

**Jades Fernando Hammes**

***DASHBOARD* E UM MODELO DE ANÁLISE PREDITIVA PARA  
DOENÇAS CEREBROVASCULARES NA ATENÇÃO  
PRIMÁRIA EM SAÚDE**

Dissertação submetida ao Programa de Mestrado Profissional em Informática em Saúde da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestre em Informática em Saúde.

**Área de Concentração:** Tecnologia de Informação e comunicação em Saúde.

**Orientadora:** Prof<sup>ª</sup> Dr<sup>ª</sup> Grace Teresinha Marcon Dal Sasso

**Florianópolis  
2018**

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Hammes, Jades Fernando

DASHBOARD E UM MODELO DE ANÁLISE PREDITIVA PARA  
DOENÇAS CEREBROVASCULARES NA ATENÇÃO PRIMÁRIA EM SAÚDE /  
Jades Fernando Hammes ; orientadora, Prof<sup>a</sup> Dr<sup>a</sup> Grace  
Teresinha Marcon Dal Sasso, 2018.  
105 p.

Dissertação (mestrado profissional) - Universidade  
Federal de Santa Catarina, Centro de Ciências da Saúde,  
Programa de Pós-Graduação em Informática em Saúde,  
Florianópolis, 2018.

Inclui referências.

1. Informática em Saúde. 2. Atenção Primária. 3. e-SUS.  
4. Machine Learning. 5. Doenças Cerebrovasculares. ? I. Dal  
Sasso, Prof<sup>a</sup> Dr<sup>a</sup> Grace Teresinha Marcon. II. Universidade  
Federal de Santa Catarina. Programa de Pós-Graduação em  
Informática em Saúde. III. Título.


Jades Fernando Hammes

**DASHBOARD E UM MODELO DE ANÁLISE PREDITIVA  
PARA DOENÇAS CEREBROVASCULARES NA ATENÇÃO  
PRIMÁRIA EM SAÚDE**

Esta Dissertação foi julgada adequada para obtenção do título de:

**MESTRE EM INFORMÁTICA EM SAÚDE**

e aprovada em 26 de Setembro de 2018 em sua forma final pelo Programa de Pós-Graduação Mestrado Profissional em Informática em Saúde.




Prof.ª Dr.ª Grace T. Marcon Dal Sasso  
Coordenadora do Curso

**Banca Examinadora:**

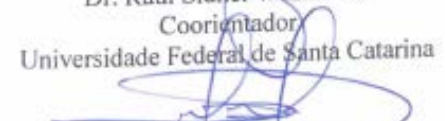


Prof.ª Dr.ª Grace T. Marcon Dal Sasso  
Orientador  
Universidade Federal de Santa Catarina



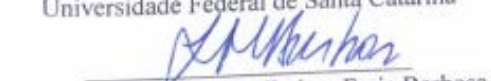
Dr. Raul Sidnei Wazlawick  
Coorientador

Universidade Federal de Santa Catarina



Prof. Dr. Eduardo Monguilhott Dalmarco  
Membro

Universidade Federal de Santa Catarina



Prof.ª Dr.ª Sayonara de Fatima Faria Barbosa  
Membro

Universidade Federal de Santa Catarina



Dedico este trabalho ao meu  
sobrinho Pedro Hammes do  
Amaral para que busque sempre  
o conhecimento e sabedoria.



## **AGRADECIMENTOS**

Agradeço a Deus, pelo dom da vida.

Agradeço meus pais, Aldori e Helizabeth, pelo amor incondicional e por sempre me incentivarem a estudar.

Agradeço a minha irmã Hyala por acreditar no meu potencial.

Agradeço aos professores do programa de mestrado pelo conhecimento compartilhado.

Agradeço a todos os colaboradores do Bridge pela compreensão durante o desenvolvimento deste trabalho.





HAMMES, Jades Fernando. **Dashboard e um modelo de análise preditiva para doenças cerebrovasculares na atenção primária em saúde**. 2018. 105f. Dissertação (Mestrado Profissional em Informática em Saúde) – Programa de Mestrado Profissional em Informática em Saúde, Universidade Federal de Santa Catarina, Florianópolis, 2018.

**Orientadora:** Prof<sup>a</sup> Dr<sup>a</sup> Grace Teresinha Marcon Dal Sasso.

## RESUMO

As doenças cerebrovasculares e hipertensivas estão entre as principais causas de óbito no mundo. Somente em 2010, foram responsáveis por quase 35 milhões de óbitos, 30% a mais do ocorrido em 1990. Este comportamento global se observa também no Brasil. Entender o comportamento de uma doença a longo prazo é fundamental para aprimorar seu tratamento e reduzir o número de óbitos por ela causados. Comparando o Brasil com os Estados Unidos, a doença cerebrovascular como principal causa de morte permanece em segundo lugar desde o ano de 2005 comparando até o ano de 2016, com um aumento de 12,1%. O objetivo deste trabalho é desenvolver um *dashboard* com indicadores e mapas georeferenciados de doenças cerebrovasculares e, principalmente, testar um modelo de análise preditiva que determine probabilisticamente a chance de um cidadão desenvolver estas doenças ponderando seus fatores de riscos em saúde. Foi utilizado o método de regressão logística otimizada por um algoritmo de otimização (Gradiente Descendente Estocástico). Após o aprendizado em dados históricos, foi possível prever a chance de um cidadão se deparar com esta patologia. O resultado foi apresentado em forma de relatório da ferramenta de dashboard desenvolvida também neste trabalho.

**Palavras-chave:** Atenção Primária. e-SUS AB. Análise Preditiva. Machine Learning. Regressão Logística. Mapa Georreferenciado. Doenças Cerebrovasculares.



HAMMES, Jades Fernando. **Dashboard and a model of predictive analysis for cerebrovascular diseases in primary health care.** 2018. 105f. Dissertation (Professional Master in Health Informatics) – Professional Master's Program in Health Informatics, Federal University of Santa Catarina, Florianópolis, 2018.

**Supervisor:** Prof<sup>a</sup> Dr<sup>a</sup> Grace Teresinha Marcon Dal Sasso.

### **ABSTRACT**

Cerebrovascular and hypertensive diseases are among the main causes of death in the world. In 2010 only, it was responsible for almost 35 million deaths, 30% more of that occurred in 1990. This global behaviour is also observed in Brazil. Understand the behaviour of a disease in the long-term is fundamental to improve its treatment and cut the number of deaths it causes. Comparing Brazil with the United States, the cerebrovascular disease as the leading cause of death remains in second place since the compared to the year of 2016, with an increase of 12.1%. The purpose of this work is to develop a dashboard with georeferenced indicators and maps, to test a model of predictive analysis probabilistically the chance of a citizen developing these diseases by pondering their factors of health risks. The optimized logistic regression method was used for an optimization algorithm (Stochastic Descending Gradient). After learning in data it was possible to predict the chance of a citizen facing this pathology. Zero results were presented in the form of a report of the developed dashboard tool also in this work.

**Keywords:** Primary Attention. e-SUS AB. Predictive Analysis. Machine Learning. Logistic Regression. Geometric map. Cerebrovascular Diseases.



## LISTA DE FIGURAS

<b>Figura 1</b> - Variação percentual das principais causas de morte entre 2005 e 2016 no Brasil.....	26
<b>Figura 2</b> - Variação percentual das principais causas de morte entre 2005 e 2016 nos EUA. ....	27
<b>Figura 3</b> - Fluxo de Processo.....	45
<b>Figura 4</b> - Tabela do banco de dados com registros fictícios. ....	58
<b>Figura 5</b> - Colunas de Dados para a Análise Preditiva. ....	59
<b>Figura 6</b> - Modelagem do Banco de Dados da Ferramenta DDCV.....	62
<b>Figura 7</b> - View da Ficha de Atendimento Individual.....	63
<b>Figura 8</b> - View da Ficha de Cadastro Domiciliar e Territorial .....	64
<b>Figura 9</b> - View da Ficha de Cadastro Individual .....	64
<b>Figura 10</b> - View Ficha de Visita Domiciliar e Territorial .....	65
<b>Figura 11</b> - Preparação para receber o arquivo de entrada .....	72
<b>Figura 12</b> - Comando para leitura do arquivo CSV no Python .....	72
<b>Figura 13</b> - Código em Python para normalização da idade .....	73
<b>Figura 14</b> - Conjuntos Fuzzy de Idade (normalização) .....	73
<b>Figura 15</b> - - Mapa de calor da correlação no aprendizado .....	75
<b>Figura 16</b> - Mapa de calor da correlação no teste.....	75
<b>Figura 17</b> - Execução do algoritmo para obter a probabilidade em todo o conjunto de dados.....	76
<b>Figura 18</b> - DDCV - Risco de Doenças Cerebrovasculares (análise_preditiva).....	77



## LISTA DE TELAS

<b>Tela 1</b> – Login.....	48
<b>Tela 2</b> – Menu Principal.....	49
<b>Tela 3</b> - Distribuição de Doenças Cerebrovasculares – Dashboard .....	50
<b>Tela 4</b> - Distribuição de Doenças Cerebrovasculares - Mapa de Calor.	51
<b>Tela 5</b> - Distribuição de Doenças Cerebrovasculares - Mapa Georreferenciado.....	52
<b>Tela 6</b> - Cidadãos com Doenças Cerebrovasculares .....	53
<b>Tela 7</b> - Doenças Cerebrovasculares por CID10.....	54
<b>Tela 8</b> - Doenças Cerebrovasculares por CID10 – Detalhe Fonte: elaborado pelo autor .....	55
<b>Tela 9</b> - Risco de Doenças Cerebrovasculares (Análise Preditiva).....	56
<b>Tela 10</b> - DDCV - Distribuição de Doenças Cerebrovasculares – <i>Dashboard</i> .....	66
<b>Tela 11</b> - DDCV - Distribuição de Doenças Cerebrovasculares - Mapa de Calor.....	67
<b>Tela 12</b> - DDCV - Distribuição de Doenças Cerebrovasculares - Mapa Georreferenciado.....	68
<b>Tela 13</b> - DDCV – Cidadãos com Doenças Cerebrovasculares.....	69
<b>Tela 14</b> - DDCV - Doenças Cerebrovasculares por CID10 – Listagem	70
<b>Tela 15</b> - DDCV - Doenças Cerebrovasculares por CID10 – Detalhe..	71





## **LISTA DE TABELAS**

<b>Tabela 1</b> - Grupo 12 da Lista Brasileira de Interações por Condições Sensíveis à Atenção Primária.....	34
--	----



## LISTA DE ABREVIATURAS E SIGLAS

<b>AP</b>	Atenção Primária
<b>API</b>	Application Programming Interface
<b>APS</b>	Atenção Primária à Saúde
<b>AVC</b>	Acidente Vascular Cerebral
<b>BD</b>	Banco de Dados
<b>CDS</b>	Coleta de Dados Simplificada
<b>CID10</b>	Classificação Internacional de Doenças
<b>CSAP</b>	Casos Sensíveis a Atenção Primária
<b>CSV</b>	Comma-separated values
<b>DAB</b>	Departamento de Atenção Básica
<b>DAC</b>	Doenças do Aparelho Circulatório
<b>DATASUS</b>	Departamento de Informática do Sistema Único de Saúde
<b>DCBV</b>	Doenças Cerebrovasculares
<b>DDCV</b>	Dashboard de Doenças Cerebrovasculares
<b>DIC</b>	Doenças Isquêmicas do Coração
<b>e- Gov</b>	Governo Eletrônico
<b>e-SUS AB</b>	e-SUS Atenção Básica
<b>ESF</b>	Estratégias de Saúde da Família
<b>IHME</b>	Institute for Health Metrics and Evaluation
<b>MS</b>	Ministério da Saúde
<b>PEC</b>	Prontuário Eletrônico do Cidadão
<b>PNAB</b>	Política Nacional da Atenção Básica
<b>SIG</b>	Sistemas de Informações Geográficas
<b>SISAB</b>	Sistema de Informação em Saúde para a Atenção Básica
<b>SUS</b>	Sistema Único de Saúde



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>23</b>
<b>2</b>	<b>OBJETIVOS.....</b>	<b>31</b>
2.1	OBJETIVOS ESPECÍFICOS .....	31
<b>3</b>	<b>REVISÃO DE LITERATURA .....</b>	<b>33</b>
3.1	CONDIÇÕES SENSÍVEIS NA ATENÇÃO PRIMÁRIA EM SAÚDE .....	33
3.2	DASHBOARD.....	36
3.3	SISTEMA GEOREFERENCIADO .....	36
<b>3.4.1</b>	<b>Regressão Logística .....</b>	<b>40</b>
<b>4</b>	<b>METODOLOGIA DO ESTUDO.....</b>	<b>43</b>
4.1	NATUREZA DO ESTUDO .....	43
4.2	TIPO DO ESTUDO.....	43
4.3	DESENVOLVIMENTO DO ESTUDO .....	43
<b>4.3.1</b>	<b>Fases do Desenvolvimento.....</b>	<b>46</b>
<b>4.3.2</b>	<b>Fichas CDS do e-SUS AB.....</b>	<b>47</b>
<b>4.3.3</b>	<b>Prototipação</b>	<b>47</b>
<b>4.3.4</b>	<b>Banco de Dados.....</b>	<b>57</b>
<b>4.3.5</b>	<b>Dados do Sistema.....</b>	<b>57</b>
<b>4.3.6</b>	<b>Análise Preditiva.....</b>	<b>58</b>
4.4	ASPECTOS ÉTICOS DO ESTUDO.....	60
<b>5</b>	<b>RESULTADOS.....</b>	<b>61</b>
5.1	ESTRUTURAR UM BANCO DE DADOS PARA ARMAZENAR AS INFORMAÇÕES RELATIVAS ÀS DOENÇAS CEREBROVASCULARES DE FORMA CONSOLIDADA PARA A ATENÇÃO PRIMÁRIA EM SAÚDE.....	61
5.2	REALIZAR MIGRAÇÃO DOS DADOS DO E-SUS AB PARA O NOVO BANCO DE DADOS.....	63
5.3	APRESENTAR OS DADOS E AS INFORMAÇÕES CONSOLIDADAS NUM DASHBOARD .....	65
5.4	APRESENTAR OS DADOS E AS INFORMAÇÕES ANALISADAS EM UM MAPA A PARTIR DA TÉCNICA DE GEOPROCESSAMENTO .....	67

5.5 APRESENTAR OS CIDADÃOS QUE JÁ TIVERAM OU TEM DOENÇAS CEREBROVASCULARES .....	68
5.6 APRESENTAR OS INDICADORES DAS DOENÇAS CEREBROVASCULARES .....	69
5.7 APRESENTAR UMA ANÁLISE PREDITIVA COM CIDADÃOS QUE POSSUEM RISCOS DE DESENVOLVEREM DOENÇAS CEREBROVASCULARES AOS PROFISSIONAIS DE SAÚDE .....	71
<b>6 CONSIDERAÇÕES FINAIS .....</b>	<b>79</b>
<b>REFERÊNCIAS.....</b>	<b>81</b>
<b>APÊNDICES .....</b>	<b>87</b>
<b>APÊNDICE A – FICHAS DE CADASTROS.....</b>	<b>89</b>
<b>ANEXOS.....</b>	<b>93</b>
<b>ANEXO A .....</b>	<b>95</b>
<b>ANEXO B .....</b>	<b>101</b>
<b>ANEXO C .....</b>	<b>105</b>

## 1 INTRODUÇÃO

Segundo a Organização Mundial da Saúde, a aplicação das Tecnologias de Informação e Comunicação à Saúde, denomina-se e-Saúde. O Ministério da Saúde (MS) está trabalhando nas diretrizes em relação a e-Saúde, alinhada com os princípios do Sistema Único de Saúde (SUS) e as políticas de Governo Eletrônico (e-Gov). O documento Estratégia e-Saúde para o Brasil propõe uma visão até 2020 e tem como objetivo aumentar a qualidade e ampliar o acesso à atenção à saúde, além de consolidar a e-Saúde como Política de Estado relevante para o SUS (BRASIL, 2014).

Dentre os benefícios esperados, trazidos pela adoção do e-Saúde, estão o acesso à informação pelo cidadão sobre a sua saúde e aos dados agregados relativos à comunidade em que vive, bem como sobre as doenças que os afetam; apoio ao profissional de saúde no planejamento do cuidado ao paciente, diagnóstico, tratamento ou prognóstico; permitir ao cidadão acesso às informações de gastos e resultados de saúde; dar acesso online ao gestor de saúde às informações necessárias para a gestão clínica, financeira e administrativa; monitoração e avaliação de análises de tendências dos processos e resultados de saúde; e detecção oportuna de eventos de saúde pública por meio de sistemas de informação que permitem a análise sistemática de dados clínicos, epidemiológicos e de fluxo de paciente (BRASIL, 2014).

O e-SUS Atenção Básica (e-SUS AB) é uma estratégia do DAB (Departamento de Atenção Básica) para reestruturar as informações da Atenção Primária (AP) em nível nacional. Esta ação está alinhada com a proposta mais geral de reestruturação dos Sistemas de Informação em Saúde do MS, entendendo que a qualificação da gestão da informação é fundamental para ampliar a qualidade no atendimento à população. A estratégia e-SUS AB, faz referência ao processo de informatização qualificada do SUS em busca de um SUS eletrônico (BRASIL, 2016a).

Em dezembro de 2017, a estratégia e-SUS AB estava implantado em quase 100% dos municípios do Brasil. Os municípios podem adotar a estratégia e-SUS AB pelo uso dos sistemas de Prontuário Eletrônico do Cidadão (PEC), pelo sistema de Coleta de Dados Simplificada (CDS) ou pela integração de outros sistemas, sejam próprios ou terceiros. Todas essas instalações enviam dados de saúde para um repositório de dados nacional chamado de Centralizador Nacional. O uso do prontuário eletrônico nesta data era de 2.060 municípios. O plano nacional é

instalar prontuário eletrônico em 100% dos municípios brasileiros. Só em 2017 foram registrados 105 milhões de procedimentos, 72 milhões de consultas médicas, 48 milhões de consultas odontológicas e 232 milhões de visitas domiciliares (BRASIL, 2016b).

Para alcançar esses números, o apoio dos profissionais de saúde foi fundamental. Segundo dados da pesquisa TIC SAÚDE 2016, no setor de Saúde, realizada pelo CETIC.BR (Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação), médicos e enfermeiros possuem uma visão positiva no que se refere ao impacto da implantação de sistemas eletrônicos em seus estabelecimentos (TIC SAÚDE 2016, 2016).

Para criar uma interface desse BD (Banco de Dados) aos gestores federais, estaduais e municipais, foi instituído pela Portaria GM/MS nº 1.412, de 10 de julho de 2013, o Sistema de Informação em Saúde para a Atenção Básica (SISAB), passando a ser o sistema de informação da Atenção Básica vigente para fins de financiamento e de adesão aos programas e estratégias da Política Nacional da Atenção Básica (PNAB) (BRASIL, 2013).

Porém, o SISAB apresenta somente os dados estruturados e de forma estática, que são apresentados na forma de tabelas e gráficos, focando, até o momento, em dados de produção ambulatorial para fins de repasse de incentivos financeiros às ESFs (Estratégias de Saúde da Família). O DAB não possui uma ferramenta para preparação dos dados de forma a ter uma fonte de informação que possa dar suporte permanente e dinâmico ao processo gerencial de apoio a tomada de decisão pelos gestores federais, estaduais e municipais e pelos coordenadores de equipes e profissionais de saúde.

As organizações de saúde em todo mundo são capazes de gerar e coletar uma grande quantidade de dados. Assim, em todo o mundo as organizações enfrentam dilemas em relação ao crescimento do volume de dados em saúde. Mas será que a capacidade de tomar decisões também está crescendo?

De acordo com o Instituto SAS (2016) - *Statistical Analysis System na North Carolina State University* - tomar decisões baseadas em fatos não é dependente da quantidade de dados que se possui. Trabalhar com tantos dados pode ser assustador. Neste sentido, o sucesso na tomada de decisão dependerá da rapidez com que se descobre as informações a partir desses dados e se utiliza esses *insights* para gerar melhores práticas na organização.



O Instituto SAS (2016) ressalta ainda que desde a visualização em painéis que mostram métricas e indicadores importantes para alcançar objetivos e metas traçadas de forma visual, facilitando a compreensão das informações (*dashboards*) em saúde, até a análise preditiva, a mineração de dados, o aprendizado de máquinas e o gerenciamento de decisões são fundamentais para avaliar o que acontecerá no futuro, prever algum comportamento desses dados e informações, bem como, usar isso para transformar dados em ideias proativas, com ações que sejam utilizadas em processos operacionais, de interesse deste estudo.

Neste sentido, as empresas e os governos estão cheios de informação em seus computadores e não sabem o que fazer com ela. A importância está no conhecimento que é extraído destas informações pela prática da análise. A análise permite conhecer a verdade dos fatos, o que melhora a tomada de decisões, para a obtenção de resultados. O conhecimento por si não cria valor. O valor é criado quando o conhecimento é utilizado na construção de planos de ação cuja execução é garantida (FALCONI, 2009).

Com esse *manancial* de informações clínicas, epidemiológicas e sanitárias, o DAB e o MS poderão usufruir de indicadores sólidos em busca da melhoria da atenção à saúde, o apoio a inclusão de novos programas de saúde e até para previsões de tendências relacionadas à saúde pública. É o momento para buscar implementações de inovações tecnológicas para assistência aos serviços de saúde e para os cidadãos usuários do SUS.

Neste sentido, a PORTARIA Nº 221, DE 17 DE ABRIL DE 2008, estabelece a Lista Brasileira de Internações por Condições Sensíveis à Atenção Primária, contendo 19 grupos de doenças, considerando aspectos nacionais para refletir as diversidades das condições de saúde e doenças no Brasil. Essa lista norteia indicadores da atividade hospitalar para serem utilizados como medida indireta do funcionamento da atenção primária e da estratégia de saúde da família (BRASIL, 2008).

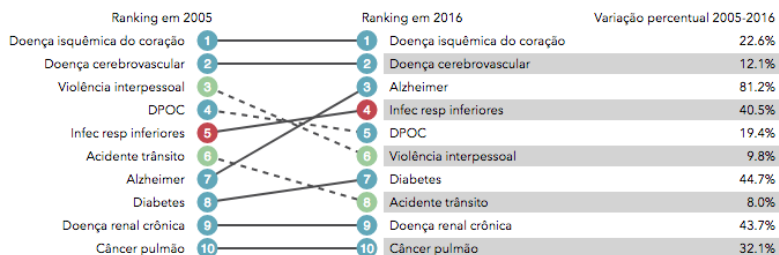
Dentro dos grupos de doenças listados nessa portaria, destacam-se as doenças cerebrovasculares, pois segundo Villela, Klein e Oliveira (2016, p.26), “As doenças cerebrovasculares e hipertensivas estão entre as principais causas de óbito no mundo”. Somente em 2010, foram responsáveis por quase 35 milhões de óbitos, 30% a mais do ocorrido em 1990. Dentre elas, as doenças isquêmicas do coração (DIC) e

doenças cerebrovasculares (DCBV), principais representantes das doenças do aparelho circulatório (DAC), foram responsáveis em 2012 por 7,4 e 6,7 milhões de mortes, respectivamente. Este comportamento global se observa também no Brasil, onde de acordo com o Departamento de Informática do Sistema Único de Saúde (DATASUS), 28,2% de todos os óbitos em 2012 tiveram sua causa básica classificada como devida a DAC (333.295). Desses óbitos por DAC, 31,1% ocorreram por DIC, 30,1% por DCBV e 13,6% por doenças hipertensivas (DHIP) (VILLELA; KLEIN; OLIVEIRA, 2016).

Entender o comportamento de uma doença a longo prazo é fundamental para aprimorar seu tratamento e reduzir o número de óbitos por ela causados. As tendências na mortalidade por DCBV foram avaliadas na Europa e em outras regiões do mundo. Porém, há poucas informações disponíveis no Brasil sobre a distribuição, comportamento e tendências de mortalidade das DCBV. Além disso, os dados publicados a respeito dessas doenças apenas apresentam a sua evolução em alguns estados ou se referem a intervalos de tempo inferiores a 10 anos (VILLELA; KLEIN; OLIVEIRA, 2016).

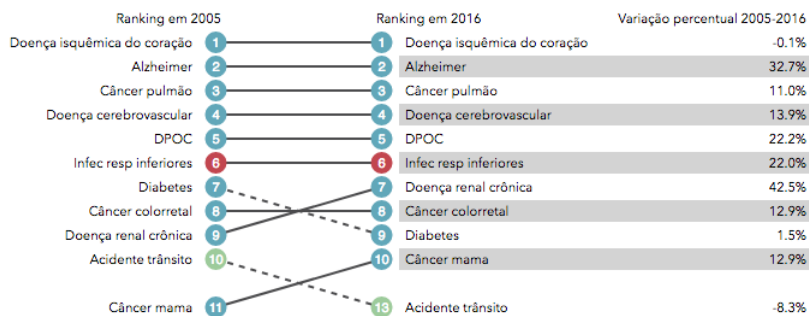
Comparando o Brasil com os Estados Unidos, a doença cerebrovascular como principal causa de morte permanece em segundo lugar desde o ano de 2005 comparando até o ano de 2016, com um aumento de 12,1%, conforme a Figura 1. Nos EUA a doença cerebrovascular como causa das maiorias das mortes é a quarta da lista, conforme a Figura 2 (IHME, 2016).

**Figura 1** - Variação percentual das principais causas de morte entre 2005 e 2016 no Brasil.



Fonte: IHME, 2016.

**Figura 2** - Variação percentual das principais causas de morte entre 2005 e 2016 nos EUA.



Fonte: IHME, 2016.

O Brasil faz pouco uso das suas bases de dados em saúde. A atenção primária conta com um banco de dados rico em informações coletadas das suas unidades básicas de saúde, mas não possui uma ferramenta para a análise inteligente desses dados. Os gestores municipais e estaduais carecem de uma ferramenta de apoio a decisão em relação aos dados de saúde. Seria importante disponibilizar painéis de indicadores de saúde, como doenças atuais da população, mostrando o estado atual, listando de forma individualizada, ou seja, apresentar o cidadão qual enfermidade ele possui e por outro lado, mostrar os índices de cada CID10 (Classificação Internacional de Doenças) no município, com um mapa epidemiológico atual e real. E o desenvolvimento de um *dashboard* específico para a atenção primária, integrado ao sistema de prontuário eletrônico do município, os gestores tem de forma visual essas informações para terem uma visão geral da saúde na sua cidade.

Um *dashboard* é “uma exibição visual das informações mais importantes necessárias para atingir determinados objetivos, que foram consolidados em uma única tela para que possam ser monitorados de uma só vez”. O termo “*dashboard*” é originado do painel de instrumentos do automóvel, onde os motoristas monitoram rapidamente as principais funções do veículo por meio de um conjunto de instrumentos. Também é conhecido como um painel de inteligência de negócios, pois são ferramentas de visualização de dados que exibem os principais indicadores de um país, organização ou projeto (BOZA, 2017).

Os sistemas georreferenciados são usados para descrever, analisar e prever padrões espaciais, dispondo de diversas aplicações em epidemiologia, como mapeamento de doenças, a investigação de surtos e análises espaciais. Seu uso depende, sobretudo, da disponibilidade de dados georreferenciados. O georreferenciamento, por sua vez, é o processo pelo qual informações textuais descritas de uma localidade, como o endereço relacionado a um evento de saúde, são convertidas em representações geográficas válidas. Ele pode ser realizado por meio da associação a um par de coordenadas geográficas ou a unidades espaciais, como setores censitários, bairros, municípios, entre outras. (SILVEIRA; OLIVEIRA; JUNGER, 2017).

Ainda, a natureza estática dos sistemas de relatórios de desempenho no setor de assistência em saúde resulta em relatórios de desempenho inconsistentes, incomparáveis, demorados e estáticos que não refletem de modo transparente a visão geral de desempenho e suporte efetivo às decisões dos gerentes de saúde. Assim, o setor de saúde precisa de ferramentas interativas de gerenciamento e visualização de dados e informações, como painéis para mostrar, medir, monitorar e gerenciar as informações em saúde com mais eficiência e assim traçar metas concretas (GHAZISAEIDI et al., 2017).

Desde o início da Revolução Industrial, sempre se colocou a possibilidade das máquinas chegarem um dia a ter iniciativa própria. Na área de análise de dados, isso significa a elaboração de algoritmos que respondam e se adaptem automaticamente aos dados sem a necessidade de intervenção humana contínua. A metodologia de *machine learning* atualmente mais utilizada em análise de dados são as árvores de decisão, que podem ser usadas quando a variável dependente assume valores finitos ou contínuos (CHIAVEGATTO FILHO, 2015).

Um estudo promissor é utilizar-se dessa base histórica para descobrir informações úteis, sobretudo na detecção de doenças a um determinado cidadão. Com um conjunto de estratégias, técnicas e ferramentas podemos aplicar um algoritmo de predição de dados para mostrar a tendência de uma patologia ocorrer numa pessoa.

Diante deste contexto, este estudo se justifica pela ausência de uma ferramenta no Ministério da Saúde que possibilite a avaliação epidemiológica da atenção primária, neste estudo, das doenças cerebrovasculares.

Este trabalho se propõe a desenvolver, por meio da integração com e-SUS AB, um sistema de apresentação em um mapa de

georreferenciamento, assim como relatório por cidadão e por CID10, obtendo uma ferramenta de apoio à decisão para análise histórica dos dados de doenças cerebrovasculares na atenção primária em saúde no Brasil. Ainda, se propõe a construir um relatório baseado em análise preditiva, através de um algoritmo de *machine learning*, para listar o risco dos cidadãos em desenvolver doenças cerebrovasculares. Com essa ferramenta, espera-se que os gestores de saúde tenham informações para apoio à decisão de forma mais eficaz e eficiente em relação a ações de saúde com base em dados clínicos e epidemiológicos.

Para tanto as questões norteadoras que este estudo se propõe a responder são:

1. Quais os critérios de desenvolvimento e a demonstração de um mapa georeferenciado, de relatórios nominais de doenças cerebrovasculares, de indicadores quantitativos de doenças cerebrovasculares na atenção primária em saúde como apoio a tomada de decisão às condições sensíveis em saúde a partir de dados individualizados integrados com o e-SUS AB?
2. Qual a contribuição da análise preditiva na prevenção das doenças cerebrovasculares no e-SUS AB?



## 2 OBJETIVOS

Desenvolver um mapa georeferenciado das doenças cerebrovasculares na atenção primária em saúde como apoio a tomada de decisão às condições sensíveis em saúde a partir da integração com dados do e-SUS AB.

Apresentar um modelo de relatório preditivo de cidadãos propensos a desenvolver doenças cerebrovasculares para possível integração com o e-SUS AB.

### 2.1 OBJETIVOS ESPECÍFICOS

- Estruturar um banco de dados para armazenar as informações relativas às doenças cerebrovasculares de forma consolidada para a atenção primária em saúde;
- Realizar migração dos dados do e-SUS AB para o novo banco de dados;
- Apresentar os dados e as informações consolidadas num *Dashboard*<sup>1</sup>;
- Apresentar os dados e as informações analisadas em um mapa a partir da técnica de geoprocessamento;
- Apresentar os cidadãos que já tiveram ou tem doenças cerebrovasculares;
- Apresentar os indicadores das doenças cerebrovasculares;
- Apresentar uma análise preditiva com cidadãos que possuem riscos de obterem doenças cerebrovasculares aos profissionais de saúde.

---

<sup>1</sup> Dashboard ou Painel de Controle é a apresentação visual das informações mais importantes e necessárias para alcançar um ou mais objetivos de um processo ou negócio, consolidadas e ajustadas em uma tela para fácil acompanhamento (FEW, 2004).





### 3 REVISÃO DE LITERATURA

Este capítulo visa aprofundar o conhecimento de temas pertinentes ao desenvolvimento deste estudo destacando as condições sensíveis na atenção primária, *business intelligence*, sistema georeferenciado e análise preditiva.

#### 3.1 CONDIÇÕES SENSÍVEIS NA ATENÇÃO PRIMÁRIA EM SAÚDE

No contexto internacional, um dos indicadores observados na atividade hospitalar como medida da efetividade da atenção primária à saúde são as condições sensíveis, que representam um conjunto de problemas de saúde para os quais a efetiva ação na atenção primária diminuiria o risco de internações. Essas atividades como a prevenção de doenças, o diagnóstico e o tratamento precoce de patologias agudas, e controle e acompanhamento de patologias crônicas, tem como consequência a redução das internações hospitalares por esses problemas. As altas taxas de internações por condições sensíveis à atenção primária em uma população podem indicar sérios problemas de acesso ao sistema de saúde ou de seu desempenho. Esse excesso de hospitalizações representa um sinal de alerta, que pode acionar mecanismos de análise e busca de explicações para a sua ocorrência. As altas taxas de internações por condições sensíveis à atenção primária estão associadas a deficiências na cobertura dos serviços e/ou à baixa resolutividade da atenção primária para determinados problemas de saúde (ALFRADIQUE et al., 2009).

As internações por casos sensíveis a atenção primária (CSAP) vem sendo amplamente utilizada em diversos países. A Espanha, por possuir um sistema nacional de saúde universal, com base na Atenção Primária à Saúde (APS), apresenta aplicação dos CSAP um pouco diferente do original, passando a serem utilizadas como indicador de efetividade do nível primário da assistência à saúde. No Brasil, a aplicação deste conceito ainda é recente, podendo-se notar uma ampliação das publicações nacionais a partir do ano de 2008 (FERREIRA; DIAS; MISHIMA, 2012).

O grupo 12 da lista de condições sensíveis à atenção primária são as doenças cerebrovasculares, objeto deste estudo, e que é composto por um pela lista de CID's específicos, conforme a Tabela 1.

**Tabela 1 - Grupo 12 da Lista Brasileira de Internações por Condições Sensíveis à Atenção Primária.**

CÓDIGO	DESCRIÇÃO
I630	Infarto cerebral devido a trombose de artérias pré-cerebrais
I631	Infarto cerebral devido a embolia de artérias pré-cerebrais
I632	Infarto cerebral devido a oclusão ou estenose não especificadas de artérias pré-cerebrais
I633	Infarto cerebral devido a trombose de artérias cerebrais
I634	Infarto cerebral devido a embolia de artérias cerebrais
I635	Infarto cerebral devido a oclusão ou estenose não especificadas de artérias cerebrais
I636	Infarto cerebral devido a trombose venosa cerebral não-piogênica
I638	Outros infartos cerebrais
I639	Infarto cerebral não especificado
I64	Acidente vascular cerebral, não especificado como hemorrágico ou isquêmico
I650	Oclusão e estenose da artéria vertebral
I651	Oclusão e estenose da artéria basilar
I652	Oclusão e estenose da artéria carótida
I653	Oclusão e estenose de artérias pré-cerebrais múltiplas e bilaterais
I658	Oclusão e estenose de outra artéria pré-cerebral
I659	Oclusão e estenose de artérias pré-cerebrais não especificadas
I660	Oclusão e estenose da artéria cerebral média
I661	Oclusão e estenose da artéria cerebral anterior
I662	Oclusão e estenose da artéria cerebral posterior
I663	Oclusão e estenose de artérias cerebelares
I664	Oclusão e estenose de artérias cerebrais, múltiplas e bilaterais
I668	Oclusão e estenose de outra artéria cerebral
I669	Oclusão e estenose de artéria cerebral não especificada
I670	Dissecção de artérias cerebrais, sem ruptura
I671	Aneurisma cerebral não-roto
I672	Aterosclerose cerebral
I673	Leucoencefalopatia vascular progressiva
I674	Encefalopatia hipertensiva
I675	Doença de Moyamoya
I676	Trombose não-piogênica do sistema venoso intracraniano
I677	Arterite cerebral não classificada em outra parte
I678	Outras doenças cerebrovasculares especificadas
I679	Doença cerebrovascular não especificada
I690	Seqüelas de hemorragia subaracnoídea
I691	Seqüelas de hemorragia intracerebral
I692	Seqüelas de outras hemorragias intracranianas não traumáticas
I693	Seqüelas de infarto cerebral
I694	Seqüelas de acidente vascular cerebral não especificado como hemorrágico ou isquêmico
I698	Seqüelas de outras doenças cerebrovasculares e das não especificadas
G450	Síndrome da artéria vértebro-basilar
G451	Síndrome da artéria carotídea (hemisférica)
G452	Síndrome das artérias pré-cerebrais, múltiplas e bilaterais
G453	Amaurose fugaz
G454	Amnésia global transitória
G458	Outros acidentes isquêmicos cerebrais transitórios e síndromes correlatas
G459	Isquemia cerebral transitória não especificada
G460	Síndrome da artéria cerebral média
G461	Síndrome da artéria cerebral anterior
G462	Síndrome da artéria cerebral posterior
G463	Síndromes vasculares do tronco cerebral
G464	Síndrome vascular cerebelar
G465	Síndrome lacunar motora pura
G466	Síndrome lacunar sensorial pura
G467	Outras síndromes lacunares
G468	Outras síndromes vasculares cerebrais em doenças cerebrovasculares

Fonte: BRASIL (2008).

No Brasil, óbitos por doença cerebrovascular superaram em números absolutos os causados pela doença coronariana. Esse padrão resulta da maior proporção da doença cerebrovascular em moradores das regiões Norte e Nordeste do País e em mulheres e pela hipertensão arterial ter maior impacto na doença cerebrovascular do que na coronariana. Taxas de incidência, letalidade e mortalidade por grupo étnico são maiores em negros em relação a brancos nos Estados Unidos, explicadas por diferenças nas médias da pressão arterial sistólica e por determinantes socioeconômicos. A única pesquisa no Brasil que abordou a relação sexo, raça e mortalidade cerebrovascular foi no município de São Paulo (1999-2001) com mortalidade maior entre os negros, ficando os pardos em situação intermediária, próxima à dos brancos (LOTUFO; BENSON, 2013).

Outro estudo, realizado em cidades com mais de 100 mil habitantes, ou seja, 300 municípios, mostrou que no ano de 2014 foram registradas 1.116.917 de internações por condições sensíveis à atenção primária (ICSAP). Altos valores de ICSAP podem representar fragilidades e/ou baixa resolutividade da porta de entrada do sistema, que é a atenção primária. As principais causas de internação foram representadas pelo grupo das pneumonias (22,72%), seguida das doenças cerebrovasculares (DCBV) (9,63%) e insuficiência cardíaca (IC) (9,32%); e com menores valores, anemia (0,30%) e doenças preveníveis por imunização (0,52%) (PEREIRA; SILVA; LIMA NETO, 2015).

No ano de 2014 o Ministério da Saúde tinha a intenção de investir R\$ 437 milhões para ampliar a assistência a vítimas de Acidente Vascular Cerebral (AVC), uma das doenças cerebrovasculares. Cerca de R\$ 370 milhões seria destinado ao financiamento de leitos hospitalares em 151 cidades. O restante seria aplicado no tratamento trombolítico (uso de medicação para desfazer o coágulo e normalizar o fluxo sanguíneo que chega ao cérebro) (BRASIL, 2012).

Só no ano de 2014, o estado de Santa Catarina, gastou R\$ 93.406.791,39, com internações por condições sensíveis à atenção primária, que representa 16% de todas as internações no estado (SES-SC, 2014)

Para reduzir o número de casos na população, o melhor tratamento é o preventivo, que engloba o controle de vários fatores de risco vasculares como a pressão arterial, diabetes, colesterol,

triglicérides, doenças cardíacas, além da necessidade de não fumar, ter uma alimentação saudável e praticar exercícios físicos (BRASIL, 2012).

### 3.2 DASHBOARD

Como uma ferramenta de gerenciamento de desempenho interativo, o *dashboard* é um sistema de entrega de informações em camadas que apresenta em numa tela informações importantes sobre objetivos estratégicos permitindo gerentes monitorar e gerenciar de forma mais eficaz (GHAZISAEIDI et al., 2017).

Em termos simples, um *dashboard* é uma exibição eletrônica composta de *Key Performance Indicators* (KPIs). A primeira etapa para a definição de um *dashboard* é identificar o público alvo e quais informações são de interesse dessas pessoas. Na área da saúde existem 3 tipos de indicadores: financeiros, clínicos e operacionais. Os indicadores financeiros abordam a saúde financeira da organização, enquanto os indicadores clínicos mostram tanto o atendimento de qualidade quanto com os aspectos operacionais de um departamento clínico (HIMSS, 2009).

### 3.3 SISTEMA GEOREFERENCIADO

Os Sistemas de Informações Geográficas (SIG) são usados para descrever, analisar e prever padrões espaciais, dispendo de diversas aplicações em epidemiologia, como o mapeamento de doenças, a investigação de surtos e análises espaciais. Seu uso depende, sobretudo, da disponibilidade de dados georreferenciados. O georreferenciamento, por sua vez, é o processo pelo qual informações textuais descritivas de uma localidade, como o endereço relacionado a um evento de saúde, são convertidas em representações geográficas válidas. Ele pode ser realizado por meio da associação a um par de coordenadas geográficas ou a unidades espaciais, como setores censitários, bairros, municípios, entre outras (SILVEIRA; OLIVEIRA; JUNGER, 2017).

Dados georeferenciados possuem informações necessárias (latitude e longitude) para identificar localizações e na superfície da Terra num geográfico. Em pesquisas epidemiológicas esse tipo de informação é principalmente para investigar a relação entre os dados georeferenciados, tanto para eventos de saúde e aspectos relacionados como para características individuais. Pode ser apresentado um

mapeamento da distribuição de doenças, um agrupamento espacial da doença, etc. (WERNECK, 2008).

As técnicas de geoprocessamento vêm sendo utilizadas no planejamento, monitoramento e avaliação das ações de saúde, além de serem consideradas como ferramentas importantes de análise das relações entre o ambiente e eventos relacionados à saúde. O geoprocessamento pode ser utilizado na análise dinâmica de difusão espacial das doenças e suas relações com o ambiente com alta resolução gráfica, na avaliação da situação de saúde de populações e na identificação de regiões e grupos sob alto risco de adoecer. O SIG possui grande capacidade para acessar e integrar diferentes níveis de informações, permite a apresentação e associação dos dados de diferentes formas (tabelas, gráficos e mapas temáticos) e possibilita o mapeamento, a exibição e a análise espacial dos dados relevantes para o processo de territorialização. Além disso, o sistema manipula os dados como os demais sistemas de informação e possui uma característica marcante de interligar atributos não espaciais a dados espaciais (MULLER; CUBAS; BASTOS, 2010).

A associação entre eventos de saúde e a sua localização geográfica (georreferenciamento) tem sido usada com sucesso na área de saúde e constitui uma ferramenta importante para a tomada de decisão pelos gestores de saúde (SA et al., 2012).

### 3.4 ANÁLISE PREDITIVA

Dentre as principais técnicas de Data Mining destacam-se: a estatística, a inteligência artificial e a aprendizagem de máquinas ou *machine learning*. Esta integração entre a inteligência artificial e a estatística, a *machine learning*, permite o design e o desenvolvimento de algoritmos que permitem que os computadores aprendam com base em dados. Assim, a partir disso é possível automatizar o aprendizado para reconhecer padrões complexos e tomar decisões inteligentes baseadas em dados (HAN; KAMBER; PEI, 2012; GRAHAM et al., 2018).

Desta forma como a mineração de dados é comumente classificada pela sua capacidade em realizar determinadas tarefas, podemos ressaltar que as tarefas mais comuns são, de acordo com Han, Kamber e Pei, 2012:

- Descrição: utilizada para descrever os padrões e tendências revelados pelos dados. A descrição geralmente oferece uma

provável interpretação para os resultados obtidos. A tarefa de descrição é utilizada em conjunto com as técnicas de análise exploratória de dados, para comprovar a influência de certas variáveis no resultado obtido.

- Classificação: uma das mais comuns, visa identificar a qual classe um determinado registro pertence. Nesta tarefa, o modelo analisa o conjunto de registros fornecidos, com cada registro já contendo a indicação à qual classe pertence, a fim de aprender como classificar um novo registro (aprendizado supervisionado). A tarefa de classificação pode ser usada por exemplo para: diagnosticar onde uma determinada doença pode estar presente.
- Estimação ou Regressão: a estimação é similar à classificação, porém é usada quando o registro é identificado por um valor numérico e não um categórico. Assim, pode-se estimar o valor de uma determinada variável analisando-se os valores das demais. A tarefa de estimação pode ser usada por exemplo para: estimar a pressão arterial ideal de um paciente baseando-se na idade, sexo e IMC.
- Predição: é similar às tarefas de classificação e estimação, porém ela visa descobrir o valor futuro de um determinado atributo. Por exemplo: prever a probabilidade de uma determinada doença se manifestar no indivíduo e ou grupo.
- Agrupamento (*Clustering*): visa identificar e aproximar os registros similares. Um agrupamento (ou cluster) é uma coleção de registros similares entre si, porém diferentes dos outros registros nos demais agrupamentos. Esta tarefa difere da classificação, pois não necessita que os registros sejam previamente categorizados (aprendizado não-supervisionado). Além disso, ela não tem a pretensão de classificar, estimar ou prever o valor de uma variável, ela apenas identifica os grupos de dados similares. Por exemplo na detecção de comportamentos atípicos para uma determinada doença.
- Associação: consiste em identificar quais atributos estão relacionados. Apresentam a forma: SE atributo X ENTÃO atributo Y. Por exemplo: determinar os casos onde um novo medicamento pode apresentar efeitos colaterais.

Desta forma, os métodos de mineração de dados são divididos em aprendizado supervisionado (preditivo) e não-supervisionado

(descritivo). Neles, os métodos são classificados de acordo com as tarefas que realizam assim, técnicas como redes neurais, árvores de decisão, regras de associação, raciocínio baseado em casos e algoritmos genéticos são as mais usadas na construção deste processo (HAN; KAMBER; PEI, 2012; NATARAJAN; FRENZEL; SMALTZ, 2017).

A diferença entre os métodos de aprendizado supervisionados e não-supervisionados está no fato de que os métodos não-supervisionados não precisam de uma pré-categorização para os registros, isto é, não é necessário um atributo alvo. As tarefas de agrupamento e associação são consideradas como não-supervisionadas. Por outro lado, no aprendizado supervisionado, os métodos são providos com um conjunto de dados que possuem uma variável alvo pré-definida e os registros são categorizados em relação a ela. As tarefas mais comuns de aprendizado supervisionado são a classificação (que também pode ser não-supervisionado) e a regressão. Durante o processo de mineração, diversas técnicas devem ser testadas e combinadas a fim de que comparações possam ser feitas e então a melhor técnica (ou combinação de técnicas) seja utilizada (HAN; KAMBER; PEI, 2012; GRAHAM et al., 2018).

Os autores classificam os diversos métodos de formas diferentes. Neste estudo, optou-se por usar a classificação adotada por Han, Kamber, Pei (2012) pois nela os métodos são classificados de acordo com as tarefas que realizam, conforme segue:

Na técnica de classificação pode ser usada a mineração de itens frequentes com a criação de um conjunto de itens frequentes, respeitando-se um valor mínimo de frequência para os itens; árvores de decisão em que cada nó (não folha) indica um teste feito sobre um valor. As ligações entre os nós representam os valores possíveis do teste do nó superior, e as folhas indicam a classe (categoria) a qual o registro pertence. Isto posteriormente segue o fluxo da árvore montada, para valores e ou registros novos; classificação Baysean, ao qual a partir do teorema de Bayes, é possível encontrar a probabilidade de um certo evento ocorrer, dada a probabilidade de um outro evento que já ocorreu; classificação baseada em regras: segue a estrutura: SE condição ENTÃO conclusão (semelhante as regras de associação); redes neurais, que simulam o comportamento dos neurônios; SVM (Support Vector Machines) utilizada tanto para técnicas de classificação quanto de regressão pois obtém altos índices de assertividade, permite modelar situações não lineares complexas gerando modelos de interpretação

simples, pode ser usada para relações lineares e não lineares, entre outros; classificação por regras de associação que visam buscar padrões de associações; aprendizado tardio que somente realizam o aprendizado quando solicitado para a classificação de um novo registro; algoritmo genético; conjuntos aproximados, técnica que consegue realizar a classificação mesmo com dados imprecisos ou errados e é utilizada para valores discretos; e conjuntos nebulosos que utiliza a lógica fuzzy.

Na técnica de predição numérica o objetivo é descobrir um possível valor futuro de uma variável. As predições numéricas visam prever valores para variáveis contínuas. Para a predição de variáveis discretas, as técnicas de classificação já apresentadas podem ser aplicadas. Os métodos mais conhecidos para predição numérica são as regressões: regressão linear; regressão não linear.

A técnica de agrupamento (clustering), considera-se como não supervisionada. Dado um conjunto de registros, são gerados agrupamentos (ou cluster), contendo os registros mais semelhantes. Nos agrupamentos são utilizados métodos de posicionamento; métodos hierárquicos; métodos baseados na densidade; métodos baseados em grades e métodos baseados em modelos.

### **3.4.1 Regressão Logística**

Os modelos de regressão constituem uma das ferramentas estatísticas mais importantes na análise estatística de dados quando se pretende modelar relações entre variáveis. O principal objetivo destes modelos é explorar a relação entre uma ou mais variáveis explicativas (ou independentes) e uma variável resposta (ou dependente). Um dos casos particulares dos modelos lineares generalizados são os modelos onde a variável resposta apresenta apenas duas categorias ou que de alguma forma foi dicotomizada assumindo valores 0 ou 1 sendo o modelo de regressão logística o mais popular desses modelos. A regressão logística é uma técnica estatística que tem como objetivo modelar, a partir de um conjunto de observações, a relação “logística” entre uma variável resposta dicotômica e uma série de variáveis explicativas numéricas (contínuas, discretas) e/ou categóricas. Na regressão logística, a variável resposta é dicotômica, atribuindo-se o valor 1 ao acontecimento de interesse (sucesso) e 0 ao acontecimento complementar (insucesso) (CABRAL, 2013).



Um dos modelos lineares generalizados mais utilizados na área de saúde é o modelo de regressão logística binária, onde a variável resposta do modelo tem distribuição de Bernoulli (ou Binomial) e a função de ligação é a função logística. Na área de saúde, o referido modelo poderia ser adotado, por exemplo, para estimar a probabilidade do paciente: aderir ao tratamento medicamentoso (adesão=1; não adesão=0); reportar um estado de saúde não bom (não bom=1; bom=0); ter uma determinada doença crônica (ter DC=1; não ter DC=0) (MORAES, 2016).



## 4 METODOLOGIA DO ESTUDO

### 4.1 NATUREZA DO ESTUDO

Estudo de natureza quantitativa, pois visa desenvolver e mostrar em um mapa georeferenciado das doenças cerebrovasculares na atenção primária em saúde como apoio a tomada de decisão às condições sensíveis em saúde a partir da integração com dados do e-SUS AB e apresentar um modelo de relatório preditivo de cidadãos propensos a desenvolver doenças cerebrovasculares para possível integração com o e-SUS AB.

### 4.2 TIPO DO ESTUDO

Produção tecnológica inovadora, pois, inexistente estrutura informacional desta natureza no Brasil. É também um estudo preditivo quantitativo porque busca prever condições sensíveis em saúde com foco em doenças cerebrovasculares a partir de um algoritmo de predição. Visa também apresentar de forma georeferenciada as condições sensíveis em saúde para a atenção primária no Brasil para doenças cerebrovasculares.

### 4.3 DESENVOLVIMENTO DO ESTUDO

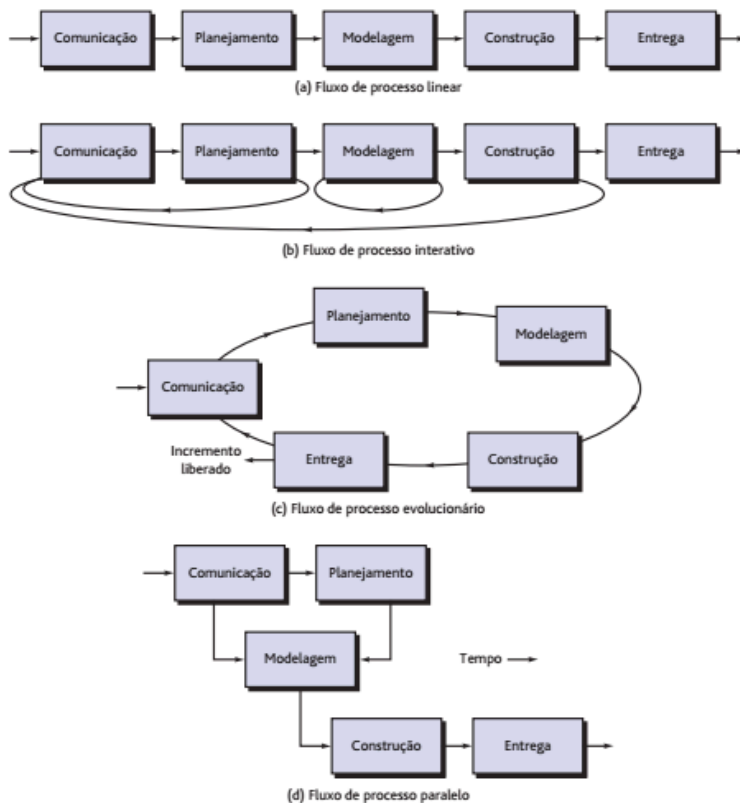
O estudo se divide em duas etapas. A primeira é desenvolvimento de um *dashboard* com relatórios específicos de doenças cerebrovasculares na atenção primária em saúde, apresentando de forma individualizada o cidadão que já teve ou tem uma DCBV. Um outro relatório são indicadores de cada CID10 do grupo de doenças cerebrovasculares, com quantidade de casos, quantidade por sexo, média de idade e distribuição por período. Ainda no *dashboard*, os casos de cidadãos que já tenham sido acometidos por esta patologia, são apresentados de forma georeferenciada num mapa. A segunda parte deste trabalho é aplicar um modelo de análise preditiva, onde um algoritmo aprende com dados históricos de características e fatores de risco para DCBV com pessoas que já tem uma doença do grupo e prediz a chance de um cidadão ter, por exemplo, um AVC. O resultado desta análise preditiva é apresentado no *dashboard* na forma de um relatório.

Para a definição do *dashboard* foram analisadas as fichas de Coletas de Dados Simplificada (CDS) do e-SUS AB para compor o modelo de informação adequado para a construção de um novo banco de dados para a nova ferramenta.

O nome dado a ferramenta é Dashboard de Doenças Cerebrovasculares (DDCBV), que visa apresentar de forma visual relatórios e gráficos, assim como, uma análise preditiva de pessoas com risco de desenvolver doenças cerebrovasculares.

A metodologia para construção do *software* foi baseada em Pressman e Maxim (2016), onde ele diz que os métodos da engenharia de software fornecem as informações técnicas para desenvolver software. Os métodos envolvem uma ampla variedade de tarefas, que incluem: comunicação, análise de requisitos, modelagem de projeto, construção de programa, testes e suporte. Os métodos da engenharia de software se baseiam em um conjunto de princípios básicos que governam cada área da tecnologia e incluem atividades de modelagem e outras técnicas descritivas. As ferramentas da engenharia de software fornecem suporte automatizado ou semiautomatizado para o processo e para os métodos. Quando as ferramentas são integradas, de modo que as informações criadas por uma ferramenta possam ser utilizadas por outra, é estabelecido um sistema para o suporte ao desenvolvimento de software, denominado engenharia de software com o auxílio do computador (PRESSMAN; MAXIM, 2016).

Ainda segundo o autor, um fluxo de processo linear executa cada uma das cinco atividades metodológicas em sequência, começando com a comunicação e culminando com a entrega (**Erro! Fonte de referência não encontrada.a**). Um fluxo de processo iterativo repete uma ou mais das atividades antes de prosseguir para a seguinte (**Erro! Fonte de referência não encontrada.b**). Um fluxo de processo evolucionário executa as atividades de forma “circular”. Cada volta pelas cinco atividades conduz a uma versão mais completa do software (**Erro! Fonte de referência não encontrada.c**). Um fluxo de processo paralelo (**Erro! Fonte de referência não encontrada.d**) executa uma ou mais atividades em paralelo com outras (por exemplo, a modelagem para um aspecto do software poderia ser executada em paralelo com a construção de outro aspecto do software) (PRESSMAN; MAXIM, 2016).

**Figura 3** - Fluxo de Processo

Fonte: PRESMANN (2016)

A ferramenta foi desenvolvida pelo próprio mestrando, desde a prototipação, modelagem do banco de dados, processo de migração, construção da ferramenta e a análise preditiva utilizando tecnologias gratuitas. Para a aplicação foi utilizado JAVA, o banco de dados PostgreSQL. Para a análise preditiva foi utilizado *Python* usando uma regressão logística otimizada por um algoritmo de otimização, *Gradiente Descendente Estocástico* (SGD).

### 4.3.1 Fases do Desenvolvimento

Após o estudo das variáveis e condições de saúde para as doenças cerebrovasculares, foi identificado nas fichas CDS do sistema e-SUS AB os dados que são coletados e que podiam ser utilizados para a montagem dos relatórios, do gráfico georeferenciado e da análise preditiva. Depois de identificados esses campos, foi feita a prototipação do sistema, por meio de rascunhos de tela, para que fosse então definida a modelagem dos dados e o processo de migração. Em seguida, foi iniciado a definição da arquitetura e componentes gráficos para a implementação do sistema, chamado de DDCV (*Dashboard* de Doenças Cerebrovasculares).

Após saber as informações que poderiam fazer parte dos relatórios, foi feita a prototipação das telas, usado para ensaio, demonstração e avaliação de design. Para um projeto rápido se concentra em uma representação dos aspectos do *software* que serão visíveis para os usuários (por exemplo, o layout da interface com o usuário ou os formatos de exibição na tela). O projeto rápido leva à construção de um protótipo. O protótipo é entregue e avaliado pelos envolvidos, os quais fornecem feedback que é usado para refinar ainda mais os requisitos. A iteração ocorre conforme se ajusta o protótipo às necessidades de vários envolvidos e, ao mesmo tempo, possibilita a melhor compreensão das necessidades que devem ser atendidas. Na sua forma ideal, o protótipo atua como um mecanismo para identificar os requisitos do software. Caso seja necessário desenvolver um protótipo operacional, pode-se utilizar partes de programas existentes ou aplicar ferramentas que possibilitem gerar rapidamente tais programas operacionais (PRESSMAN; MAXIM, 2016). A ferramenta utilizada para esta fase foi o moqups®, da S.C Evercoder Software S.R.L.

Para os mapas georreferenciados, foi utilizado uma API (Application Programming Interface - Interface de Programação de Aplicativos) do Google® para processar os endereços em ponto geográfico para plotagem no componente de mapas, tanto para o mapa de pontos quanto para o mapa de calor. Um mapa de calor é uma visualização usada para representar a intensidade dos dados em pontos geográficos, por padrão, as áreas de maior intensidade serão coloridas em vermelho e as áreas de menor intensidade serão exibidas em verde (GOOGLE, 2018).

Para a análise preditiva, foi exportado um arquivo CSV (Comma-Separated Values) para que pudesse ser aplicado o algoritmo de predição sendo o retorno outro CSV, onde foi importado na ferramenta para a apresentação do relatório de risco de doenças cerebrovasculares.

#### **4.3.2 Fichas CDS do e-SUS AB**

Após análise de todo o modelo de dados do e-SUS AB, são nas fichas do CDS que convergem todos os atendimentos e é o modelo de informação único para todos os municípios, ou seja, mesmo que um município use um sistema próprio, a integração é feita a partir das fichas CDS. Por esse motivo, a ferramenta se baseou nos campos dessas fichas. O dicionário de dados das fichas CDS está disponível publicamente no Portal do DAB/MS.

Atualmente o e-SUS AB CDS possui 12 fichas e foi analisada cada uma delas, resultando em informações de 4 delas:

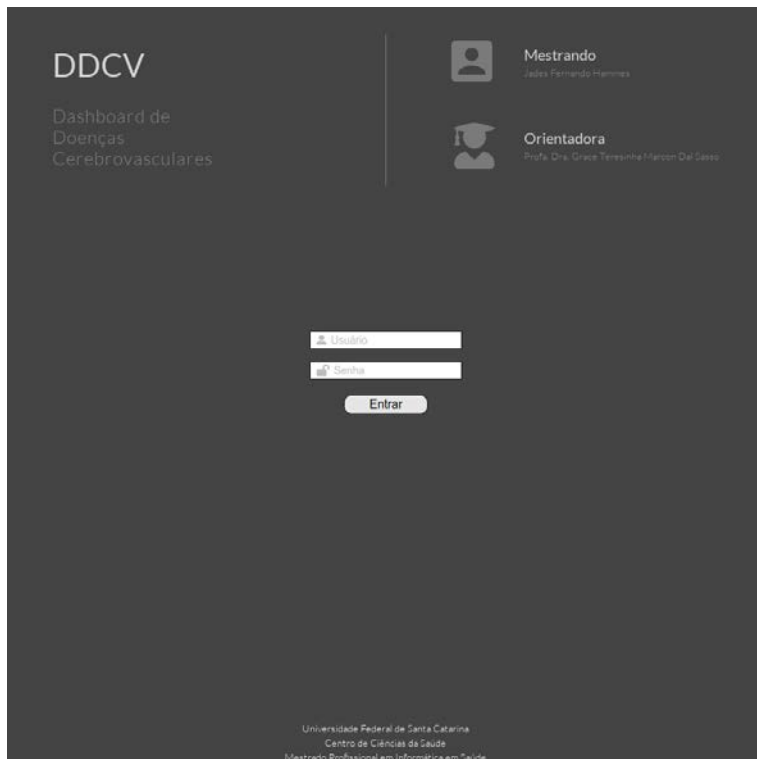
- Ficha de Cadastro Individual
- Ficha de Cadastro Domiciliar e Territorial
- Ficha de Atendimento Individual
- Ficha de Visita Domiciliar e Territorial

Para cada uma dessas fichas, foi avaliado a necessidade do campo para o estudo e uso nos gráficos e relatórios, como por exemplo, o endereço para o mapa georeferenciado. As informações escolhidas estão no Apêndice A.

#### **4.3.3 Prototipação**

Na fase de prototipação foram desenhadas as telas da ferramenta de acordo com a análise realizada.

Primeiramente, a necessidade de autenticação na ferramenta, por questões de segurança e privacidade dos dados, conforme Tela 1.

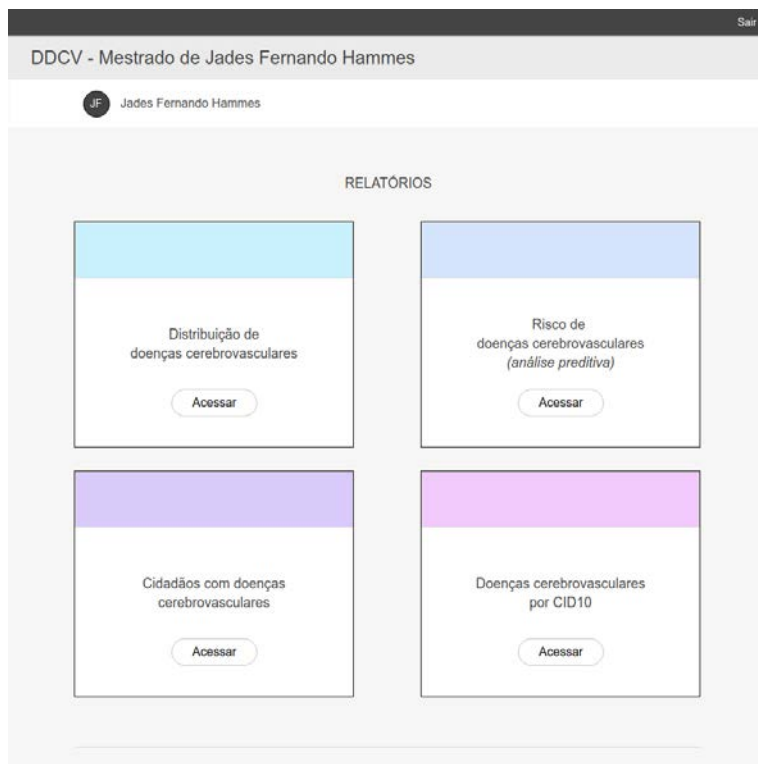
**Tela 1 – Login**

Fonte: Elaborado pelo Autor

A segunda página é o menu principal, onde é possível ver e selecionar os relatórios disponíveis no DDCV, conforme Tela 2.



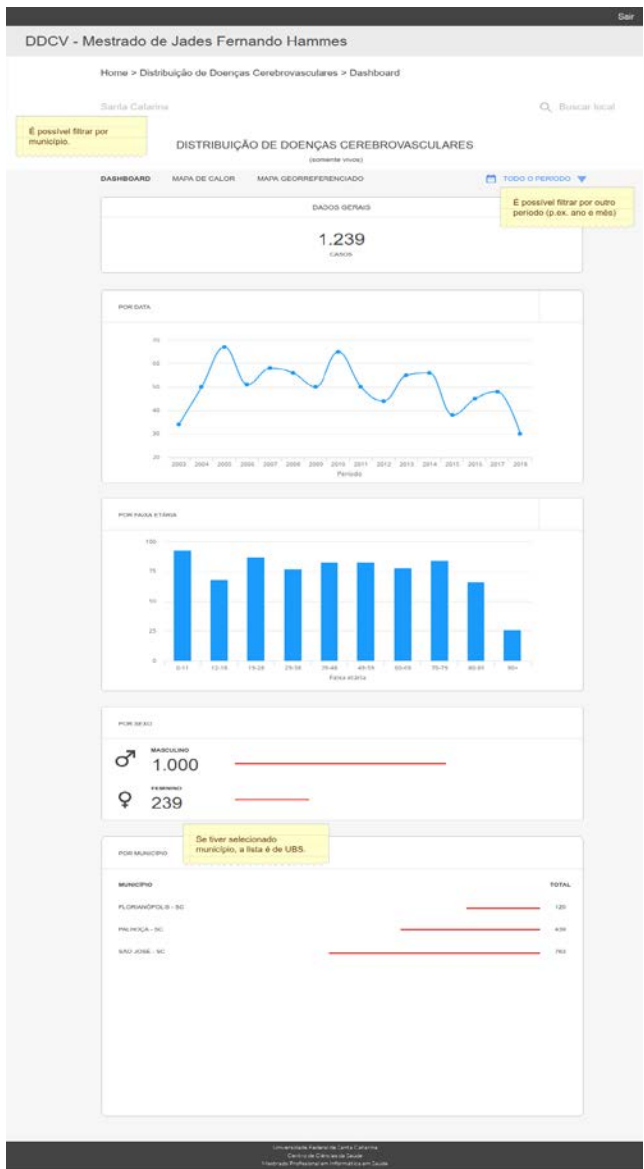
## Tela 2 – Menu Principal



Fonte: Elaborado pelo Autor

A funcionalidade de Distribuição de Doenças Cerebrovasculares está dividida em três abas, sendo um Dashboard geral, um Mapa de Calor e um Mapa Georreferenciado. No Dashboard são apresentados os indicadores gerais de todas as doenças cerebrovasculares, como o número de casos, distribuição por data, quantidade por faixa etária, separação por sexo e distribuição por município e UBS. A Tela 3 ilustra o relatório.

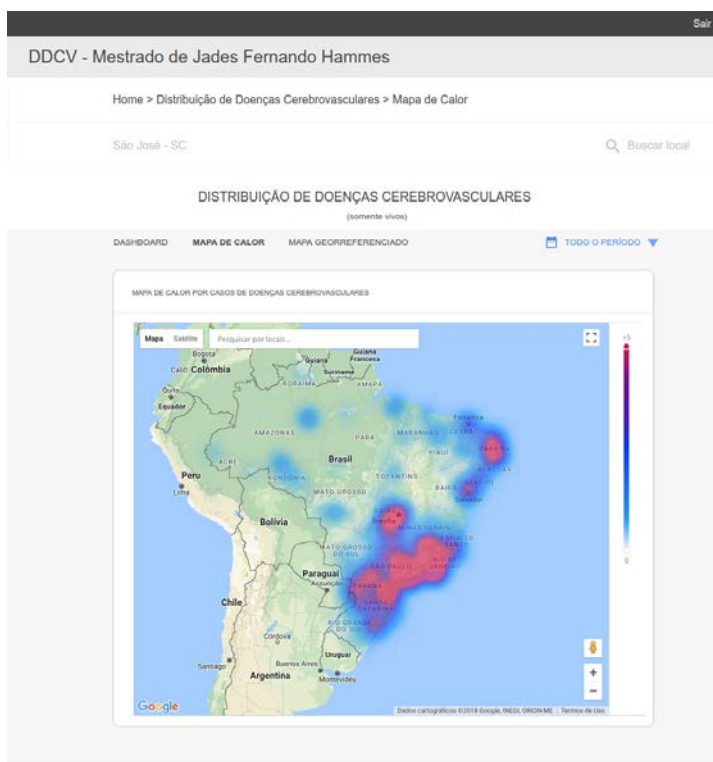
Tela 3 - Distribuição de Doenças Cerebrovasculares – Dashboard



Fonte: Elaborado pelo Autor

Ainda na funcionalidade de Distribuição de Doenças Cerebrovasculares, é apresentada em um mapa de calor, indicando aonde se concentram os casos de doenças cerebrovasculares no território, ilustrado pela Tela 4.

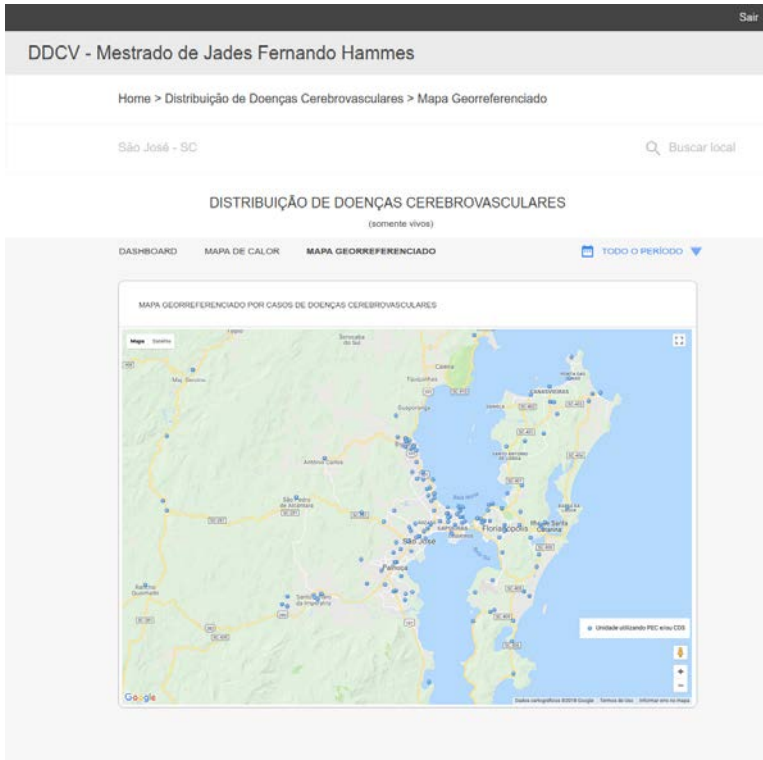
**Tela 4** - Distribuição de Doenças Cerebrovasculares - Mapa de Calor.



Fonte: Elaborado pelo Autor

O mapa georeferenciado apresenta as residências de pessoas com doenças cerebrovasculares identificadas por um ponto no mapa como podemos ver na Tela 5.

## Tela 5 - Distribuição de Doenças Cerebrovasculares - Mapa Georreferenciado



Fonte: Elaborado pelo Autor

Neste protótipo é listado cada cidadão que possui doença cerebrovascular, assim como sua idade e os problemas que são fatores de risco desta patologia, a Tela 6 mostra como será apresentada a listagem.

## Tela 6 - Cidadãos com Doenças Cerebrovasculares

DDCV - Mestrado de Jades Fernando Hammes



Home > Cidadãos com doenças cerebrovasculares

São José - SC 🔍 Buscar local

### CIDADÃOS COM DOENÇAS CEREBROVASCULARES

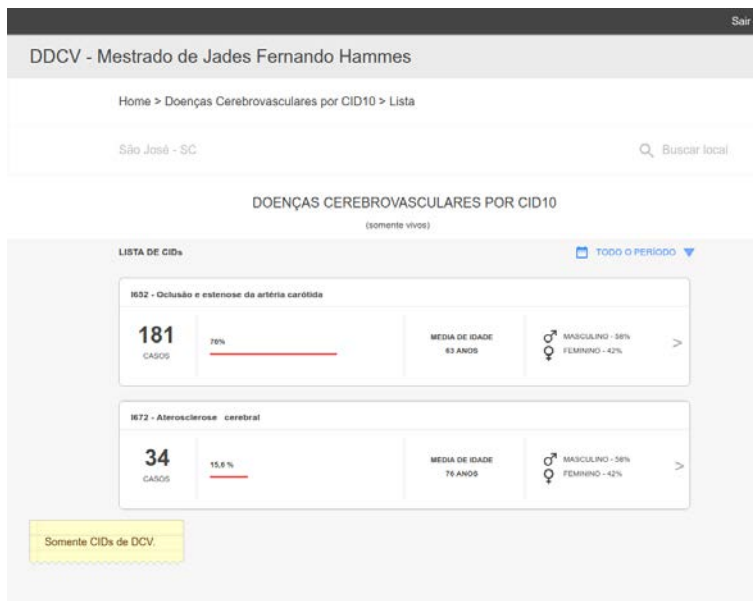
(somente vivos)

LISTA DE CIDADÃOS 📅 TODO O PERÍODO ▼

 <b>MASCULINO</b>	<b>JÓÃO DA SILVA</b> <span style="float: right;">52 ANOS</span> 12/04/2018 CID10 - I630 - Infarto cerebral devido a trombose de artérias pré-cerebrais <b>PROBLEMAS</b> Obesidade   Diabetes <b>MUNICÍPIO E UNIDADE DE REFERÊNCIA</b> São José - SC   UBS Prata Comporta
 <b>FEMININO</b>	<b>MARIANA COSTA</b> <span style="float: right;">39 ANOS</span> 21/05/2018 CID10 - I638 - Outros infartos cerebrais <b>PROBLEMAS</b> Não especificado. <b>MUNICÍPIO E UNIDADE DE REFERÊNCIA</b> São José - SC   UBS Prata Comporta

Fonte: Elaborado pelo Autor

Outro indicador importante, é saber a quantidade de casos na população por CID10, ou seja, para cada doença do grupo de DCBV, qual o número de pessoa, qual a porcentagem em relação aos demais CIDs, quantos masculinos e feminos, qual a média de idade que ocorre esta patologia. Ainda nesta funcionalidade, é possível visualizar não só em números, mas também em forma de gráfico, visualizando também a quantidade de casos por período. As telas 7 e 8 ilustram esta opção no sistema.

**Tela 7 - Doenças Cerebrovasculares por CID10**

Fonte: Elaborado pelo Autor

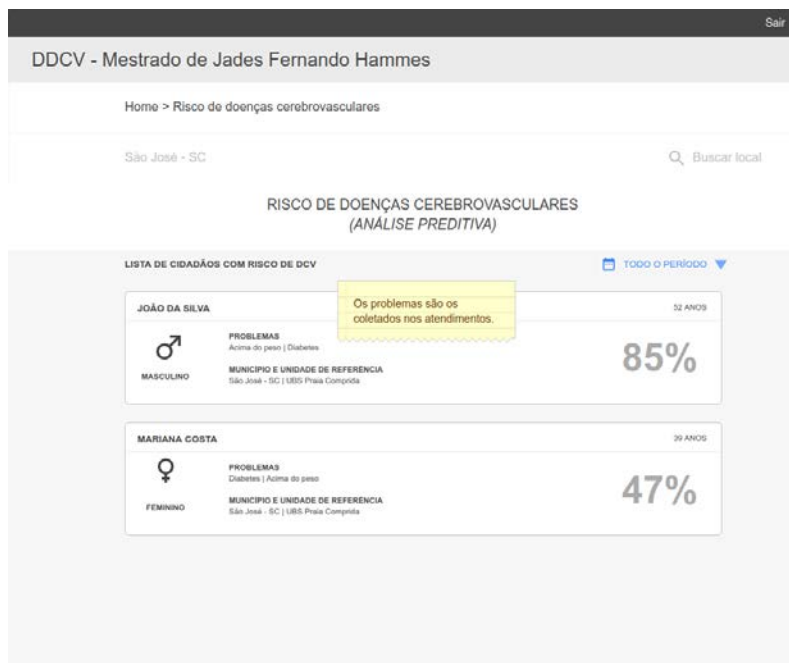
**Tela 8 - Doenças Cerebrovasculares por CID10 – Detalhe**  
 Fonte: elaborado pelo autor



Fonte: Elaborado pelo Autor

Um dos principais objetivos deste trabalho é a análise preditiva, que está detalhada mais abaixo, e para visualizar o resultado da aplicação do algoritmo, a ferramenta dispõe de uma tela para apresentar a porcentagem da chance de um cidadão adquirir uma patologia de DCBV. O nome da funcionalidade no sistema é Risco de Doenças Cerebrovasculares (Análise Preditiva), conforme Tela 9.

**Tela 9** - Risco de Doenças Cerebrovasculares (Análise Preditiva)



Fonte: Elaborado pelo Autor



#### **4.3.4 Banco de Dados**

Para que a ferramenta ficasse independente da aplicação do e-SUS AB, optou-se por criar um novo bando de dados, apenas com os campos necessários para os relatórios e para predição. O banco de dados escolhido foi o PostgreSQL®, um sistema gerenciador de bando de dados, de código aberto. O novo banco de dados foi modelado para receber os dados já processados do e-SUS AB, que ao migrar de um banco de dados para outro, realiza um processamento que faz cálculo da idade (data do atendimento – data de nascimento) e determina se o cidadão é obeso ou não, pelo cálculo do IMC.

Como a proposta foi trabalhar com dados fictícios, uma vez que o dado seria individualizado, foi criada uma rotina para popular o e-SUS AB de forma automatizada, que gerou uma quantidade finita de registros para simular os atendimentos no sistema.

#### **4.3.5 Dados do Sistema**

A proposta da ferramenta deste trabalho, apresenta dados individualizados do cidadão pois um dos objetivos é apresentar qual cidadão possui doença, além de predizer a chance de uma pessoa obter uma DCBV e esta foi a grande motivação de trabalhar esse modelo com dados fictícios, para não expor a privacidade do usuário do SUS. Foi criada uma rotina para popular as tabelas do e-SUS AB de forma automatizada, que gerou uma quantidade finita de registros para simular os atendimentos no sistema. O código-fonte da geração de dados fictícios está no Anexo B e a Figura 4 ilustra a tabela já com os registros criados.

**Figura 4** - Tabela do banco de dados com registros fictícios.

id	id_origem	data	cnes	unidade_saude	ano	cns	cns_re	nome	nome_social	
PK bigint	bigint	date	character varying	character varying (200)	integer	character varying (15)	character varying (255)	character varying (255)	character varying (255)	
1	1	141971	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898004284405316	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
2	2	142027	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	709801073398693	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
3	3	142057	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	702501309655533	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
4	4	142112	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	704700715189833	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
5	5	142148	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898004280312836	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
6	6	142165	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	700403981586345	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
7	7	142199	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	707009823932320	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
8	8	142242	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	700001884401605	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
9	9	142272	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898004806862373	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
10	10	142288	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898002917307457	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
11	11	142313	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898004675531859	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
12	12	142316	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	702509307201338	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
13	13	142327	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	706908195073536	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
14	14	142328	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	700406974107049	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
15	15	142343	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	704608161464530	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
16	16	142478	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	898004284413793	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
17	17	142529	2017...	6633765	ESF JOVINO INACIO CARDO...	2017	705004084732953	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES
18	18	142611	2017...	2660741	ESF APONSO RABE	2017	700409915740647	[null]	JOSE ANTONIO SILVA	JOSE ANTONIO SILVA
19	19	142618	2017...	2660741	ESF APONSO RABE	2017	708906798247115	[null]	MARIA FERNANDA SOARES	MARIA FERNANDA SOARES

Fonte: Elaborado pelo Autor

### 4.3.6 Análise Preditiva

O modelo de predição utilizado neste trabalho foi a regressão logística otimizada por um algoritmo de otimização, o Gradiente Descendente Estocástico (SGD). Em termos simples, o gradiente descendente estocástico não usa todos os dados mas sim uma fração dos dados. Esse subconjunto dos dados é obtido com reamostragem dos dados originais. A reamostragem é o componente estocástico do algoritmo. Com isso, o traço da função objetivo deixa de ser suave, o que impossibilita o uso de critérios de parada baseados em diferenças de valores consecutivos. Então o mais comum é executar o algoritmo até exceder o número máximo de iterações. O número máximo de iterações deve ser escolhido de forma a garantir suficiente proximidade com o ponto de ótimo da função (ZEVIANI; FERREIRA, 2017).

Para a ferramenta desenvolvida neste trabalho, foram gerados aleatoriamente dados de 170.567 cidadãos, dos quais 561 possuem DCV. Essas pessoas serviram de referência para o algoritmo aprender os padrões e as relações as quais o cidadão desenvolve uma doença cerebrovascular. Os seguintes fatores de risco foram analisados como variáveis do estudo conforme Figura 5.

**Figura 5** - Colunas de Dados para a Análise Preditiva.

```
columns = ["id", "cns", "nome", "idade", "sexo", "raca_cor", "obeso",
"ex_obeso", "fumante", "ex_fumante", "alcool", "ex_alcool", "drogas",
"ex_drogas", "hipertensao_arterial", "diabetes", "dcv", "gestante"]
```

Fonte: Elaborado pelo Autor

Após a definição das variáveis, é necessário definir o tipo de cada uma das variáveis de entrada, ou seja, deve-se preparar o arquivo para receber os dados de forma que o algoritmo possa analisar corretamente, evitando, por exemplo, variáveis booleanas.

Como a idade é uma variável, é necessário verificar os conjuntos Fuzzy<sup>2</sup>, nesse caso, para normalizar a informação da idade, que é contínua com intervalo [0, inf], e é diferente de outras classes. Com o objetivo de definir grupos de pertinência, para cada idade, elas passaram a variar entre [0,1], e possuem a informação de jovem, adulto e idoso.

Antes ainda da execução do algoritmo, é essencial a preparação dos dados, criando, por exemplo, colunas binárias para o sexo (sexo\_MASCULINO e sexo\_FEMININO). Nesse momento, aplica-se também os grupos de idade, que é separar as idades em novos grupos.

Depois de normalizado e preparado os dados, eles são convertidos, treinados e testados com os registros que já possuem DCV que demonstra ao algoritmo quais fatores contribuem para a patologia.

É importante observar o Plot de correlação, que é a dependência ou associação de qualquer relação estatística (causal ou não causal) entre duas variáveis, pois se elas forem muito diferentes, há um possível erro no algoritmo.

A última etapa é o OUTPUT que prepara o algoritmo treinado para executar em todo o conjunto de dados e obter o resultado de probabilidade.

---

<sup>2</sup> Esta teoria foi introduzida em 1965 pelo matemático Lotfi Asker Zadeh, com a intenção de dar um tratamento matemático a certos termos linguísticos subjetivos como: “aproximadamente”, “em torno de”, dentre outros. Pode-se dizer que a Teoria dos Conjuntos Fuzzy representa um primeiro passo no sentido de se programar e armazenar conceitos vagos em computadores, tornando possível a produção de cálculos com informações imprecisas, a exemplo do que faz o ser humano (SOUZA, 2018).

#### 4.4 ASPECTOS ÉTICOS DO ESTUDO

Este trabalho está de acordo com a Resolução n° 466/12, do Conselho Nacional de Saúde que traz termos e condições a serem seguidos em todas as pesquisas que envolvem seres humanos. Aborda requisitos do Sistema de avaliação ética brasileiro, compondo um sistema que utiliza mecanismos, ferramentas e instrumentos próprios de inter-relação que visa à proteção dos participantes de pesquisa (BRASIL, 2013).

Este estudo foi dispensado da apreciação pelo Comitê de Ética em Pesquisa, pois utiliza exclusivamente dados fictícios, uma vez que o propósito do trabalho é identificar nominalmente o cidadão, portanto, para manter o sigilo e respeito ao paciente, optou-se por construir os dados aleatoriamente.

Para a prova do modelo em questão, foi criada uma base de dados fictícia, de forma automatizada, para popular dos dados necessários para o estudo. Os registros foram adicionados no modelo das fichas de coleta de dados simplificada: Cadastro Individual, de Cadastro Domiciliar e Territorial, Ficha de Atendimento Individual e Ficha de Visita Domiciliar e Territorial. Essas fichas são as que contém os campos necessários referentes ao diagnóstico e variáveis de risco para doenças cerebrovasculares.

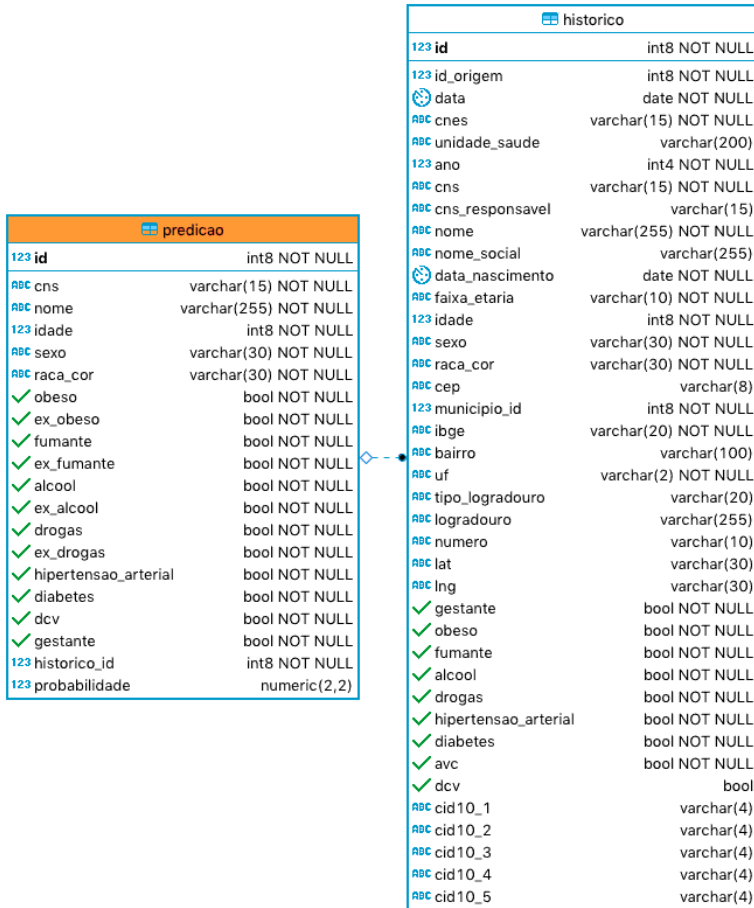
## 5 RESULTADOS

Com base no objetivo geral e específicos definidos neste trabalho, os resultados são apresentados abaixo.

### 5.1 ESTRUTURAR UM BANCO DE DADOS PARA ARMAZENAR AS INFORMAÇÕES RELATIVAS ÀS DOENÇAS CEREBROVASCULARES DE FORMA CONSOLIDADA PARA A ATENÇÃO PRIMÁRIA EM SAÚDE

Com base nos dados necessários para a ferramenta DDCV, analisado anteriormente, foi feita a modelagem dos dados e criadas duas tabelas no banco de dados PostgreSQL®. A tabela histórico recebeu os dados dos atendimentos realizados pelo CDS, armazenando, principalmente, se o cidadão já teve DCV e os campos de riscos de saúde para esta patologia. A tabela de predição armazenou os riscos de saúde e também se o cidadão deixou de ter um risco, mas um dia ele já existiu, que são: se já foi obeso, se já foi fumante, se já foi alcoólatra, se já foi usuário de drogas e se já teve um DCV prévio. Essas informações serão usadas na análise preditiva posteriormente. As tabelas de histórico e predição estão representadas na Figura 6.

**Figura 6 - Modelagem do Banco de Dados da Ferramenta DDCV**



Fonte: Elaborado pelo Autor

## 5.2 REALIZAR MIGRAÇÃO DOS DADOS DO E-SUS AB PARA O NOVO BANCO DE DADOS

Para migrar os dados do e-SUS AB para o DDCV foram criadas inicialmente *views* de banco de dados, que é uma tabela virtual, somente dos campos selecionados para o estudo.

Foram feitas uma visão do banco de dados do e-SUS AB para cada Ficha CDS, mas somente com os campos necessários para o novo banco de dados. As figuras abaixo representam as visões de cada Ficha CDS.

**Figura 7** - View da Ficha de Atendimento Individual

esusab_atendimento_individual	
123	co_seq_fat_atd_ind int8
	dt_atend_individual date
ABC	nu_cnes varchar(20)
ABC	nu_cns bpchar(15)
	dt_nascimento date
ABC	ds_sexo text
123	st_gestante int4
123	nu_peso float8
123	nu_altura float8
ABC	ds_filtro_cids varchar(4000)
123	st_fumante int8
123	st_alcool int8
123	st_drogas int8
123	st_hipertensao_arterial int8
123	st_diabete int8

Fonte: Elaborado pelo Autor

**Figura 8** - View da Ficha de Cadastro Domiciliar e Territorial

esusab_cadastro_domiciliar		
123	co_seq_fat_cad_dom_familia	int8
ABC	nu_cns_responsavel	bpchar(15)
ABC	no_tipo_logradouro	varchar(144)
ABC	no_logradouro	varchar(256)
ABC	nu_domicilio	varchar(255)
ABC	nu_cep	varchar(255)
ABC	no_bairro	varchar(256)
ABC	co_ibge	varchar(7)
ABC	no_localidade	varchar(72)
ABC	no_uf	varchar(2)

Fonte: Elaborado pelo Autor

**Figura 9** - View da Ficha de Cadastro Individual

esusab_cadastro_individual		
123	co_seq_fat_cad_individual	int8
	dt_cad_individual	timestamp
ABC	nu_cnes	varchar(20)
ABC	no_unidade_saude	varchar(500)
ABC	no_bairro_unidade_saude	varchar(500)
ABC	co_ibge_municipio_unidade	varchar(10)
ABC	nu_cns	bpchar(15)
ABC	nu_cns_responsavel	bpchar(15)
ABC	no_cidadao	varchar(255)
ABC	no_social_cidadao	varchar(255)
	dt_nascimento	date
ABC	dssexo	text
ABC	ds_raca_cor	text
123	st_gestante	int4
ABC	ds_condicao_peso	text
123	st_fumante	int4
123	st_alcool	int4
123	st_outra_droga	int4
123	st_hipertensao_arterial	int4
123	st_diabete	int4
123	st_avc	int4

Fonte: Elaborado pelo Autor



**Figura 10** - View Ficha de Visita Domiciliar e Territorial

esusb_visita_domiciliar		
123	co_seq_fat_visita_domiciliar	int8
	dt_registro	date
ABC	nu_cnes	varchar(20)
ABC	nu_cns	bpchar(15)
	dt_nascimento	date
ABC	ds_sexo	text
123	nu_peso	float8
123	nu_altura	float8
123	st_gestante	int4
123	st_fumante	int4
123	st_alcool	int4
123	st_drogas	int4
123	st_hipertensao	int4
123	st_diabete	int4

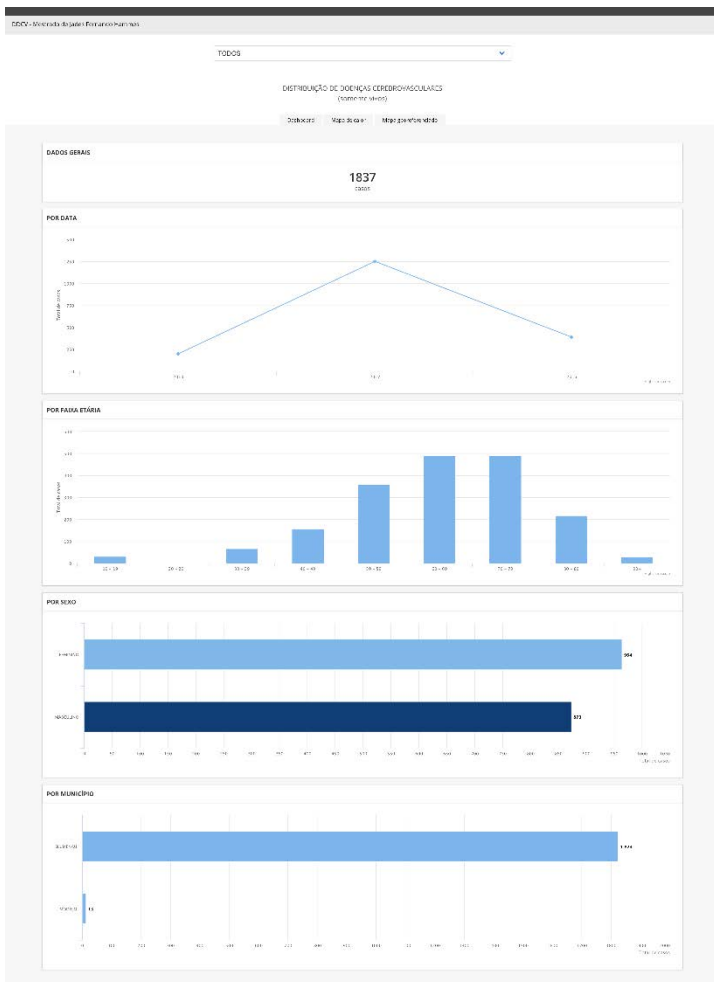
Fonte: Elaborado pelo Autor

Foi realizado um processo de migração na nova aplicação que fez a leitura dessa *view* e gravou no banco de dados da ferramenta DDCV, fazendo os processamentos necessários. O código desse processamento está no Anexo A.

### 5.3 APRESENTAR OS DADOS E AS INFORMAÇÕES CONSOLIDADAS NUM DASHBOARD

Utilizando a base de dados processada bem como a implementação da base na linguagem JAVA com Spring, são apresentados os indicadores dos dados de doenças cerebrovasculares. A Tela 10 mostra o resultado já na ferramenta DDCV.

**Tela 10 - DDCV - Distribuição de Doenças Cerebrovasculares – Dashboard**

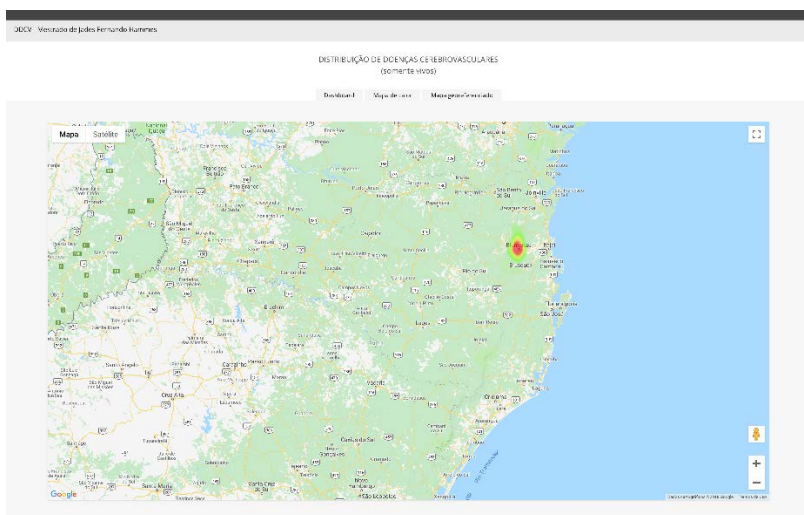


Fonte: Elaborado pelo Autor

## 5.4 APRESENTAR OS DADOS E AS INFORMAÇÕES ANALISADAS EM UM MAPA A PARTIR DA TÉCNICA DE GEOPROCESSAMENTO

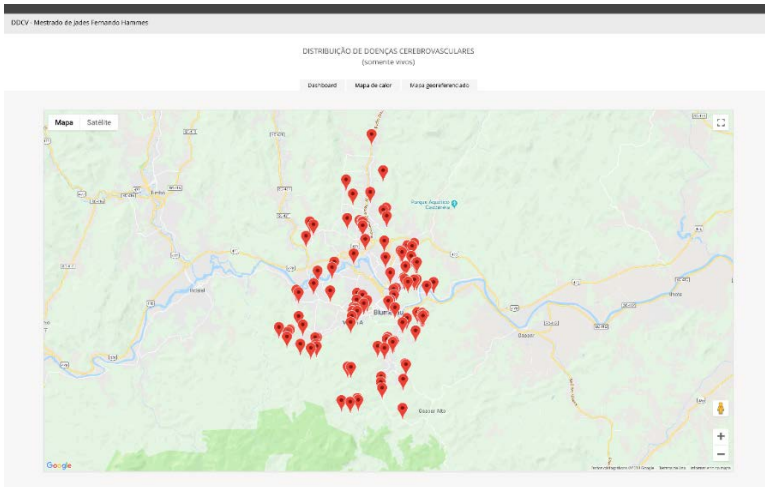
A partir do *Google® Geocoding API*, oferecido pela companhia Google®, foi realizada a conversão de endereços da base de dados por coordenadas geográficas, esse processo é chamado de geocodificação. (GOOGLE, 2018b). Após essa fase, foi utilizado o *GoogleMaps®* para mostrar na ferramenta o mapa de calor (conforme Tela 11) e o mapa georeferenciado (**Erro! Fonte de referência não encontrada.**).

**Tela 11** - DDCV - Distribuição de Doenças Cerebrovasculares - Mapa de Calor



Fonte: Elaborado pelo Autor

## Tela 12 - DDCV - Distribuição de Doenças Cerebrovasculares - Mapa Georreferenciado



Fonte: Elaborado pelo Autor

### 5.5 APRESENTAR OS CIDADÃOS QUE JÁ TIVERAM OU TEM DOENÇAS CEREbroVASCULARES






Este relatório apresenta de forma individualizada, os cidadãos que já tiveram doenças cerebrovasculares, para acompanhamento e gestão do cuidado em saúde. É mostrado também os problemas de saúde que o indivíduo possui. A **Erro! Fonte de referência não encontrada.** 13 mostra o resultado da implementação.

## Tela 13 - DDCV – Cidadãos com Doenças Cerebrovasculares

DDCV - Menuário de Jader Perinardi Hummes

MODELO

CIDADÃOS COM DOENÇAS CEREBROVASCULARES  
(somente vivos)

 MASCULINO	<b>JOSE ANTONIO SILVA - 7284 - 57 ANOS</b> 07/06/2018 I64 ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO <b>PROBLEMAS</b> ELIMINATE USO DE ÁLCOOL <b>MUNICÍPIO E/OU UNIDADE DE REFERÊNCIA</b> MODELO - UNIDADE SANITARIA SEDE DE MODELO
 MASCULINO	<b>JOSE ANTONIO SILVA - 149931 - 48 ANOS</b> 21/03/2018 I64 ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO <b>PROBLEMAS</b> NENHUM PROBLEMA REGISTRADO <b>MUNICÍPIO E/OU UNIDADE DE REFERÊNCIA</b> MODELO - UNIDADE SANITARIA SEDE DE MODELO
 MASCULINO	<b>JOSE ANTONIO SILVA - 164443 - 44 ANOS</b> 27/04/2018 I64 ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO <b>PROBLEMAS</b> NENHUM PROBLEMA REGISTRADO <b>MUNICÍPIO E/OU UNIDADE DE REFERÊNCIA</b> MODELO - UNIDADE SANITARIA SEDE DE MODELO
 MASCULINO	<b>JOSE ANTONIO SILVA - 164445 - 44 ANOS</b> 27/04/2018 I64 ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO <b>PROBLEMAS</b> NENHUM PROBLEMA REGISTRADO <b>MUNICÍPIO E/OU UNIDADE DE REFERÊNCIA</b> MODELO - UNIDADE SANITARIA SEDE DE MODELO
 FEMININO	<b>MARIA FERNANDA SOARES - 256225 - 43 ANOS</b> 23/04/2018 I64 ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO <b>PROBLEMAS</b> NENHUM PROBLEMA REGISTRADO <b>MUNICÍPIO E/OU UNIDADE DE REFERÊNCIA</b> MODELO - UNIDADE SANITARIA SEDE DE MODELO

Fonte: Elaborado pelo Autor

## 5.6 APRESENTAR OS INDICADORES DAS DOENÇAS CEREBROVASCULARES

Outro relatório importante é o indicador de cada CID10 do grupo de doenças cerebrovasculares. Esta funcionalidade apresenta uma tela com uma listagem e a quantidade de casos por CID10, a média de idade e a distribuição por sexo. Ao clicar na linha, é apresentada outra tela com os gráficos de casos por período, faixa etária e sexo novamente, mas em forma mais visual. As telas 14 e 15 mostram como ficou a implementação da funcionalidade.

## Tela 14 - DDCV - Doenças Cerebrovasculares por CID10 – Listagem

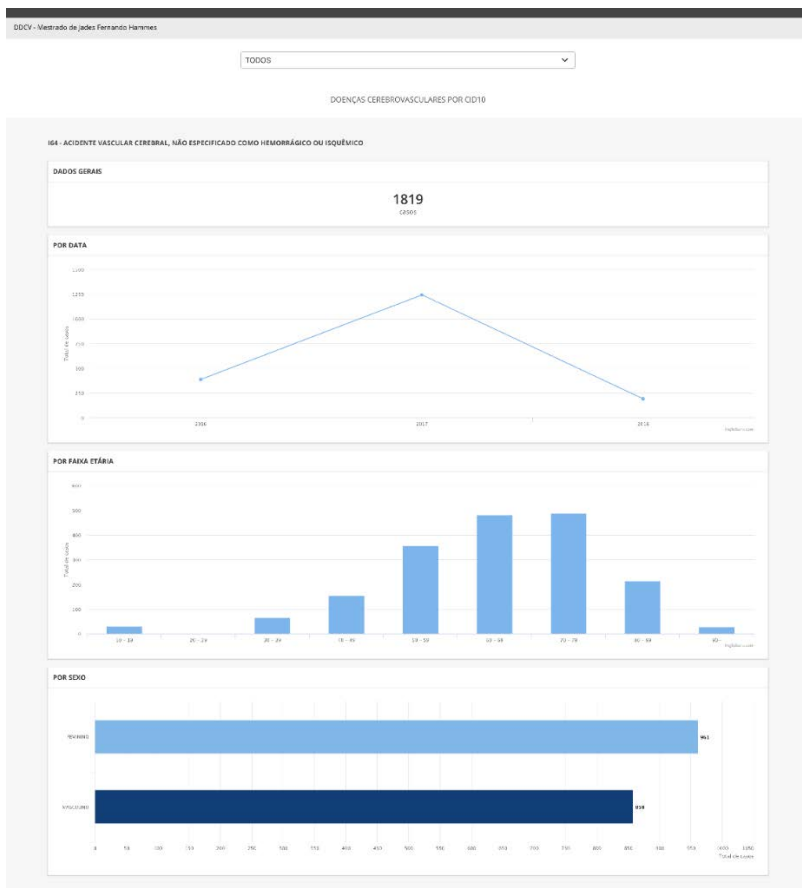
DDCV - Moszardo de Jádys Fernando Nunes

## DOENÇAS CEREBROVASCULARES POR CID10

<b>I64 - ACIDENTE VASCULAR CEREBRAL, NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO</b>		
1819 CASOS	Medida de idade: 65	MASCULINO 47,17% FEMININO 52,83%
<b>I694 - SEQUELAS DE ACIDENTE VASCULAR CEREBRAL NÃO ESPECIFICADO COMO HEMORRÁGICO OU ISQUÊMICO</b>		
9 CASOS	Medida de idade: 70	MASCULINO 77,78% FEMININO 22,22%
<b>I693 - SEQUELAS DE INFARTO CEREBRAL</b>		
4 CASOS	Medida de idade: 66	MASCULINO 50,00% FEMININO 50,00%
<b>I692 - OCLUSÃO E ESTENOSE DA ARTÉRIA CARÓTIDA</b>		
3 CASOS	Medida de idade: 65	MASCULINO 100,00% FEMININO 0,00%
<b>G498 - OUTROS ACIDENTES ISQUÊMICOS CEREBRAIS TRANSITÓRIOS E SÍNDROMES CORRELATAS</b>		
2 CASOS	Medida de idade: 88	MASCULINO 100,00% FEMININO 0,00%
<b>I671 - ANEURISMA CEREBRAL NÃO ROTO</b>		
1 CASOS	Medida de idade: 57	MASCULINO 0,00% FEMININO 100,00%
<b>G499 - ISQUEMIA CEREBRAL TRANSITÓRIA NÃO ESPECIFICADA</b>		
1 CASOS	Medida de idade: 90	MASCULINO 100,00% FEMININO 0,00%

Fonte: Elaborado pelo Autor

## Tela 15 - DDCV - Doenças Cerebrovasculares por CID10 – Detalhe.



Fonte: Elaborado pelo Autor

### 5.7 APRESENTAR UMA ANÁLISE PREDITIVA COM CIDADÃOS QUE POSSUEM RISCOS DE DESENVOLVEREM DOENÇAS CEREBROVASCULARES AOS PROFISSIONAIS DE SAÚDE

Para aplicação do algoritmo de *machine learning*, foi utilizado o Python v3.7.0. Inicialmente foi exportada a tabela de predição que foi

carregada com os dados do e-SUS AB para o DDCV, após o processamento apresentado anteriormente. Foi utilizado um arquivo de *Comma-Separated Values* (CSV) tanto para levar os dados para o Python quanto para retornar o resultado com a probabilidade para a aplicação DDCV e atualizada a tabela no campo probabilidade.

Na preparação para receber os dados de origem, foram definidas as colunas de entrada, ou seja, as variáveis que o algoritmo de predição vai usar para aprender o padrão e para fazer a probabilidade posteriormente. Esse ajuste inicial pode ser visto na Figura 11.

**Figura 11** - Preparação para receber o arquivo de entrada

```
columns = ["id", "cns", "nome", "idade", "sexo", "raca_cor", "obeso", "ex_obeso", "fumante", "ex_fumante", "alcool", "ex_alcool", "drogas", "ex_drogas", "hipertensao_arterial", "diabetes", "dcv", "gestante"]
dtype = {name: int for name in columns}
dtype['cns'] = str
dtype['nome'] = str
dtype['sexo'] = str
dtype['raca_cor'] = str

NO_DCV_CLASS = 0
DCV_CLASS = 1

BINARY = True
```

Fonte: Elaborado pelo Autor

Em seguida, foi realizada a leitura do arquivo, neste caso um CSV específico para a predição Figura 11 e como já era previsto, conforme mencionado na metodologia, foi preciso normalizar a idade por meio de Conjuntos Fuzzy (Figura 12 e 13):

**Figura 12** - Comando para leitura do arquivo CSV no Python

```
df = pd.read_csv('predicao_201809021851.csv', sep=',', header=0, index_col=None, dtype=dtype)
```

Fonte: Elaborado pelo Autor



**Figura 13** - Código em Python para normalização da idade

```

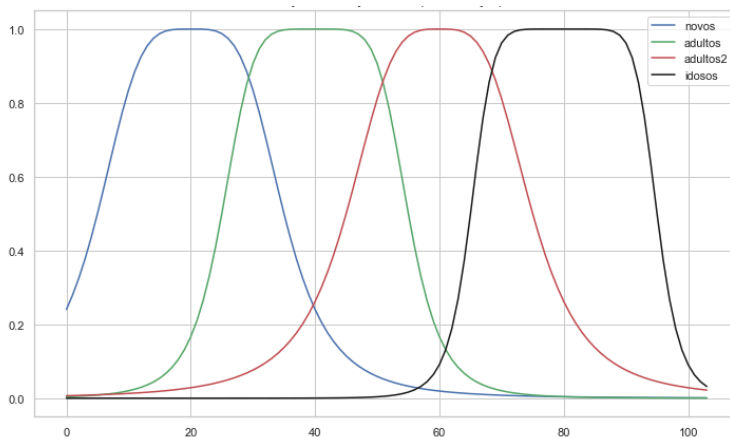
x_age = np.arange(0, df.idade.max(), 1)

young = fuzz.gbellmf(x_age,a=15, b=2, c=20)
mid = fuzz.gbellmf(x_age,a=15, b=2.8, c=40)
old = fuzz.gbellmf(x_age,a=15, b=1.8, c=60)
very_old = fuzz.gbellmf(x_age,a=15, b=4., c=80)

fig = plt.figure(figsize=(10,6))
plt.plot(x_age, young, 'b', label='novos')
plt.plot(x_age, mid, 'g', label='adultos')
plt.plot(x_age, old, 'r', label='adultos2')
plt.plot(x_age, very_old, 'k', label='idosos')
plt.legend()
plt.tight_layout()
plt.title("Conjuntos Fuzzy de idade (normalização)")
plt.savefig('normalizacao_idade', tight_layout=True)

```

Fonte: Elaborado pelo Autor

**Figura 14** - Conjuntos Fuzzy de Idade (normalização)

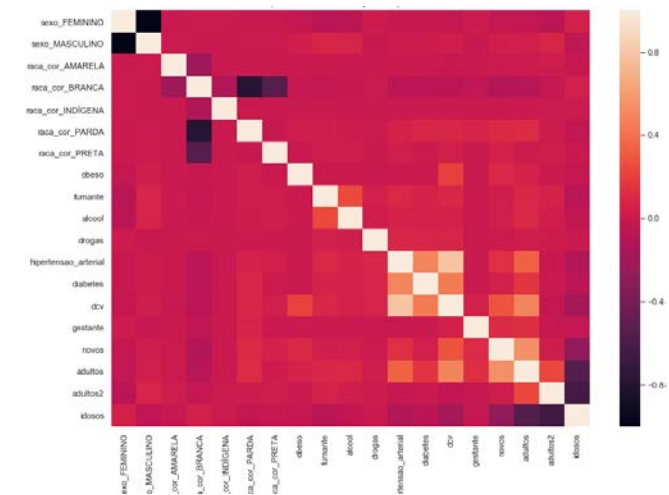
Fonte: Elaborado pelo Autor

Para a execução do algoritmo, os dados preparados foram convertidos em  $X_{train}$  e  $X_{test}$  e foram executados usando regressão logística otimizada por um algoritmo de otimização, Gradiente Descendente Estocástico (SGD). O registro que possui  $DCV = true$  serve para demonstrar para o algoritmo quais fatores contribuem para a

doença e o grupo de pessoas com mais idade e sem a doença, serve para posicionar o algoritmo de quais características levam a uma baixa chance de ter a doença.

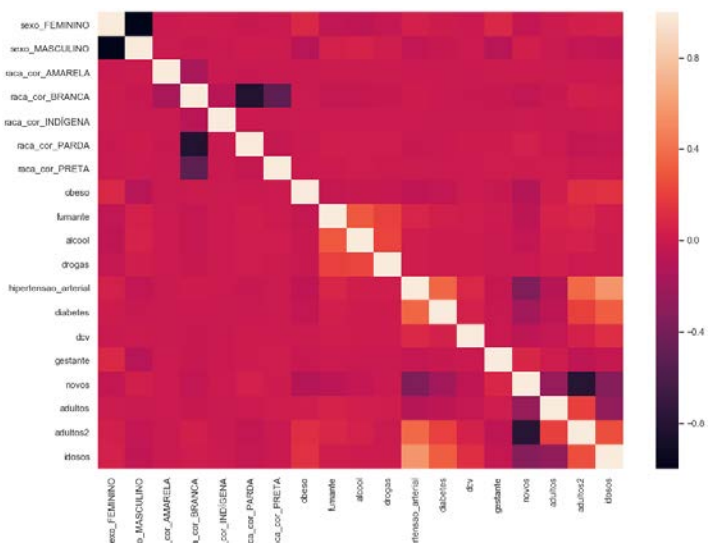
É possível observar a correlação entre as variáveis de treinamento e de teste para descobrir falhas nos dados, pois se elas diferem muito, seria um problema para o algoritmo, mas neste caso, elas são semelhantes. Correlação é a dependência ou associação de qualquer relação estatística (causal ou não causal) entre duas variáveis. Se a correlação for positiva, elas são diretamente proporcionais, se for negativa, são inversamente proporcionais e se for zero, elas não têm relação alguma. Os mapas de calor dessa correlação no aprendizado (**Erro! Fonte de referência não encontrada.**) e no teste (**Erro! Fonte de referência não encontrada.**) mostram de forma visual a relação negativa entre sexo masculino e feminino, o que valida que os dados não têm erro em relação a essas colunas. Já DCV apresentou correlação positiva com hipertensão, com isso, é esperado que essas duas variáveis impactem de forma semelhante no resultado tanto para o sexo masculino quanto feminino.

**Figura 15** - Mapa de calor da correlação no aprendizado



Fonte: Elaborado pelo Autor

**Figura 16** - Mapa de calor da correlação no teste



Fonte: Elaborado pelo Autor

Por fim, preparou-se o algoritmo treinado para executar em todo o conjunto de dados e obter o resultado de probabilidade, conforme comando na **Erro! Fonte de referência não encontrada..**

**Figura 17** - Execução do algoritmo para obter a probabilidade em todo o conjunto de dados

```
R = data.copy()
del R['y']

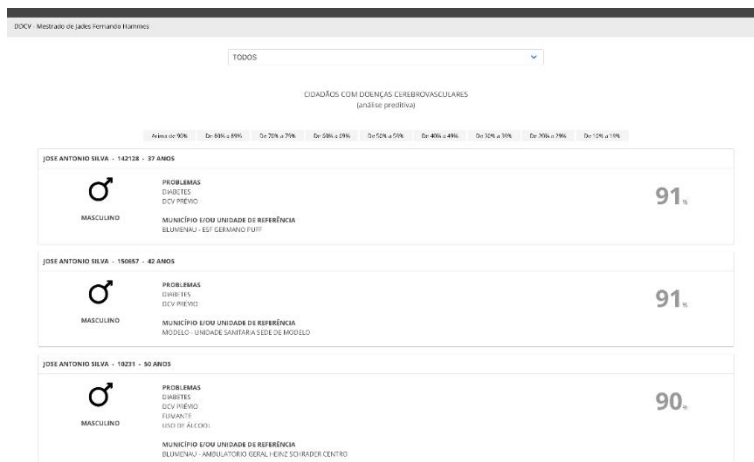
def notetrigger(row):
    count = ""
    for c, col in enumerate(row):
        if col == 1:
            count+=R.columns[c]+" & "
    return count[:-3]

y_r = clf_grid.predict(R)
y_proba = clf_grid.predict_proba(R)

R['triggers'] = R.apply(notetrigger, axis=1)
R['cns'] = df.cns[R.index]
R['idade'] = df.idade[R.index]
R['dcv_proba'] = y_proba[:,1]
R.to_csv(path_or_buf="result.csv", index=False, sep=';', columns=['cns', 'idade',
'dcv', 'dcv_proba', 'triggers'], float_format='%.2f')
```

Fonte: Elaborado pelo Autor

**Figura 18** - DDCV - Risco de Doenças Cerebrovasculares (análise\_preditiva)



Fonte: Elaborado pelo Autor



## 6 CONSIDERAÇÕES FINAIS

O desenvolvimento da ferramenta de dashboard e análise preditiva do presente estudo possibilitou verificar que é possível oferecer aos profissionais de saúde e aos gestores de saúde municipais, estaduais e federal relatórios demográficos georeferenciados para visualizar as condições de saúde no território, assim como ter relatórios de acompanhamento dos cidadãos com doenças cerebrovasculares e visualizar indicadores de cada doença na sua região de saúde. O relatório de Risco de Doenças Cerebrovasculares (Análise Preditiva), apresenta para o cidadão, através do seu profissional de saúde referência no território, a chance dele ser surpreendido por uma patologia de DCBV, pois foi aplicado inteligência artificial para prever uma doença de acordo com seus parâmetros de riscos de saúde. Isto agrega não só um apoio a gestão do cuidado como pode diminuir o número de casos, aumentando consequentemente a qualidade de vida do cidadão.

Além disso, um relatório de riscos de doenças cerebrovasculares de cidadãos propensos a obter essas patologias, baseados em machine learning e análise preditiva, possibilitará que ações de prevenção sejam executadas antes da ocorrência da condição clínica, diminuindo custos de internação hospitalar, fornecer subsídios para os profissionais de saúde na tomada de decisão e planejamento em saúde e, principalmente, poderá salvar vidas e/ou evitar piora na qualidade de vida.

Nesse sentido, a utilização de recursos tecnológicos como o que foi desenvolvido neste trabalho, tem fundamental importância na gestão e prevenção de doenças com chance de sobrevida ainda maior e auxiliar na gestão do cuidado.

Contudo, importante ressaltar que somente dados da atenção básica não são suficientes para fazer a predição com alto índice de confiabilidade uma vez que os hospitais e UPAs têm mais dados relacionados a doenças cerebrovasculares. A interoperabilidade das informações em saúde entre os diferentes níveis dos serviços de saúde seja públicos ou privados, pode fortalecer a confiabilidade na predição de condições de saúde.

O uso de dados fictícios pode não representar a realidade da situação de saúde do município, estado ou Brasil, mas o grande propósito do estudo é a validação da predição dos dados, o que aconteceu, pois, o algoritmo aprendeu com a base histórica e fez a

predição dos riscos de acontecimentos de doenças cerebrovasculares nos cidadãos fictícios. Tal análise fornece maiores subsídios para que o algoritmo seja aplicado em diferentes realidades dos municípios com base em dados reais

Alguns fatores importantes para as condições de risco de DCV não puderam ser avaliadas na análise preditiva, pois não existem no modelo de informação do e-SUS AB CDS, como por exemplo, hábitos alimentares, prática de atividades físicas e uso de anticoncepcional.

Como perspectiva futura, seria de grande valia a aplicação do estudo numa base de produção, certamente com acesso restrito a profissionais de saúde destacando-se o quão delicado é utilizar somente do aprendizado de máquina para informar ao cidadão a chance de adquirir um possível AVC ou outro CID10 do mesmo grupo sensível as condições de saúde. Esse artefato não deve ser usado como única fonte de alertar riscos da doença, mas sim como mais uma forma de analisar as chances de adquirir qualquer doença cerebrovascular, não supre, de forma alguma, o conjunto de fatores e avaliações clínicas dos profissionais de saúde de forma contínua.



## REFERÊNCIAS

ALFRADIQUE, M. E. et al. Internações por condições sensíveis à atenção primária: a construção da lista brasileira como ferramenta para medir o desempenho do sistema de saúde (Projeto ICSAP - Brasil). **Cad. Saúde Pública**, Rio de Janeiro, v. 25, n. 6, p. 1337-1349. 2009. Disponível em:

<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-311X2009000600016&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2009000600016&lng=en&nrm=iso)>. Acesso em: 08 fev.2018.

BOZA. M. *SDG DASHBOARDS*. The role of information tools in the implementation of the 2030 Agenda. learning from experiences in Latin America. 2017. Disponível em:

<<https://pt.slideshare.net/UNDPasiapacific/sdg-dashboards-the-role-of-information-and-information-tools-in-the-implementation-of-the-2030-agenda-learning-from-experiences-in-latin-america>>. Acesso em: 23 mar. 2018.

BRASIL. Acidente vascular cerebral (AVC). 2012. Disponível em: <<http://www.brasil.gov.br/saude/2012/04/acidente-vascular-cerebral-avc>>. Acesso em: 23 mar. 2018.

BRASIL. Conselho Nacional de Saúde. **Resolução nº 466, de 12 de dezembro de 2012**. Aprova normas regulamentadoras de pesquisas envolvendo seres humanos. Brasília: Diário Oficial da União, 2013.

BRASIL. Ministério da Saúde. **Diretoria de Atenção Básica. e-SUS Atenção Básica**. 2016a.. Disponível em: <<http://dab.saude.gov.br/portaldab/esus.php>>. Acesso em: 23 mai. 2016.

BRASIL. Ministério da Saúde. Ministério da Saúde. **Estratégia e-Saúde para o Brasil**. Brasília. 2014. 136 p.

BRASIL. **Portaria nº 1.412, de 10 de julho de 2013**. Institui o Sistema de Informação em Saúde para a Atenção Básica (SISAB). 2013. Disponível em:

<[http://bvsmis.saude.gov.br/bvsmis/saudelegis/gm/2013/prt1412\\_10\\_07\\_2013.html](http://bvsmis.saude.gov.br/bvsmis/saudelegis/gm/2013/prt1412_10_07_2013.html)>. Acesso em: 23 mai. 2016.

BRASIL. **Portaria nº 221, de 17 de abril de 2008.** Portaria Nº 221, de 17 de Abril de 2008. Lista Brasileira de Internações por Condições Sensíveis à Atenção Primária. 2008. Disponível em: <[http://bvsm.s.saude.gov.br/bvs/saudelegis/sas/2008/prt0221\\_17\\_04\\_2008.html](http://bvsm.s.saude.gov.br/bvs/saudelegis/sas/2008/prt0221_17_04_2008.html)>. Acesso em: 12 mar. 2017..

BRASIL. **Prontuário Eletrônico chega a 57 milhões de brasileiros.** 2016b. Disponível em: <<http://datasus.saude.gov.br/noticias/atualizacoes/1073-prontuario-eletronico-chega-a-57-milhoes-de-brasileiros>>. Acesso em: 23 mar. 2018.

CHIAVEGATTO FILHO, A. D. P. Uso de big data em saúde no Brasil: perspectivas para um futuro próximo. **Epidemiol. Serv. Saúde**, Brasília, v. 24, n. 2, p. 325-332, June 2015. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2237-96222015000200325&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222015000200325&lng=en&nrm=iso)>. Acesso em: 13 Set. 2018. <http://dx.doi.org/10.5123/S1679-49742015000200015>.

FALCONI, V. O. Verdadeiro Poder: práticas de gestão que conduzem a resultados revolucionários. 1ª Edição. **Nova Lima: INDG Tecnologia e Serviços Ltda**, 2009.

FERREIRA, M. DIAS, B. M.; MISHIMA, S. M. Internações por condições sensíveis: possibilidade de avaliação na atenção básica. **Rev. Eletr. Enf. [Internet]**. 2012. Disponível em: <<http://www.fen.ufg.br/revista/v14/n4/v14n4a03.htm>>. Acesso em: 08 fev.2018.

FEW, S. **Dashboard Confusion**. 2004. Disponível em: <[http://www.perceptualedge.com/articles/ie/dashboard\\_confusion.pdf](http://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf)>. Acesso em: 08 fev.2018.

GHAZISAEIDI, M. et al. Development of performance dashboards in healthcare sector: key practical issues. **Acta Informatica Medica**, v. 23, n. 5, p. 317, 2015. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4639357/pdf/AIM-23-317.pdf>> Acesso em: 23 mar. 2018.

GOOGLE. **Maps Platform**. Disponível em:  
<<https://developers.google.com/maps/documentation/javascript/heatmap-layer>> Acesso em: 23 mar. 2018.

GOOGLE. **Maps Platform**. Disponível em:  
<https://developers.google.com/maps/documentation/geocoding/start>.  
Acesso em: 23 mar. 2018b.

GRAHAM, B. et al. **Using Data Mining to Predict Hospital Admissions From the Emergency Department**. IEEEAccess. V.6, 2018. p 10458 - 10468. <Disponível em:  
[http://www.ieee.org/publications\\_standards/publications/rights/index.html](http://www.ieee.org/publications_standards/publications/rights/index.html)> Acesso em: 23 mar. 2018.

HAN, J.; KAMBER, M.; ZEI, J. **Data mining: concepts and techniques**. Third edition. Amsterdam: Elsevier, 2012.740p.

HIMSS. Transforming health through information and technology. **Executive Dashboard Development Guide** - Healthcare Information and Management Systems Society. April. 2009.

IHME. **Institute for Health Metrics and Evaluation**. 2016. Disponível em: <<http://www.healthdata.org/united-states?language=129>>. Acesso em: 08 fev.2018.

LOTUFO, P. A.; BENSENOR, I. J. B. Raça e mortalidade cerebrovascular no Brasil. **Rev. Saúde Pública**, São Paulo, v. 47, n. 6, p. 1201-1204. 2013. Disponível em:  
<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0034-89102013000601201&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102013000601201&lng=en&nrm=iso)>. Acesso em: 12 mar. 2018.

MORAES, J. R. **Modelo de regressão logística binária**. 2016. Disponível em: < <http://www.estadisticacomr.uff.br/?p=598>> Acesso em: 23 mar. 2018.

MULLER, E. P. L.; CUBAS, M. R.; BASTOS, L. C. Georreferenciamento como instrumento de gestão em unidade de saúde da família. **Rev. bras. enferm.**, Brasília , v. 63, n. 6, p. 978-982. 2010. Disponível em:

<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0034-71672010000600017&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-71672010000600017&lng=en&nrm=iso)>. Acesso em: 23 mar. 2018.

PEREIRA, F. J. R.; SILVA, C. C.; LIMA NETO, E. A. Perfil das Internações por Condições Sensíveis à Atenção Primária subsidiando ações de saúde nas regiões brasileiras. **Saúde debate**, Rio de Janeiro, v. 39, n. 107, p. 1008-1017. 2015. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-11042015000401008&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-11042015000401008&lng=en&nrm=iso)>. Acesso em: 23 mar. 2018.

PRESSMAN, R.; MAXIM, B. **Engenharia de Software-8ª Edição**. McGraw Hill Brasil, 2016.

SA, J. H. G. et al. Geohealth: A Georeferenced system for health data analysis in primary care. **IEEE Latin America Transactions**, v. 10, n. 1, p. 1352-1356, 2012.

SES-SC. Secretaria de Estado da Saúde de Santa Catarina. **Internações por Condições Sensíveis à Atenção Primária – ICSAP**. 2014. Disponível em:

<[http://portalses.saude.sc.gov.br/index.php?option=com\\_content&view=article&id=2622&Itemid=489](http://portalses.saude.sc.gov.br/index.php?option=com_content&view=article&id=2622&Itemid=489)>. Acesso em: 23 mar. 2018..

SILVEIRA, I. H.; OLIVEIRA, B. F. A.; JUNGER, W. L. Utilização do Google Maps para o georreferenciamento de dados do Sistema de Informações sobre Mortalidade no município do Rio de Janeiro, 2010-2012\*. **Epidemiol. Serv. Saúde**, Brasília, v. 26, n. 4, p. 881-886. 2017. Disponível em

<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2237-96222017000400881&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222017000400881&lng=en&nrm=iso)>. Acesso em: 23 mar. 2018.

SOUZA, O. N. **Introdução à Teoria dos Conjuntos Fuzzy**. Disponível em: <<http://www.ime.unicamp.br/~valle/PDFfiles/osmar10.pdf>> Acesso em: 23 mar. 2018.

TIC SAÚDE 2016. **Pesquisa Sobre o Uso das Tecnologias de Informação e Comunicação nos Estabelecimentos de Saúde Brasileiros**. Disponível em:

<[http://cetic.br/media/docs/publicacoes/2/tic\\_saude\\_2016\\_livro\\_eletronico.pdf](http://cetic.br/media/docs/publicacoes/2/tic_saude_2016_livro_eletronico.pdf)>. Acesso em: 23 mar. 2018..

VILLELA, P. B.; KLEIN, C. H.; OLIVEIRA, G. M. M. Evolução da Mortalidade por Doenças Cerebrovasculares e Hipertensivas no Brasil entre 1980 e 2012. **Arq. Bras. Cardiol.**, São Paulo , v. 107, n. 1, p. 26-32, July 2016 . Disponível em:

<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0066-782X2016004000026&lng=en&nrm=isso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2016004000026&lng=en&nrm=isso)>. Acesso em: 12 dez. 2017.

WERNECK, G. L. Georeferenced data in epidemiologic research. **Ciênc. saúde coletiva**, Rio de Janeiro , v. 13, n. 6, p. 1753-1766, 2008. Disponível em:

<[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-81232008000600010&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232008000600010&lng=en&nrm=iso)>. Acesso em: 23 mar. 2018.

ZEVIANI, W. M.; FERREIRA, E. V. **Métodos de Gradiente Descendente**. 2017. Disponível em:

<<http://leg.ufpr.br/~walmes/ensino/ML/tutorials/02-gradient-methods.html#gradiente-descendente-estocastico>>. Acesso em: 23 mar. 2018.



## **APÊNDICES**





## APÊNDICE A – FICHAS DE CADASTROS

### Ficha de Cadastro Individual

Nome Campo	Tipo	Unidade Medida
CNES	-	-
Data	Data	-

Nome Campo	Tipo	Unidade Medida
CNS do cidadão	Numérico	-
Cidadão é o responsável familiar?	Boolean	-
CNS do responsável familiar	Numérico	
Nome completo	Texto	-
Nome social	Texto	-
Data de nascimento	Data	-
Sexo	-	-
Raça / Cor	-	-
Data do óbito	Data	-
Está gestante	Boolean	-
Sobre seu peso, você se considera	-	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
Está fumante?	Boolean	-
Faz uso de álcool?	Boolean	-
Faz uso de outras drogas?	Boolean	-
Tem hipertensão arterial?	Boolean	-
Tem diabetes?	Boolean	-
Teve AVC / derrame?	Boolean	-

### **Ficha de Cadastro Domiciliar e Territorial**

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CNES	-	-
Data	Data	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CEP	Numérico	-
Município	-	-
UF	-	-
Bairro	Alfanumérico	-
Tipo de logradouro	-	-
Nome do logradouro	Alfanumérico	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
Nº	Alfanumérico	-
Sem número	Boolean	-
Complemento	Alfanumérico	-
CNS do responsável	Numérico	-

### **Ficha de Atendimento Individual**

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CNES	-	-
Data	Data	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CNS do cidadão	Numérico	-
Racionalidade em saúde (Exceto alopátia / convencional)	-	-
Perímetro cefálico (cm)	Numérico	Cm
Peso (kg)	Numérico	Kg
Altura (cm)	Numérico	Cm
DUM	Data	-
Gravidez planejada	Boolean	-
Problemas / Condições avaliadas	-	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CID10 - 01	-	-
CID10 - 02	-	-

### **Ficha de Visita Domiciliar e Territorial**

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CNES	-	-
Data	Data	-

<b>Nome Campo</b>	<b>Tipo</b>	<b>Unidade Medida</b>
CNS do cidadão	Numérico	-
Acompanhamento	-	-
Peso (kg)	Numérico	Kg
Altura (cm)	Numérico	Cm

## **ANEXOS**



## ANEXO A

```

2 ProcessamentoController.java 1.09 KB
3
4 package br.ufsc.ddc.processamento;
5
6 import java.util.Date;
7 import java.util.HashMap;
8 import java.util.Map;
9
10 import org.springframework.beans.factory.annotation.Autowired;
11 import org.springframework.stereotype.Controller;
12 import org.springframework.web.bind.annotation.GetMapping;
13
14 import br.ufsc.ddc.entidades.Historico;
15 import br.ufsc.ddc.services.CidIdService;
16 import br.ufsc.ddc.services.GeocodeService;
17
18 @Controller
19 public class ProcessamentoController {
20
21     @Autowired
22     private CadastroIndividualProcessamentoService cadastroIndividualProcessamentoService;
23
24     @Autowired
25     private CadastroDomiciliarProcessamentoService cadastroDomiciliarProcessamentoService;
26
27     @Autowired
28     private AtendimentoIndividualProcessamentoService atendimentoIndividualProcessamentoService;
29
30     @Autowired
31     private VisitaDomiciliarProcessamentoService visitaDomiciliarProcessamentoService;
32
33     @Autowired
34     private CidIdService cidIdService;
35
36     @Autowired
37     private PredicaoService predicaoService;
38
39     @Autowired
40     private GeocodeService geocodeService;
41
42     private Map<String, Historico> cache = new HashMap<String, Historico>();
43
44     @GetMapping("/processamento")
45     public void indexAction() throws InterruptedException {
46         System.out.println("Inicio: " + new Date());
47         System.out.println("Inicio Cadastro Individual: " + new Date());
48         this.cadastroIndividualProcessamentoService.processar();
49         System.out.println("Inicio Cadastro Domiciliar: " + new Date());
50         this.cadastroDomiciliarProcessamentoService.processar(cache);
51         System.out.println("Inicio Atendimento Individual: " + new Date());
52         this.atendimentoIndividualProcessamentoService.processar(cache);
53         System.out.println("Inicio Visita Domiciliar: " + new Date());
54         this.visitaDomiciliarProcessamentoService.processar(cache);
55         System.out.println("Inicio Predicoes: " + new Date());
56         this.predicaoService.processar();
57         System.out.println("Inicio CID10: " + new Date());
58         this.cidIdService.processar();
59         System.out.println("Fim: " + new Date());
60         this.geocodeService.processar();
61     }
62 }

```

```

1  package br.ufsc.ddc.processamento;
2
3  import java.time.temporal.ChronoUnit;
4  import java.util.List;
5  import java.util.Map;
6
7  import javax.persistence.EntityManager;
8  import javax.persistence.PersistenceContext;
9
10 import org.springframework.beans.factory.annotation.Autowired;
11 import org.springframework.stereotype.Service;
12 import org.springframework.transaction.annotation.Transactional;
13
14 import com.querydsl.jpa.impl.JPAQuery;
15
16 import br.ufsc.ddc.entities.AtendimentoIndividual;
17 import br.ufsc.ddc.entities.Historico;
18 import br.ufsc.ddc.entities.AtendimentoIndividual;
19 import br.ufsc.ddc.services.CidIdService;
20
21 @Service
22 @Transactional
23 public class AtendimentoIndividualProcessamentoService {
24
25     @PersistenceContext
26     private EntityManager em;
27
28     @Autowired
29     private CidIdService cidIdService;
30
31     @Autowired
32     private FaixaEtariaService faixaEtariaService;
33
34     public void processar(Map<String, Historico> cache) {
35         @AtendimentoIndividual atendimentoIndividual = @AtendimentoIndividual.atendimentoIndividual;
36         JPAQuery<AtendimentoIndividual> q = new JPAQuery<AtendimentoIndividual>(this.em);
37         q.from(AtendimentoIndividual);
38         q.where(AtendimentoIndividual.cns.isNotNull());
39         List<AtendimentoIndividual> items = q.fetch();
40
41         for (AtendimentoIndividual atendimentoIndividual : items) {
42             Historico historicoCadIndividual = cache.get(atenidmentoIndividual.getCns());
43
44             if (historicoCadIndividual != null) {
45                 Historico historico = new Historico();
46                 historico.setCns(atenidmentoIndividual.getCns());
47                 historico.setData(atenidmentoIndividual.getDataAtendimentoIndividual());
48                 historico.setAno(atenidmentoIndividual.getDataAtendimentoIndividual().getYear());
49                 historico.setNomeHistoricoCadIndividual(getNome());
50                 historico.setNomeSocial(historicoCadIndividual.getNomeSocial());
51                 historico.setCns(atenidmentoIndividual.getCns());
52                 historico.setCns(atenidmentoIndividual.getCns());
53                 historico.setCnsResponsavel(historicoCadIndividual.getCnsResponsavel());
54                 historico.setDataNascimento(atenidmentoIndividual.getDataNascimento());
55                 historico.setDataIdade(ChronoUnit.YEARS.between(atenidmentoIndividual.getDataNascimento(), atendimentoIndividual.getDataAtendimentoIndividual()));
56                 historico.setFaixaEtaria(this.faixaEtariaService.getFaixaEtaria(historico.getIdade()));
57                 historico.setSexo(atenidmentoIndividual.getSexo());
58                 historico.setRacaCor(historicoCadIndividual.getRacaCor());
59                 historico.setGestante(atenidmentoIndividual.getGestante() != null ? atendimentoIndividual.getGestante() : false);
60                 historico.setObeso(atenidmentoIndividual.getObeso());
61                 historico.setImune(atenidmentoIndividual.getImune());
62                 historico.setAlcoo(atenidmentoIndividual.getAlcoo() != null ? atendimentoIndividual.getAlcoo() : false);
63                 historico.setDrogas(atenidmentoIndividual.getDrogas() != null ? atendimentoIndividual.getDrogas() : false);
64                 historico.setHipertensaoArterial(atenidmentoIndividual.getHipertensaoArterial() != null ? atendimentoIndividual.getHipertensaoArterial() : false);
65                 historico.setDiabetes(atenidmentoIndividual.getDiabetes() != null ? atendimentoIndividual.getDiabetes() : false);
66                 historico.setCidIdService.hasCidV(atenidmentoIndividual.getCids());
67                 historico.setCep(historicoCadIndividual.getCep());
68                 historico.setIbge(historicoCadIndividual.getIbge());
69                 historico.setMunicipio(historicoCadIndividual.getMunicipio());
70                 historico.set(historicoCadIndividual.getI());
71                 historico.setTppolograduouo(historicoCadIndividual.getTppolograduouo());
72                 historico.setLogradouro(historicoCadIndividual.getLogradouro());
73                 historico.setBairro(historicoCadIndividual.getBairro());
74                 historico.setNumero(historicoCadIndividual.getNumero());
75
76                 List<String> cids = cidIdService.getCids(atenidmentoIndividual.getCids());
77
78                 for (int i = 0; i < cids.size(); i++) {
79                     if (i == 0) {
80                         historico.setCidId_1(cids.get(i));
81                     }
82                     if (i == 1) {
83                         historico.setCidId_2(cids.get(i));
84                     }
85                     if (i == 2) {
86                         historico.setCidId_3(cids.get(i));
87                     }
88                     if (i == 3) {
89                         historico.setCidId_4(cids.get(i));
90                     }
91                     if (i == 4) {
92                         historico.setCidId_5(cids.get(i));
93                     }
94                 }
95                 this.em.merge(historico);
96             }
97             this.em.clear();
98         }
99     }
100 }

```



```

1 package br.ufsc.ddc.processamento;
2
3 import java.util.HashMap;
4 import java.util.List;
5 import java.util.Map;
6
7 import javax.persistence.EntityManager;
8 import javax.persistence.PersistenceContext;
9
10 import org.springframework.beans.factory.annotation.Autowired;
11 import org.springframework.stereotype.Service;
12 import org.springframework.transaction.annotation.Transactional;
13
14 import com.querydsl.jpa.impl.JPAQuery;
15
16 import br.ufsc.ddc.entidades.CadastroDomiciliar;
17 import br.ufsc.ddc.entidades.Historico;
18 import br.ufsc.ddc.entidades.ICadastroDomiciliar;
19 import br.ufsc.ddc.entidades.IHistorico;
20 import br.ufsc.ddc.servicos.MunicipioService;
21
22 @Service
23 @Transactional
24 public class CadastroDomiciliarProcessamentoService {
25
26     @PersistenceContext
27     private EntityManager em;
28
29     @Autowired
30     private MunicipioService municipioService;
31
32     private Map<String, CadastroDomiciliar> cacheCadastroDomiciliar = new HashMap<String, CadastroDomiciliar>();
33
34     public void processar(Map<String, Historico> cache) {
35         @Historico @Historico = @Historico;
36         JPAQuery<Historico> q = new JPAQuery<Historico>(this.em);
37         q.from(q@Historico);
38         List<Historico> items = q.fetch();
39
40         for (Historico historico : items) {
41             CadastroDomiciliar cadastroDomiciliar = this.getCadastroDomiciliar(historico.getCnsResponsavel() != null ? historico.getCnsRes
42             if (cadastroDomiciliar != null) {
43                 historico.setCep(cadastroDomiciliar.getCep());
44                 historico.setInscricao(cadastroDomiciliar.getInscricao());
45                 historico.setMunicipio(this.municipioService.getMunicipioByIBGE(cadastroDomiciliar.getInscricao()));
46                 historico.setTabelaLogradouro(cadastroDomiciliar.getTabelaLogradouro());
47                 historico.setLogradouro(cadastroDomiciliar.getLogradouro());
48                 historico.setBairro(cadastroDomiciliar.getBairro());
49                 historico.setNumero(cadastroDomiciliar.getNumeroDomicilio());
50                 historico.setUf(cadastroDomiciliar.getUf());
51                 cache.put(historico.getCns(), historico);
52                 this.em.merge(historico);
53                 this.em.flush();
54                 this.em.clear();
55             }
56         }
57     }
58
59     public CadastroDomiciliar getCadastroByCnsResponsavel(String cnsResponsavel) {
60         if (this.cacheCadastroDomiciliar.containsKey(cnsResponsavel)) {
61             return this.cacheCadastroDomiciliar.get(cnsResponsavel);
62         } else {
63             CadastroDomiciliar qCadastroDomiciliar = @CadastroDomiciliar;
64             JPAQuery<CadastroDomiciliar> q = new JPAQuery<CadastroDomiciliar>(this.em);
65             q.from(qCadastroDomiciliar);
66             q.where(qCadastroDomiciliar.cnsResponsavel.eq(cnsResponsavel));
67             CadastroDomiciliar cadastroDomiciliar = q.fetchFirst();
68             this.cacheCadastroDomiciliar.put(cnsResponsavel, cadastroDomiciliar);
69             return cadastroDomiciliar;
70         }
71     }
72 }

```

```

1 package br.ufsc.ddc.processamento;
2
3 import java.time.temporal.ChronoUnit;
4 import java.util.List;
5
6 import javax.persistence.EntityManager;
7 import javax.persistence.PersistenceContext;
8
9 import org.springframework.beans.factory.annotation.Autowired;
10 import org.springframework.stereotype.Service;
11 import org.springframework.transaction.annotation.Transactional;
12
13 import com.querydsl.jpa.impl.JPAQuery;
14
15 import br.ufsc.ddc.entities.CadastroIndividual;
16 import br.ufsc.ddc.entities.Historico;
17 import br.ufsc.ddc.entities.Municipio;
18 import br.ufsc.ddc.entities.QCadastroIndividual;
19 import br.ufsc.ddc.services.MunicipioService;
20
21 @Service
22 @Transactional
23 public class CadastroIndividualProcessamentoService {
24
25     @PersistenceContext
26     private EntityManager em;
27
28     @Autowired
29     private FaixaEtariaService faixaEtariaService;
30
31     @Autowired
32     private MunicipioService municipioService;
33
34
35     public void processar() {
36         QCadastroIndividual qCadastroIndividual = QCadastroIndividual.cadastroIndividual;
37         JPAQuery<CadastroIndividual> q = new JPAQuery<CadastroIndividual>(this.em);
38         q.from(qCadastroIndividual);
39         q.where(qCadastroIndividual.datnascimento.isNotNull());
40         q.where(qCadastroIndividual.cns.isNotNull());
41         q.where(qCadastroIndividual.cns.notEqualIgnoreCase("0"));
42         List<CadastroIndividual> itens = q.fetch();
43
44         for (CadastroIndividual cadastroIndividual : itens) {
45             Historico historico = new Historico();
46             historico.setIdorigem(cadastroIndividual.getId());
47             historico.setData(cadastroIndividual.getDataCadastroIndividual());
48             historico.setUnidadeSaude(cadastroIndividual.getUnidadeSaude());
49             historico.setCns(cadastroIndividual.getCns());
50             historico.setAno(cadastroIndividual.getDataCadastroIndividual().getYear());
51             historico.setNome(cadastroIndividual.getNomeCidadao());
52             historico.setNomeSocial(cadastroIndividual.getNomeCidadao());
53             historico.setCns(cadastroIndividual.getCns());
54             historico.setCnsResponsavel(cadastroIndividual.getCnsResponsavel());
55             historico.setDataNascimento(cadastroIndividual.getDataNascimento());
56             historico.setData(ChronoUnit.YEARS.between(cadastroIndividual.getDataNascimento(), cadastroIndividual.getDataCadastroIndividual()));
57             historico.setFaixaEtaria(faixaEtariaService.getFaixaEtaria(historico.getData()));
58             historico.setSexo(cadastroIndividual.getSexo());
59             historico.setRaccor(cadastroIndividual.getRaccor());
60             historico.setEstante(cadastroIndividual.getEstante() != null ? cadastroIndividual.getEstante() : false);
61             historico.setObeso(cadastroIndividual.getObesidade());
62             historico.setFumante(cadastroIndividual.getFumante() != null ? cadastroIndividual.getFumante() : false);
63             historico.setAlco(cadastroIndividual.getAlco() != null ? cadastroIndividual.getAlco() : false);
64             historico.setDrops(cadastroIndividual.getDrops() != null ? cadastroIndividual.getDrops() : false);
65             historico.setHipertensaoArterial(cadastroIndividual.getHipertensaoArterial() != null ? cadastroIndividual.getHipertensaoArterial() : false);
66             historico.setDiabete(cadastroIndividual.getDiabete() != null ? cadastroIndividual.getDiabete() : false);
67             historico.setVec(cadastroIndividual.getVec() != null ? cadastroIndividual.getVec() : false);
68             Municipio municipio = this.municipioService.getMunicipioByIBGE(cadastroIndividual.getIbgeUnidadeSaude());
69             historico.setIbge(cadastroIndividual.getIbgeUnidadeSaude());
70             historico.setMunicipio(municipio);
71             historico.setIrr(cadastroIndividual.getIrradiacaoUnidadeSaude());
72             historico.setU(municipio.getU().getSigla());
73             this.em.merge(historico);
74
75         }
76     }

```

```

1 package br.ufsc.ddc.processamento;
2
3 import java.time.temporal.ChronoUnit;
4 import java.util.List;
5 import java.util.Map;
6
7 import javax.persistence.EntityManager;
8 import javax.persistence.PersistenceContext;
9
10 import org.springframework.beans.factory.annotation.Autowired;
11 import org.springframework.stereotype.Service;
12 import org.springframework.transaction.annotation.Transactional;
13
14 import com.querydsl.jpa.impl.JPAQuery;
15
16 import br.ufsc.ddc.entidades.Historico;
17 import br.ufsc.ddc.entidades.Visitadomiciliar;
18 import br.ufsc.ddc.entidades.Visitadomiciliario;
19
20 @Service
21 @Transactional
22 public class VisitadomiciliarProcessamentoService {
23
24     @PersistenceContext
25     private EntityManager em;
26
27     @Autowired
28     private FaixaEtariaService faixaEtariaService;
29
30     public void processar(Map<String, Historico> cache) {
31         @Visitadomiciliar @visitadomiciliar = @Visitadomiciliar.visitadomiciliario(this.em);
32         JPAQuery<Visitadomiciliario> q = new JPAQuery<Visitadomiciliario>(this.em);
33         q.from(Visitadomiciliario);
34         q.where(Visitadomiciliario.cns.isNotNull());
35         List<Visitadomiciliario> items = q.fetch();
36
37         for (Visitadomiciliario visitadomiciliario : items) {
38             Historico historicoCadIndividual = cache.get(visitadomiciliario.getCns());
39
40             if (historicoCadIndividual != null) {
41                 Historico historico = new Historico();
42                 historico.setCodigo(visitadomiciliario.getId());
43                 historico.setData(visitadomiciliario.getDataVisitadomiciliario());
44                 historico.setData(visitadomiciliario.getDataVisitadomiciliario().getYear());
45                 historico.setNome(historicoCadIndividual.getNome());
46                 historico.setNomeSocial(historicoCadIndividual.getNomeSocial());
47                 historico.setCnes(visitadomiciliario.getCnes());
48                 historico.setCns(visitadomiciliario.getCns());
49                 historico.setCnsResponsavel(historicoCadIndividual.getCnsResponsavel());
50                 historico.setDataNascimento(historicoCadIndividual.getDataNascimento());
51                 historico.setData(ChronoUnit.YEARS.between(historicoCadIndividual.getDataNascimento(), visitadomiciliario.getDataVisitadomiciliario()));
52                 historico.setFaixaEtaria(this.faixaEtariaService.getFaixaEtaria(historico.getIdade()));
53                 historico.setSexo(visitadomiciliario.getSexo());
54                 historico.setRacaCor(historicoCadIndividual.getRacaCor());
55                 historico.setGestante(visitadomiciliario.getGestante());
56                 historico.setNome(visitadomiciliario.getNome());
57                 historico.setFumante(visitadomiciliario.getFumante() != null ? visitadomiciliario.getFumante() : false);
58                 historico.setAlcool(visitadomiciliario.getAlcool() != null ? visitadomiciliario.getAlcool() : false);
59                 historico.setDrogas(visitadomiciliario.getDrogas() != null ? visitadomiciliario.getDrogas() : false);
60                 historico.setHipertensaoArterial(visitadomiciliario.getHipertensaoArterial() != null ? visitadomiciliario.getHipertensaoArterial() : false);
61                 historico.setDiabetes(visitadomiciliario.getDiabetes() != null ? visitadomiciliario.getDiabetes() : false);
62                 historico.setAvc(false);
63                 historico.setCep(historicoCadIndividual.getCep());
64                 historico.setIge(historicoCadIndividual.getIge());
65                 historico.setMunicipio(historicoCadIndividual.getMunicipio());
66                 historico.setUf(historicoCadIndividual.getUf());
67                 historico.setUnidadeSaude(historicoCadIndividual.getUnidadeSaude());
68                 historico.setTipoLogradouro(historicoCadIndividual.getTipoLogradouro());
69                 historico.setLogradouro(historicoCadIndividual.getLogradouro());
70                 historico.setBairro(historicoCadIndividual.getBairro());
71                 historico.setNumero(historicoCadIndividual.getNumero());
72                 this.em.merge(historico);
73                 this.em.flush();
74                 this.em.clear();
75             }
76         }
77     }
78 }

```





```

GeradorDataPassada.java 109 KB
package br.gov.essus.generator;
import java.util.Calendar;
import java.util.concurrent.ThreadLocalRandom;
import lombok.AccessLevel;
import lombok.NoArgsConstructor;

@ArgsConstructor(access = AccessLevel.PRIVATE)
public class GeradorDataPassada {

    public static long difHancimentoEntre1000Anos() {
        Calendar dezAnosAtras = Calendar.getInstance();
        dezAnosAtras.add(Calendar.YEAR, -10);

        Calendar noventaAnosAtras = Calendar.getInstance();
        noventaAnosAtras.add(Calendar.YEAR, -90);

        return ThreadLocalRandom.current().nextLong(noventaAnosAtras.getTimeInMillis(), dezAnosAtras.getTimeInMillis());
    }

    public static long ultimos30Dias(long timeInMillis) {
        Calendar calendar = Calendar.getInstance();
        calendar.setTimeInMillis(timeInMillis);
        calendar.add(Calendar.DATE, -ThreadLocalRandom.current().nextInt(1, 30));
        return calendar.getTimeInMillis();
    }

    public static long ultimosCincoAnos() {
        Calendar startDate = Calendar.getInstance();
        startDate.add(Calendar.YEAR, -5);

        return ThreadLocalRandom.current().nextLong(startDate.getTimeInMillis(), Calendar.getInstance().getTimeInMillis());
    }
}

```

```

GeradorThriftExemplo.java 100 KB
package br.gov.esus.executor;
2
3 import br.gov.esus.executor.util.ZipWriter;
4 import br.gov.esus.thrift.DefaultValues;
5 import br.gov.esus.thrift.FichaAtendimentoIndividualThriftCreator;
6 import br.gov.esus.thrift.FichaCadastroBomniciilarThriftCreator;
7 import br.gov.esus.thrift.FichaCadastroIndividualThriftCreator;
8 import br.gov.esus.thrift.FichaVisitaBomniciilarThriftCreator;
9
10 public class GeradorThriftExemplo {
11
12     private static final int FCI_ = 1000;
13     private static final int FCDT = 1000;
14     private static final int FAL_ = 1000;
15     private static final int PVDF = 1000;
16
17     public static void main(String[] args) {
18         long currentTimeMillis = System.currentTimeMillis();
19         createdSystem.getProperties().put("user.home") + "/Downloads/FCI_FCDT_FAL_PVDF_1000_" + DefaultValues.unidadeDefinido().get(0).getLocalidadeFunm().n
20         System.out.println("total = " + (System.currentTimeMillis() - currentTimeMillis));
21     }
22
23     public static void create(String path) {
24
25         ZipWriter zipWriter = new ZipWriter(path);
26
27         long currentTimeMillis = System.currentTimeMillis();
28         for (int i = 0; i < FCI_; i++) {
29             zipWriter.add(FichaCadastroIndividualThriftCreator.create());
30         }
31         System.out.println("FCI_ = " + (System.currentTimeMillis() - currentTimeMillis));
32
33         currentTimeMillis = System.currentTimeMillis();
34         for (int i = 0; i < FCDT; i++) {
35             zipWriter.add(FichaCadastroBomniciilarThriftCreator.create());
36         }
37         System.out.println("FCDT = " + (System.currentTimeMillis() - currentTimeMillis));
38
39         currentTimeMillis = System.currentTimeMillis();
40         for (int i = 0; i < FAL_; i++) {
41             zipWriter.add(FichaAtendimentoIndividualThriftCreator.create());
42         }
43         System.out.println("FAL_ = " + (System.currentTimeMillis() - currentTimeMillis));
44
45         currentTimeMillis = System.currentTimeMillis();
46         for (int i = 0; i < PVDF; i++) {
47             zipWriter.add(FichaVisitaBomniciilarThriftCreator.create());
48         }
49         System.out.println("PVDF = " + (System.currentTimeMillis() - currentTimeMillis));
50
51         zipWriter.close();
52     }
53
54 }
55 }

```





## ANEXO C

```

1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200

```