

**UMA PLATAFORMA DE EXTRAÇÃO E ANÁLISE DE DADOS DE PROGRAMAS
SOCIAIS DO GOVERNO BRASILEIRO**

Tassio Sirqueira¹

Jessica Facioli²

Lucas Vianna³

Caroline Batista⁴

RESUMO

Estudos econômicos normalmente envolvem a análise de dados complexos, fazendo uso de ferramentas estatísticas. Entretanto essas ferramentas são independentes e não auxiliam os pesquisadores na coleta e extração de amostra dos dados a serem analisados, o que leva os pesquisadores a despenderem um longo tempo e esforço físico para realizar essas atividades, mesmo que para responder questões simples de pesquisa. Com base nisso, a proposta deste trabalho é apresentar a plataforma *EconomiC Analyzer* (ECA), sendo uma ferramenta de captura, extração e análise de dados público de programas sócias do Governo Brasileiro. Atualmente a ECA busca auxiliar na extração de amostras dos dados e em algumas análises voltadas para os economistas, focando em dados dos beneficiários e de algum tipo de programa social do Governo Federal. Esse trabalho apresentará detalhes técnicos da ECA,

¹Doutorando em Informática pela PUC-Rio, mestre em Ciência da Computação pela UFJF, bacharel em Sistemas de Informação pelo CES-JF e Professor dos cursos de ADS e SI nas Faculdades Integradas Vianna Júnior. E-mail: tmsirqueira@vianna.edu.br

²Doutoranda em Economia pela UFJF, mestre em Economia Aplicada pela UFJF, bacharel em Ciências Econômicas pela UFJF. E-mail: jessica.facioli@economia.ufjf.br

³Tecnólogo em Análise e Desenvolvimento de Sistemas pelas Faculdades Integradas Vianna Júnior. E-mail: lucashkd@gmail.com

⁴Aluna do curso de Sistemas para Internet pelas Faculdades Integradas Vianna Júnior. E-mail: caroline.batista@viannasempre.com.br

além de sua arquitetura, estrutura dos dados, os desafios e alguns resultados alcançados até o momento.

PALAVRAS-CHAVE: ECONOMIA COMPUTACIONAL. PROGRAMA BOLSA FAMÍLIA. DADOS PÚBLICOS. PROGRAMAS SOCIAIS. PLATAFORMA ECA.

INTRODUÇÃO

Mensalmente o Governo Federal do Brasil disponibiliza no portal da transparência⁵ dados sobre os programas sociais, vinculados ao Cadastro Único (CadÚnico⁶), contendo diversas informações, permitindo assim que pesquisadores utilizem esses dados publicamente.

De acordo com o IBGE (IBGE, 2018), o CadÚnico foi criado em julho de 2001 com a finalidade de formar uma base de dados única para os programas sociais do Governo Federal, e atualmente já possui mais de 80 milhões de pessoas cadastradas, onde a partir do mesmo é que o poder público formula a gestão de políticas públicas para atender a população da população em situação de vulnerabilidade social.

O Governo Federal possui atualmente 22 programas sociais que fazem uso da base de dados do CadÚnico, sendo seu principal foco, de acordo com a CAIXA (CAIXA, 2018), as famílias que estão em situação de extrema pobreza e pobreza; sendo essa parcela da população atendida pelo programa Bolsa Família (PBF), conforme (BRASIL, 2018). O fato da família estar inscrita no CadÚnico não garante seu vínculo a um ou mais programas sociais, entretanto, para participar de qualquer programa social é necessário que se esteja cadastrado no CadÚnico.

⁵Portal da Transparência - <http://www.portaldatransparencia.gov.br/>

⁶Cadastro Único - <http://mds.gov.br/assuntos/cadastro-unico>

Além de o governo disponibilizar os dados das pessoas beneficiárias no Portal da Transparência, atualmente o mesmo está disponibilizando também dados de famílias e pessoas, preservando o anonimato destas famílias/pessoas dentro do CadÚnico. Com esses dados disponibilizados, surgem diversas oportunidades de pesquisas principalmente para os economistas, porém para alcançar os resultados são necessários recursos computacionais, visto que esses dados são volumosos (Big Data), dificultando a manipulação para extração de amostras e sua análise.

Segundo Sosa Escurdero (2008), a revolução do Big Data está mudando radicalmente a forma como os dados são produzidos, gerenciados, armazenados e analisados pelos economistas, tornando mais acessíveis abordagens estatísticas e computacionais comumente utilizadas em outros campos, mas pouco exploradas na economia. Além disso, Sirqueira & Dalpra (2018) abordam que os desafios do Big Data vão além, devido o volume de dados a ser analisado, a velocidade necessária para o processamento dos mesmos e em muitos casos a falta de estruturas fixas dos dados, dificultando a extração de informações úteis.

Nesse sentido, o objetivo desse trabalho é apresentar a plataforma de gestão de dados para os programas sociais do Governo Federal, denominada *EconomiC Analyzer* (ECA), de modo a permitir que pesquisadores possam capturar dados do portal da transparência do governo e extrair amostras, permitindo analisar as informações dos programas sociais de modo mais simples e fácil, dado que atualmente todos os dados são disponibilizados no Portal da Transparência em arquivos CSV (*Comma-separated values*), com periodicidade mensal para cada programa social do Governo Federal. Outro detalhe importante é com relação a estrutura do arquivo CSV, onde cada arquivo de um programa social diferente apresenta os dados de acordo com o programa social ao qual pertence, não existindo o padrão nos dados disponibilizados.

A plataforma ECA encontra-se em teste pelo programa de pós-graduação em Economia da Universidade Federal de Juiz de Fora, auxiliando nas pesquisas que envolvem dados sobre o programa Bolsa Família.

Além desta introdução, na seção 2 apresentaremos algumas informações dos programas sociais do Governo Federal, em especial do programa Bolsa Família (PBF). Na seção 3, abordaremos as características do Big Data e seu impacto na área de economia, desenvolvendo uma nova área denominada economia computacional. Na seção 4 apresentaremos detalhes dos dados disponibilizados pelo Governo Federal e os motivos que levam a adoção de sistemas gerenciadores de banco de dados (SGBD) não relacionais, também conhecidos como SGBDs *No SQL*. A seção 5 apresentamos a arquitetura e os detalhes da plataforma ECA. Já na seção 6 abordaremos alguns resultados obtidos, nossas considerações finais e os trabalhos futuros que estão sendo desenvolvidos.

1 FUNDAMENTAÇÃO TEÓRICA

No ano de 2017 o CadÚnico possuía informações de aproximadamente 26 milhões de famílias brasileiras, englobando os seguintes programas sociais: i) Tarifa Social de Energia Elétrica, ii) Programa Minha Casa Minha Vida, iii) Carteira do Idoso, iv) Aposentadoria para Pessoas de Baixa Renda, v) Telefone Popular, vi) Isenção de Pagamento de Taxa de Inscrição em Concursos Públicos, vii) Programas Cisternas, viii) Água para Todos, ix) Bolsa Verde (Programa de Apoio à Conservação Ambiental), x) Bolsa Estiagem, xi) Programa de Fomento às Atividades Produtivas Rurais/ Assistência Técnica e Extensão Rural, xii) Programa Nacional de Reforma Agrária, xiii) Programa Nacional de Crédito Fundiário, xiv) Crédito Instalação, xv) Carta Social, xvi) Serviços Assistenciais, xvii) Programa Brasil Alfabetizado, xviii) Programa de Erradicação do Trabalho Infantil (Peti), xix) Identidade Jovem, xx) ENEM, xxi) Benéfico de Prestação Continuada e xxii) Programa Bolsa Família.

Dentre todos os programas que englobam a base de dados do CadÚnico, o principal é o programa Bolsa Família (PBF), que foi criado em 2003 e integrou

diversos outros programas já existentes, tais como o Bolsa Escola (Valente, 2003), Bolsa Alimentação (SEGALL-CORRÊA, 2008), Auxílio-Gás (LAVINAS, 2004) e Fome Zero (HALL, 2006).

O Bolsa Família é destinado às famílias que vivem em situação de extrema pobreza e pobreza, onde a elegibilidade da família consiste em uma renda *per capita* mensal de até R\$85,00, ou com renda *per capita* de R\$85,01 à R\$170,00 mensais, caso tenham crianças ou adolescentes de 0 a 17 anos (BRASIL, 2016).

Como explicado por Facioli & Sirqueira (2017), o programa Bolsa Família atua com 4 linhas de benefícios, sendo: i) benefício básico (BB), ii) benefício variável (BV), iii) benefício variável vinculado ao adolescente (BVA), iv) benefício para superação de extrema pobreza (BSP). Devido esta combinação de benefícios, cada família recebe um valor diferente, dependendo da sua renda mensal por pessoa e quantidade de membros na família. Esse valor e quais famílias participam do programa são definidas pelo Governo Federal junto à CAIXA, que são os gestores do programa.

Mensalmente são beneficiadas pelo programa Bolsa Família aproximadamente 13 milhões de famílias, sendo que cada município possui um limite baseado no número de pessoas vivendo em situação de pobreza, conforme Camargo (2013). O objetivo do programa é que seja o mais focalizado possível, transferindo recursos para as pessoas elegíveis e somente a elas (KERSTENETZKY, 2009).

Alguns trabalhos como os de Soares (2003), Campello (2013) e Camargo (2013), buscam verificar a eficiência do programa Bolsa Família no atendimento das famílias em estado de vulnerabilidade social. Entretanto, esses trabalhos apontam como problema os dados existentes do programa, que formam a base do CadÚnico, visto que são autodeclarados e preenchidos via um formulário *online*. Além disso, a gestão do CadÚnico é compartilhada e descentralizada entre a União, os estados e os municípios, o que dificulta a obtenção de resultados confiáveis, uma vez que se torna necessário cruzar estes dados com outras fontes de informações.

Nas ciências econômicas, os pesquisadores da área social propõem a verificação se os programas sociais estão focalizados para pessoas que atendem os critérios de elegibilidades, dado que esses programas são dispendiosos para o governo, como exemplo, o programa Bolsa Família em 2017 teve um orçamento anual de 28,5 bilhões de reais. Dentre os fatores investigados que podem influenciar a focalização dos programas sociais, estuda-se o clientelismo; questionário fraco; fraudes; erros de medida na renda; volatilidade da renda e o fato da própria renda ser autodeclarada pelos beneficiários.

Segundo MDS⁷, os dados contidos no CadÚnico e no Portal da Transparência estão divididos entre: i) dados que identificam famílias e pessoas cadastradas, contendo nome, documentos pessoais, Número de Identificação Social (NIS), código familiar, endereço e telefone; ii) dados que caracterizam pessoas, famílias e domicílios cadastrados, apresentando características específicas do domicílio de uma família, total de membros de uma família, despesas mensais, rendimentos, escolaridade, situação no mercado de trabalho, entre outros dados; e iii) dados não identificados, sendo aqueles que podem ser agregados, conformando números absolutos e porcentagens, bem como aqueles que, quando correlacionados, não identificam uma pessoa ou família.

Tendo essa visão, a proposta da plataforma ECA é possibilitar ao pesquisador coletar estes dados advindos de diferentes fontes e contendo diferentes informações e possibilitar a extração de amostras concisas, que podem ser confiáveis e de fácil manipulação para extração de informações.

⁷ <http://mds.gov.br/assuntos/cadastro-unico/dados>

2 BIG DATA E ECONOMIA COMPUTACIONAL

Big Data corresponde não apenas a ideia de grandes volumes de dados, mas também a grandes variabilidades de formatos, necessidade de velocidade de processamento e a complexidade envolvida ao dado, requerendo tecnologias e técnicas avançadas em engenharia de dados e de software para capturar, armazenar e analisar as informações, conforme Sirqueira & Dalpra (2018).

No Big Data foram atribuídas características ao conceito consideradas os 5 V's do Big Data: i) volume, ii) variedade, iii) velocidade, iv) veracidade e v) valor. O volume refere-se à quantidade de dados gerados como um todo, onde segundo McAfee *et al.* (2012), a cada segundo circulam na internet um número de dados superiores as informações armazenadas na rede há vinte anos atrás, o que permite empresas e o governo trabalharem com várias informações de diversas fontes.

Gandomi & Haider(2015) definem que a variedade consiste na heterogeneidade estrutural do conjunto de dados, ou seja, é a utilização de vários tipos de dados, englobando os estruturados, semiestruturados e não estruturados. Para os mesmos autores, a velocidade refere-se à velocidade na qual os dados são gerados e disponibilizados para análise.

A velocidade impacta na qualidade dos dados e informações geradas, com o objetivo de trazer informações autênticas que contribuam para uma correta tomada de decisão ao final de um estudo (LOBO, 2017).

Por fim, o valor pode ser dividido em valor do custo da tecnologia e valor derivado do uso do Big Data e valor do dado. O valor do custo da tecnologia é o custo necessário para obter uma certa tecnologia de hardware ou software, e o custo do Big Data refere-se a custo de capital, como custo da eficiência operacional e melhorias de procedimentos de negócios, e o valor do dado o quanto aquele dado é importante para o estudo e o que ele representa.

Para trabalhar com o Big Data, na ciências econômicas está surgindo o campo denominado economia computacional, é definida pela *Society for*

*Computational Economics*⁸ (SCE), como a exploração da interseção da economia com a computação, visando a modelagem computacional baseada em sistemas multiagentes, econometria computacional, estatística, finanças computacionais, modelagem computacional de sistemas macroeconômicos dinâmicos, ferramentas computacionais para o design de mercados automatizados, ferramentas de programação especificamente projetadas para economia computacional e ferramentas pedagógicas para o ensino de economia da computação.

Como supramencionado, as pesquisas em economia utilizam-se normalmente de ferramentas estatísticas, contudo essas são limitadas em relação a coleta e extração de amostras, o que dificulta o trabalho dos pesquisadores.

Visando auxiliar o pesquisador, a plataforma *EconomiC Analyzer* possibilita a coleta e seleção dos dados, além da geração de análises básicas, dispensando o uso de ferramentas atualmente utilizadas pelos economistas como SPSS, R e STATA.

Até o início do desenvolvimento desta plataforma, não foram identificados trabalhos semelhantes, demonstrando ser um campo em aberto e que pode gerar resultados surpreendentes, além dos já obtidos.

3 CARACTERÍSTICAS DOS DADOS

Os dados disponibilizados no portal da transparência pelo Governo Federal seguem o formato CVS, separados por ‘;’ ou tabulação, por programa social, com periodicidade mensal. Cada arquivo possui tamanho variável, tendo em média 1,5Gb de dados, onde sua estrutura interna muda de acordo com o programa social, impactando na leitura e importação dos dados para o banco de dados da ECA.

⁸ SCE: <http://comp-econ.org/>.

Todos os dados antes da importação para a plataforma passam por uma normalização, onde são padronizados os formatos dos dados e a sua separação quando necessário, visto que alguns dados são disponibilizados multivalorados.

Na próxima seção será apresentada a plataforma *EconomiC Analyzer*, com sua arquitetura para operar sobre os dados do Governo Federal e seus detalhes técnicos.

4 APLICAÇÃO

A *EconomiC Analyzer* é uma plataforma de coleta, extração e análise de dados econômicos *open source*, que tem por objetivo auxiliar os pesquisadores por meio da centralização de seus estudos sobre programas sociais do Governo Federal em um único local.

Atualmente a plataforma conta com 3 módulos disponíveis, sendo eles: i) *ECA-Importer*; ii) *ECA-Analyzer* e; iii) *ECA-Network*, os quais serão apresentados a seguir.

4.1 Arquitetura

A plataforma ECA trabalha de forma modular, sua arquitetura pode ser vista na Figura 1.

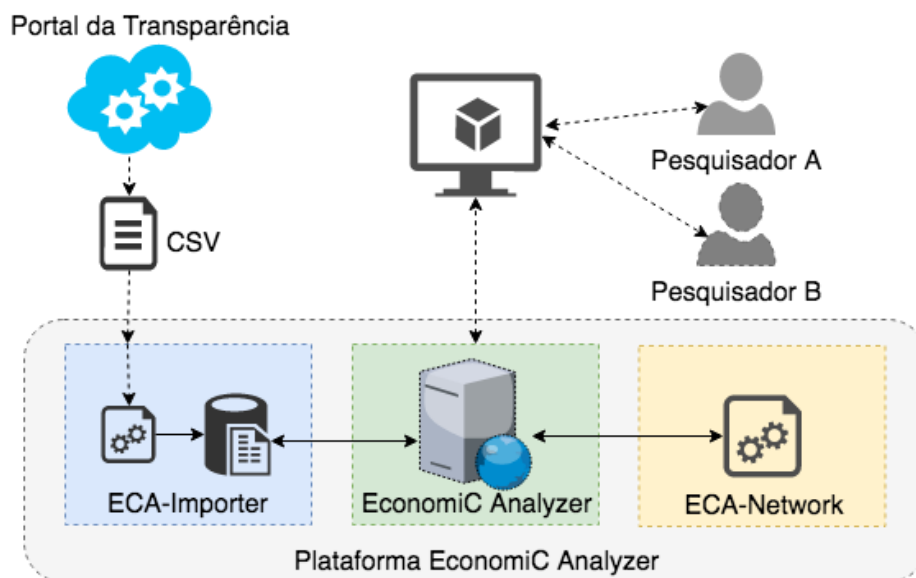


Figure 1. Arquitetura da plataforma ECA.

4.1.1 ECA-Importer

O *ECA-Importer* é responsável por coletar os arquivos CSV, disponibilizados no portal da transparência do Governo Federal, e transmiti-los para o banco de dados. Uma apresentação da sua interface pode ser vista na Figura 2. Existem duas versões disponíveis do *ECA-Importer*, uma para banco de dados relacional⁹ (MySQL) e outra para banco de dados NoSQL (MongoDB), disponível em Java¹⁰ e em Python¹¹.

No início do projeto o objetivo era desenvolver o importador apenas para banco de dados relacional, utilizando o MySQL¹² como SGBD (Sistema Gerenciador de Banco de Dados) padrão. Seguindo as regras de modelagem de banco de dados relacional, foi desenvolvido o modelo apresentado na Figura 3.

⁹*ECA-Importer* para MySQL - <https://github.com/FIVJ/ECA-Importer>

¹⁰*ECA-Importer* para MongoDB em Java - <https://github.com/FIVJ/ECA-Importer-MongoDB>

¹¹*ECA-Importer* para MongoDB em Python - <https://github.com/FIVJ/ECA-Importer-MongoDB-Python>

¹²MySQL - <https://www.mysql.com/>

Contudo, devido ao grande número de arquivos CSV, cada qual com aproximadamente 13 milhões de dados, totalizando de janeiro de 2011 a junho de 2017, mais de 1 bilhão de registros, tornou-se impossível realizar a importação desses dados respeitando as propriedades de um SGBD relacional, em um banco com as devidas formas normais. Esse problema advém do fato que para importação de cada novo registro deve-se verificar se o mesmo já não existe no banco, evitando a redundância nos dados.



Figure 2. Interface do ECA-Importer.

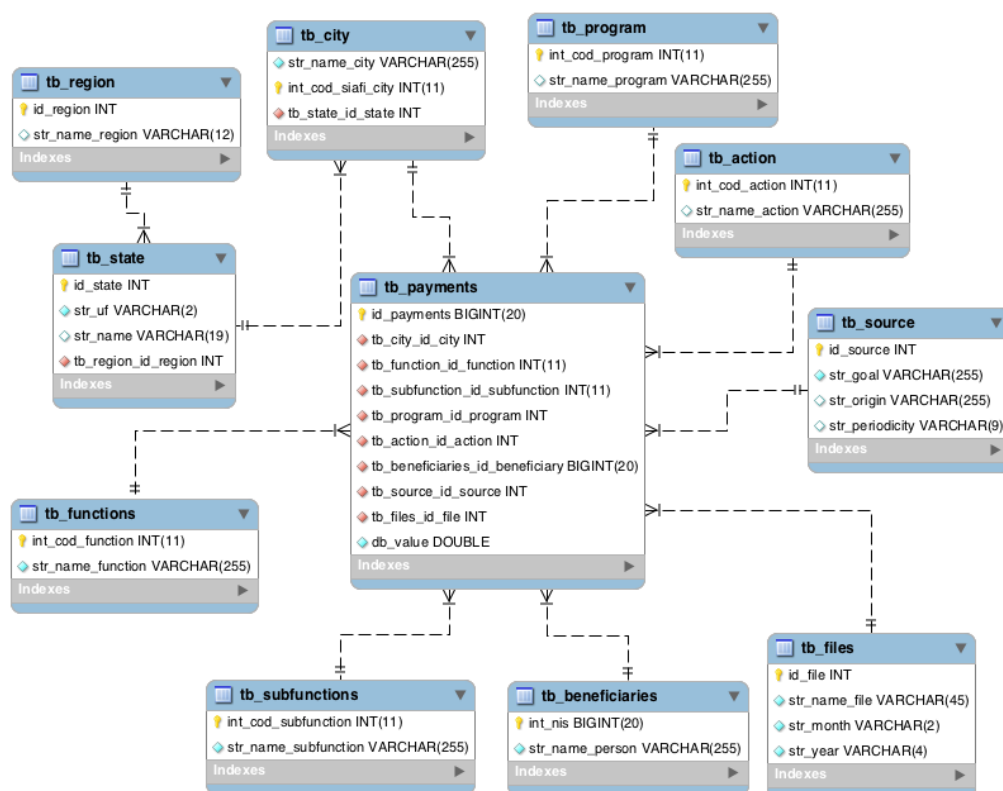


Figure 3. Diagrama Entidade-Relacionamento da plataforma ECA.

Dessa forma, optou-se pela utilização de um SGBD NoSQL. O termo “NoSQL” significa “*Not only SQL*”, onde neste pode-se aplicar o conceito SQL (*Structured Query Language*, ou Linguagem de Consulta Estruturada).

Diferente dos SGBDs relacionais, no NoSQL não há dependência de tabelas e colunas fixas, algo bastante útil nesse projeto, uma vez que os dados disponibilizados pelos programas sociais no portal da transparência são semiestruturados.

Dentre os projetos NoSQL mais notáveis até o momento, podemos citar o projeto de software livre MongoDB¹³, que nada mais é que um banco de dados orientado a documentos, o qual armazena dados em coleções de documentos

¹³MongoDB - <https://www.mongodb.com/>

semelhantes a JSON (*JavaScript Object Notation*) compostos por nomes de campos e um tipo específico de valor.

O NoSQL é uma tecnologia que vem ganhando mercado, principalmente quando trata-se de *Big Data*, seja pelo grande volume de dados, falta de estrutura nos dados, descentralidade ou estabilidade dos dados persistidos. Por esses motivos criou-se o *ECA-Importer* para MongoDB.

4.1.2 EconomiC Analyzer

O *EconomiC Analyzer* é a interface *web* disponibilizada para extração de amostras e visualização dos dados. Através dela é possível acessar todas as coleções, extrair dados e gerar gráficos ou relatórios em PDF.

O objetivo do *EconomiC Analyzer* é facilitar aos pesquisadores o acesso aos dados, de forma que eles não precisem lidar direto com o banco ou conhecer a sintaxe de manipulação dos SGBDs. Essa característica é importante devido a plataforma ser voltada para economistas e pesquisadores da área econométrica, que não possuem conhecimento em SQL, No SQL e programação.

Assim como o *ECA-Importer*, o *EconomiC Analyzer* também encontra-se disponível no GITHUB¹⁴, sendo tal versão destinada a SGBDs relacionais. A versão para SGBDs NoSQL encontra-se em desenvolvimento.

Exemplos da interface *web* podem ser vistos nas figuras 4 e 5, representando respectivamente a tela de consulta aos dados dos beneficiários dos programas sociais e a tela do *dashboard*, com informações básicas sobre os dados disponíveis no sistema.

¹⁴*Economic Analyzer* - <https://github.com/FIVJ/EconomiC-Analyzer>

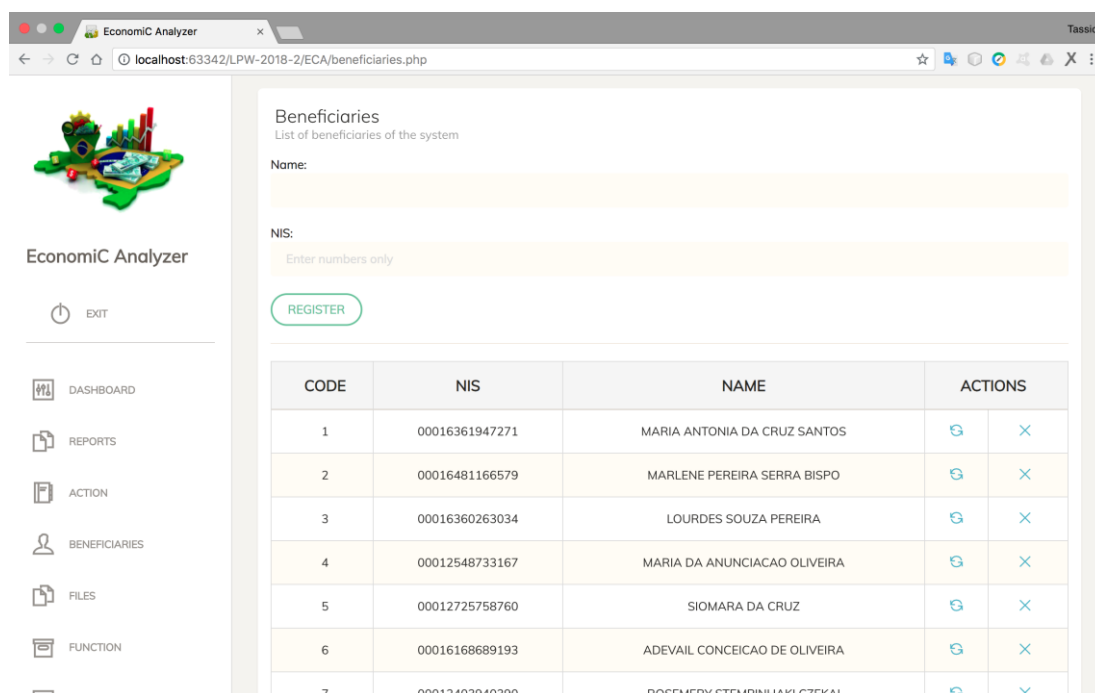


Figure 4. Dados dos beneficiários no Economiz Analyzer.

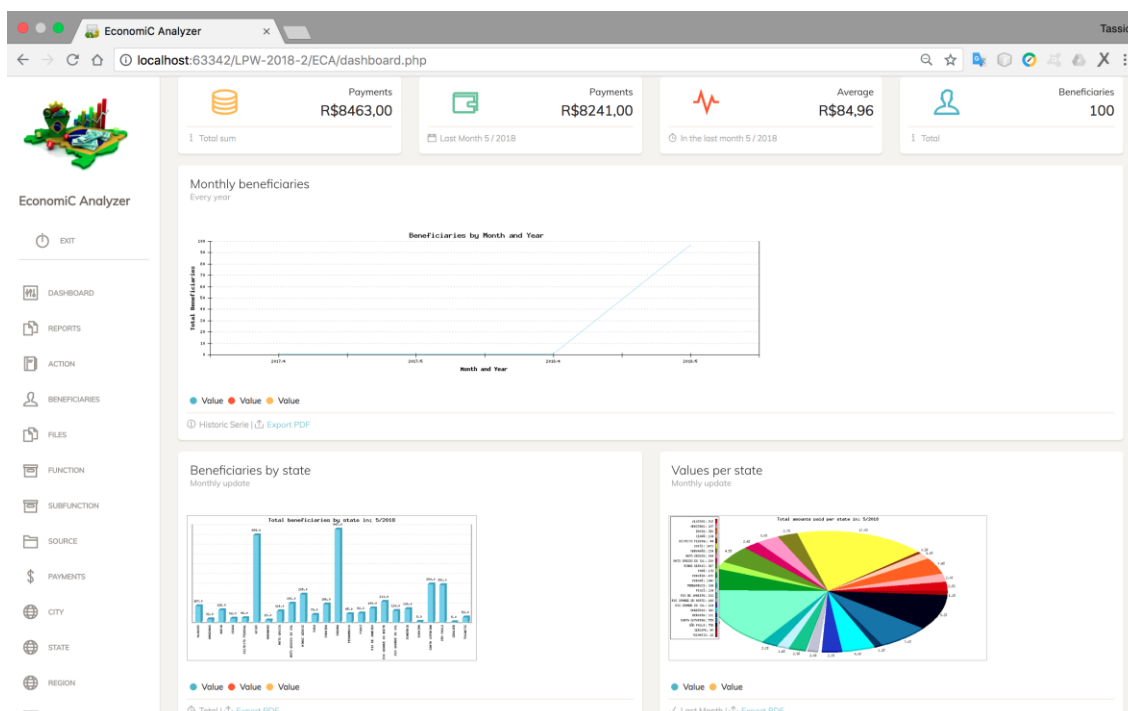


Figure 5. Dashboard do Economiz Analyzer.

4.1.3 ECA-Network

Já o *ECA-Network* é o módulo em constante desenvolvimento para identificação de redes sociais e agrupamento de indivíduos dentro dos programas sociais do Governo Federal. Através de modelos econômicos são definidas como as redes são modeladas e com base nesses modelos são geradas as saídas de texto e grafo. Exemplos produzidos por esse módulo podem ser visto nas figuras 6 e 7 respectivamente.

Na Figura 6 foi utilizado um modelo simples, onde os beneficiários dos programas sociais são agrupados em redes por cidade. Além da representação por grafo, também são apresentados os resultados via texto, facilitando a identificação dos dados. Neste exemplo temos todas os beneficiários separados por cor, onde cada cor representa uma cidade do Brasil. Contudo, essa é uma representação densa, de difícil entendimento.

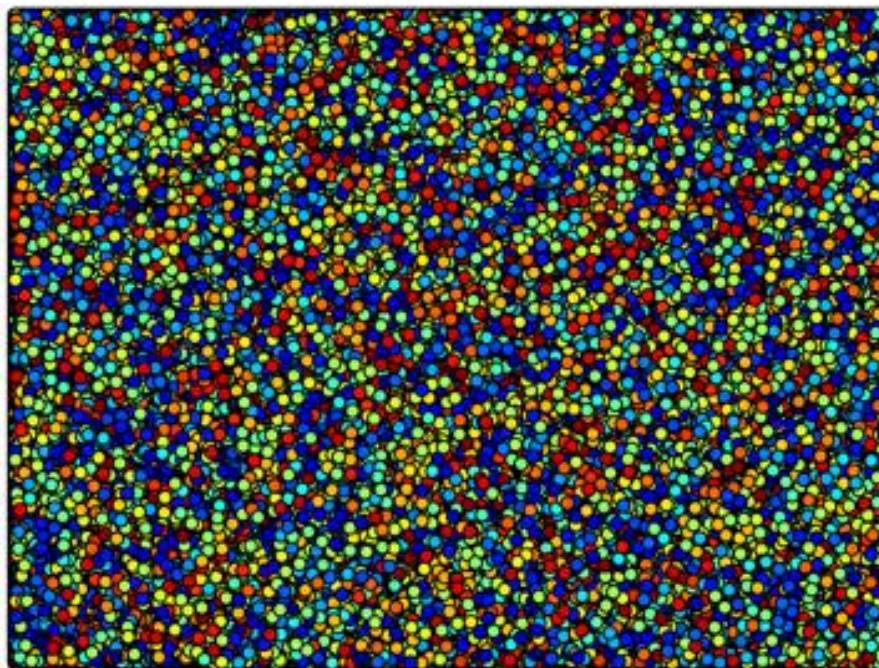


Figure 6. Redes de beneficiários de programas sociais separados por cidades.

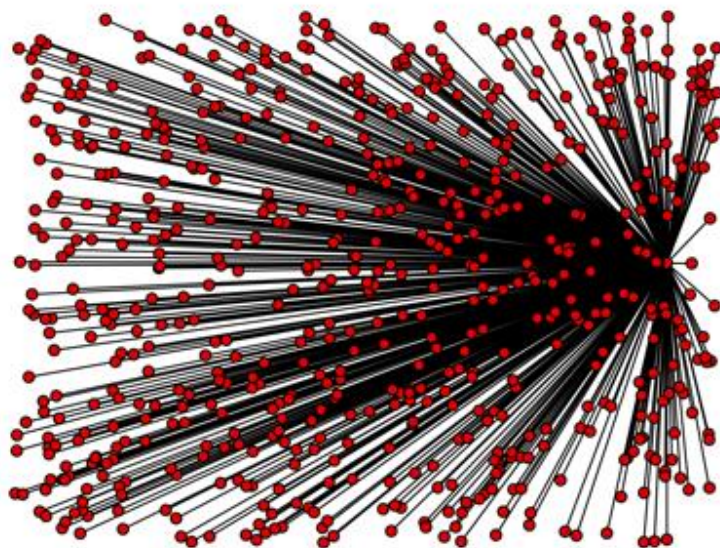


Figure 7. Rede de beneficiários do PBF considerando uma raça e uma cidade específica.

Já na Figura 7 temos a mesma representação do modelo, onde selecionou-se apenas os beneficiários residentes em uma mesma cidade, sendo essa representação mais legível e também conta com a saída em texto da lista dos beneficiários que resultaram neste grafo.

4.2 Banco de Dados

Conforme supramencionado, a *Economic Analyzer* utiliza o MongoDB como SGBD padrão, e sua base está estruturada em 5 coleções, sendo elas: i) Pagamentos PBF; ii) Saques PBF; iii) Pescador Artesanal; iv) Outros Programas e v) Beneficiários. Dessa maneira, cada coleção possui todos os dados disponibilizados sobre os programas sociais, por meio do portal da transparência do governo.

Cada documento do MongoDB contém todos os dados do registro e todas as coleções possuem índices para otimização das buscas, que também possibilitam agregação¹⁵ ou MapReduce¹⁶.

¹⁵ Agregação no MongoDB - <https://docs.mongodb.com/manual/aggregation/>

Esses dados contemplam o período de janeiro de 2011 a dezembro de 2017, sendo as coleções Pagamentos PBF e Saques PBF as maiores coleções. Só a coleção Pagamentos PBF no ano de 2011 possui 156.554.819 registros, mantendo-se em torno desse valor nos anos seguintes.

CONCLUSÃO

A plataforma *EconomiC Analyzer* começou a ser desenvolvida durante uma iniciação científica, no Instituto Vianna Júnior, e teve como objetivo auxiliar estudantes de pós-graduação em Economia Aplicada, da Universidade Federal de Juiz de Fora, a realizarem seus estudos sobre *social network* no programa Bolsa Família, e como as formações de redes no programa podem influenciar a focalização do mesmo.

O desenvolvimento da plataforma ECA continua ativa e ganhando novas funcionalidade desde 2017. A cada semestre a plataforma agrega novos recursos, novos modelos econômicos e amplia seus horizontes, até então voltados ao campo de estudo econométrico.

Mesmo sendo uma solução relativamente nova, já apresenta grande potencial para a área de economia computacional, visto seus ganhos frente as dificuldades da área de economia em trabalhar com grandes arquivos de dados semiestruturados.

Entre os objetivos futuros, além dos já supracitados, buscamos avançar na análise dos dados disponíveis e na geração de relatórios gráficos e em PDF, extração dos dados em formatos voltados para ferramentas estatísticas e de mineração de dados.

¹⁶ MapReduce no MongoDB - <https://docs.mongodb.com/manual/core/map-reduce/>

AGRADECIMENTOS

Os autores agradecem todo apoio e incentivo a pesquisa dados pelo Instituto Vianna Junior e pela Universidade Federal de Juiz de Fora.

REFERÊNCIAS

BRASIL. Ministério do Desenvolvimento Social e Combate à Fome (MDS). 2018. Programas que atendem aos inscritos. Disponível em: <<http://mds.gov.br/assuntos/cadastro-unico/o-que-e-e-para-que-serve/programas-e-beneficios>>. Acessado em 31 de jan. de 18.

BRASIL. Ministério do Desenvolvimento Social e Combate à Fome (MDS). 2016. Bolsa Família. Disponível em: <<http://www.mds.gov.br/bolsafamilia>>. Acesso em: 14 de fev. 2018.

CAIXA. 2018. Programas e benefícios sociais utilizam o Cadastro Único. Disponível em: <<http://www.caixa.gov.br/cadastros/cadastro-unico/Paginas/default.aspx>>. Acessado em 31 de jan. de 18.

CAMARGO, Camila Fracaro et al. Perfil socioeconômico dos beneficiários do programa Bolsa Família. **IPEA**. Programa Bolsa Família: uma década de inclusão e cidadania. Brasília: Governo Federal, 2013.

CAMPELLO, Tereza. Uma década derrubando mitos e superando expectativas. **BOLSA FAMÍLIA**, p. 15, 2013.

FACIROLI, Jéssica; SIRQUEIRA, Tássio Ferenzini Martins. Análise espacial da taxa de cobertura do Programa Bolsa Família para os municípios brasileiros. **Desenvolvimento Regional em debate: DRd**, v. 7, n. 2, p. 118-137, 2017.

GANDOMI, Amir; HAIDER, Murtaza. Beyond the hype: Big data concepts, methods, and analytics. **International Journal of Information Management**, v. 35, n. 2, p. 137-144, 2015.

HALL, Anthony. From Fome Zero to Bolsa Família: social policies and poverty alleviation under Lula. **Journal of Latin American Studies**, v. 38, n. 4, p. 689-709, 2006.

IBGE. 2018. Cadastro Único dos Programas Sociais – **CadÚnico**. Disponível em: <<https://ces.ibge.gov.br/base-de-dados/metadados/mds/cadastro-unico-dos-programas-sociais-cadunico.html>>. Acessado em 31 de jan. de 18.

KERSTENETZKY, Celia Lessa. Redistribuição e desenvolvimento? A economia política do programa bolsa família. **Dados-Revista de Ciências Sociais**, v. 52, n. 1, 2009.

LAVINAS, Lena. Universalizando direitos. **Observatório da Cidadania**, v. 59, p. 67-74, 2004.

LOBO, Rodrigo Maciel. O uso de grande volume e variedade de informações-**Big Data**. 2017.

MCAFEE, Andrew et al. Big data: the management revolution. **Harvard business review**, v. 90, n. 10, p. 60-68, 2012.

SEGALL-CORRÊA, Ana Maria et al. Transferência de renda e segurança alimentar no Brasil: análise dos dados nacionais. **Revista de Nutrição**, 2008.

SIRQUEIRA, Tassio; DALPRA, Humberto. NoSQL e a Importância da Engenharia de Software e da Engenharia de Dados para o Big Data. **37º Jornada de Atualização da Informática (JAI)**. Congresso da Sociedade Brasileira de Computação (CSBC). Cap. 2. 2018.

SOARES, Sergei et al. Os impactos do benefício do Programa Bolsa Família sobre a desigualdade e a pobreza. **Bolsa família**, v. 2010, p. 27-52, 2003.

SOSA ESCUDERO, Walter. Big data: desafíos para la docencia en Ciencias Económicas. **Econo**, v. 8, 2017.

VALENTE, Ana Lucia. Espaço Aberto O Programa Nacional de Bolsa Escola e as ações afirmativas no campo educacional. **Revista Brasileira de Educação**, n. 24, p. 0, 2003.