em desenvolvimento traz vários desafios para as políticas públicas. Em sua primeira fase, observa-se o aumento da oferta potencial de força de trabalho, resultante do aumento da população em idade ativa como proporção da população total. Havendo aceleração do crescimento econômico com melhoria da distribuição funcional e pessoal da renda, os efeitos predominantes tendem a ser positivos. Mas, sem a geração concomitante de postos de trabalho compatíveis com esse incremento, o resultado tende a ser o desemprego e o rebaixamento das condições de vida da população. A depender do desempenho macroeconômico do país, as mudanças na estrutura etária da população podem ter impactos positivos ou negativos sobre o mercado de trabalho, além dos efeitos sobre a formação e alocação da poupança agregada e do investimento produtivo. Em uma segunda fase da transição demográfica, observa-se o processo de envelhecimento populacional que traz novos desafios para as políticas públicas, especialmente para os sistemas previdenciários e de saúde. O Brasil se insere nesse contexto com sua economia apresentando grande dificuldade para absorver uma taxa de atividade crescente, decorrente tanto da transição demográfica, quanto da busca por melhores oportunidades de emprego e renda, em postos de trabalho de maior qualidade. Este projeto busca analisar os efeitos de interação das dinâmicas macroeconômica e demográfica brasileiras considerando as transformações na estrutura produtiva e distributiva da economia face às mudanças rápidas na estrutura etária da população. Mobiliza a análise estatística

e econométrica para a construção de cenários que integram as tendências da produtividade, da ocupação e do rendimento médio real das famílias brasileiras, variáveis fundamentais para o aproveitamento da janela de oportunidade demográfica e que devem nortear as ações governamentais de políticas públicas para o desenvolvimento social e econômico do país.

---

### Equipe

#### Coordenação

Miguel Bruno

#### Discente

Leandro Maia Gonçalves

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Qualidade em Big Data

### Palavras-chave

Qualidade de dados, Erros em pesquisa, Big Data

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 -

O uso de big data no apoio à produção de estatísticas oficiais tem se tornado uma realidade cada vez mais premente. Embora promissora, essa fonte de dados apresenta desafios como representatividade, incompletude, confiabilidade, ética, entre outros. Logo, requer um olhar metodológico e analítico atento. O objetivo da linha de pesquisa Qualidade em Big Data é explorar alguns problemas estatísticos relacionados ao uso de big data, debatendo possíveis soluções e contribuindo para o aumento do seu potencial de uso.

### Equipe

#### Coordenação

Elizabeth Belo Hypólito

#### Demais docente

Alinne de Carvalho Veiga

#### Discentes

Déborah Bruna Cardoso Vaz Coelho

Victor Arno Fazolato Lindenblatt

Marcus André Alves Zimmermann Vieir

#### Outro pesquisador

Natália da Silva Castro

### Outra instituição ou área do IBGE envolvida no projeto

Coordenação da Central de Entrevistas Telefônicas Assistidas por Computador, do IBGE

### Financiamento

Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES  
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq



## Teoria Algébrica dos Grafos e Aplicações

### Palavras-chave

grafos, matrizes associadas a grafos, rede

### Área de conhecimento

Teoria dos Grafos, Álgebra Linear, Otimização Combinatória

### Início Término

2023 2025

A Teoria dos Grafos, uma área clássica e sedimentada da Pesquisa Operacional e da Ciência da Computação, e a Álgebra Linear, parte da Álgebra, quando aplicada à Teoria dos Grafos, utilizando, por exemplo, as matrizes de adjacência, incidência, Laplaciana e Laplaciana sem sinal, constituem a base para a Teoria Algébrica dos Grafos. Este projeto trata de problemas envolvendo autovalores de matrizes associadas a grafos e suas aplicações em redes.

### Equipe

#### Coordenação

Carla Silva Oliveira

#### Demais docentes

Claudia Marcela Justel (IME)

Celso Marques da Silva Junior (CEFET-RJ)

André Brondani (UFF)

Francisca França (UFF)

#### Discente

Matheus Iozzi

### Outra instituição ou área do IBGE envolvida no projeto

Centro Federal de Educação  
Tecnológica - CEFET

Instituto Militar de Engenharia - IME

Universidade Federal Fluminense - UFF

### Financiamento

Fundação Carlos Chagas Filho de Amparo  
à Pesquisa do Estado do RJ - FAPERJ

## Uso de Big Data em Institutos de Estatística na América Latina e Caribe

### Palavras-chave

Big data, América Latina e Caribe, Estatísticas oficiais

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 -

O objetivo é acompanhar a trajetória do uso de big data em institutos nacionais de estatística na América Latina e Caribe. Essa atividade é parte do plano de trabalho do Hub Regional da ONU para Big Data para Estatísticas Oficiais. Desde 2021 são realizadas consultas anuais aos institutos nacionais de estatística. A consulta incluir perguntas sobre o uso de big data para a produção de estatísticas oficiais, estatísticas experimentais, estudos em curso e interesse no uso de big data.

### Equipe

#### Coordenação

Andrea Diniz da Silva

#### Demais docentes

Elizabeth Belo Hypolito

Fábio Lucas Pimentel de Oliveira

#### Discente

Marcus André Alves Zimmermann Vieira

### Outra instituição ou área do IBGE envolvida no projeto

United Nations.

### Financiamento

Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

## Custos sociais dos agrotóxicos no Brasil

### Palavras-chave

Externalidades, regulação dos agrotóxicos e valoração ambiental

### Área de conhecimento

Economia ambiental e bem-estar social

### Início Término

2022 2024

O projeto busca a elaboração de um livro que reúne nos últimos 20 anos os trabalhos do docente, cujo tema de pesquisa centra-se nos custos sociais dos agrotóxicos, reunindo desde artigos publicados em periódicos como a valoração dos impactos dos agrotóxicos à saúde, e outras publicações científicas mais recentes que buscam valorar os impactos para além do campo da saúde pública, como os custos atinentes ao meio ambiente, custos com regulação, bem como as desonerações tributárias conferidas aos agrotóxicos. Recentemente, incluímos uma mensuração do uso de agrotóxicos ilegais no país. Deve-se ainda ressaltar que no último ano, a pesquisa teve alguns dos seus resultados discutidos em fóruns, audiências públicas e congressos sobre o tema e pretende-se em breve submeter publicações à periódicos científicos a fim de melhor divulgá-los.

### Equipe

#### Coordenação

Wagner Lopes Soares

#### Discente

Fernando Falcão

#### Outra instituição ou área do IBGE envolvida no projeto

Escola Nacional de Saúde Pública  
Sérgio Arouca - ENSP da FIOCRUZ

Associação Brasileira de Saúde  
Coletiva - ABRASCO

## Dinâmica populacional e mudança do clima no Brasil: trajetórias e cenários para a adaptação

### Palavras-chave

Projeções populacionais, mudanças ambientais, adaptação

### Área de conhecimento

Demografia

### Início Término

2022 2025

Os avanços recentes na modelagem da mudança do clima fomentaram a incorporação mais ampla das ciências sociais, e, principalmente, das dinâmicas populacionais. A atual geração de cenários socioeconômicos utilizada no campo, aqui denominada como trajetórias socioeconômicas compartilhadas, ampliou o papel da demografia ao incorporar projeções populacionais baseadas no comportamento futuro da fecundidade, mortalidade e migração considerando as variáveis sexo, idade e nível educacional. Os cenários foram elaborados como um conjunto de narrativas, que na ausência de políticas de adaptação e mitigação, traçam as dinâmicas socioeconômicas futuras em relação aos desafios à adaptação e mitigação. Atualmente, ocupam um papel central nas projeções das mudanças do clima, embora ainda sejam pouco explorados em termos regionais e nos estudos de adaptação. Nesse escopo, a presente pesquisa parte da concepção das trajetórias socioeconômicas compartilhadas para traçar e analisar os cenários da dinâmica socioeconômica e populacional no Brasil para o período 2020-2050, especificamente em relação a adaptação à mudança do clima. Para tal, serão consideradas as heterogêneas vulnerabilidades e dinâmicas populacionais no país e os desafios à adaptação, principalmente na interface população, urbanização e ambiente.

### Equipe

#### Coordenação

César Augusto Marques da Silva

#### Demais docentes

Samir KC

Erich Striessnig

#### Discentes

Juliana Trece

Luiggia Girardi Bastos Reis de Araújo

#### Outra instituição ou área do IBGE envolvida no projeto

Universidade de Viena

#### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Estimadores de mobilidade, instabilidade de renda e transições com amostras longitudinais complexas

### Palavras-chave

Distribuição de renda, mobilidade, instabilidade, volatilidade, desigualdade e bem-estar

### Área de conhecimento

Probabilidade e estatística

### Início

2022

### Término

2023

Embora haja uma extensa literatura sobre desigualdades no Brasil, estudos sobre mobilidade e instabilidade de renda são mais raros, pois dependem de dados retrospectivos ou longitudinais, bem mais escassos do que as bases transversais que permitem estimar desigualdades. Há um amplo conjunto de conhecimentos acumulados e difundidos sobre indicadores de desigualdade, suas propriedades, formas de estimá-los e decompô-los, além de suas relações com outros indicadores de bem-estar. O objetivo da pesquisa é comparar a variedade de indicadores empregados em diferentes estudos brasileiros e estrangeiros sobre mobilidade e instabilidade de renda, suas propriedades, possíveis relações com outros indicadores de bem-estar e questões ligadas à sua estimação a partir de amostras complexas.

### Equipe

#### Coordenação

Marcos Dantas Hecksher

#### Demais docentes

Pedro Luis do Nascimento Silva

Elizabeth Belo Hypólito

#### Outra instituição ou área do IBGE envolvida no projeto

Instituto de Pesquisa Econômica Aplicada - IPEA

## Métodos de otimização aplicados a problemas de estatística: melhorando a eficácia e a eficiência no levantamento e divulgação das estatísticas oficiais

### Palavras-chave

Amostragem, análise de agrupamentos, sigilo, metaheurísticas e programação matemática

### Área de conhecimento

Otimização e estatística

### Início Término

2022 2025

Neste projeto serão desenvolvidos algoritmos que combinam técnicas de otimização e estatística, com o objetivo de melhorar a eficácia e eficiência no levantamento e divulgação das Estatísticas Oficiais. Mais especificamente, tais algoritmos serão aplicados à resolução de problemas de agrupamento e estratificação que surgem no âmbito de pesquisas realizadas mediante aplicação de amostragem probabilística.

### Equipe

#### Coordenação

José André de Moura Brito

#### Demais docentes

Flávio Marcelo Tavares Montenegro

Gustavo Silva Semaan

Nelson Maculan

#### Discente

Victória Vargas

#### Outro pesquisador

Augusto Cesar Fadel

### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Métodos e Qualidade, do IBGE

Universidade Federal Fluminense - UFF

Universidade Federal do Rio de Janeiro - UFRJ

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (Projeto Universal)

## Panorama da população brasileira a partir dos recortes de Gênero, Orientação Sexual e Identidade de Gênero: análise de vulnerabilidades nos campos da saúde sexual e reprodutiva, arranjos familiares, conjugais e parentais e exposição a violências a partir dos dados populacionais existentes.

### Palavras-chave

Minorias Sexuais, Saúde Sexual e Reprodutiva, Violências, Estatísticas Públicas, Vulnerabilidades

#### Área de conhecimento

Demografia

#### Início Término

2022 2024

A presente proposta tem como tema global a elaboração, a partir dos dados oficiais existentes e previstos, bem como por meio de levantamentos complementares que se fizerem necessários, de um panorama sociodemográfico das minorias sexuais, a saber: pessoas com orientação sexual e pessoas com identidade de gênero que diferem do padrão cisheteronormativo. A partir do tema global proposto, temos como temas específicos analisar e trazer um raio-x dessa população alvo a partir dos seguintes pontos: 1) saúde sexual e reprodutiva e questões relacionadas à saúde mental dessa população; 2) situação conjugal e parental dessa população; 3) acesso dessa população à aparelhos de saúde pública; 4) situação de escolaridade, renda e emprego dessa população e; 5) exposição à violências em decorrência de orientação sexual e identidade de gênero. Como estratégia metodológica pretende-se analisar as bases de dados populacionais existentes como a Pesquisa Nacional de Saúde – 2019 que trouxe de forma inédita no Brasil a inclusão da variável orientação sexual permitindo o início de investigações específicas da população LGB (Lésbicas, Gays e Bissexuais). De forma complementar pretende-se verificar dados relacionados a pessoas que possuem relação conjugal com parceiros de mesmo sexo, o que é possível pela presença da variável “possui cônjuge de mesmo sexo” em pesquisas como o Censo Demográfico e a PNAD.

### Equipe

#### Coordenação

Angelita Alves de Carvalho

Rafael Chaves Vasconcelos Barreto

## Violência contra mulheres no Brasil: permanências entre discursos e práticas

### Palavras-chave

Violência de gênero, violência doméstica, discursos, políticas públicas

### Área de conhecimento

Ciências Sociais

### Início Término

2022 -

A violência contra mulheres está imbricada a outros aspectos complexos como desigualdades socioeconômicas, inacessibilidade a direitos e preconceitos raciais. Nas últimas décadas, houve avanços significativos em direção ao enfrentamento a esse problema, inclusive com o estabelecimento de políticas públicas de assistência às vítimas. Destacam-se, nesse cenário, a Lei 11.340/2006 e a Lei 13.104/2015. Não obstante, as transformações legais não foram suficientes para modificar em definitivo a realidade brasileira em relação aos crimes cometidos no âmbito da violência doméstica. Decerto, o problema não é novo. No passado, os delitos cometidos em "defesa da honra" deram a tônica dos crimes violentos contra a mulher, assim como os chamados "crimes passionais" permearam as páginas dos jornais nas décadas de 1980 e 1990. A infeliz permanência é a manutenção do mesmo padrão de crimes e vítimas, ainda que os discursos tenham se alterado em alguma medida. Neste projeto, desenvolvido no escopo do Grupo de Pesquisa Gênero, Sexualidades, Reprodução e suas Interseccionalidades (Generis), pretende-se analisar de que forma os casos de violência contra mulheres foram abordados nas páginas dos jornais no período entre 1980 e 2020, tendo como fonte prioritária a imprensa periódica da cidade do Rio de Janeiro. Em tal investigação, buscar-se-á evidenciar as modificações existentes no plano discursivo e sua possível correlação com as transformações sociais experimentadas pela sociedade nas últimas décadas - como o crescimento significativo da participação de mulheres na educação superior, o aumento nos anos de escolarização feminina em relação a indivíduos do sexo masculino, as políticas públicas de redistribuição de renda, as medidas de atenção e assistência a mulheres em situação de violência, dentre outros elementos. Postula-se que a superação efetiva do cenário atual, em que as mulheres brasileiras permanecem se destacando entre as vítimas de violência doméstica na



América Latina e no mundo, poderá ser alcançada a partir de um amplo e multifacetado conjunto de políticas públicas que tenham como norte desde a melhoria das condições socioeconômicas até a abordagem das questões de gênero em ambientes educacionais. Nesse sentido, importará caracterizar as formas de violência mais comuns ao longo do período analisado, evidenciando as possíveis assimetrias existentes entre o perfil majoritário das vítimas e os casos retratados nos jornais. Por fim, buscar-se-á compreender e explicitar as lógicas culturais e comportamentais que ainda subsistem na sociedade brasileira e perduram como legitimadoras desse tipo de violência.

---

### Equipe

#### **Coordenação**

Joice de Souza Soares

#### **Demais docente**

Juscelino Bezerra dos Santos

## Análise de dados amostrais complexos utilizando redes neurais

### Palavras-chave

Redes neurais, superpopulação e amostras complexas e ponderação

### Área de conhecimento

Probabilidade e estatística

### Início

2021

### Término

2023

O ajuste de um modelo de Redes Neurais Artificiais (RNA) considera que os dados são provenientes de amostragem aleatória simples com reposição (AASC). Entretanto, na prática, pesquisas amostrais raramente utilizam a seleção de amostras via AASC, e esquemas amostrais mais complexos são empregados. A amostragem complexa reflete estruturas mais complexas da população. Estas estruturas consideradas no plano amostral precisam ser incorporadas também quando ajustamos uma RNA aos dados resultantes. Na literatura estatística, existem diferentes abordagens para modelagem de dados amostrais complexos. Entretanto, na literatura relacionada a RNA não existe menção de como proceder quando os dados são provenientes de um esquema amostral complexo. O objetivo da pesquisa é aprofundar a avaliação da abordagem de superpopulação proposta em Pereira (2017) para ajuste de modelos de RNA com dados amostrais complexos, considerando também as questões de identificabilidade dos modelos propostos.

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Outro pesquisador

Savano Sousa Pereira

#### Outra instituição ou área do IBGE envolvida no projeto

Mobi2Buy

## Arranjos institucionais de gestão territorial e planejamento em regiões metropolitanas: complexidade local - regional

### Palavras-chave

Região metropolitana, municípios, planejamento urbano e regional e gestão territorial

### Área de conhecimento

Geografia

### Início Término

2021 -

Pesquisa destinada ao monitoramento (observatório) da gestão territorial metropolitana, com avaliação dos instrumentos de planejamento municipal à luz da complexidade de relações político-institucionais em escala regional, especialmente quanto às diretrizes de política urbana, desenvolvimento e gestão, identificando as tensões e conflitos entre os interesses local e regional. Busca também contribuir para o desenvolvimento de uma metodologia visando uma governança metropolitana a partir do planejamento (municipal e regional) e das práticas e instrumentos vigentes de regulação e fomento do desenvolvimento urbano sustentável.

### Equipe

#### Coordenação

Romay Conde Garcia

#### Discentes

Miguel Castilho Junior

Bruno Bonsanto Dias

Adriene dos Santos Duarte

André Victor Mendes Rosa

#### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Meio Ambiente,  
Coordenação de Estruturas Territoriais e  
Coordenação de População e Indicadores  
Sociais, do IBGE

## Parâmetros espectrais de grafos e suas aplicações

### Palavras-chave

Álgebra linear, grafos e redes

### Área de conhecimento

Teoria espectral dos grafos

### Início

2021

### Término

2024

Os principais temas abordados neste projeto se encontram dentro da linha de pesquisa Teoria Espectral dos Grafos, área que utiliza os conceitos da Álgebra Linear em problemas da Teoria dos Grafos com aplicações em Química, Ciência da Computação e Otimização Combinatória. Nesta linha, os problemas abordados estão ligados a limites inferiores e superiores envolvendo parâmetros da Teoria dos Grafos e autovalores de matrizes associadas a grafos; e na determinação de grafos extremais. Outro tema envolve aplicações em redes. A confiabilidade de uma rede é uma medida importante para diversas aplicações reais na análise de redes, como, por exemplo, redes de transporte, transmissão de energia e comunicação, cuja necessidade de manter todos seus vértices conectados é fundamental para o pleno funcionamento do sistema modelado por um grafo. Neste tema, os problemas abordados estão relacionados com a confiabilidade da rede, a qual é uma medida que envolve a probabilidade de falha nos vértices e/ou arestas das redes e alguns parâmetros da Teoria Espectral dos Grafos.

### Equipe

#### Coordenação

Carla Silva Oliveira

#### Discentes

Victor Melquiades

Uilton Cesar Peres Junior

Rafaela Soares

João Domingos Gomes da Silva Junior

### Outra instituição ou área do IBGE envolvida no projeto

Centro Federal de Educação  
Tecnológica - CEFET

### Financiamento

Conselho Nacional de Desenvolvimento  
Científico e Tecnológico - CNPq

## Percepções sobre tatuagem e sua relação com vivências no mercado de trabalho no Brasil

### Palavras-chave

Pesquisa quantitativa, estatística social, tattoo, preconceito e questionário eletrônico

### Área de conhecimento

Probabilidade e estatística

### Início Término

2021 -

A tatuagem está inserida em vários contextos históricos. Sociedades antigas usavam esse tipo de arte em ritos de passagens e eventos especiais. Marinheiros e aventureiros possuíam o costume de marcar seus corpos para simbolizar pertencimento nacional e identidade religiosa. Mesmo disseminada entre pessoas de todas as crenças, gêneros e idades, ainda há incerteza sobre como a sociedade julga pessoas tatuadas em determinados espaços sociais. No vital ambiente do trabalho podem ser encontrados relatos sobre a possível existência de preconceito. Considerando os desafios que jovens profissionais precisam enfrentar no início da carreira profissional, e tendo como base o debate atual sobre o tema, este projeto teve como objetivo realizar uma pesquisa quantitativa para investigar a percepção e o posicionamento quanto à posse de tatuagem em questões associadas ao mercado de trabalho de brasileiros(as) com idade a partir de 16 anos. A coleta de dados ocorreu entre os dias 31 de outubro e 13 de novembro de 2021 através de plataformas digitais obtendo 938 respostas válidas. O questionário incorporou perguntas relacionadas às características socioeconômicas dos respondentes, à posse de tatuagem, a percepções sociais e ao mercado de trabalho, bem como investigou situações vivenciadas por pessoas tatuadas que apontassem indícios de reprovação num período de referência de 5 anos. A pesquisa também fomentou debates sobre o tema nas redes sociais, ratificando a relevância do estudo e o interesse de profissionais nas informações coletadas. Seus resultados indicam que ainda há preconceito sobre a posse de tatuagem no mercado de trabalho e revelou indícios de preconceito mesmo entre pessoas tatuadas. Destaca-se que projeto deu

origem a um pôster que foi premiado na 5ª Semana de Iniciação Científica da ENCE realizada em 2021 e também foi apresentado no 24º Simpósio Nacional de Probabilidade e Estatística em 2022.

---

## Equipe

### **Coordenação**

Denise Britz do Nascimento Silva

### **Discentes**

Camilla C. Alves

Júlia S. de Araújo

Sofia S. H. de Brito

## Uma linguagem unificada na modelagem de sinais do mundo

### Palavras-chave

Modelagem de dados, arquitetura de retroalimentação e processos estocásticos

### Área de conhecimento

Probabilidade e estatística

### Início

2021

### Término

2024

Estatística é, indiscutivelmente, uma grande ferramenta na modelagem de dados. É auxiliada pela Probabilidade, que providencia uma série de modelos paramétricos para tal. No entanto, e devido aos conjuntos de dados disponíveis atualmente, modelos que representam os dados necessitam de um incremento na sua estrutura, se beneficiando da distinção dos dados e suas representações, e das ambiguidades neste âmbito, não-unicidade de representação, e não-unicidade de dados de uma representação. Lidar com esses dados requer estruturas adicionais, em particular, adaptáveis ao ingresso de novos dados. Para efeitos de uma abordagem unificadora, pode-se nomear os dados como sinais do mundo, e suas representações de conceitos do mundo. A proposta deste projeto é considerar uma arquitetura retroalimentada envolvendo os sinais e os conceitos do mundo, mostrar como ela se aplica a diferentes modelagens de dados, exibindo o papel central da estatística e da probabilidade neste contexto.

### Equipe

#### Coordenação

Maysa S. de Magalhães

Francisco Moura Neto

#### Demais docentes

Ursula Andrea Rohrer

Monique Carmona

#### Outra instituição ou área do IBGE envolvida no projeto

Universidade do Estado do Rio de Janeiro - UERJ

Universidade Federal do Rio de Janeiro - UFRJ

## Uso de big data em estatísticas oficiais na América Latina e Caribe

### Palavras-chave

Big data, estatísticas oficiais e estatísticas experimentais

Área de conhecimento	Início	Término
Probabilidade e estatística	2021	2022

A demanda por mais e melhores estatísticas cresce na direção oposta à disponibilidade de recursos financeiros destinados à produção de estatísticas públicas, em especial as oficiais, e à cooperação dos respondentes, principalmente entre os que vivem nos grandes centros urbanos ou estão inseridos no mercado de trabalho. Para garantir a oferta de estatísticas oficiais e atender emergentes demandas nacionais e internacionais, institutos nacionais de estatística de vários países vêm considerando o uso de big data. Para apoiar essas iniciativas, as Nações Unidas vêm, desde 2013, promovendo o uso de big data para produção de estatísticas oficiais por meio de várias ações como a composição de forças-tarefa, a estruturação de uma plataforma contendo métodos, ferramentas e especialistas no uso de big data para a produção estatística assim como a instalação de Hubs Regionais para o Uso de Big Data. Atualmente são quatro hubs regionais situados em quatro países: China, Ruanda, Emirados Árabes e Brasil. O Hub Regional no Brasil está sediado na Escola Nacional de Ciências Estatísticas - ENCE/Instituto Brasileiro de Geografia e Estatística - IBGE. Sua área de abrangência é a América Latina e Caribe. O presente projeto é parte integrante do plano de trabalho do Hub Regional para Big Data no Brasil. O objetivo do trabalho é reunir conhecimento sobre iniciativas dos institutos nacionais de estatística da América Latina e do Caribe relacionadas com o uso de big data para produzir estatísticas oficiais e estatísticas experimentais. O estudo inclui a realização de duas pesquisas, uma diretamente na página de cada instituto de estatística e outra por meio de consulta com o uso de questionário a ser preenchido por representantes de cada instituto. É considerado big data aquele proveniente



de redes sociais, web scraping, scanner, sinal telefone celular, imagem de satélite, cartão de crédito, dados de sensores, registros de saúde, câmera de segurança, entre outras. Além de investigar as fontes utilizadas, também é de interesse saber os tópicos para os quais estatísticas estão sendo produzidas e ou planejadas. Adicionalmente, na consulta, são abordados os desafios identificados para o uso de big data para a produção de estatísticas oficiais.

---

## Equipe

### Coordenação

Andrea Diniz da Silva

### Discentes

Beatriz Menezes Marques de Oliveira

Ísis Gonçalves Peixoto

Katia Aragão Eusébio

Lidiane Braga Sales de Souza

Mariza Rayanne da Silva Pereira

### Outra instituição ou área do IBGE envolvida no projeto

Divisão de Estatística das Nações Unidas - UNSD

### Financiamento

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Instituto Brasileiro de Geografia e Estatística - IBGE

## Combinando sub-amostras de cotas e probabilísticas selecionadas por setores para produzir estimativas mais confiáveis

### Palavras-chave

Amostragem probabilística, ponderação, pareamento de amostras e quase-aleatorização

### Área de conhecimento

Probabilidade e estatística

### Início Término

2020 2022

Pesquisas tradicionais enfrentam desafios crescentes devido ao aumento das taxas de não resposta e à diminuição dos recursos disponíveis para as organizações de pesquisa. Uma solução recentemente proposta envolve a combinação de pesquisas amostrais não probabilísticas (muitas vezes mais baratas) com pesquisas amostrais probabilísticas (geralmente mais caras), usando estas últimas como referência para atribuir pesos às unidades das amostras não probabilísticas. Considerando um caso especial em que uma única pesquisa foi projetada e realizada usando simultaneamente as duas abordagens amostrais em uma operação de campo única, este artigo comparou o uso de métodos de quase-randomização e de pareamento de amostras para atribuir pesos à parte não-probabilística da amostra. O método de quase-randomização forneceu as estimativas pontuais mais próximas e erros padrão menores (em média) quando comparadas com as estimativas do método de referência – amostra probabilística.

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Outros pesquisadores

Marcelo Trindade Pitta

Isabela Bertolini Coelho

#### Outra instituição ou área do IBGE envolvida no projeto

Núcleo de Informação e Coordenação do Ponto BR- NIC.br

#### Financiamento

Núcleo de Informação e Coordenação do Ponto BR -NIC.br

## Algoritmo de otimização para definição de áreas de ponderação do Censo Demográfico 2020

### Palavras-chave

Otimização, grafos, árvore geradora, metaheurísticas e busca local

### Área de conhecimento

Estatística e otimização

### Início

2019

### Término

2024

Neste projeto serão desenvolvidos algoritmo de otimização para a definição das Áreas de Ponderação (Aponds) do Censo 2020. As áreas de ponderação são formadas a partir de  $k$  grupos mutuamente exclusivos de setores censitários, observando-se, obrigatoriamente, os critérios de contiguidade e total de domicílios (critérios de viabilidade) e um critério de homogeneidade: (i) Contiguidade: Os setores agregados em cada uma das APONDS devem ser vizinhos (possuir fronteira em comum), ou deve ser possível sair de um setor A e chegar a um setor B, ambos em uma mesma APOND; passando apenas por setores que também estejam alocados nessa mesma APOND; (ii) Total de Domicílios: a soma de unidades domiciliares associadas aos setores que definem cada uma das APONDS deve ser maior ou igual a um total C pré-estabelecido.

### Equipe

#### Coordenação

José André de M. Brito

Marcus Vinicius Morais Fernandes

#### Demais docente

Flávio Marcelo Tavares Montenegro

#### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Métodos e Qualidade, do IBGE

## Análise espaço-temporal e circunstancial dos óbitos de policiais militares do Estado do Rio de Janeiro (2015-2019)

### Palavras-chave

Violência, polícia, mortalidade e indicadores de saúde

### Área de conhecimento

Demografia

### Início Término

2019 2021

A pesquisa realizou, a partir de dados do Instituto de Segurança Pública, uma análise da evolução dos óbitos não-naturais de Policiais Militares no Estado. Além disso, a pesquisa realizou uma avaliação da distribuição espacial da mortalidade violenta desses profissionais, associando tal indicador a outros de segurança, como número de homicídios e apreensão de armas.

### Equipe

#### Coordenação

Ana Carolina Soares Bertho

#### Discente

Matheus Fernandes Mendonça

### Financiamento

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

## Informação estatística e políticas públicas no Brasil

### Palavras-chave

Estatísticas públicas, políticas públicas e sistema estatístico

### Área de conhecimento

Estatísticas públicas

### Início

2019

### Término

-

A abrangência temática e a granularidade territorial/populacional das Estatísticas Públicas refletem o estágio de estruturação do Estado de Bem Estar e das políticas, programas e outras iniciativas implementadas para sua efetivação. De um lado, as Políticas Públicas pautam as demandas de informação para sua formulação e acompanhamento; de outro, o Sistema Estatístico contribui para aprimorar a ação governamental, pelo desvelamento de realidades socioeconômicas nos diagnósticos em diferentes escalas e pela avaliação dos efeitos, defeitos ou não efeitos das políticas e programas. As Estatísticas também respondem à complexidade da base econômica, às demandas dos agentes econômicos e ao conjunto de preocupações sociais a cada momento histórico, verbalizado pela academia, movimento social ou movimento cultural. Refletir sobre como esse contexto político-institucional pós-Constituição de 1988 até o momento atual tem impactado na produção das Estatísticas Públicas no Brasil é a principal motivação da proposição desse projeto de pesquisa. Procura-se trazer contribuições para análise situacional das instituições públicas dedicadas à produção de estatísticas e pesquisas e, em perspectiva comparativa internacional, trazer insumos para o debate sobre Sistema Estatístico e Políticas Públicas. Mais especificamente, a partir do levantamento de documentos formativos de políticas e programas públicos a partir dos anos 1990- diagnósticos e avaliações- em especial os relativos às áreas setoriais de Desenvolvimento Social, Educação, Saúde e Trabalho-, além de entrevistas qualitativas com técnicos de instituições estatísticas e gestores públicos, procura-se evidenciar o uso, relevância e

apropriação de evidências, dados e estatísticas produzidas pelo Instituto Brasileiro de Geografia e Estatística - IBGE, Instituto de Pesquisa Econômica Aplicada, INEP (Instituto Nacional de Estudos e Pesquisa Educacionais), INPE (Instituto Nacional de Pesquisas Espaciais), Fundação Osvaldo Cruz e outros centros de pesquisas na Políticas Públicas.

---

### Equipe

#### Coordenação

Paulo de Martino Jannuzzi

#### Outra instituição ou área do IBGE envolvida no projeto

Escola Nacional de Administração Pública - ENAP

#### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Preferências reprodutivas, fecundidade não realizada e contradições na América Latina

### Palavras-chave

Fecundidade abaixo do nível de reposição, fecundidade não realizada, mismatch e América Latina

### Área de conhecimento

Demografia

### Início

2019

### Término

-

Já é bastante explorada em estudos quantitativos o Gap de fecundidade nos países desenvolvidos, em que os indivíduos possuem em média menos filhos do que gostariam. Na América Latina poucos estudos têm mensurado a existência da fecundidade não realizada e seus determinantes. Sendo assim o objetivo deste projeto é analisar a discrepância de fecundidade na região, a partir de diferentes indicadores disponíveis na literatura. Além disso, o projeto busca discutir os problemas nos dados e as possíveis contradições existentes nos mesmos. Para isso, se utiliza de dados quantitativos das DHS e MICs disponíveis para os diversos países latino-americanos e Brasil. O objetivo é expandir os conhecimentos acerca da fecundidade não realizada em contextos que não os países desenvolvidos.

### Equipe

#### Coordenação

Angelita Alves de Carvalho

#### Demais docentes

Ingrid Gomes Dias

Gabriela M. O. Bonifácio

#### Discentes

Lívia Fernandes

Tiago Abreu de Andrade

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio Grande do Norte - UFRN

Universidade Federal de Viçosa - UFV

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Produção de estatísticas experimentais para indicadores de mercado de trabalho baseados em modelos usando dados da PNAD Contínua

### Palavras-chave

Estatísticas oficiais, pesquisas repetidas, modelos de espaço de estados, séries temporais, erro amostral e taxa de desocupação

### Área de conhecimento

Probabilidade e estatística

### Início Término

2019 -

Muitos indicadores socioeconômicos usados por setores da sociedade, tal como a taxa de desocupação, são provenientes de pesquisas repetidas. Esse tipo de pesquisa, que em geral investiga as mesmas unidades amostrais em diversas ocasiões, é base para a produção de informações que permitem o estudo da evolução de um fenômeno ao longo do tempo. Há interesse na análise das séries temporais de pesquisas repetidas, em especial na decomposição das séries observadas em componentes de tendência e sazonalidade. No caso de pesquisas amostrais repetidas, entretanto, deve-se considerar o efeito que o plano amostral e o desenho da pesquisa exercem sobre as séries observadas. Ambos influenciam diretamente a autocorrelação da série e, portanto, o modelo de decomposição adequado ao caso. A Pesquisa Nacional por Amostragem de Domicílios Contínua – PNAD Contínua – é a principal pesquisa amostral domiciliar do país a partir da qual são divulgados alguns dos principais indicadores socioeconômicos nacionais, em especial aqueles referentes à força de trabalho, como a taxa de desocupação. O Instituto Brasileiro de Geografia e Estatística - IBGE divulga mensalmente indicadores nacionais do mercado de trabalho com base em dados de trimestres móveis, isto é, dados de três meses consecutivos combinados. Adicionalmente, quatro vezes por ano, divulga indicadores em níveis geográficos mais detalhados. Todos esses são produzidos utilizando métodos de estimação baseados no desenho amostral da pesquisa (conhecidos na literatura estatística como design-based methods). Entretanto, outros tipos de estimadores, que incorporam métodos de modelagem estatística (model-based estimators), são recomendados para atender algumas demandas históricas e outros desafios recentes. Dentre as principais demandas estão a dessazonalização



da série da taxa de desocupação; a produção de estatísticas baseadas apenas na amostra mensal; e o uso de fontes de dados alternativos, como as chamadas de big data. Outro desafio é a estimação do efeito da mudança do modo de coleta ocorrido em 2020 devido à pandemia (de CAPI – computer-assisted personal interviewing – para CATI – computer-assisted telephone interviewing). O projeto de pesquisa visa desenvolver modelos de espaços de estados para a estimação e análise da série temporal da PNADC incorporando efeitos do plano amostral da pesquisa, efeitos de mudança de modo de coleta e também informações produzidas por fontes alternativas de dados. O projeto busca aprimorar as estimativas trimestrais, bem como possibilitar a produção de estimativas mensais para estatísticas do mercado de trabalho a partir das amostras mensais da PNADC, utilizando métodos de estimação baseados em modelos.

---

## Equipe

### Coordenação

Denise Britz do Nascimento Silva

### Discente

Caio Gonçalves

### Outros pesquisadores

Luna Hidalgo

Jan van den Brakel

### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Trabalho e Rendimento,  
do IBGE

Statistics Netherlands

## Ajuste de modelos multivariados multiníveis sob amostragem informativa

### Palavras-chave

Modelos multiníveis, resposta multivariada, amostragem informativa e avaliação de proficiência

### Área de conhecimento

Probabilidade e estatística

### Início Término

2018 2022

Desenvolve-se uma abordagem para o ajuste de modelos multivariados multiníveis que consideram a amostragem informativa de unidades populacionais tanto em nível de grupo como de unidade. A abordagem proposta envolve a extração do modelo multivariado multinível para os dados da amostra selecionada em função do modelo populacional correspondente e das probabilidades de seleção da amostra. Em seguida, o modelo amostral resultante é ajustado usando métodos bayesianos. Um estudo de simulação baseado em modelos foi realizado para estudar o desempenho dessa abordagem em um cenário motivado por um estudo de avaliação de proficiência nacional brasileiro. Os resultados indicam que nossa abordagem permite uma melhor inferência sob a amostragem informativa considerada.

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Demais docente

Fernando Antonio da Silva Moura

#### Discente

Natália Raquel de Souza Pires

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio de Janeiro - UFRJ

### Financiamento

Sociedade para o Desenvolvimento da Pesquisa Científica - SCIENCE

## Doenças transmitidas pelo *Aedes aegypti* e a Agenda 2030: um desafio para o Brasil

### Palavras-chave

*Aedes aegypti*, mortalidade, morbidade e ODS

#### Área de conhecimento

Demografia

#### Início

2018

#### Término

2020

A pesquisa teve como objetivo fazer um levantamento bibliográfico, bem como dos indicadores de mortalidade e morbidade por doenças transmitidas pelo *Aedes aegypti* a fim de avaliar se o Brasil conseguirá atingir a meta prevista nos Objetivos de Desenvolvimento Sustentável (ODS) de acabar com tais enfermidades até 2030.

### Equipe

#### Coordenação

Ana Carolina Soares Bertho

#### Discente

Natália da Silva Fernandes

### Financiamento

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Fatores de risco associados à ocorrência de maus-tratos em pessoas idosas

### Palavras-chave

Pessoas idosas, maus-tratos, violência e vulnerabilidade

### Área de conhecimento

Demografia

### Início Término

2018 2020

A pesquisa teve como objetivo identificar os fatores associados à ocorrência de maus-tratos contra pessoas idosas. Para isso, foi usada a pesquisa SABE (Saúde, Bem-Estar e Envelhecimento) realizada em 2015 no município de São Paulo.

### Equipe

#### Coordenação

Ana Carolina Soares Bertho

#### Demais docente

Marília Miranda Forte Gomes

#### Discente

Bruna Matter dos Santos

### Outra instituição ou área do IBGE envolvida no projeto

Universidade de Brasília - UnB

Universidade de São Paulo - USP

## Mercado de trabalho e fecundidade

### Palavras-chave

Mercado de trabalho, fecundidade, gênero e juventude

### Área de conhecimento

Demografia

### Início

2017

### Término

-

Tem-se visto um aumento da inserção das mulheres no mercado de trabalho e, paralelamente, uma constante redução da taxa de fecundidade, existindo uma relação negativa entre participação feminina no mercado de trabalho e o número de filhos. Contudo, em alguns países desenvolvidos, esta relação está se invertendo e tornando-se positiva. Sendo assim, o objetivo do trabalho é analisar a relação entre inserção feminina no mercado de trabalho e a fecundidade nas últimas décadas na América Latina, especialmente no Brasil. Pretende-se ainda analisar a relação de outros indicadores tais como a taxa de ocupação, taxa de desemprego e indicadores masculinos no mercado de trabalho sobre a taxa de fecundidade total. Temas como desigualdades de gênero e NEETs também permeiam este projeto.

### Equipe

#### Coordenação

Angelita Alves de Carvalho

#### Demais docentes

Daniela Polessa

Gustavo da Silva Ferreira

#### Discentes

Marcos Filgueiras de Sousa

Ana Elisa de Mello Britto

### Financiamento

Instituto Brasileiro de Geografia e Estatística - IBGE

## Perfil da saúde sexual e reprodutiva das mulheres brasileiras

### Palavras-chave

Saúde reprodutiva, indicadores e avaliação

#### Área de conhecimento

Demografia

#### Início Término

2015 2020

Este projeto busca descrever os perfis de saúde sexual e reprodutiva da mulher brasileiras em seus diferentes aspectos. Este se divide nas seguintes temáticas: uso de contraceptivos e demanda não atendida por planejamento da fecundidade entre mulheres de diferentes raça/cor; acesso e cobertura de exames preventivos de câncer de colo do útero e mama entre mulheres de diferentes idades; determinantes do tipo de parto realizado e do tempo de aleitamento materno. Como fontes de dados estão sendo exploradas as informações disponíveis na DHS 1986, 1996 e PNDS 2006 (Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher), bem como na PNS 2013 (Pesquisa Nacional de Saúde).

### Equipe

#### Coordenação

Angelita Alves de Carvalho

#### Demais docentes

Daniela Polessa

Gustavo da Silva Ferreira

#### Discentes

Rafael Molina Guimarães

André Avila Schols

Gabriela de Freitas Vieira

Joana Guedes

Matheus Ludolf

Letícia Baptista de Paula Barros

### Financiamento

Instituto Brasileiro de Geografia e Estatística - IBGE

## Modelos de espaços latentes para análise de redes sociais

### Palavras-chave

Espaços latentes, redes sociais, modelos, fluxos, marketing e demografia

### Área de conhecimento

Probabilidade e estatística

### Início

2013

### Término

-

Esta pesquisa busca desenvolver modelos de espaços latentes para análise de dados de redes sociais. Adicionalmente, esta pesquisa busca também desenvolver aplicações destes métodos nas mais diversas áreas, como fluxos populacionais (migração, pendularidade, turismo) e sistemas de recomendação de produtos e serviços.

### Equipe

#### Coordenação

Gustavo da Silva Ferreira

#### Discentes

Marianna Moreira do Nascimento

Pedro Henrique Sodre Puntel

Bianca Tiemi Matsui

Teresa Villanueva Caballero

Izabel Nolau de Souza

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio de Janeiro - UFRJ

## Utilização da capacidade produtiva instalada e produto potencial no Brasil: alternativas não convencionais de mensuração e construção de indicadores para o nível da economia nacional

### Palavras-chave

Taxa de utilização da capacidade produtiva; estoque de capital fixo produtivo, PIB, PIB potencial, crescimento e mudança tecnológica, abordagens clássica, keynesiana e kaleckiana, regime de crescimento e acumulação de capital, Econometria das séries temporais, método de cointegração, economia brasileira

### Área de conhecimento

Área de crescimento e desenvolvimento socioeconômico

### Início Término

2012 2014

As medidas de utilização da capacidade produtiva instalada (UCI) constituem informação-chave em pesquisas aplicadas em Economia. No âmbito das macroanálises, a UCI surge como variável fundamental, seja em um horizonte de longo prazo, típico dos modelos de crescimento econômico, quanto nas análises de conjuntura, necessárias à formulação e monitoramento das políticas fiscal e monetária. Por esta razão, entre os indicadores de evolução da demanda agregada (e do potencial de oferta das estruturas produtivas vigentes) para fins de controle inflacionário, o nível de utilização da capacidade produtiva tem ocupado um lugar proeminente, apesar da pluralidade de formas, muitas vezes contraditórias, de se interpretar seus verdadeiros impactos sobre preços e quantidades produzidas. Para a teoria keynesiana e demais correntes heterodoxas, os incrementos da taxa de utilização da capacidade produtiva instalada são decisivos para a promoção do crescimento econômico, pois expressam a expansão da demanda efetiva que baliza as decisões de investimento. Economias que sistematicamente subutilizam o estoque de capital fixo produtivo disponível, permanecem presas a trajetórias de baixo e instável crescimento, reduzindo seu potencial de geração de renda e ocupação. Por outro lado, a disponibilidade de uma medida robusta para a UCI, não apenas para o setor industrial, mas também para as atividades de serviço e agricultura, é de suma importância para a eficiência de toda política econômica anti-inflacionária e, mais amplamente, para o alcance da estabilidade macroeconômica. Entretanto, no Brasil não existem estatísticas oficiais de UCI para nenhum setor de



atividade econômica. Os dados disponíveis são produzidos por entidades privadas, mediante sondagens industriais ou enquetes junto às firmas respectivas. Todavia, diversas limitações metodológicas têm sido documentadas pela literatura econômica sobre o tema, justificando-se a busca por metodologias alternativas de mensuração da UCI. Partindo dos aportes de Shaikh e Moudud (2004) para uma medida alternativa às sondagens e às metodologias baseadas em funções de produção, essa pesquisa estima econometricamente a taxa de utilização da capacidade produtiva instalada para o total da economia brasileira, no período 1953-2013, com dados em frequência anual e, em seguida, para o subperíodo 1992-2013, com dados em frequência trimestral. Os resultados mostraram-se promissores e compatíveis com o de outras variáveis macroeconômicas observadas. As séries obtidas foram comparadas com os dados obtidos pelas sondagens industriais e são relevantes tanto por suas implicações para estudos sobre os problemas do crescimento econômico brasileiro quanto para a implementação e monitoramento da política econômica.

## Equipe

### Coordenação

Miguel Antonio Pinho Bruno

### Demais docentes

Antonio Ricardo Dantas Caffé (UFBA)

Carlos Eduardo Iwai Drumond (UESC)

Cleiton Silva de Jesus (UEFS)

Eduardo de Lima Campos (ENCE)

Gervásio Santos Ferreira (UFBA)

Patrick Franco Alves (UnB)

Ricardo Severo de Camargo

Macedo (IBMEC - RJ)

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal da Bahia - UFBA

Universidade Estadual de Santa Cruz - UESC

Universidade Estadual de Feira de Santana - UEFS

Universidade de Brasília - UnB

Faculdades Ibmecc - IBMEC/RJ

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Planejamento amostral e amostragem preferencial em Geoestatística

### Palavras-chave

Geoestatística, amostragem preferencial e amostragem

### Área de conhecimento

Probabilidade e estatística

### Início Término

2011 -

Este projeto busca desenvolver e aplicar metodologias para obtenção do planejamento amostral ótimo em Geoestatística utilizando-se abordagens baseadas em maximização de funções utilidades. Este planejamento busca avaliar situações onde a decisão do pesquisador baseia-se na redução da variância preditiva do processo subjacente e nos casos em que existe o interesse no monitoramento de eventos extremos. Adicionalmente busca-se entender e mitigar os efeitos da amostragem preferencial em dados Geoestatísticos.

### Equipe

#### Coordenação

Gustavo da Silva Ferreira

PESQUISA > Programa Institucional  
de Bolsas de Iniciação  
Científica - PIBIC

## Análises de dados de Violências por Sexo, Orientação Sexual e Espaços de Prevalência: Uma Análise a partir da Pesquisa Nacional de Saúde – PNS 2019

### Palavras-chave

Pesquisa Nacional de Saúde, Minorias Sexuais, Violências

### Área de conhecimento

Demografia

### Início Término

2023 2024

Esta pesquisa tem como objetivo principal analisar a relação entre a ocorrência de violência e a orientação sexual das vítimas no Brasil. Em face da variedade de variáveis disponíveis na Pesquisa Nacional de Saúde – 2019 são elencados como objetivos específicos: identificar se há diferença no local de ocorrência de violência de acordo com a orientação sexual; observar os principais autores da violência, analisando se há associação de perfil de autor com vítima de determinada orientação sexual. Em paralelo a isso, pretende-se controlar as análises por variáveis sociodemográficas, a fim de se verificar o perfil da população alvo e, por meio de dados relacionados à saúde disponíveis na pesquisa, buscar possíveis correlações relacionadas à saúde psicológica desses sujeitos.

### Equipe

#### Coordenação

Angelita Alves de Carvalho

#### Demais docente

Rafael Chaves Vasconcelos Barreto

#### Discente

Fernanda Castilho Gomes De Oliveira

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Proposta de Algoritmo de Agrupamento Fuzzy

### Palavras-chave

Agrupamento; Fuzzy; Otimização; Metaheurística

### Área de conhecimento

Computação e Estatística

### Início Término

2023 2024

Planejar e desenvolver um algoritmo de agrupamento fuzzy, baseado em um método de otimização global.

### Equipe

#### Coordenação

José André de Moura Brito

#### Discente

Gabriel Ludovice

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Análise de dados musicais no R

### Palavras-chave

Análise de sentimentos, R, programação, webscraping

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2024

A análise estatística de dados musicais tem se tornado uma área de pesquisa crescente, impulsionada pela disponibilidade de grandes volumes de dados musicais digitais e pelo desenvolvimento de técnicas de análise de dados cada vez mais sofisticadas. "Dados musicais" é um tipo de conjunto de dados relacionado à música, que permite, por exemplo, realizar análise de gêneros musicais, detecção de padrões melódicos, análise de acordes, análise de sentimentos (com base em letras de músicas), classificação de gêneros, recomendação de músicas, entre outras aplicações.

### Equipe

#### Coordenação

Alinne de Carvalho Veiga

#### Demais docente

Guilherme Santos (UFRJ)

#### Discente

Victor Sano Manabe

### Financiamento

Instituto Brasileiro de Geografia e Estatística - IBGE

## Estudo sobre qualidade de Big Data em estatísticas públicas

### Palavras-chave

Qualidade de dados, Erros em pesquisa, Big Data

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2024

O Estudo sobre qualidade de big data em estatísticas públicas busca contribuir para o avanço da discussão sobre a qualidade em big data, traçando um histórico sobre o tema e buscando compreender como novos quadros e abordagens desenvolvidos para fontes secundárias de dados podem auxiliar na produção de estatísticas públicas baseadas em big data, garantindo que as mesmas estejam em consonância com as melhores práticas nacionais e internacionais.

### Equipe

#### Coordenação

Elizabeth Belo Hypólito

#### Discente

Déborah Bruna Cardoso Vaz Coelho

### Financiamento

Instituto Brasileiro de Geografia e Estatística - IBGE

## Inferência para localizações faltantes em Geoestatística associadas à ocorrência de valores extremos

### Palavras-chave

Processo gaussiano, outliers, Geoestatística, dados faltantes, software R

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2024

Este projeto tem o objetivo de implementar rotina(s) no software R associadas à inferência para localizações faltantes no espaço associadas à valores extremos. Utilizando conhecimentos de Inferência Estatística, em especial de Inferência Bayesiana, serão estudadas situações onde deseja-se encontrar a localização associada a um determinado valor extremo. Baseando-se na inclusão de um parâmetro de escala em cada localização do espaço, o trabalho estende a metodologia atual baseada em Processos Gaussianos, permitindo um melhor ajuste e melhores previsões para locais faltantes na presença de outliers.

### Equipe

#### Coordenação

Gustavo da Silva Ferreira

Renata Souza Bueno

#### Discente

Marcelo do Nascimento Maximiano Filho

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq



## Medidas de Centralidade em grafos e o autovetor associado ao raio espectral da matriz Aalpha

### Palavras-chave

centralidade, autovetor, rede

### Área de conhecimento

Teoria dos Grafos, Otimização Combinatória

### Início Término

2023 2024

Redes são sistemas tecnológicos, físicos, biológicos ou sociais caracterizados por um conjunto grande de entidades bem definidas que interagem dinamicamente entre si, como por exemplo: a Internet e a Web. As redes físicas incluem, por exemplo, as redes de distribuição de energia e de água, as redes de transporte; as redes sociais podem ser vistas como redes de relacionamento pessoal/profissional, de comunidades, de pesquisadores, de publicações e as redes biológicas como cadeias alimentares e de transmissão de doenças. Por isso, compreender as suas estruturas, comunidades, funções, interações e propriedades torna-se de fundamental importância para se pensar como pode se manter a funcionalidade dessas redes. Este projeto tem como objetivos introduzir uma nova medida espectral de centralidade, fazer uma comparação dessa medida com as medidas já existentes na literatura e realizar algumas aplicações em redes reais e em algumas famílias de grafos.

### Equipe

#### Coordenação

Carla Silva Oliveira

#### Discente

Matheus Iozzi

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Mulheres sob ataque na Baixada Fluminense (2018-2022): Gênero, raça, classe e território como catalisadores da violência

### Palavras-chave

Violência de gênero, territórios, desigualdades

### Área de conhecimento

Ciências Sociais

Início Término

2023 2024

A violência é um tema que atiga o interesse de pesquisa em diferentes áreas. No campo dos estudos de gênero, as análises têm sido importantes aliadas para dar visibilidade à situação de violência experimentada por milhões de mulheres. No Brasil, a violência contra a mulher adquire um lugar de destaque não apenas pela alta incidência, mas também porque parece representar a expressão mais bem acabada de uma série de desigualdades e injustiças sociais. O Rio de Janeiro é a Unidade da Federação com a terceira maior taxa de feminicídios em relação aos homicídios dolosos cometidos contra mulheres no país (FBSP, 2022). Entretanto, o estado possui assimetrias populacionais, bem como desigualdades sociais, econômicas e culturais que podem se relacionar à distribuição territorial da violência. Desse modo, pretende-se averiguar os casos de violência contra mulheres tendo como foco territorial de investigação a Baixada Fluminense. Atualmente, treze cidades compõem o grupo: Belford Roxo, Duque de Caxias, Guapimirim, Itaguaí, Japeri, Magé, Mesquita, Nilópolis, Nova Iguaçu, Paracambi, Queimados, São João de Meriti e Seropédica. Frequentemente abordados como um bloco, visão que privilegia a visão construída a partir capital fluminense, tais municípios apresentam peculiaridades sob variados aspectos. Nesta pesquisa, desenvolvida no âmbito do Projeto de Iniciação Científica (Pibic/IBGE), busca-se averiguar as especificidades locais no tocante aos registros de violência contra a mulher. Objetiva-se com esta pesquisa lançar luz sobre a dinâmica da violência contra mulheres nos territórios que formam a Baixada Fluminense. Importa, desse modo, evidenciar as principais características sociodemográficas das vítimas dos crimes; das pessoas suspeitas de autoria; e dos espaços de ocorrência dos delitos. Relevante, ainda, é demonstrar em que medida há similitudes e dissonâncias entre os diferentes municípios; e entre estes e a capital metropolitana. Por fim, busca-se evidenciar os fatores asso-

ciados e/ou potencializadores dos casos de violência contra mulheres nos diferentes territórios analisados; e em que medida tais elementos ajudam a ratificar ou desconstruir a imagem de violência endêmica associada às cidades da Baixada Fluminense.

---

### Equipe

#### **Coordenação**

Joice de Souza Soares

#### **Demais docente**

Juscelino Bezerra

#### **Discente**

Giulianna Affonso da Silva

### Financiamento

Instituto Brasileiro de Geografia e  
Estatística - IBGE

## Projeção da população a partir de cenários: estimativas do nível educacional e da fecundidade

### Palavras-chave

Projeções multidimensionais, fecundidade, educação.

### Área de conhecimento

Demografia

### Início Término

2023 2024

A compreensão do futuro da dinâmica populacional é central para a elaboração de projeções populacionais coerentes e confiáveis. Embora tenham níveis de incerteza e imprecisão inerentes, tais projeções são fundamentais ao planejamento de políticas públicas e da economia, já que implicam em uma série de mudanças na estrutura populacional. Neste âmbito, recentemente foram elaboradas uma série de projeções populacionais globais a partir de um conjunto de cenários socioeconômicos, denominados trajetórias socioeconômicas compartilhadas. Estes incluem projeções populacionais baseadas na dinâmica da fecundidade, mortalidade e migração considerando as variáveis sexo, idade e nível educacional, segundo potenciais dinâmicas socioeconômicas futuras. A proposta da presente pesquisa é utilizar tal arcabouço para construir estimativas populacionais no Brasil e em suas Unidades da Federação, especificamente das trajetórias da fecundidade e dos níveis educacionais nas Unidades da Federação do país para o período 2020-2050. A pesquisa terá como plataforma de manipulação de dados e construção dos indicadores o software R, com uso de bases de dados censitárias e de registros administrativos. Serão utilizadas técnicas estatísticas e demográficas para estruturação das bases de dados e cômputo das estimativas presentes e futuras. Visando a disseminação da pesquisa, também serão priorizadas a publicização dos códigos, pressupostos das estimativas e resultados em plataformas de acesso pública.

### Equipe

#### Coordenação

César Augusto Marques da Silva

#### Demais docente

Gustavo Henrique Mitraud Assis Rocha

#### Discente

Júlia Neotti do Nascimento

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Uso de big data para o estudo de assentamentos humanos precários: o potencial de imagens de satélite

### Palavras-chave

Big data, imagem de satélite, assentamentos humanos precários, Indicador ODS 11.1.1

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2024

Pretende-se investigar o potencial de big data gerado por meio de imagens de satélite para compreender a dinâmica de ocupação territorial em assentamentos precários. Por meio de estudo de áreas selecionadas, será desenvolvido um método que possa ser aplicado a áreas com dinâmica social similar, além de gerada discussão que possibilite a adaptação do método para aplicação em áreas com características distintas. Ao longo de 12 meses, a equipe de pesquisa vai avaliar a disponibilidade de imagens de satélite de qualidade e gratuitas para áreas selecionadas de assentamento precário, realizar levantamento e avaliar a aplicabilidade de ferramentas gratuitas para a manipulação de imagens de satélite, experimentar e selecionar modelos de classificação de imagens que ajudem na identificação de assentamentos precário em imagens de satélite.

### Equipe

#### Coordenação

Andrea Diniz da Silva

Marcelo Rodrigues de Albuquerque  
Maranhão

#### Discente

Lívia Médice de Souza Lima

### Financiamento

Conselho Nacional de Desenvolvimento  
Científico e Tecnológico - CNPq

## A saúde do homem brasileiro por perfil de atividade sexual e orientação sexual: uma análise a partir da Pesquisa Nacional de Saúde - PNS 2019

### Palavras-chave

Orientação sexual, saúde sexual e PNS 2019

### Área de conhecimento

Demografia

### Início

2022

### Término

2023

Esta pesquisa busca identificar e analisar o perfil de saúde sexual dos homens cis-gêneros brasileiro que se autodeclararam homossexuais ou bissexuais. Além disso, pretende-se controlar as análises por variáveis sociodemográficas, a fim de se verificar o perfil dessa população e entender possíveis impactos sociais e de acesso à serviços de saúde que a orientação sexual diferente do padrão heteronormativo (heterossexual e bissexual) pode acarretar ao sujeito.

### Equipe

#### Coordenação

Angelita Alves de Carvalho

Rafael Chaves Barreto

#### Discente

Uriel Teixeira

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Estudo sobre o uso de big data para a gestão do serviço de coleta de resíduos sólidos urbanos

### Palavras-chave

Resíduos sólidos, big data e imagem de satélite

### Área de conhecimento

Probabilidade e estatística

### Início Término

2022 2023

Está em debate, na Câmara dos Deputados, o texto-base que irá pautar o estabelecimento de um Plano Nacional de Cidades Inteligentes no Brasil. Trata-se de uma discussão que ocorre na esteira da elaboração da Carta Brasileira Cidades Inteligentes (BRASIL, 2021), iniciativa coordenada pelo Ministério do Desenvolvimento Regional e que envolve, além de outras instituições do Governo Federal, o apoio e a expertise de agentes internacionais, como a Agência Alemã de Cooperação Internacional (GIZ) e a União Internacional de Telecomunicações (ITU). Sem dispor de dados capazes de captar as transformações socioeconômicas e/ou as modificações em padrões de uso e ocupação do solo ocorridas em períodos como o atual (marcados pelo atraso na realização do Censo Demográfico 2020), o dimensionamento da prestação de serviços – como o de coleta de resíduos – termina por comprometer o planejamento orçamentário e a cobertura territorial do que se pretende oferecer à população. Em vista dessa leitura prévia, este projeto se presta a averiguar se a utilização de big data, à luz do referencial de cidades inteligentes e sustentáveis, vem a ser uma alternativa informacional capaz de refletir demandas coletivas vinculadas a serviços como a coleta de resíduos sólidos e, portanto, melhor orientar a tomada de decisão à escala municipal. Postos os intrínsecos desafios relacionados às limitações metodológicas desse tipo de informação, elegeu-se o município de Maricá-RJ enquanto objeto específico de análise. Influenciado pela proximidade com Niterói e, especialmente, com a capital fluminense, esse município tem apresentado desafios consideráveis à gestão de resíduos sólidos e à

conservação. O presente Projeto busca analisar o potencial de utilização de imagens de satélite para o dimensionamento socioterritorial da demanda de serviços de coleta de resíduos sólidos, utilizando como caso o município de Maricá-RJ.

---

## Equipe

### Coordenação

Andrea Diniz da Silva  
Elizabeth Belo Hypolito

### Discentes

Caivin Macedo Ribeiro Borges  
Dimitrio dos Santos Gregório  
Fábio Lucas Pimentel de Oliveira  
Fernanda Castilho Gomes de Oliveira  
Laira Zopellaro Machado Miranda Ferreira

### Outro pesquisador

Álvaro de Moraes Frota

## Financiamento

Conselho Nacional de Desenvolvimento  
Científico e Tecnológico - CNPq



## Identidade de gênero e orientação sexual em estatísticas oficiais

### Palavras-chave

População LGBTQIA+ e inquéritos populacionais

### Área de conhecimento

Demografia

### Início

2022

### Término

2023

Esta pesquisa busca realizar um levantamento bibliográfico sobre a coleta de informações oficiais sobre sexo, identidade de gênero e orientação sexual em pesquisas domiciliares (nacionais e internacionais) e registros administrativos (nacionais). Por meio do desenvolvimento do projeto será ampliado e consolidado conhecimentos sobre a coleta de informações sobre sexo, identidade de gênero e orientação sexual na produção de estatísticas oficiais, constituindo um acervo de materiais sobre o tema e será composta uma base de metadados sobre os materiais selecionados, sumarizar os principais resultados encontrados em termos de dimensionamento da população LGBTQIA+.

### Equipe

#### Coordenação

Barbara Cobo

Angelita Alves de Carvalho

#### Discente

Antonia Xavier

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Música no R: revisão sistemática e aplicação

### Palavras-chave

Dados musicais, modelos de markov, meta-análise, R

### Área de conhecimento

Probabilidade e estatística

### Início

2022

### Término

2023

O objetivo deste projeto é responder à principal pergunta: quais metodologias de análise estatística têm sido empregadas em análises de dados musicais? Como dados musicais, entende-se como dados não estruturados, ou semiestruturados, de composição musical (progressão entre acordes e notas), letras musicais e informações sobre músicas em geral. Busca-se realizar uma revisão sistemática na literatura para que uma das técnicas seja aplicada com o uso do R, que apresenta cerca de 62 pacotes dedicados a lidar com dados musicais em seu ambiente. Busca-se, também, um maior entendimento da aplicação e adequabilidade destes pacotes. O projeto tem como finalidade uma aplicação aos dados musicais de um compositor, a definir, com objetivo de gerar publicação científica para congressos da área.

### Equipe

#### Coordenação

Alinne de Carvalho Veiga

#### Discente

Lucas Uchoa Moreira Gomes

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## O mercado de pinturas no Brasil entre 2000 e 2022: um estudo a partir dos leilões da Bolsa de Arte

### Palavras-chave

Mercado de arte, economia da arte, leilões de arte e preços hedônicos

### Área de conhecimento

Econometria

### Início Término

2022 2023

O projeto tem como objetivo primeiro criar uma base de dados para vendas de pinturas no Brasil. Para isso utilizaremos os dados provenientes dos leilões realizados pela Bolsa de Arte. O segundo objetivo é duplo. Criar um índice de preços para obras de arte no Brasil. O outro é criar um índice de preços para artistas brasileiro. Para isso, utilizaremos a metodologia dos preços hedônicos.

### Equipe

#### Coordenação

Eduardo Lima Campos

#### Demais docente

Luiz Andrés Ribeiro Paixão

#### Discentes

A definir

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Estudo sobre o uso de big data em estatísticas oficiais

### Palavras-chave

Big data, estatísticas públicas, estatísticas oficiais e pesquisa bibliográfica

### Área de conhecimento

Probabilidade e estatística

### Início

2021

### Término

2022

A demanda por mais e melhores estatísticas cresce em direção oposta à disponibilidade de recursos destinados à produção de estatísticas públicas, em especial oficiais. Além disso, decresce a cooperação de respondentes de pesquisas, principalmente entre os que vivem nos grandes centros ou estão inseridos no mercado de trabalho, aumentando o risco de viés nas estatísticas baseadas nos dados coletados para apenas uma parte da população. Para melhorar a oferta de estatísticas oficiais e atender demandas nacionais e internacionais, Institutos de Estatística de vários países têm investido cada vez mais no uso de fontes alternativas de dados como, por exemplo, big data. Embora promissor, big data apresenta limitações metodológicas que devem ser observadas. Seu uso vem acompanhado de desafios que incluem a necessidade de buscar estratégias para a coleta de dados que sejam representativos da população de interesse, de identificar métodos estatísticos confiáveis e precisos para estimar as quantidades de interesse, e ferramentas para análise dos dados que possibilitem reduzir viés decorrente da forma de obtenção dos dados. O presente projeto busca contribuir para o avanço da discussão sobre essa temática, traçando um histórico e possibilitando criação de um ponto de partida para pensar novos usos para essa fonte de dados em estatísticas oficiais. Pretende ampliar e consolidar conhecimentos sobre o uso de big data na produção de estatísticas oficiais, constituindo um acervo de materiais sobre o tema, por meio de pesquisa bibliográfica. Dentre as atividades previstas estão a seleção e catalogação

de referências bibliográficas, segundo o tipo de fonte, ano de publicação, finalidade, uso de método estatístico, dentre outros atributos; a composição de uma base de metadados sobre os artigos selecionados e a preparação de um artigo sobre o estado da arte do uso de big data na produção estatística.

---

## Equipe

### Coordenação

Andrea Diniz da Silva  
Elizabeth Belo Hypolito

### Discentes

Antonia Xavier  
Áttila Kopplin Chiquito  
Beatriz Menezes Marques de Oliveira  
Ísis Gonçalves Peixoto  
Lucas Uchoa Moreira Gomes

### Outro pesquisador

Álvaro de Moraes Frota

## Financiamento

Coordenação de Aperfeiçoamento de  
Pessoal de Nível Superior - CAPES  
Instituto Brasileiro de Geografia e  
Estatística - IBGE

## Similaridade Semântica de Nomes de Produtos e Serviços Utilizando Grafos de Conhecimento

### Palavras-chave

Similaridade semântica, pareamento de textos, grafo de conhecimento e processamento de linguagem natural

### Área de conhecimento

Ciência da computação

### Início Término

2021 2022

O presente projeto de pesquisa visa investigar o uso de grafos de conhecimento para resolver o problema do casamento semântico das descrições de produtos e serviços constantes na classificação do SNIPC (Sistema Nacional de Índice de Preços ao Consumidor do IBGE) com descrições usadas pelos demais sistemas do IBGE e de outras instituições. O SNIPC produz o Índice Nacional de Preços ao Consumidor Amplo (IPCA) que tem por objetivo medir a inflação de um conjunto de produtos e serviços comercializados no varejo, referentes ao consumo pessoal das famílias brasileiras. As descrições da classificação SNIPC são organizadas de maneira hierárquica. Em seu nível mais agregado, os produtos e serviços são divididos em nove grupos (Alimentação e Bebidas, Habitação, Artigos de Residência, Vestuário, Transportes, Saúde e Cuidados Pessoais, Despesas Pessoais, Educação e Comunicação). Além dos grupos, há mais três outros níveis: subgrupo, item e subitem. O nível de subitem é o mais desagregado, em que a inflação é divulgada, sendo composto por 2.570 itens (nomes) distintos. O objetivo final do projeto de pesquisa é o desenvolvimento de um algoritmo capaz de receber como entrada um nome de um produto ou serviço qualquer e, como saída, apresentar a descrição SNIPC correspondente em qualquer um de seus quatro níveis hierárquicos (grupo, subgrupo, item e subitem), obtida através de um processo de casamento semântico. O algoritmo poderá ser aplicado, entre outras finalidades, para o desenvolvimento de conversores automáticos de tabelas (programas para associar automaticamente as descrições

usadas no SNIPC com nomes de produtos/serviços constantes em tabelas de outras instituições) e no aprimoramento dos sistemas de consulta do IBGE. Para que seja possível desenvolver o algoritmo, será preciso realizar a construção de um grafo de conhecimento de produtos e serviços, na língua portuguesa.

---

## Equipe

### Coordenação

Eduardo Corrêa Gonçalves

### Discentes

Leonardo Souza Gomes de Lima

Thiago Pereira Meirelles

### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Índice de Preços,  
do IBGE

## Tabulação de Dados Amostrais - PNAD-Contínua

### Palavras-chave

R, estatísticas públicas, package, amostragem complexa e Pesquisa Nacional por Amostra de Domicílios - PNAD

### Área de conhecimento

Probabilidade e estatística

### Início Término

2021 2022

É incontestável o papel das estatísticas públicas para a formulação e avaliação de políticas públicas. Considerando as dimensões do Brasil, manter o sistema estatístico nacional em funcionamento é uma tarefa complexa. O Instituto Brasileiro de Geografia e Estatística - IBGE tem cumprido o seu papel disponibilizando informações estatísticas de qualidade com boa cobertura e periodicidade. Em um território tão grande, os levantamentos amostrais são necessários para reduzir custos e atender as necessidades de periodização mantendo a boa qualidade. Para atender as necessidades destas pesquisas, os desenhos amostrais são sofisticados (complexos) e as análises a partir destas pesquisas amostrais devem levar em consideração o desenho amostral nos cálculos das estatísticas. Atualmente, mesmo com a grande disseminação de informação pela internet e com as facilidades dos recursos computacionais, o usuário comum, um analista, que utiliza estatísticas das pesquisas amostrais enfrenta dificuldades para proceder os cálculos corretamente. Considerando a relevância das estatísticas públicas e a dificuldade dos usuários para fazer tabulações e interpretar os dados de pesquisas amostrais complexas é relevante a iniciativa de propor instrumentos que facilitem o acesso e proporcionem uma interpretação adequada destas estatísticas. Este projeto tem por objetivo gerar um método simplificado para tabulações da PNAD-Contínua e explorar os resultados dando ênfase às medidas e incertezas em tabelas e gráficos.

### Equipe

#### Coordenação

Cassio Freitas Pereira de Almeida

Sandra Canton Cardoso

#### Discente

Miguel do Nascimento Faria Conforto



## Avaliação da iniciação sexual de estudantes no município do Rio de Janeiro: Uma análise da PeNSE 2015

### Palavras-chave

análise de sobrevivência, desenho amostral complexo, iniciação sexual, adolescentes

### Área de conhecimento

Probabilidade e estatística

### Início Término

2020 2021

Doenças sexualmente transmissíveis, gestações não planejadas e aborto são problemas sociais e de saúde pública enfrentados por diversos países, atingindo diversas faixas etárias. Neste estudo, investigou-se fatores que influenciam no tempo até a iniciação sexual dos estudantes no município do Rio de Janeiro. Para esse fim, técnicas de análise de sobrevivência foram utilizadas, mais especificamente o estimador de Kaplan-Meier e o teste de Linear-Rank adaptados para o plano amostral complexo da edição da Pesquisa Nacional de Saúde do Escolar (PENSE) realizada pelo IBGE. Verificou-se que as variáveis "sexo", "Dependência administrativa da escola", "Consumo de álcool", "Frequência em que os pais sabiam o que o jovem estava fazendo em seu tempo livre", "Com quem o jovem mora", "Se o jovem trabalha", "Uso de cigarro", "Declaração do estado de saúde", "Se já foi forçado a ter relação sexual" e "Se já recebeu orientação sobre doenças sexualmente transmissíveis na escola" foram significativas para explicar o tempo até a iniciação sexual.

### Equipe

#### Coordenação

Maria Luíza Guerra de Toledo

#### Demais docente

Maria Salet Ferreira Novellino

#### Discente

Lucas Dirk Gomes Ferreira

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Avaliação de classificadores interpretáveis em bases de dados de pesquisas do Instituto Brasileiro de Geografia e Estatística - IBGE

### Palavras-chave

Classificadores interpretáveis, árvores de decisão e ciência de dados

#### Área de conhecimento

Probabilidade e estatística

#### Início

2020

#### Término

2021

Classificadores interpretáveis possuem larga utilização no contexto da administração pública, onde as previsões baseadas em dados são normalmente utilizadas para apoiar os gestores no processo de tomada de decisões que podem ter um efeito profundo nas pessoas. Esta pesquisa visa realizar uma análise comparativa de diferentes algoritmos para a construção de classificadores interpretáveis utilizando base de dados do Instituto Brasileiro de Geografia e Estatística - IBGE.

### Equipe

#### Coordenação

Eduardo Corrêa Gonçalves

#### Discente

Vinícius Rogério da Silva

## Percepção da população sobre os impactos da pandemia da COVID-19 nos Direitos Humanos

### Palavras-chave

Direitos Humanos, COVID-19, pesquisa web, ranking generalizado

### Área de conhecimento

Probabilidade e estatística

### Início Término

2020 2021

A consolidação dos Direitos Humanos na cultura popular e institucional do país requer prática efetiva e conjunta do Estado e da Sociedade. Para que isso ocorra, o reconhecimento dos Direitos Humanos como tal, juntamente com a identificação pela população de práticas institucionais garantidoras de tais direitos, é um importante ponto de partida. Organismos nacionais e internacionais têm promovido uma série de intervenções que dão visibilidade à temática dos Direitos Humanos, em especial pela via normativa. Entretanto, as práticas emanadas pelo Poder Público não contribuem para a unificação de conceitos e práticas sobre o tema. Apesar das intensas e permanentes lutas pelos direitos humanos, certas práticas estruturais de violação ainda persistem, pondo em risco o acesso a esses direitos, especialmente por parte das pessoas e dos grupos mais vulneráveis, seja por razões econômicas, culturais ou políticas. Como agravante, em adição às restrições estruturais, ainda podem sobrevir circunstâncias excepcionais que fragilizam mais ainda o acesso das pessoas aos direitos humanos. Esse é o caso de crises econômicas agudas, guerras, desastres naturais e epidemias que podem acometer cidades, países ou todo o planeta. Esse é o caso da COVID-19, que se alastrou por todos os continentes e se constituiu uma ameaça à saúde e à vida das populações em geral, mas especialmente daqueles que possuem menos recursos para adotarem medidas preventivas e curativas eficazes. O enfrentamento à COVID-19, em função de seu alto grau de contágio e significativo nível de letalidade, exigiu medidas de isolamento e distanciamento social, as quais foram recomendadas pela Organização Mundial da Saúde e endossadas por governos nacionais ou locais. Tais medidas foram adotadas em níveis mais ou menos rigorosos por diferentes países e localidades dentro de cada país. Como resultado, houve mudança significativa na vida social, com a suspensão de atividades não essenciais que implique aglomeração, ou seja, quase tudo. Como consequência direta disso, ocorreu uma diminuição do volume atividade produtiva e um enfraqueci-

mento da economia. Um impacto importante, o qual constitui o foco desse projeto de pesquisa, foi a restrição de acesso aos direitos humanos. Para efeito dessa pesquisa, será adotado o entendimento mais inclusivo de direitos humanos. Uma teoria geral dos direitos humanos, explica que esses direitos possuem distintas dimensões. É comum se falar em direitos civis e políticos, caracterizados como direitos ligados às esferas da integridade, segurança, acesso à justiça e liberdades individuais e públicas. Também é comum se falar em direitos econômicos e sociais, caracterizados como direitos ligados às esferas da saúde, educação, previdência e bem-estar social. Essas duas dimensões (dos direitos civis e políticos e dos direitos econômicos e sociais) estão definidas claramente pelo processo histórico de sua afirmação, porém ainda há outras possíveis dimensões sendo debatidas e estudadas, pois decorrem de movimentos historicamente recentes (como a dimensão dos direitos planetários e dos direitos transgeracionais). Sendo assim, a proposta da pesquisa é que na dimensão dos direitos civis e políticos estejam focados nos seguintes direitos: 1) direito de ir e vir; 2) direito ao voto. E na dimensão dos direitos econômicos e sociais que sejam contemplados: 1) direito à educação; 2) direito à saúde; 3) direito ao trabalho e renda. O objetivo da pesquisa é conhecer a percepção da população sobre o aumento ou a redução das dificuldades para o exercício de direitos humanos durante a epidemia da COVID-19 no Brasil, apontando diferenciais associados com as dimensões geográfica do planeta e socioeconômica da população.

---

## Equipe

### Coordenação

Andrea Diniz da Silva  
Elizabeth Belo Hypolito

### Demais docente

Jose Ricardo Ferreira Cunha

### Outro pesquisador

Antônio Etevaldo Junior

### Discentes

Ana Clara Cancelo Cruz  
Bernardo Braga Martins  
Bruno F. A. Paz  
Fausto M. Pinheiro Jr  
Raphaella S. R. C. Lima  
Vitor Hugo C. Souza

### Outra instituição ou área do IBGE envolvida no projeto

Universidade do Estado do Rio de Janeiro - UERJ

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq  
Instituto Brasileiro de Geografia e Estatística - IBGE

## Acompanhamento dos indicadores de desenvolvimento sustentável para a saúde no Brasil

### Palavras-chave

Desenvolvimento sustentável, ODS, Agenda 2030, indicadores de saúde, mortalidade e morbidade

### Área de conhecimento

Demografia

### Início Término

2019 2025

Em 2015, os Estados Membros da Organização das Nações Unidas (ONU) definiram a Agenda 2030 para o Desenvolvimento Sustentável, um conjunto de programas, ações e diretrizes. O documento firmado pelos países define um 17 Objetivos do Desenvolvimento Sustentável (ODS) e 169 metas correspondentes que deverão ser alcançados até 2030. Os objetivos incluem, entre outras coisas, a erradicação da pobreza e da fome, igualdade de gênero e educação de qualidade. O objetivo número 3 se relaciona diretamente à boa saúde e bem-estar, visa “Assegurar uma vida saudável e promover o bem-estar para todos, em todas as idades”. Tendo isso em vista, o objetivo central deste projeto é verificar se é viável que o Brasil atinja as metas definidas pelos ODS relacionadas à saúde dentro do prazo estabelecido.

### Equipe

#### Coordenação

Ana Carolina Soares Bertho

#### Demais docentes

Larissa Carvalho Alves

Gustavo da Silva Ferreira

Daniel Takata Gomes

Cássio Maldonado Turra

#### Discentes

Theo Nascimento Soares da Cunha

Paulo César Silva Andrade dos Santos

Mariam Trilce Martinto Galvez

Natália da Silva Fernandes

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal de Minas Gerais - UFMG

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES

Instituto Brasileiro de Geografia e Estatística - IBGE

## Ensinando Python para alunos do ensino fundamental II através de um jogo 2D

### Palavras-chave

Jogos educacionais, jogos digitais, programação de computadores e python

### Área de conhecimento

Ciência da computação

### Início

2019

### Término

2020

Python é uma linguagem de programação livre, de sintaxe simplificada e extremamente poderosa, especialmente projetada para permitir a construção de programas simples e legíveis. A linguagem possui um grande ecossistema de pacotes, sendo muitos deles destinados ao desenvolvimento de jogos 2D e 3D. Este pode ser considerado um aspecto favorável para a utilização do Python como primeira linguagem, uma vez que diferentes pesquisas têm apontado que ensinar jovens a desenvolver os seus próprios jogos faz crescer o interesse e o rendimento destes em disciplinas que envolvem o aprendizado de conceitos de programação e desenvolvimento de sistemas. Devido a estas características, nas últimas duas décadas a linguagem tem sido largamente adotada para o ensino de programação, tanto em cursos de nível superior como no ensino médio. Porém, ainda há poucos estudos na literatura que tiveram por foco avaliar se Python também se demonstra eficaz para introduzir conceitos de programação a alunos do ensino fundamental. Neste projeto, a programação de jogos de computador foi utilizada como ferramenta para o ensino de conceitos de Computação e Estatística para crianças e jovens. Mais especificamente, o estudo teve por objetivo responder à seguinte questão: “Utilizando a linguagem Python e a pygame, sua principal biblioteca para criação de jogos, é possível desenvolver um jogo pequeno, porém totalmente funcional e gráfico, que cubra todos os conceitos básicos de programação (constantes, variáveis, desvio, repetição e modularização) e que possa servir como ferramenta para o ensino de programação e estatística para estudantes do Ensino Fundamental II?”. O jogo escolhido para testar o potencial da dupla Python e pygame é o tradicional jogo da memória (neste trabalho batizado como “PyMemória!”), pois foi identificado que o desenvolvimento do mesmo explora com relativa facilidade de implementação os

conceitos de programação de interesse relacionados anteriormente. Como um dos objetivos do projeto também foi o aprendizado de estatística por um dos grupos de estudantes, foram introduzidas algumas noções introdutórias de estatística, como por exemplo, noção de aleatoriedade e taxas relativas.

---

## Equipe

### Coordenação

Eduardo Corrêa Gonçalves

### Discente

André Henrique de Barros Gomes

## Estudo dos modelos de regressão assimétricos

### Palavras-chave

Modelos de regressão, assimetria e curtose

#### Área de conhecimento

Probabilidade e estatística

#### Início

2019

#### Término

2020

A classe de modelos de regressão linear, por muitos anos, foi usada para explicar a maioria dos fenômenos aleatórios. Uma suposição assumida por esta classe é de que os dados são normalmente distribuídos e, portanto, a suposição de simetria dos dados é assumida também. Porém, em diversas áreas do conhecimento, o afastamento da simetria impossibilita a aplicação deste tipo de abordagem. Uma solução alternativa para resolver os problemas de assimetria dos dados tem sido o uso de transformação de variáveis. Uma consequência imediata do uso de variáveis transformadas é que as conclusões que serão obtidas se referirão tão somente às observações da variável modificada quando, na verdade, o que se deseja é uma conclusão para a variável na sua escala original. Em muitas situações práticas, propõe-se como alternativa a utilização de uma família mais geral de distribuições, de forma que se consiga modelar a assimetria dos dados e além disso, incluir a distribuição normal como um caso particular. Esta família de distribuições é denominada normal assimétrica que foi introduzida formalmente por Azzalini (1985). Outra possível abordagem para incorporar a presença de assimetria nos dados é o uso da distribuição t de Student assimétrica que pertence a classe de distribuições de mistura de escala normal assimétrica e possui uma cauda mais pesada que a distribuição normal assimétrica. Uma outra possibilidade para considerar maior flexibilidade na modelagem dos dados é o modelo de mistura normal-log-normal apresentada em Bueno (2016) no contexto da estatística espacial. Neste cenário, o desenvolvimento de estudos das propriedades dos modelos de regressão assimétricos bem como maneiras de fazer validação e diagnóstico é de vital importância para situações onde a suposição de normalidade não é satisfeita. O objetivo do trabalho é realizar um estudo sobre



as propriedades teóricas das distribuições assimétricas e avaliar em que condições os modelos de regressão assimétricos possuem melhor adequação aos conjuntos de dados. As distribuições que serão consideradas para avaliação serão a normal assimétrica, t de Student assimétrica e a mistura normal-log-normal.

---

### Equipe

#### Coordenação

Renata Souza Bueno

#### Demais docente

Gustavo Henrique Mitraud Assis Rocha

#### Discente

Isaac Victor Silva Rodrigues

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Seleção de ordem em modelos autorregressivos

### Palavras-chave

Modelos autorregressivos, seleção de ordem, métodos de seleção, lasso bayesiano e seleção de variáveis via busca estocástica

### Área de conhecimento

Probabilidade e estatística

### Início

2019

### Término

2020

Considere a modelagem de séries temporais estacionárias em que os dados são realizações de processos autorregressivos (AR). O modelo autorregressivo pode ser visto como um modelo de regressão linear no qual as variáveis regressoras são os valores do processo em  $p$  instantes anteriores (Box et al., 1994). Assim será utilizado o modelo autorregressivo de ordem  $p$  (AR( $p$ )). Geralmente, na estimação dessa classe de modelos assume-se que a ordem do processo AR é conhecida antecipadamente, ou são utilizados procedimentos de seleção de modelos que, sequencialmente, ajustam modelos de dimensões crescentes. Segundo a perspectiva bayesiana, Litterman (1986b) apresenta uma estratégia para solucionar o problema em uma extensão natural dos modelos AR para séries temporais multivariadas (modelos vetoriais autorregressivos). Esta proposta pode ser facilmente particularizada para o caso de séries univariadas. Muitos estudos foram desenvolvidos com o intuito de realizar simultaneamente a estimação dos coeficientes e seleção da ordem do AR. Para isto são utilizadas as técnicas para seleção de preditores comumente aplicadas em modelos de regressão. Neste contexto e sob uma perspectiva frequentista, Nardi e Rinaldo (2011) utilizam o método Lasso para ajustar modelos autorregressivos de ordem  $p$  (AR( $p$ )). Este projeto teve por objetivo utilizar métodos de seleção de variáveis, como o Lasso (Park e Casella, 2009) e o Elastic Net (Li e Lin, 2010), para a selecionar a ordem de modelos autorregressivos. Os métodos foram comparados por meio de simulações e foi realizada uma aplicação a dados reais. Para isto utilizou-se o software R.

### Equipe

#### Coordenação

Larissa de Carvalho Alves

#### Discente

Luiza do Bonfim Arce

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Amostras complexas e painéis rotativos: o uso do R na construção de um pacote para a estimação de modelos lineares multivariados longitudinais multiníveis

### Palavras-chave

Pwigs, cran, modelos multiníveis, modelos longitudinais e dados amostrais complexos

### Área de conhecimento

Probabilidade e estatística

### Início Término

2018 2019

Pesquisas amostrais oficiais geralmente possuem plano amostral complexo, que inclui estratificação e conglomeração das unidades amostrais. Os métodos de estimação disponíveis na maioria dos pacotes estatísticos supõem que o processo de amostragem gerador do dados se dá por amostragem aleatória simples. O simples uso destes métodos levam a subestimação ou sobrestimação dos erros padrão das estimativas. A Pesquisa Nacional por Domicílios Contínua (PNADc) contém um desenho amostral complexo e um esquema de painéis rotativos que permite análises longitudinais. No R, o pacote survey permite analisar dados amostrais complexos mas em sua opções não existe estimação de um modelo hierárquico longitudinal. Esta metodologia foi apresentada em Veiga et al. (2014), para o caso específico da Pesquisa Mensal de Emprego (PME), que possuía plano amostral complexo com esquema de painéis rotativos diferente do da PNADc. O objetivo deste projeto é aprofundar o uso do R como linguagem de programação. Primordialmente, visa-se o desenvolver a competência da construção de um pacote no R que adapte a metodologia apresentada em Veiga et al. (2014) para o caso da PNADc, mas que também automatize para casos mais gerais com diferentes esquemas de amostragem e de rotação de painéis.

### Equipe

#### Coordenação

Alinne Veiga

#### Discente

João Victor Aguiar de Oliveira Messa

## Classificação multirrótulo de documentos texto utilizando a relevância binária e o algoritmo naïve bayes

### Palavras-chave

Classificação multirrótulo, naïve bayes, relevância binária e aprendizado de máquina

### Área de conhecimento

Probabilidade e estatística

### Início

2018

### Término

2019

Classificação multirrótulo (CMR) é um dos tópicos de pesquisa mais relevantes na área de Aprendizado de Máquina. Neste problema, o objetivo é realizar a associação automática de objetos a uma ou mais classes, pertencentes a um conjunto pré-definido de classes. Existem muitas aplicações modernas e importantes para a CMR, tais como a genômica funcional (determinar as funções biológicas de genes e proteínas) e a categorização de textos (associar documentos texto a tópicos). Este trabalho desenvolveu um classificador multirrótulo no ambiente R, avaliado com uma base de dados contendo informações sobre mais de 264.000 filmes.

### Equipe

#### Coordenação

Eduardo Corrêa Gonçalves

#### Discentes

José Maurício Karl Ururahy

José Lucas Godoy Viana Bastos

## Modelos dinâmicos para dados censurados assimétricos e com caudas pesadas

### Palavras-chave

Modelo Tobit

### Área de conhecimento

Probabilidade e estatística

### Início Término

2018 2019

Projeto de iniciação científica que teve como objetivo estender o modelo Tobit dinâmico para o caso no qual os erros das observações seguem uma distribuição assimétrica ou com cauda pesada. Foram propostos dois modelos e realizados estudos de simulação em vários cenários.

### Equipe

#### Coordenação

Gustavo Henrique Mitraud Assis Rocha

#### Demais docente

Renata Souza Bueno

#### Discente

Roberto da Fonseca Rutz

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Produtividade, terciarização e emprego no Brasil: desafios para o crescimento econômico inclusivo com melhora sustentável das condições de vida da população

### Palavras-chave

Emprego, produtividade, terciarização, condições de vida

### Área de conhecimento

Desenvolvimento socioeconômico

### Início Término

2015 2015

Este projeto de pesquisa analisa as relações entre produtividade, geração de emprego e distribuição de renda na economia brasileira atual, caracterizada por seu crescimento lento e instável, sob expansão precoce do setor serviços, na medida em que o Brasil ainda não ingressou no rol dos países desenvolvidos. As principais perguntas a serem investigadas são: a) de que modo os baixos ganhos de produtividade do trabalho, com expansão do emprego em postos de baixa qualidade, num contexto de rápido crescimento dos serviços, expressariam antes a existência de problemas estruturais não reconhecidos enquanto tais e, conseqüentemente, não resolvidos? b) quais as políticas sociais que poderiam contribuir para reverter essa dinâmica contraditória ou esse “círculo vicioso” que implica redução da competitividade externa e um menor ritmo de melhoria das condições de vida da população? c) quais as conseqüências dos efeitos de interação das dinâmicas macroeconômica e demográfica nesse contexto? Os dados disponíveis mostram comportamentos díspares por idade, escolaridade e região do país no que concernem ao mercado de trabalho. Parte expressiva da população parece estar adiando a busca por um posto de trabalho e deixa, portanto, de integrar a população economicamente ativa. O governo tem promovido, através da ampliação de diversos programas de apoio financeiro, bolsas e projetos acadêmicos, o prolongamento do tempo de estudo dos jovens e jovens adultos (PROUNI, FIES, Ciências sem Fronteiras), etc.). Essas políticas, inquestionavelmente meritórias em seus objetivos essenciais, não deixam de cumprir outro papel relevante numa economia de baixo crescimento: alivia as pressões e demanda sociais sobre o Estado brasileiro e terminam integrando seus gastos de legitimação e reforçando suas condições de

governabilidade. Entretanto, se essas políticas podem ser prolongadas, a priori, indefinidamente, seu público-alvo certamente não pode e não pretende desfrutar dos programas para sempre como alternativa aos postos de trabalho que a economia tem de gerar. Após o período de treinamento ou de qualificação, esse contingente populacional comporá imediatamente a oferta de força de trabalho, elevando as pressões pela ocupação disponível.

---

### Equipe

#### Coordenação

Miguel Antonio Pinho Bruno

#### Discente

Gabriel Azevedo Fioretti

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq

## Distribuição funcional da renda e evolução da wage share no Brasil - determinantes estruturais e seus impactos sobre os rendimentos do trabalho no período 1995-2013

### Palavras-chave

wage share, distribuição funcional da renda, salário médio real, nível geral de emprego

### Área de conhecimento

Desenvolvimento socioeconômico

### Início Término

2013 2014

O presente projeto tem por objetivo estimar os determinantes da distribuição funcional da renda para o total da economia brasileira, no período de 1995-2013. Para isso, são utilizadas a wage share – participação da massa salarial real no PIB –, que é um indicador sintético, pois depende de um conjunto de variáveis que se influenciam mutuamente. Em consequência, sua adequada interpretação exige a consideração das variáveis que o determinam e uma metodologia adequada que permita construí-lo para o nível nacional, em séries anuais e trimestrais e não apenas para as seis grandes regiões metropolitanas. A pesquisa possibilitou a obtenção das séries da parcela salarial e de seus determinantes estruturais tanto em frequência anual quanto trimestral, no período em questão. Os dados mostraram que, a partir de 2004, a concentração funcional da renda no Brasil reduziu-se, mas ainda está longe dos padrões internacionais que caracterizam uma nação desenvolvida. Políticas públicas articuladas com estratégias consistentes de desenvolvimento socioeconômico são necessárias para manter esse indicador em elevação, como resultado tanto da geração de empregos quanto do aumento do salário médio real no Brasil. Trabalhos futuros que necessitem desses dados podem ser beneficiados por essas novas séries.

### Equipe

#### Coordenação

Miguel Antonio Pinho Bruno

#### Discente

Natália Z. Magno

### Financiamento

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq



**PESQUISA > Programa Institucional  
de Projeto de Pesquisa  
para Alunos de  
Graduação - PPAG**

## Análise de dados musicais no R: uma análise da progressão de acordes

### Palavras-chave

Dados musicais, Modelos de Classificação, R, Padrões musicais

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2024

A análise estatística de dados musicais tem se mostrado uma área de pesquisa crescente, impulsionada pelo acesso a grandes volumes de dados musicais digitais e pelo desenvolvimento de técnicas avançadas de análise de dados. Neste projeto de pesquisa, propomos a investigação da progressão de acordes em dados musicais utilizando a linguagem de programação estatística R, bem como a aplicação de modelos de classificação para identificar padrões harmônicos em diferentes gêneros musicais. A análise de progressão de acordes é de grande importância na música, uma vez que acordes são blocos fundamentais para a harmonia e estruturação musical. Através dessa análise, podemos entender um pouco mais sobre padrões e características musicais que são relevantes para diversos domínios, como composição, arranjo, análise musical e recomendação de músicas.

### Equipe

#### Coordenação

Alinne Veiga

#### Demais docente

Guilherme Santos (UFRJ)

#### Discente

Bruno Fernandes Abelaira Paz

André Vicente Priedols e Souza

# Aplicação e comparação do método CHAID em amostragem aleatória e complexa: um estudo exploratório

## Palavras-chave

teste qui-quadrado, árvore de classificação

## Área de conhecimento

Probabilidade e estatística

## Início Término

2023 2024

O método CHAID (Chi-squared Automatic Interaction Detection) é uma técnica estatística usada para identificar padrões e associações em dados categóricos. Ele constrói árvores de classificação que dividem os dados em grupos homogêneos, permitindo a análise de interações complexas entre variáveis categóricas. O CHAID baseia-se em testes qui-quadrado para avaliar associações significativas e determinar as variáveis de maior impacto na variável de interesse. Apesar de amplamente usado com dados de amostragem aleatória, esse técnica é pouco explorada no contexto de amostragem complexa. O objetivo do projeto é estudar e aplicar o método CHAID em conjuntos de dados provenientes de amostragem aleatória e amostragem complexa. Além disso, busca-se realizar uma comparação dos resultados obtidos com o modelo logístico, que já está bem estabelecido no contexto de amostragem complexa.

## Equipe

### Coordenação

Gustavo Henrique M. A. Rocha

### Discente

Leonardo Ferreira de Araújo Sá

## E depois da formatura? Uma análise do perfil dos egressos da Graduação em Estatística da ENCE

### Palavras-chave

ensino superior, acompanhamento de egressos, evasão, avaliação, destinos ocupacionais

### Área de conhecimento

Educação - Planejamento Educacional

### Início Término

2022 2023

Conforme Regimento da ENCE, seu curso de Graduação destina-se a formar bacharéis em Estatística, estimular e promover a iniciação à pesquisa científica, assegurando formação abrangente, eficaz e de alto padrão técnico em nível de graduação. No que se refere à missão da ENCE de formação de pessoas, sabe-se que a relação entre aluno e IES não termina com a colação de grau. Apesar do momento importante, os egressos continuarão sendo os representantes da proposta pedagógica do curso em sua atuação futura. O egresso de um curso de graduação vive o impacto de suas escolhas acadêmicas. Informações sobre carreira, educação permanente e continuada, mudanças profissionais, métricas de sucesso, impacto do repertório da IES no dia a dia, entre outras, trazem grande valor para o planejamento dos gestores educacionais. Desse modo, entender o egresso, suas vitórias e suas dificuldades, ajuda a traçar o perfil de futuros alunos. Além disso, os egressos podem avaliar os cursos que fizeram e apontar seus níveis de satisfação com a faculdade. Esta reflexão pode auxiliar as instituições de ensino na avaliação de suas práticas e até mesmo motivar uma nova atuação junto aos graduandos. De todo o modo, ambas as partes se beneficiam de uma relação mais direta. Saber identificar, de modo objetivo, que tipo de profissional a ENCE tem formado, e como esse está inserido no mercado de trabalho é essencial e fundamental para que a Coordenação de Graduação mantenha o seu Projeto Pedagógico e a grade curricular atualizadas de acordo com tais atuações no mercado de trabalho. O entendimento do perfil do aluno egresso, e das áreas em que atua é, ainda, um dos critérios de avaliação do curso pelo MEC.

Deste modo, este projeto está centrado em uma das metas do PDI, traçada na versão de 2016-2020, e mantida no PDI 2021-2025: “Estabelecer mecanismos de acompanhamento de egressos da graduação e da pós-graduação”. Com foco nos egressos

da graduação, identificou-se, em um um primeiro momento, as principais lacunas a respeito da atividade dos alunos após a colação de grau. Essas lacunas serviram de base para a construção de um questionário que foi aplicado, de modo virtual, aos ex-alunos. Os resultados da pesquisa compõem um material que auxilia a viabilizar o atingimento da meta do PDI, no que se refere à graduação.

---

## Equipe

### Coordenação

Maria Luíza Guerra de Toledo

### Discente

Luis Filipe Tavares de Viveiros Junior

Julia Neotti do Nascimento

Leticia Carvalho de Oliveira

## Análise dos determinantes do tempo de amamentação: Uma abordagem de análise de sobrevivência utilizando o Random Survival Forest

### Palavras-chave

amamentação, análise de sobrevivência, machine learning, random survival forest

### Área de conhecimento

Probabilidade e estatística

### Início Término

2021 2022

A prática do aleitamento materno traz inúmeros benefícios, uma vez que se constitui em uma medida importante para estabelecer uma boa condição de saúde para a criança. Visto a importância desta temática e a quantidade de fatores que influenciam a duração do aleitamento materno, é de extrema relevância entender quais e como os determinantes afetam o aleitamento materno, seja agindo de forma a estender a sua duração seja influenciando sua interrupção. Este trabalho tem como foco a análise estatística de uma base de dados apresentada por Klein e Moeschberger (2013), referente a tempos de desmame de bebês amamentados no peito, e teve como objetivo analisar os determinantes da duração do aleitamento materno na população pesquisada. Pra tal, utilizou-se a técnica Random Survival Forest, de aprendizado de máquina. Estima-se que com apenas 3 meses de vida, metade das crianças já terão desmamado, e por volta dos 6 meses de vida, apenas 21,5% das crianças ainda estarão sendo amamentadas. Vale ressaltar que a recomendação da OMS é que todos os bebês sejam alimentados exclusivamente com leite materno até os 6 meses de idade. Mulheres mais jovens e mais velhas têm menores probabilidades de alcançarem a meta de 6 meses de amamentação do que mulheres que têm entre 20 e 25 anos. De todo modo, independente da idade da mãe, essas probabilidades são baixas, não ultrapassando 0,3. Estima-se que essa probabilidade tem o seu menor valor para mulheres com cerca de 9 anos de escolaridade, aumenta linearmente atingindo seu máximo para mulheres com 15 anos de estudo, e a partir disso diminui novamente. Uma vez superadas as dificuldades iniciais da amamentação, o risco de desmame vai se tornando cada vez menor, até que se estabiliza por volta de aproximadamente dois

anos. Mulheres que amamentam por mais tempo têm o seguinte perfil: De cor/raça branca, e que não utilizaram cigarro ou álcool na gravidez.

---

### Equipe

#### **Coordenação**

Maria Luíza Guerra de Toledo

#### **Discente**

Fernando Pires dos Santos

## Investigação das percepções sobre os impactos da COVID-19 nos Direitos Humanos: construção de um banco de perguntas

### Palavras-chave

Banco de perguntas, pesquisa, direitos humanos e COVID

### Área de conhecimento

Probabilidade e estatística

### Início Término

2020 2021

A discussão sobre os direitos humanos, o que e quais são, além das percepções individuais associadas a eles é ampla e complexa. Enquanto algumas pesquisas utilizam conceitos construídos com base nos trechos da Declaração Universal dos Direitos Humanos (DUDH), outras utilizam escolhas particulares abrangendo, por exemplo, o uso de drogas, eutanásia e aborto, que não constam na DUDH. Para garantir, em alguma medida, a comparabilidade de diferentes estudos sobre o mesmo tema, o reuso de perguntas é uma estratégia muito utilizada. Bancos de perguntas, muitas vezes acompanhadas de protocolos para adaptação ou tradução, são disponibilizados, principalmente por grupos acadêmicos. Porém, mudanças de cenário impõem a necessidade de novos ajustes ou mesmo de novas perguntas, para que os fenômenos sejam medidos em seu novo contexto. Essa é uma das imposições da pandemia da COVID-19. O principal objetivo do projeto foi construir um banco de perguntas elaborado, inicialmente, para atender à pesquisa Percepções sobre os Impactos da COVID-19 nos Direitos Humanos: uma pesquisa com moradores da cidade do Rio de Janeiro, realizada por professores e alunos da Escola Nacional de Ciências Estatísticas e da Universidade do Estado do Rio de Janeiro, com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq e do Instituto Brasileiro de Geografia e Estatística - IBGE, no ano de 2021 e, posteriormente, a diferentes usuários interessados no tema. As fontes utilizadas para a construção do acervo, denominado Banco de Perguntas DHC-19, foram artigos científicos pré-selecionados sobre direitos humanos, pesquisas diversas sobre direitos humanos; e perguntas de autoria do próprio grupo de pesquisas. O banco fornece o enunciado da pergunta em português e em sua língua original (português ou inglês), a classificação da pergunta em aberta ou fechada e, no caso de pergunta fechada,



a informação referente à admissão de resposta única ou de respostas múltiplas e as categorias de resposta. As demais informações que compõem o banco são: o propósito da pergunta (concepção de direitos humano, reconhecimento dos direitos humanos como tal, percepções sobre a limitação dos direitos humanos, avaliação do governo, adesão ao isolamento social durante a pandemia da COVID-19, acesso à informação), a origem (artigo e, nesse caso, o nome do artigo; elaboração do grupo de pesquisa; outro banco de perguntas); se passou por teste e, em caso afirmativo, o tipo de teste realizado (teste informal, revisão por especialistas, grupo focal, entrevista cognitiva, testes de campo), se foi implementada em pesquisa e, se sim, se a mesma era nacional ou internacional, assim como o ano da implementação.

---

### Equipe

#### Coordenação

Elizabeth Belo Hypólito

Andrea Diniz da Silva

#### Discentes

Bruno F. A. Paz

Fausto M. Pinheiro Jr.

Raphaella S. R. C. Lima

#### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio de Janeiro - UFRJ

## Comparação do ajuste de modelos de regressão censurados em diferentes pacotes estatísticos

### Palavras-chave

Dados censurados, mcmc, aproximação determinística e simulação

### Área de conhecimento

Probabilidade e estatística

### Início

2018

### Término

2019

O efeito de censura é um fenômeno comumente observado na estatística e pode ser definido como a condição em que um conjunto de valores da uma amostra não é observado. Usualmente, a modelagem tradicional (modelos lineares) é pautada em amostras completamente observadas, ao passo que sob o efeito de censura, a estimação pode ser comprometida. Uma possível abordagem é o modelo Tobit, que associa um conjunto de covariáveis completamente observadas a uma variável resposta censurada através de uma relação linear. Neste contexto, assume-se o zero como valor limitante, logo apenas valores positivos são efetivamente observados, todavia, o modelo Tobit é capaz de realizar inferência considerando a presença da censura. Sob o paradigma bayesiano, o ajuste de modelos de regressão censurada usualmente é conduzido pelo método Monte Carlo via Cadeias de Markov (MCMC). Face ao método, alguns pacotes do software R facilitam a sua utilização, por outro lado, o MCMC pode apresentar dificuldades quando submetido a uma alta taxa de censura ou a um grande conjunto amostral. Neste sentido, a Aproximação Gaussiana é uma alternativa metodológica considerada computacionalmente mais econômica. A pesquisa tem por objetivo avaliar o ajuste de modelos de regressão censurados através dos pacotes Bugs, Jags, Stan e Nimble, representando a abordagem por MCMC, em comparação a alternativa metodológica sob a forma da Aproximação Gaussiana, com base em um estudo simulado através da criação de cenários de dificuldade.

### Equipe

#### Coordenação

Gustavo Henrique Mitraud Assis Rocha

#### Demais docente

Renata Souza Bueno

#### Discente

Rafael Cabral Fernandez

## Desenvolvimento de um aplicativo web com o Shiny para o ajuste de modelos de regressão censurados

### Palavras-chave

Dados censurados

### Área de conhecimento

Probabilidade e estatística

### Início Término

2018 2019

Com o avanço da tecnologia, a utilização de interfaces digitais, interativas e dinâmicas tem se tornado uma alternativa inovadora e cada vez mais presente na Estatística. Neste trabalho, foi desenvolvido uma aplicação via web por meio do pacote Shiny do software estatístico R, visando facilitar a visualização e análise dos resultados da modelagem de dados censurados. O aplicativo foi estruturado a partir do estudo da classe de modelos de regressão censurados, de suas propriedades e métodos para a realização do procedimento de inferência através do ponto de vista bayesiano. Para demonstrar as funcionalidades do aplicativo foram utilizados dados simulados com o objetivo de validar a construção da modelagem.

### Equipe

#### Coordenação

Gustavo Henrique Mitraud Assis Rocha

#### Demais docente

Renata Souza Bueno

#### Discente

Raquel Ribeiro de Sant'ana

## Estudo e aplicações da classe de modelos fatoriais bayesianos: uma análise de indicadores sociais

### Palavras-chave

Análise fatorial, inferência bayesiana, mcmc e indicadores sociais

#### Área de conhecimento

Probabilidade e estatística

#### Início

2018

#### Término

2019

Este projeto tem por objetivo estudar a classe e as propriedades dos modelos fatoriais sob enfoque bayesiano. Realizar estudos com dados simulados para compreensão do procedimento de inferência e por fim aplicar o modelo para análise de indicadores sociais.

### Equipe

#### Coordenação

Larissa de Carvalho Alves

Renata Souza Bueno

#### Discente

Beatriz Rodrigues Pinna

## Estudo e aplicações da classe de modelos fatoriais bayesianos: uma análise temporal

### Palavras-chave

Modelagem temporal, análise fatorial, inferência bayesiana e MCMC

### Área de conhecimento

Probabilidade e estatística

### Início Término

2018 2019

Este projeto tem por objetivo estudar a classe e as propriedades dos modelos fatoriais sob enfoque bayesiano. Realizar estudos com dados simulados para compreensão do procedimento de inferência e por fim aplicar o modelo para análise temporal.

### Equipe

#### Coordenação

Larissa de Carvalho Alves

Renata Souza Bueno

#### Discente

Danielle Rodrigues Pinna

# EXTENSÃO



## Aprendizado colaborativo: a matemática como via de comunicação entre escola e faculdade

### Palavras-chave

matemática, ensino, extensão, escola, oficina

### Área de conhecimento

Matemática

Início Término

2023 2023

Tomando como ponto de partida a dificuldade inerente à disciplina de matemática no ambiente escolar de modo geral, o projeto propõe atividades em formato de oficinas onde os(as) alunos(as) da ENCE, já cientes da importância do saber matemático em sua vida acadêmica e, supervisionados(as)/orientados(as) pelos(as) professores(as) envolvidos(as) no projeto, dialogam sobre suas experiências como alunos(as) do nível superior e partilham dúvidas dos(as) alunos(as) do 9º ano da Escola Municipal Orlando Villas Boas - EMOVB. A dinâmica das atividades terá como objetivo aprimorar a apresentação dos conceitos básicos de matemática através de exemplos que se utilizam do cotidiano de interesse dos adolescentes com intuito de melhorar o rendimento e colaborar para o aprendizado de matemática destes(as). Além disso, as atividades também vão promover o contato com uma instituição de ensino superior, oportunizando a universidade pública como uma perspectiva de futuro para estes adolescentes e como um lugar de todos e todas, desmistificando o acesso ao ensino público superior e aos cursos de exatas. A ideia original deste projeto em 2019 partiu de uma necessidade apresentada em reuniões com a coordenadora pedagógica da EMOVB. Dentre as dificuldades da escola, destacou-se o rendimento abaixo do esperado dos alunos do 9º ano na disciplina matemática em avaliações diagnósticas da Secretaria Municipal de Educação. Como o problema da dificuldade em matemática é persistente torna-se relevante a manutenção deste projeto de extensão onde os alunos da ENCE podem ajudar nos conceitos básicos de matemática junto aos alunos do 9º ano da EMOVB. Esta troca de saberes permanente será viabilizada através de reuniões semanais da equipe técnica da ENCE na própria faculdade e dos encontros também semanais com adolescentes do nono ano que acontecerá na escola.

Este projeto, cuja temática geral é educação, tem a especificidade da educação matemática que envolve alguns desafios para que o aprendizado conjunto seja algo atraente e acessível aos alunos das mais diversas realidades. Para isso o projeto prevê oficinas de matemática baseadas na utilização de materiais concretos tornando mais acessível os conceitos envolvidos nas cinco unidades temáticas – números, álgebra, geometria, grandezas e medidas, probabilidade e estatística – que devem ser trabalhadas no nono ano. Acreditamos que este trabalho complementar e dinâmico no contra turno da EMO-VB possa gerar uma mudança comportamental nos adolescentes que apresentam dificuldades em matemática de modo a gerar uma nova postura de interesse pela disciplina.

---

## Equipe

### Coordenação

Gustavo da Silva Ferreira

### Demais docentes

Juscelino Bezerra dos Santos

Lúcia Maria dos Santos Pinto

### Discentes

Ana Carolina Wanderley Rodrigues dos Santos

Bruno Henrique de Oliveira Araújo

Flavia Cadilhe Gomes da Silva

Lucca Lemos Nolasco

Maiara Rodrigues de Medeiros Ouro

Nicole Miranda Tuisan da Costa

### Outra instituição ou área do IBGE envolvida no projeto

Escola Municipal Orlando Villas Boas



## Educação estatística: a estatística como aliada na construção do pensamento crítico

### Palavras-chave

Educação estatística; ensino básico

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2023

O projeto tem como objetivo estabelecer uma parceria da ENCE com o IBGE educa para que alunos e professores da ENCE colaborem na construção de materiais que poderão ser disponibilizados através do portal do IBGE educa para promover a educação estatística. Além de criar ações junto às escolas municipais do entorno da ENCE para promover e divulgar o conhecimento estatístico.

### Equipe

#### Coordenação

Renata Souza Bueno

#### Demais docente

Larissa de Carvalho Alves

#### Discentes

Carolina Pereira Barcelos

Heitor Xavier Barros

Julio Cesar dos Santos Magalhães

Victor Melquiades Xavier Leal

Arthur Pacheco Raposo

Calvin Macedo Ribeiro Borges

Fernanda Castilho Gomes de Oliveira

Tiago Abreu de Andrade

Uriel Jose Gomes Teixeira

### Outra instituição ou área do IBGE envolvida no projeto

Centro de Documentação e Disseminação de Informações do IBGE

## Ferramentas computacionais no ensino e aprendizado escolar

### Palavras-chave

extensão, ferramentas, computação, ensino, escolas

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2023

O uso de ferramentas computacionais é de extrema importância no ensino e aprendizagem nas mais variadas disciplinas escolares, em especial no ensino da matemática. Adicionalmente, o letramento computacional e estatístico também se mostra como uma habilidade desejada no mercado de trabalho atual, tornando bastante atrativo o ensino e o desenvolvimento desses conteúdos desde o ensino fundamental. Partindo desse ponto de partida, o projeto ora apresentado propõe a capacitação de professores e alunos da Escola Municipal Orlando Villas Boas por meio do oferecimento de minicursos preparados por alunos da Escola Nacional de Ciências Estatísticas - ENCE. O projeto pretende oferecer 1 minicurso para professores e 1 minicurso para alunos da Escola com periodicidade semestral com o intuito de perenizar esta iniciativa como um dos projetos de extensão da Escola Nacional de Ciências Estatísticas junto à comunidade do seu entorno. Em relação ao minicurso de capacitação de docentes, a ideia é preparar um minicurso voltado para a elaboração de gráficos e outras ferramentas de visualização de dados. Já em relação ao minicurso voltado para os alunos, a ideia é oferecer um minicurso voltado para conceitos básicos de programação e análise descritiva de textos e dados usando o software R. Este trabalho será coordenado por um(a) professor(a) da ENCE e contará com a participação de uma equipe de, no mínimo, 8 alunos da ENCE. A coordenação do projeto manterá reuniões periódicas com os(as) alunos(as) da ENCE participantes do projeto e dará orientações sobre escopo e didática dos minicursos. Como resultado do trabalho, espera-se a capacitação de até 15 professores e 30 alunos semestralmente nos laboratórios de informática da ENCE.

Especificamente em relação aos professores, pretende-se estudar a possibilidade de extensão da iniciativa para capacitação de professores de outras escolas no entorno da ENCE, a partir de 2024, por meio da escola virtual do IBGE.

---

## Equipe

### Coordenação

Alinne de Carvalho Veiga

### Discentes

Ana Clara dos Santos Justiniano Teixeira

Antonia Xavier

Diogo Costa Martins Pizaneschi

Frampton Galhardo de Oliveira

Giulia Killer Soares de Souza

Julio Cesar dos Santos Magalhães

Leone Magno Torres Pereira

Leticia Soto Chaves

Lucas Cardoso Polido

Lucas Rezende da Silva Gaudencio

Paulo Fernando Ferreira Corrado

Rhuan Andrade da Costa

Rodrigo Viana Pereira

Samuel Brenno Mendes Figueiredo

Thiago Minussi Salem de Souza

Vivian Luize Leocardo da Silva Vaz

### Outra instituição ou área do IBGE envolvida no projeto

Escola Municipal Orlando Villas Boas

## Retratos socioeconômicos da favela do Jacarezinho: Etapas de planejamento e campo

### Palavras-chave

Pesquisa amostral domiciliar, indicadores de qualidade de vida

### Área de conhecimento

Probabilidade e estatística

### Início Término

2023 2023

Retratos do Jacarezinho é uma pesquisa por amostragem probabilística de domicílios, desenvolvida em parceria entre a ENCE/IBGE e o LaJaca. Seu objetivo é produzir informações que permitam avaliar as condições de vida das famílias do Jacarezinho. No ano de 2023, foram realizadas as etapas de planejamento da pesquisa, as quais envolve: a definição dos objetivos e temas a serem captados; dos principais indicadores da pesquisa; do plano tabular; e do questionário, composto por temas como moradia, educação, saúde, trabalho, transporte, lazer, meio ambiente e segurança; o treinamento dos entrevistadores; a coleta dos dados e a etapa de crítica e preparação dos dados para a análise.

### Equipe

#### Coordenação

Elizabeth Belo Hypólito

#### Demais docente

Maurício Franca Lila

#### Discentes

Davi Henrique Serra da Silva

Kayo Moura da Silva

Taís da Silva Figueiredo

#### Outros pesquisadores

Poema Eurístenes Portela (LabJaca)

Pedro Paulo dos Santos da Silva (LabJaca)

#### Outra instituição ou área do IBGE envolvida no projeto

Associação Laboratório sobre o Jacaré - LabJaca

## Rio em dados 2023

### Palavras-chave

letramento estatístico, ensino, divulgação, estatística, redes sociais

### Área de conhecimento

Probabilidade e estatística

Início Término

2023 2023

No dia onze de março de 2020, a Organização Mundial da Saúde (OMS) declarou que o mundo vivia em uma pandemia do novo coronavírus. Durante a pandemia, todos se depararam com notícias e divulgações científicas feitas pela imprensa e pelas redes sociais que eram de difícil entendimento. Termos pouco usados pela população ganharam popularidade naquele momento por conta do que estávamos vivendo, entre os termos estavam “média móvel”, “achatar a curva” e “segunda onda”. Diante deste contexto, sentimos uma necessidade de ajudar a esclarecer as informações que eram divulgadas nos meios de comunicação e assim o grupo Rio em Dados foi criado. O grupo percebeu que seria muito importante esclarecer os conceitos estatísticos por trás das notícias e tornar a informação mais acessível, para isso foi criado um perfil na rede social Instagram para a divulgação da estatística. O grupo é formado por alunos, ex-alunos e professores do curso de graduação em estatística da ENCE e dos cursos de graduação e pós-graduação em estatística da UFRJ. Ao passar do tempo, sentimos necessidade de expandir os temas abordados para além da pandemia. Neste sentido, foram formadas diferentes frentes de trabalho no nosso perfil para abordarmos temas além da pandemia como curiosidades da estatística, divulgar análises de dados da atualidade e ensinar de maneira simples como programar e como usar a estatística. Essas frentes foram nomeadas como “Estatística na Sociedade”, “ABC da Estatística” e “Curiosidades da Estatística”. O sucesso do grupo Rio em Dados nos impulsionou a criar este projeto de extensão que visa ampliar a divulgação da estatística de diferentes formas. A ideia é expandir o uso das redes sociais para levar a informação de qualidade e de fácil entendimento para as pessoas. O objetivo principal do grupo Rio em Dados é promover o letramento Estatístico, tornando mais fácil o entendimento de conceitos teóricos para a sociedade em geral. A estratégia do projeto se caracteriza

por ações que visam estimular o pensamento crítico e aproximar a sociedade do curso de graduação em Estatística. Tais ações contribuem para a elevação da qualidade da formação acadêmica dos alunos de graduação e pós-graduação.

---

## Equipe

### Coordenação

Gustavo da Silva Ferreira

### Demais docentes

Ana Carolina Soares Bertho

João Batista de Moraes Pereira

Larissa de Carvalho Alves

Kelly Cristina Mota Gonçalves

Renata Souza Bueno

### Discentes

Ana Clara dos Santos Justiniano Teixeira

André Lessa Silva

Beatriz Rodrigues Pinna

Bruno Pereira Pinto

Camila Rodrigues Ramos

Camilla Carvalho Alves

Eduardo Corrêa Limp

Fillipe Martins Torres Gonçalves

Francisco Rubens Karkow do Amaral

Gabriel Alvarenga Rocha

Guilherme dos Santos

Jedielso Sales de Souza

João Vitor Prisco da Silva

Juan Pablo Ferreira Cavalcante

Júlia Silva de Araújo

Leandro de Freitas Rodrigues

Leonardo Mansor Leite Meira

Leonardo Souza Gomes de Lima

Lucas Uchoa Moreira Gomes

Luís Vicente Guimarães Figueira

Luiza do Bonfim Arce

Maria Sylvia Guimarães Vellozo Martins

Marianna Monteiro Canedo

Marina Carvalho Mangini

Matheus de Castro Brito

Matheus Vasconcellos Costa Bom

Paulo Pladena de Carvalho Garcia Neto

Renan Bezerra Dantas dos Santos

Rodolfo Bramont Eiriz de Souza

Ruan Francisco Ribeiro Saback

Samuel Brenno Mendes Figueiredo

Sofia Swerts Henriques de Brito

Thais Santos de Mesquita

Vinícius Rogério da Silva

Vitor Cunha Pereira

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio de Janeiro - UFRJ

## Estudo dos impactos da COVID-19 no Jacarezinho

### Palavras-chave

COVID-19, desigualdade e pesquisa amostral

#### Área de conhecimento

Probabilidade e estatística

#### Início

2021

#### Término

2022

Em março de 2020, a Organização Mundial da Saúde (OMS) decretou que a COVID-19 havia se tornado uma pandemia. Um ano e meio depois desse acontecimento, o Brasil registrou mais de 600 mil mortes pela doença, marca que o coloca entre os três países com maior número de mortes causadas pela COVID-19. As desigualdades raciais, sociais, territoriais, tal como as de gênero e sexualidade, figuram como “fatores de risco” centrais para compreensão dos impactos da pandemia na população. A frase “Nem bala, nem fome, nem COVID”, mobilizada por ativistas negras(os), faveladas(os) e periféricas(os), cristaliza que determinadas pessoas foram simultaneamente submetidas à violência policial, à insegurança alimentar e à pandemia. Ao passo que a inflação e o desemprego culminaram no crescimento da fome no país, as operações policiais demasiadamente violentas não deixaram de ocorrer em favelas, como no caso da chacina do Jacarezinho que vitimou 28 pessoas em maio de 2021 e é considerada a ação policial mais violenta na história do estado do Rio de Janeiro. Nessa perspectiva, o estudo objetiva compreender como a vida da população da favela do Jacarezinho foi modificada durante a pandemia, procurando identificar alguns dos seus impactos na vida dos moradores, especialmente em vista das desigualdades destacadas acima. É a partir desse cenário, de sobreposição de violências e agravamento de vulnerabilidades, que investigaremos as mudanças nas condições de vida da população da favela a partir de quatro dimensões: i) saúde física e mental; ii) renda, trabalho

e hábitos de consumo; iii) educação; iv) lazer e uso de tecnologias. A investigação será baseada nas respostas obtidas por meio de uma pesquisa amostral sobre o impacto da COVID-19 realizada com os moradores da favela do Jacarezinho.

---

## Equipe

### Coordenação

Andrea Diniz da Silva

Maurício Franca Lila

### Discentes

Kayo Moura da Silva

### Outros pesquisadores

Poema Eurístenes Portela

Pedro Paulo dos Santos da Silva

### Outra instituição ou área do IBGE envolvida no projeto

Associação Laboratório sobre o Jacaré - LabJaca



## Rio em Dados 2020-2022

### Palavras-chave

Ensino da estatística, estatística na sociedade e curiosidades da estatística

### Área de conhecimento

Probabilidade e estatística

### Início

2020

### Término

2022

O projeto foi criado em 2020 e tem como objetivo divulgar informações relevantes para a sociedade, explorando os diversos aspectos da Estatística. A ideia do projeto é trabalhar continuamente e usar as redes sociais para a divulgação. Atualmente estamos divididos em 3 grupos: abc da estatística, curiosidades da estatística e estatística na sociedade.

### Equipe

#### Coordenação

Larissa de Carvalho Alves  
Renata Souza Bueno  
Gustavo da Silva Ferreira  
Kelly Cristina Mota Gonçalves  
João Batista de Moraes Pereira

#### Discentes

André Lessa  
Beatriz Pinna  
Camilla Carvalho Alves  
Fernando Santos Barbosa  
Francisco Rubens do Amaral  
Gabriel Machado  
Guilherme Santos  
Guilherme Willian Conde Simões  
Isabella Marreco Guedes  
Jedielso Sales

João Logatto  
Julia de Araújo  
Júlia Silva de Araújo  
Leonardo Lima  
Lucas Uchoa  
Luis Henrique Toledo  
Luis Vicente Figueira  
Luiza do Bonfim Arce  
Sofia Swerts  
Tais Figueiredo  
Vinícius Rogério da Silva  
William Conde Simões

### Outra instituição ou área do IBGE envolvida no projeto

Universidade Federal do Rio de Janeiro - UFRJ

## Aprendizado colaborativo: a matemática como via de comunicação entre escola e faculdade - 2022

### Palavras-chave

Extensão, ensino fundamental e escola municipal Orlando Villas Boas

#### Área de conhecimento

Matemática básica

#### Início

2019

#### Término

2020

Tomando como ponto de partida a dificuldade inerente à disciplina de matemática no ambiente escolar de modo geral, vamos propor “aulas de reforço” onde os(as) alunos(as) da ENCE, já cientes da importância do saber matemático em sua vida acadêmica e, supervisionados(as)/orientados(as) pelo(as) professor(as) envolvido(as) no projeto, irão dialogar sobre suas experiências como alunos(as) do nível superior e partilhar as dúvidas dos(as) alunos(as) do 9º ano da Escola Municipal Orlando Villas Boas em aulas dinâmicas de reforço escolar com intuito de melhorar o rendimento e colaborar para o aprendizado de matemática destes(as) adolescentes. O projeto pretende atender, especificamente, as três turmas de 9º ano da Escola Orlando Villas Boas. Sendo que apenas os(as) alunos(as) que tirarem nota inferior a 5,0 na avaliação bimestral, participariam do projeto. Isto daria aproximadamente 15 alunos(as) de cada turma, totalizando, em média, 45 alunos(as) atendidos no segundo semestre de 2019 e outros novos(as) 45 alunos(as) no primeiro semestre de 2020. A ideia é que seja um trabalho conjunto entre uma equipe de professores da ENCE, os professores de matemática da Escola Municipal, alguns alunos da graduação e os(as) adolescentes da Escola Municipal Orlando Villas Boas que seriam atendidos(as) em três diferentes dias da semana numa média de 15 alunos(as) por cada dia. A equipe técnica da ENCE responsável por este “reforço escolar” será formada pela Coordenação do projeto, professora Lúcia Maria dos Santos Pinto, pela professora Daniela Polessa Paula, pelo professor Juscelino Bezerra dos Santos e por uma equipe mínima de seis alunos(as) da ENCE. A professora coordenadora deste projeto se reuniria semanalmente com os(as) alunos(as) da

ENCE interessados(as) no projeto e discutiria com eles(as) formas de transmitir alguns conceitos básicos contidos nos descritores de matemática da Prefeitura. A ideia seria utilizar o máximo possível de materiais concretos, tais como, dobradura, papel quadriculado, material dourado, tangran, dentre outros, com a expectativa de que tais atividades gerem mais confiança nos(as) adolescentes do 9º ano que neste momento se encontram com mais dificuldades na disciplina de matemática. Este projeto não prevê a necessidade de recursos financeiros para as partes envolvidas uma vez que os(as) alunos(as) da ENCE receberão horas complementares como retorno pela sua atuação. Espera-se com isto não somente melhorar o rendimento dos(as) adolescentes, mas também atender uma demanda da ENCE em projetos de extensão.

---

### Equipe

#### Coordenação

Lúcia M. dos S. Pinto

Juscelino Bezerra dos Santos

#### Discentes

Juliana Correia de Araujo Aquino

Júlia Vallés Marques

Ana Marta Silva de Oliveira

Camila Rodrigues Ramos

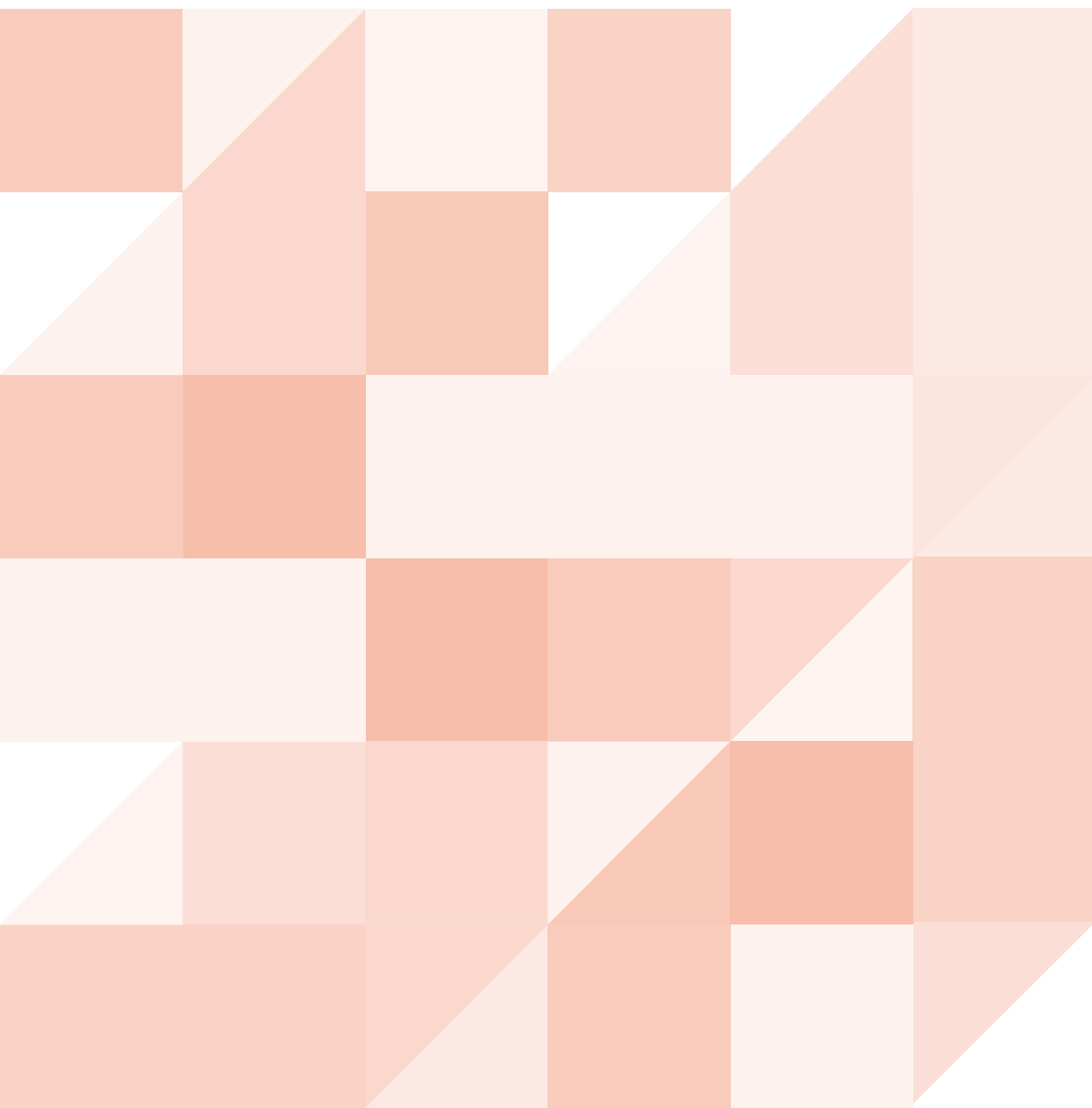
Lukas Alegretti

Melinda Santos

#### Outra instituição ou área do IBGE envolvida no projeto

Escola Municipal Orlando Villas Boas

## OUTROS PROJETOS



## Pesquisa por Amostra de Domicílios de Niterói - PMADNit

### Palavras-chave

Políticas públicas, condições de vida, planejamento, gestão municipal, evidências

### Área de conhecimento

Demografia

### Início Término

2023 2025

A PMADNit é uma pesquisa domiciliar por amostragem, tendo como população de pesquisa a população residente em domicílios particulares permanentes situados no Município de Niterói - RJ. A pesquisa visa contribuir para a compreensão da realidade da população residente do Município, monitorando os impactos das políticas públicas já realizadas e subsidiando a criação de novas estratégias.

### Equipe

#### Coordenação

Denise Britz do Nascimento Silva

#### Demais docentes

Barbara Cobo Soares

Pedro Luis do Nascimento Silva

Paulo de Martino Jannuzzi

### Outra instituição ou área do IBGE envolvida no projeto

Sociedade para o Desenvolvimento da Pesquisa Científica - SCIENCE

### Financiamento

Prefeitura Municipal de Niterói

## Pesquisa de prevalência de COVID-19 nas escolas públicas e privadas do Estado da Paraíba

### Palavras-chave

Saúde pública, pandemia e estudo sorológico

#### Área de conhecimento

Probabilidade e estatística

#### Início

2021

#### Término

2022

O objetivo dessa pesquisa é obter dados sobre a prevalência de contaminação pelo vírus causador da COVID-19 entre estudantes e professores das escolas públicas e privadas do Estado da Paraíba, e estimar quantas pessoas têm essa doença. Os dados foram coletados por meio de perguntas feitas e registradas em um questionário eletrônico, programado em um tablet, e por meio da aplicação de um teste de COVID-19. Os testes aplicados em duas etapas da pesquisa foram primeiro o teste RT-PCR, e na segunda etapa o teste de antígeno. Em ambos os casos era feita coleta de secreção do nariz e da garganta de cada participante. No caso da etapa que empregou o teste RT-PCR, o material colhido era enviado ao LACEN-PB para análise. Na etapa que empregou o teste de antígeno, os resultados eram obtidos no momento da coleta e informados imediatamente ao participante. A pesquisa se justifica pela necessidade do Governo da Paraíba de conhecer o estado da pandemia de COVID-19 junto às escolas que retomaram atividades presenciais para poder definir suas estratégias e ações de combate à pandemia. Na primeira etapa foram conduzidas coletas da pesquisa junto aos públicos: educação infantil; ensino fundamental - anos iniciais; ensino fundamental - anos finais; ensino médio. Na segunda etapa foram contemplados novamente os públicos ensino fundamental - anos finais, e ensino médio.

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Demais docente

Mauricio Teixeira Leite de Vasconcellos

#### Outra instituição ou área do IBGE envolvida no projeto

Sociedade para o Desenvolvimento da Pesquisa Científica - SCIENCE

Universidade Federal da Paraíba - UFPB

Escola de Saúde Pública da Paraíba - ESP-PB

#### Financiamento

Fundação de Apoio à Pesquisa do Estado da Paraíba - FAPESQ

## Pesquisa sobre as condições de trabalho na agricultura brasileira

### Palavras-chave

Trabalho precário, pesquisa domiciliar e histórico laboral

### Área de conhecimento

Probabilidade e estatística

### Início

2021

### Término

2024

De acordo com o Índice Global de Escravidão de 2018, estima-se que existam 369 mil vítimas de tráfico de trabalho no Brasil em qualquer dia, muitas traficadas através de extensas redes de trânsito doméstico. O setor agrícola é, de longe, responsável pela maior parcela de casos conhecidos de tráfico de trabalho no Brasil. De acordo com os casos arquivados no Observatório do Trabalho Escravo, mais de 80% dos casos conhecidos de tráfico ocorrem no setor agrícola, com cerca de 38% na pecuária, 27% na cana-de-açúcar e 22% no cultivo de arroz. Também foram encontradas vítimas em muitos outros setores agrícolas, como chá, cacau e café. A pesquisa tem por objetivo estimar a prevalência atual do trabalho em condições precárias, e conhecer características das relações de trabalho no setor agrícola em quatro estados brasileiros (Bahia, Maranhão, Piauí e Tocantins).

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Demais docente

Mauricio Teixeira Leite de Vasconcellos

Denise Britz do Nascimento Silva

### Outra instituição ou área do IBGE envolvida no projeto

Sociedade para o Desenvolvimento da Pesquisa Científica - SCIENCE

Stanford University

### Financiamento

Stanford University

## Programa de capacitação intermediário em Ciência de Dados no Instituto Brasileiro de Geografia e Estatística - IBGE

### Palavras-chave

Ciência de Dados, Instituto Brasileiro de Geografia e Estatística - IBGE e capacitação

### Área de conhecimento

Ciência de Dados

### Início

2021

### Término

2023

Capacitar os participantes para analisar contextos e problemas institucionais que envolvam coleta, exploração, análise e uso de dados para tomada de decisão e aplicação das principais técnicas em ciência de dados em processos de trabalho no Instituto Brasileiro de Geografia e Estatística - IBGE.

### Equipe

#### Coordenação

Pâmella Raquel Calazans Lopes

Aline Bezerra da Silva

#### Demais docentes

Eduardo Corrêa Gonçalves

Cassio Freitas Pereira Almeida

#### Outros pesquisadores

João Carlos Silva Rodrigues

Marcelo Rodrigues de Holanda Maia

Breno Tiago Novello Trotta de Oliveira

### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Treinamento e Aperfeiçoamento e Diretoria de Pesquisas, do IBGE



## Pesquisa Domiciliar Participativa (PDP) do Projeto Rio Doce

### Palavras-chave

Pesquisa domiciliar, avaliação de impacto, amostragem e métodos de pesquisa

### Área de conhecimento

Probabilidade e estatística

### Início

2020

### Término

2023

A PDP tem como objetivos centrais fornecer um retrato da situação socioeconômica da população da região delimitada como atingida pelo desastre do rompimento da barragem de rejeitos de Fundão, em Mariana (MG), produzir informações que permitam testar a hipótese de ocorrência de impacto econômico do desastre sobre a população da região atingida e estimar a magnitude desse impacto, caso confirmado.

### Equipe

#### Coordenação

Denise Britz do Nascimento Silva

#### Demais docentes

Pedro Luis do Nascimento Silva

Mauricio Teixeira Leite de Vasconcellos

### Outra instituição ou área do IBGE envolvida no projeto

Sociedade para o Desenvolvimento da Pesquisa Científica - SCIENCE

Fundação Getúlio Vargas - FGV

### Financiamento

Fundação Getúlio Vargas - FGV

## Métodos de pareamento de dados sobre agropecuária

### Palavras-chave

Estatísticas agropecuárias, integração de dados e pareamento de dados

### Área de conhecimento

Probabilidade e estatística

### Início

2019

### Término

2020

Indicadores importantes das condições de vida da população dependem da produção de estatísticas agropecuárias. Do Produto Interno Bruto aos indicadores de desenvolvimento sustentável para o monitoramento da erradicação da fome, do crescimento sustentado e do uso sustentável de recursos naturais, todos dependem de estatísticas agropecuárias de boa qualidade e com produção contínua. O atual sistema de pesquisas agropecuárias do Instituto Brasileiro de Geografia e Estatística - Instituto Brasileiro de Geografia e Estatística - IBGE, é baseado em censos, pesquisas em subpopulações de estabelecimentos e pesquisas subjetivas (consulta a especialistas). Apesar de produzir grande quantidade de dados com alta qualidade, limitações intrínsecas às fontes incluem a natural obsolescência dos dados na medida em que nos distanciamos da sua data de referência, subcobertura de produtos e de estabelecimentos e impossibilidade de calcular uma medida de precisão para as estimativas feitas, dependendo da pesquisa. Para melhorar a oferta de estatísticas agropecuárias e atender demandas nacionais e internacionais, o Instituto Brasileiro de Geografia e Estatística - IBGE considera a integração de dados de pesquisas e registros, da própria instituição ou de outros órgãos ligados à Administração Pública Federal. Contudo, nem todos os registros de tais fontes dispõem de um identificador único e, sendo assim, a integração dos dados depende do desenvolvimento de um método de pareamento que ajude a identificar quais registros pertencem a um mesmo estabelecimento. Estudos e experimentos com dados provenientes do Censo Agro 2017 e dados fornecidos pela Secretaria Estadual de Defesa Animal e Vegetal do Pará e do Tocantins foram iniciados em agosto de

2018. Foram preparados códigos na linguagem R para a limpeza, padronização e pareamento dos dados. Além disso, foi realizado um experimento para identificar um limite de tolerância ótimo para a diferença entre as coordenadas quando estas são utilizadas como variável de pareamento na integração de dados do Censo Agro 2017 e da Defesa Vegetal do Tocantins). Em 2019 tiveram início os estudos para o pareamento dos dados do Censo Agro 2017 com os provenientes do Cadastro Central de Empresas - CEMPRE, assim como para a geração de dados de treinamento. O objetivo do estudo é desenvolver um método de pareamento de dados que possibilite a identificação dos registros pertencentes a um mesmo estabelecimento agropecuário e, conseqüentemente, possibilite a integração de dados provenientes do Censo Agro 2017 com dados provenientes de registros administrativos.

---

### Equipe

#### Coordenação

Andrea Diniz da Silva

#### Outros pesquisadores

Raphael Molina Guimaraes

Geremias de Mattos Fontes Neto

#### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Métodos e Qualidade e  
Coordenação de Agropecuária, do IBGE

## Programa de capacitação em Ciência de Dados aplicada a Geociências

### Palavras-chave

Geociências, Ciência de Dados, machine learning e pesquisa-aplicação

### Área de conhecimento

Ciência de dados

### Início

2019

### Término

-

Atualmente, o volume de dados na área da geociência tem transformado os processos de trabalho. Há uma grande facilidade em obtenção de imagem digitais de diversos tipos e em alta periodicidade. Há uma disponibilidade de dispositivos relativamente simples equipados com diversos sensores que propiciam a coleta de muitos dados já em meio digital. Esta grande quantidade de dados é de grande interesse e importância para a geração de produtos na área de geociências. Contudo, esta produção está mais complexa, com demanda para produtos mais detalhados, com maior cobertura e periodicidade. Neste contexto, é primordial a utilização de novas tecnologias para atender a esta demanda. A integração dos processos de produção geocientíficos do Instituto Brasileiro de Geografia e Estatística - IBGE, da Diretoria de Geociências - DGC com o núcleo de pesquisa em Ciências de Dados da ENCE, possibilitará o mapeamento de demandas e identificação de tecnologias e metodologias capazes de suprir as necessidades da área. O propósito não é somente apresentar soluções, mas promover a aprendizagem contínua visando a especialização em Ciência de Dados na DGC, através da pesquisa-aplicação.

### Equipe

#### Coordenação

Cassio Freitas Pereira de Almeida

#### Outra instituição ou área do IBGE envolvida no projeto

Coordenação de Treinamento e Aperfeiçoamento e Diretoria de Geociências, do IBGE

## Revisão do livro *Análise de Dados Amostrais com R*

### Palavras-chave

Amostragem; superpopulação; pseudo-máxima verossimilhança; modelagem; efeito do plano amostral; estimação de variâncias

### Área de conhecimento

Probabilidade e estatística

### Início

2019

### Término

2024

O objetivo deste projeto é fazer uma segunda edição, revisada e ampliada, do livro *Análise de Dados Amostrais Complexos*, publicado por Pessoa e Silva (1998). Na parte 1 do livro, o material do livro original será revisado e atualizado, incluindo-se exemplos e exercícios usando dados de pesquisas amostrais mais atuais e as ferramentas do software R para análise de dados amostrais. Na segunda parte, um conjunto de autores convidados terão capítulos dedicados a temas novos que não foram tratados no livro original, expandindo assim a cobertura temática e ilustrando as ferramentas de análise disponíveis para os novos temas.

### Equipe

#### Coordenação

Pedro Luis do Nascimento Silva

#### Outros pesquisadores

Antonio José Ribeiro Dias

Sonia Albieri

Zélia Magalhães Bianchini

#### Discente

Guilherme Jacob

### Outra instituição ou área do IBGE envolvida no projeto

Núcleo de Informação e Coordenação do Ponto BR - NIC.br

Universidade Federal de Juiz de Fora - UFJF

Universidade Federal do Rio de Janeiro - UFRJ

OUTROS > Livros  
PROJETOS

## Zona Portuária do Rio de Janeiro: Múltiplos olhares sobre um espaço em mutação

### Organizadores

Letícia Giannella  
João Carlos Monteiro

### Editora

Consequência

### Palavras-chave

Zona portuária, Porto Maravilha e planejamento urbano

### Área de conhecimento

Geografia

### Edição

1ª (2022)

### Formato

Digital

O livro foi organizado em conjunto com o professor João Carlos Monteiro, como um registro das memórias do evento “Porto Maravilha 10 anos: passado, presente e futuro da zona portuária” e conta com 18 capítulos de pesquisadores renomados sobre diversas perspectivas de análise a respeito das transformações da zona portuária no contexto do Porto Maravilha e possibilidades de futuro.

## Linguagem R: Uma visão computacional

**Autor**

José André de Moura Brito

**Editora**

Ciência Moderna

**Palavras-chave**

Linguagem R, programação e simulação

**Área de conhecimento**

Computação e estatística

**Edição**

1ª (2020)

**Formato**

Impresso e digital

Este livro traz uma visão geral da linguagem R. Apresenta desde os conceitos mais básicos da linguagem, até alguns conceitos mais avançados, no que diz respeito à utilização das estruturas de seleção, de repetição, de vetores, matrizes, listas, data frames, fatores, tabelas, da família das funções apply, lapply e tapply, construção de gráficos, além de uma introdução ao paralelismo. Os sete primeiros capítulos desse livro trazem uma descrição do ambiente de trabalho do R e dos objetos básicos e especiais, que são utilizados nessa linguagem, sejam eles: vetores, matrizes, data frames, listas e fatores. Os capítulos subsequentes apresentam: (i) estruturas de repetição e seleção utilizadas no R; (ii) uma descrição detalhada do funcionamento e da criação de funções no R; (iii) manipulação de strings; (iv) introdução ao paralelismo e à criação e manipulação de vários tipos de gráficos no R; (v) descrição dos quatro tipos de funções que são utilizadas no R em distribuições de probabilidade e a proposta de alguns projetos, mais especificamente, exercícios de maior complexidade que abrangem vetores, matrizes, data frames, estruturas de repetição, funções e gráficos. Em todos os capítulos, as definições e conceitos são introduzidos com exemplos. Além disso, ao final dos capítulos de número 2 até 10, são apresentados exercícios resolvidos (mais de 70) com grau de dificuldade variado.



## Pandas Python: Data Wrangling para Ciência de Dados

**Autor**

Eduardo Corrêa Gonçalves

**Editora**

Casa do Código

**Palavras-chave**

Data wrangling, Ciência de Dados e Python

**Área de conhecimento**

Probabilidade e estatística

**Edição**

1ª (2019)

**Formato**

Impresso e digital

Ciência de dados (data science) é um processo que emprega técnicas estatísticas e computacionais para analisar grandes bases de dados, procurando extrair delas conhecimento útil para empresas, instituições científicas, governos e demais organizações. Tipicamente, os projetos de ciência de dados são divididos em quatro etapas de execução. A primeira consiste na definição do problema que será resolvido (por exemplo, criar um sistema para caracterizar o perfil dos clientes de uma empresa). A segunda é a etapa de pré-processamento, onde as bases de dados relevantes (base de vendas, base de clientes etc.) devem ser reunidas e adequadamente formatadas. Na terceira etapa, um algoritmo é aplicado sobre os dados pré-processados, com o objetivo de extrair um modelo estatístico que terá por objetivo identificar padrões de relacionamento entre os itens de dados (por exemplo, o algoritmo pode gerar um modelo que revela as características mais comuns dos compradores de cada produto). Por fim, na quarta etapa, os especialistas da empresa avaliam os resultados gerados pelo modelo, procurando determinar sua relevância. A etapa de pré-processamento costuma ser a mais trabalhosa. Nela, são realizadas as tarefas de seleção, limpeza e transformação dos dados. O objetivo da seleção de dados é reunir todos os dados relevantes para a resolução do problema definido (por exemplo, combinar dados dos sistemas corporativos da empresa com dados disponibilizados na internet). Limpeza, significa eliminar sujeira e informações irrelevantes. Transformação consiste em converter os dados de origem para

um outro formato, mais adequado para ser usado pelo algoritmo. Essas atividades são comumente referenciadas como atividades de Data Wrangling. A biblioteca pandas (Python Data Analysis Library) é um software livre, especialmente projetado para Data Wrangling, que ao longo dos últimos anos se consolidou como a biblioteca para ciência de dados mais utilizada no ambiente Python. Este livro aborda a pandas sob uma perspectiva profissional, explicando como utilizá-la para resolver problemas práticos de Data Wrangling. O livro combina a teoria com um projeto que envolve o uso da pandas como ferramenta para viabilizar a execução das atividades de seleção, estudo, limpeza e transformação de uma base de dados real que contém informações detalhadas sobre diversos países (variando desde a população e extensão de cada país até as características de suas bandeiras, entre outras informações). O projeto mostra o passo a passo para realizar o pré-processamento desta base de dados, que será então utilizada como fonte para a criação de um modelo de Machine Learning.

## Lógica de Programação com Pascal: Um curso aplicado

### Autor

José André de Moura Brito

### Editora

Ciência Moderna

### Palavras-chave

Lógica, Pascal e programação

### Área de conhecimento

Computação

### Edição

1ª (2014)

### Formato

Impresso e digital

Este livro apresenta os conceitos da lógica de programação através da linguagem Pascal. Nesse sentido, são apresentados desde os conceitos mais básicos da linguagem até os mais sofisticados, no que diz respeito à utilização das estruturas sequencial, de seleção, de repetição, de vetores, de matrizes, de funções e procedimentos e de manipulação de registros e arquivos. Definições e conceitos são introduzidos com vários exemplos e exercícios resolvidos com um grau de dificuldade variado. Além disso, ao final de cada um dos capítulos, são propostos vários exercícios classificados como fáceis, moderados e difíceis. Os programas apresentados neste livro abordam problemas das áreas de Matemática, Computação, Estatística, Física e Engenharia. O último capítulo traz um conjunto de aplicações das áreas de Matemática, Estatística e Computação.



# Equipe técnica

## **Escola Nacional de Ciências Estatísticas**

Paulo de Martino Jannuzzi

## **Núcleo de Pesquisas da ENCE**

Andrea Diniz da Silva (Gerente)

Elizabeth Belo Hypolito

Luiz Felipe d'Alberto Louzada

## **Projeto Editorial**

### **Centro de Documentação e Disseminação de Informações**

#### **Coordenação de Produção Editorial e Gráfica**

Marisa Sigolo


#### **Gerência de Editoração**

##### **Diagramação textual**

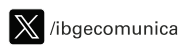
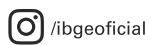
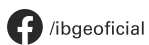
Leonardo Martins

##### **Programação visual**

Thiago Goulart da Silva



Se o assunto é **Brasil**,  
procure o **IBGE**.



**www.ibge.gov.br** 0800 721 8181