



1º Edição do DATATHON DE NUTRIÇÃO

**O suplemento nutricional pode reduzir os custos de saúde
relacionados à queda na população idosa?**

André Luiz Dalazen

Danielle Calazans Dondoni

Eric Lopes Leal

Miguel Batista da Silva Neto

Milton Nascimento dos Santos Júnior

Dezembro de 2020

SUMÁRIO

1. OBJETIVOS.....	4
2. INTRODUÇÃO.....	5
3. REVISÃO DA LITERATURA.....	5
3.1. População idosa no Brasil.....	5
3.2. Saúde Pública.....	6
3.3. Consumo alimentar da população brasileira.....	6
3.4. Riscos associados ao envelhecimento.....	7
3.4.1. Osteoporose.....	8
3.4.2. Sarcopenia.....	10
3.4.3. Quedas: Morbidade e Mortalidade.....	12
4. METODOLOGIA.....	13
4.1. Compreensão do Negócio.....	15
4.2. Entendimento dos Dados.....	15
4.3. Preparação dos dados.....	16
4.4. Modelagem.....	17
4.5. Avaliação.....	18
5. RESULTADOS E DISCUSSÃO.....	19
6. <i>INSIGHTS</i>.....	26
7. CONCLUSÕES.....	33
8. REFERÊNCIAS.....	34

LISTA DE ABREVIATURAS

BPS	Banco de Preço de Saúde
CID-10	Classificação Internacional de Doenças - 10ª Edição
CRISP-DM	Processo Padrão Inter-Indústrias para Mineração de Dados
DATASUS	Portal do Departamento de Informática do Sistema Único de Saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
MS	Ministério da Saúde
POF	Pesquisa de Orçamentos Familiares
SIGTAP	Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS
SIH	Sistema de Informações Hospitalares
SUS	Sistema Único de Saúde
TABNET	Software de Informações de Saúde

1. OBJETIVOS

- Realizar análise de dados para compreender aspectos da saúde dos idosos, com ênfase na temática nutrição para o público com idade igual ou superior a sessenta anos.
- Melhorar o entendimento sobre o universo da nutrição no Brasil, focando o olhar no impacto da nutrição, em destaque para o uso suplemento nutricional no idoso.

2. INTRODUÇÃO

Em 2010, o Brasil tinha cerca de 190,7 milhões de habitantes e o número de pessoas com 60 anos ou mais foi estimado em 20,6 milhões, quase 11 % da população total. Em 2020, a população estimada é de 211,7 milhões, sendo 14 % composta por idosos.

O crescimento da população idosa no Brasil acontece de forma acelerada e significativa. O aumento da expectativa de vida da população é resultado da modernização da medicina e da agricultura, do acesso a saneamento básico e dos avanços tecnológicos, refletindo de forma direta na saúde da população idosa, dado que ao prolongar a vida, aumenta-se a quantidade de indivíduos na terceira idade.

Considerando esse processo de envelhecimento populacional, novas demandas de saúde emergem, sendo potencializados os cuidados para a prevenção de doenças e promoção da qualidade de vida. (GUSMÃO e FILHO, 2018).

3. REVISÃO DA LITERATURA

O termo adotado neste trabalho está em concordância com o disposto no artigo primeiro do Estatuto do Idoso, Lei 10.741/2003, no qual é definida como idosa a pessoa que tem 60 anos ou mais de idade (BRASIL, 2003).

3.1. População idosa no Brasil

Na última estimativa populacional publicada no Diário Oficial da União, datada de julho de 2020, há aproximadamente 211,7 milhões de residentes no Brasil (IBGE, 2020a). Conforme as projeções populacionais, estima-se que o número de idosos seja de 30,2 milhões, 14,3% em relação à população total. Destes, 13,3 milhões são homens (44%) e 16,9 milhões são mulheres (56%). A distribuição dos idosos nas cinco regiões mantém estreita semelhança com os dados apresentados anteriormente, apresentando as seguintes projeções em quantidade e porcentagem de idoso: Norte (1,64 mi – 5%), Nordeste (7,25 mi – 24%), Sudeste (14,30 mi - 47%), Sul (5,02 mi – 17%) e Centro Oeste (1,98 mi – 4%). (IBGE, 2020b)

Nos próximos 10 anos, 2030, estima-se que a população será de aproximadamente 225 milhões de pessoas, dos quais 42,1 milhões terão idade igual ou superior a 60 anos, correspondendo a 18,7% da população. Estendendo as projeções a 2040, a população idosa será de 54 milhões, cerca de 23% da população brasileira (232 milhões) (IBGE, 2018)

Evidencia-se que o Brasil passa por um fenômeno de transição demográfica e envelhecimento populacional. Esse processo demandará cada vez mais, produtos e serviços específicos para esta parcela crescente da população além de ajustes na estrutura social, econômica, política e cultural da sociedade, uma vez que os idosos possuem demandas específicas para obtenção de adequadas condições de vida (WINGERTER *et al.* 2020).

3.2. Saúde Pública

Segundo os dados mais recentes da Agência Nacional de Saúde Suplementar (ANS), de outubro de 2020, 8,7 milhões de idosos possuíam planos de saúde de assistência médico-hospitalar no Brasil, dos quais 55% estavam na faixa de 60 a 69 anos, 29% entre 70 e 79 anos e 16% com 80 anos ou mais. O número de idosos corresponde a 12% do total de beneficiários da saúde suplementar e 29% da população brasileira idosa. (ANS, 2020)

Cabe ressaltar que o direito de usar os recursos de saúde são universais e se estendem a pessoas com planos privados de saúde. Pode-se inferir que o SUS é nominalmente responsável por mais de 70% dos cuidados de saúde dos idosos, portanto, os custos do Estado relativos à saúde tendem a ter aumento expressivo nos próximos anos.

3.3. Consumo alimentar da população brasileira

Analisando a Pesquisa de Orçamentos Familiares/IBGE(POF) pode-se identificar os hábitos alimentares da população brasileira de acordo com o rendimento familiar,

consolidados na Tabela 01 – Consumo por renda. Devida a dimensão da tabela, ela está disponibilizada no repositório do Github.

O link de acesso é: <https://github.com/Datanoninho/DataThon>

Observa-se que os alimentos mais consumidos entre os brasileiros são os laticínios, principalmente leite de vaca pasteurizado, porém esse comportamento não se repete de acordo com as rendas familiares. Famílias com rendas até R\$ 1908,00 e até R\$ 2862,00 tendem a consumir mais cereais e leguminosas (representando 5,12% e 4,33%, respectivamente, do consumo), com rendas até R\$ 5724,00 e até R\$ 9540,00 tendem a consumir mais laticínios (4,24% e 4,36%, respectivamente) e com rendas até R\$ 9540,00 e superiores tendem a consumir mais frutas (4,27% e 4,37%, respectivamente)

Através destes dados, nota-se que a dieta da população é deficiente das principais fontes de proteína, como carnes dos tipos mais variados, assim, podemos inferir que o baixo consumo dessas fontes pode estar associado a um quadro de subnutrição principalmente na população onde a renda é mais baixa e a aquisição desse tipo de alimento é comumente mais custosa.

3.4. Riscos associados ao envelhecimento

Com o envelhecimento, o corpo humano entra em processo de declínio fisiológico, declínio do processo de absorção, utilização e excreção de nutrientes além da diminuição da densidade óssea e da massa muscular, instabilidade postural, comprometimento da capacidade visual e auditiva, maior consumo de medicamentos devido à presença de inúmeras doenças comuns ao idoso, além de riscos ambientais que podem predispor à queda (ROSA *et al*, 2015 e KRAUSE *et al.*, 2011)

Dentre as inúmeras possíveis alterações fisiológicas que podem acometer uma pessoa idosa, destacamos a ocorrência da osteoporose e sarcopenia.

3.4.1. Osteoporose

A osteoporose é caracterizada por baixa massa óssea e deterioração da microarquitetura do tecido ósseo, com consequente aumento da fragilidade óssea e suscetibilidade a fraturas. Estudos indicam que após os 50 anos, 1 em cada 2 mulheres e 1 em cada 4 homens terão uma fratura relacionada à osteoporose, provenientes de uma série de fatores além da fraqueza óssea; incluindo quedas, reflexos lentos, sarcopenia e preenchimento inadequado de tecido mole. (PETERS e MARTINI, 2010)

Estima-se que no Brasil 1 a cada 3 pacientes com fraturas de colo do fêmur são diagnosticados com osteoporose e apenas 1 a cada 5 desses recebem algum tipo de tratamento. O custo apenas das fraturas no colo do fêmur onde há casos de osteoporose para as companhias de seguro é estimado em 6 milhões de dólares anuais.

A nutrição desempenha um papel fundamental na redução do risco de osteoporose através de seu efeito em todos esses fatores de fragilidade, especialmente no desenvolvimento e manutenção da massa óssea. A adequada ingestão de cálcio, vitamina D e proteína leva à redução da remodelação óssea, à maior retenção de cálcio, à redução da perda óssea relacionada à idade e à redução do risco de fraturas, podendo atingir uma redução do risco de osteoporose em até 50% com controle nutricional e de peso. (PETERS e MARTINI, 2010)

A recomendação da quantidade diária de minerais que precisam ser ingeridos varia principalmente conforme sexo e idade. Neste ponto, cabe ressaltar que cada país recomenda uma quantidade diferente, tornando difícil o comparativo mais assertivo entre as literaturas.

No Brasil, a indicação é que o consumo de cálcio seja de 1000 mg/dia para mulheres e homens adultos (ANVISA, 2005). Na Austrália são recomendados 1000 mg/dia e 800 mg/dia, para mulheres e homens idosos. Já a recomendação da FAO/WHO estabelece como 1300 mg/dia a quantidade de cálcio a ser ingerida por idosos, independente do sexo.

Os minerais magnésio e fósforo estão distribuídos no corpo humano de forma majoritária na estrutura óssea, mas também atuam em diversos processos do organismo. O magnésio participa do metabolismo da glicose e da regulação da pressão arterial enquanto o fósforo, além de ser importante para a regulação do pH, é um componente de todas as células e metabólitos importantes, incluindo o DNA, o RNA, o ATP e os fosfolipídios (KRAUSE et al, 2011). O consumo recomendado é de 260 mg/dia de magnésio e 700 mg/dia de Fósforo (ANVISA, 2005).

As proteínas exercem papel fundamental no crescimento, construção e reparação dos tecidos do nosso corpo. A faixa aceitável de ingestão de proteínas é de 10% a 35% da quantidade total de calorias. Estudos indicam que a quantidade de proteína de 0,8 g/kg por dia não é ideal para a saúde óssea em idosos, sendo recomendado de 1,0 a 1,2 g/kg por dia de proteínas (ou aproximadamente 13%-16% das calorias totais), suficientes para manter o metabolismo normal de cálcio e o equilíbrio de nitrogênio sem afetar a função renal (PETERS e MARTINI, 2010). Na resolução 269/05 da ANVISA a recomendação de consumo por adultos é de 50 g/dia. (ANVISA, 2005).

O consumo de vitamina D é essencial para o desenvolvimento e manutenção do osso, tanto pelo auxílio à absorção de cálcio da dieta, quanto pela renovação e mineralização adequada do tecido ósseo. Há inúmeras evidências de que a deficiência de vitamina D é um importante contribuinte para a osteoporose devido a uma de absorção intestinal menos eficiente de cálcio, aumento da perda óssea, fraqueza muscular e uma microestrutura óssea enfraquecida.

Apesar de não estar bem definido a concentração ideal dos níveis séricos de 25 hidroxivitamina D [25(OH)D], a Sociedade Brasileira de Endocrinologia e Metabologia (SBEM) e a Sociedade Brasileira de Patologia Clínica (SBPC) recomendam valores de referência de 25(OH)D estratificados de acordo com a idade e características clínicas individuais (MOREIRA et al, 2020), sendo apresentadas as seguintes categorias:

- Deficiência: <20 ng/mL
- Adequado para a população em geral < 65 anos: entre 20-60 ng/mL
- Adequado para indivíduos com condições vulneráveis: 30-60 ng/mL
- Risco de intoxicação: >100 ng/mL

São classificados como vulneráveis os idosos, pacientes pós-cirurgia bariátricas, gestantes, indivíduos em uso de drogas que interferem no metabolismo da vitamina D, pacientes com osteoporose, hiperparatireoidismo secundário, osteomalácia, diabetes mellitus tipo 1, câncer, doença crônica, doença renal ou má absorção.

Conforme resolução da ANVISA a quantidade recomendada para adultos é de 5 micrograma/dia, sendo 1 micrograma de calciferol igual a 40 UI, perfazendo 200 UI/dia (ANVISA, 2005).

3.4.2. Sarcopenia

Definida originalmente como a diminuição da massa muscular relacionada ao envelhecimento, o termo sarcopenia tornou-se um mais abrangente, sendo usado para definir a perda de massa e força muscular relacionadas à idade (ALEXANDRE, *et al* 2019).

A sarcopenia está associada à mobilidade reduzida, equilíbrio em pé prejudicado, declínio funcional, hospitalização e mortalidade Yeung (*et al*, 2019). Uma série de fatores associados ao envelhecimento contribuem para o desenvolvimento de sarcopenia, como perda de neurônios motores (alfa), o aumento de citocinas inflamatórias, a diminuição da atividade física e diminuição hormonal (MALAFARINA, *et al*, 2013).

A ingestão de proteínas, aminoácidos essenciais, leucina e vitamina D são reconhecidas como fatores importantes na gestão de sarcopenia (NASCIMENTO e MOREIRA, 2019 apud MORLEY *et al.*, 2010). O consumo inadequado de alimentos pode ser uma das principais causas de sarcopenia e uma dieta pobre em proteínas determina uma resposta compensatória, caracterizado por uma redução na massa magra. A quantidade recomendada de 0,8 g / kg por dia de proteínas é insuficiente em idosos, sendo entre 1,25 e 1,5 g / kg por dia a quantidade mais aconselhável (MALAFARINA, *et al*, 2013).

A recomendação atual de carboidratos é que seja aproximadamente 45% a 65% das calorias totais, ressaltando a importância no aumento do consumo de carboidratos complexos para promoção de fibras, minerais e vitaminas essenciais. O consumo de

lipídios deve se manter entre 20% a 35% e advindos preferencialmente de gorduras poli-insaturadas e monoinsaturadas (KRAUSE *et al*, 2011).

Em relação a vitaminas e minerais tem-se um desafio para os idosos, onde faz-se necessário elevar o nível dos mesmos em relação ao valor calórico total consumido, com risco aumentado de deficiência de vitamina B12 e Vitamina D que é associada à redução da inflamação e, conseqüentemente, a melhora da força muscular. Os aminoácidos essenciais (particularmente, a leucina), por sua vez, estimulam a síntese de proteína muscular em idosos (NASCIMENTO e MOREIRA, 2019)

Na revisão sistemática e de meta-análise apresentada por Yeung (*et al*, 2019), o número de idosos com sarcopenia varia entre 2% e 37%, dependendo da definição de sarcopenia aplicada. Algumas definições têm como base apenas a baixa massa muscular. Outras são fundamentadas na evidência de baixa massa muscular e baixa força muscular.

Yeung *et al* (2019) destacam a associação positiva entre indivíduos sarcopênicos com risco maior de quedas e fraturas em comparação com indivíduos não sarcopênicos; independente da população, sexo, definição, continente e qualidade do estudo.

No estudo apresentado por Alexandre e colaboradores (2019), com 1168 idosos residentes em São Paulo, foram avaliadas três condições clínicas:

- Sarcopenia: idosos com índice de massa muscular esquelética $\leq 8,90$ kg/m² para homens e $\leq 6,37$ kg/m² para mulheres
- Dinapenia: força de preensão manual < 30 kg para homens e < 20 kg para mulheres
- Sarcodinapênicos: idosos que apresentavam sarcopenia associada à dinapenia.

A condição que apresentou maior prevalência foi a dinapenia (34,4% em mulheres e 25,8% em homens), seguida pela sarcodinapenia (10,4% em mulheres e 6,9% em homens) e sarcopenia (4,3% em mulheres e 5,5% em homens) (ALEXANDRE, *et al* 2019).

A prevalência da sarcopenia pode atingir 5 a 13% em idosos, com faixa etária de 60-70 anos e 11 a 50% naqueles com idade maior de 80 anos, a depender do método adotado para diagnosticar. No Brasil foi descoberto uma prevalência de 15,4% para

população idosa residente na Cidade de São Paulo, em 2006 (ALEXANDRE, *et al* 2019).

3.4.3. Quedas: Morbidade e Mortalidade

A temática queda em idosos tem se mostrado multifatorial e complexa. Os fatores responsáveis pelas quedas podem ser classificados como intrínsecos ou extrínsecos. O primeiro grupo está relacionado a alterações fisiológicas do processo de envelhecimento, a doenças e aos efeitos causados pelo uso de fármacos, já o segundo grupo está relacionado a fatores que dependem de circunstâncias sociais e ambientais. Estudos demonstram que 29% dos idosos caem pelo menos 1 vez ao ano e 13% dos idosos caem de forma recorrente. (CELICH, *et al* 2010 e (SILVA, *et al* 2018))

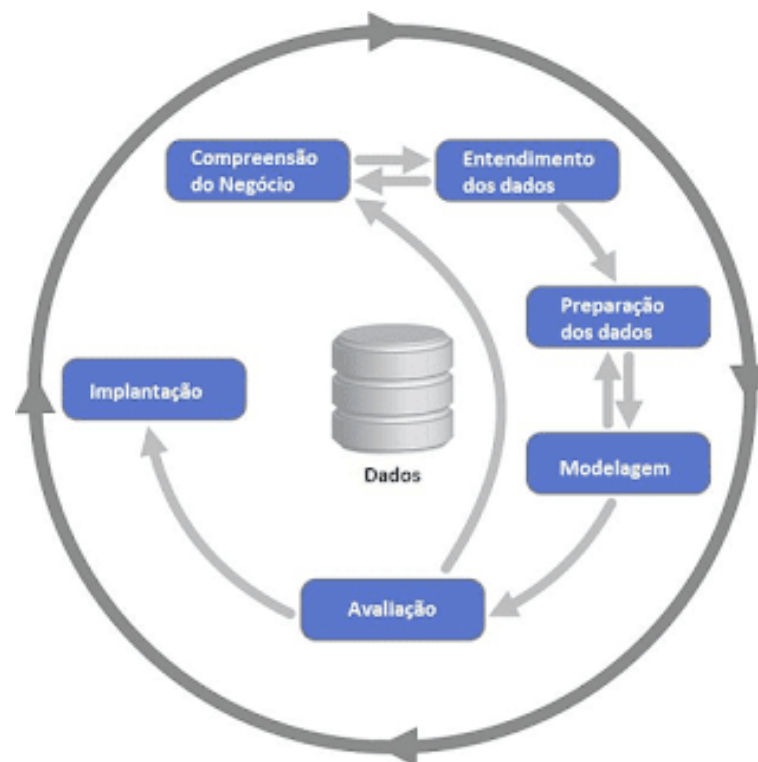
Considerando o grande número de pessoas idosas no Brasil e projeção de crescimento, associado à elevada porcentagem de idosos que sofrem queda, o levantamento de dados sobre fatores de risco possui grande importância para a adoção de medidas protetoras e preventivas. (SILVA, *et al* 2018)

O incentivo à atividade física, nutrição adequada, avaliação de riscos domésticos, revisão periódica da medicação e identificação dos fatores que aumentam os riscos em pessoas que já sofreram quedas são extremamente importantes para reduzir potencialmente a morbimortalidade, os custos hospitalares e asilamento de idoso (SILVA, *et al* 2018)

4. METODOLOGIA

A metodologia CRISP-DM (acrônimo de **C**Ross **I**ndu**S**try **P**rocess for **D**ata **M**ining - em português pode ser traduzido como Processo Padrão Inter-Indústrias para Mineração de Dados) é um processo com foco em aumentar o uso de dados nos negócios. Ele é composto em seis etapas não obrigatórias e não encadeadas, **Figura 1**, as quais devem ser implementadas, para que um projeto de Ciência de Dados possa ser bem sucedido.

Figura 1: Representação das etapas da metodologia CRISP-DM.



- **Etapa 1 - Compreensão de Negócio**

Esta pode ser considerada a etapa mais importante, pois apenas através da compreensão das regras de negócio da empresa é possível definir objetivos claros para o desenvolvimento de um projeto de ciência de dados. A má compreensão dessa etapa é o que leva ao fato de mais de 80% dos projetos de Ciência de Dados falharem, já que as perguntas dos investidores, diretores e *CEOs* para problemas enfrentados pela empresa ficam sem respostas eficientes.

- **Etapa 2 – Entendimento dos Dados**

Nessa etapa é analisado quais dados serão necessários, de quais fontes e por quais métodos serão obtidos, para que o projeto seja bem sucedido. Esta etapa compreende os requisitos, coleta e compreensão dos dados.

- **Etapa 3 - Preparação de Dados**

Uma vez que os dados tenham sido coletados, eles precisam passar por um tratamento para que possam ser efetivamente utilizados. Nesse momento define-se o que acontecerá com campos em branco (se eles serão removidos ou informados como a média dos demais dados).

- **Etapa 4 – Modelagem**

Uma vez que os dados tenham sido tratados, eles precisam ser convertidos em informação, este é o propósito de um projeto de ciência de dados. Isso pode ser realizado de diversas maneiras, entretanto as mais comuns são o uso de Inteligência Artificial e de Métodos Estatísticos para realizar previsões. É nesse momento em que se encontram padrões e estruturas as quais podem prover *insights* para os negócios da empresa.

- **Etapa 5 – Avaliação**

O modelo criado precisa ser testado e avaliado. Essa etapa ocorre com um conjunto de dados separados especificamente para analisar se o modelo é capaz de “generalizar”, ou seja, se ele consegue ter uma boa performance, assertividade, com dados diferentes dos usados para treinar o modelo. Nesse caso, os dados para treino possuem o valor esperado que o modelo apresenta. Com base na diferença entre o dado real e o previsto pelo modelo podemos determinar a assertividade do mesmo.

▪ Etapa 6 - Implantação

Nessa fase o modelo é utilizado por *stakeholders*, diretores, gerentes, CEOs, com novos dados provenientes de fora do conjunto original (usado para treino e teste). Nessa fase pode ser percebida a necessidade de refinar o modelo ou mesmo de adicionar novas variáveis.

Acima de tudo, o mais importante é lembrar que esta metodologia é cíclica, onde o foco é a comunicação entre os interessados e o cientista de dados. A qualquer momento em que seja necessário voltar para um estágio anterior deverá sê-lo feito. Essa característica é o que faz que, eventualmente, mesmo após o estágio de entrega, se houver um novo entendimento do modelo de negócio, o processo todo é reiniciado.

4.1. Compreensão do Negócio

A Danone Nutricia, criada em 2019 a partir da fusão das unidades de negócio Early Life Nutrition (ELN) e Advanced Medical Nutrition (AMN) tem como propósito transformar vidas por meio da nutrição.

A união das expertises de suas unidades de negócio, transforma a Danone Nutricia em uma potência em nutrição especializada do Brasil, com portfólio de soluções nutricionais especializadas, inovadoras e fundamentadas pela ciência. Produtos como Aptamil, Neocate, Ketocal, Nutridrink, Nutrisson e Cubitan, representam a essência dessa nova empresa, uma vez que transformam a vida de seus consumidores, pois objetivam fortalecer a saúde e o bem-estar das pessoas quando mais precisam. (DISTRITO, 2020)

4.2. Entendimento dos Dados

Para conduzir esse estudo, foram utilizados dados provenientes de diversas fontes. O Instituto Brasileiro de Geografia e Estatística (IBGE, 2018) forneceu as estimativas e projeções populacionais. Os dados referentes às internações foram obtidos do Sistema de Informação Hospitalar do Sistema Único de Saúde (SIH-SUS). No Banco

de Preços de Saúde (BPS) foram obtidos dados referentes à compra de suplemento nutricional. No SIGTAP o valor dos procedimentos realizados pelo SUS.

4.3. Preparação dos dados

- **Limpeza de dados**

O processo de limpeza de dados envolve a remoção de erros tipográficos, validação de dados e aprimoramento de dados. Este processo foi realizado até que os dados atendessem aos critérios de qualidade, que incluem; validade, precisão, integridade, consistência e uniformidade. Para realizar essa etapa utilizou-se a linguagem Python e as seguintes bibliotecas: Pandas, Matplotlib, Seaborn e Numpy.

- **Remoção de observações indesejadas**

Como um dos principais objetivos da limpeza de dados é garantir que o conjunto de dados estejam livre de observações indesejadas, isso é classificado como a primeira etapa da limpeza de dados. Existem 2 tipos: as duplicatas e irrelevâncias.

- **Dados em duplicata**

Os arquivos analisados não apresentaram dados em duplicidade

- **Observações irrelevantes**

Em algumas colunas existiam dados irrelevantes à análise proposta, estes, foram excluídos para não atrapalhar a visibilidade dos dados necessários.

- **Definição da estrutura de dados**

Foram realizadas adequações nos dados, como correção de erros tipográficos, correção de erros gramaticais e ajuste do tipo de dados (*data*, *inteiro*, *float*). Também foram realizadas a junção (*merge*) entre algumas tabelas que adequassem ao

proposto. Ao final foi feito uma concatenação das tabelas para juntar todos os dados, transformando em um único *dataframe* com mais coesão.

- **Outliers de filtragem**

Para melhorar o desempenho do nosso modelo, foram removidos valores de *outliers*, definidos como são pontos de dados que diferem significativamente de outras observações em um conjunto de dados.

- **Tratamento de dados ausentes**

Não foram observadas quantidades significativas de dados ausentes ou dados que pudéssemos atribuir um valor quando o valor disponibilizado era nulo.

4.4. Modelagem

A utilização de ferramentas de modelagem utilizadas tem uma função: a previsão. Previsão ou *forecast* é uma tarefa bastante comum na área de ciência de dados, pois através dela as empresas estabelecem suas metas e planejamentos.

Existem várias formas de realizar tais previsões, cada uma com suas aplicações para determinadas situações, dentre as mais comuns estão: Regressão Linear, Regressão Logística, Clusterização, Séries Temporais entre outras. No caso deste estudo, foi utilizado das Séries Temporais.

As séries temporais são um método estatístico destinado a prever um determinado comportamento com o passar do tempo. De forma bastante simplificada, as séries temporais analisam um comportamento durante um intervalo de tempo, de forma sequencial (dia após dia, ano após ano) e tentam prever com base nessa sazonalidade o comportamento futuro do objeto estudado. Uma outra característica importante desse método é que os valores estão correlacionados uns com os outros, o valor de ontem impacta no comportamento de hoje e assim sucessivamente.

Uma vez determinado o uso de um modelo preditivo utilizando-se de séries temporais, optamos por adotar o FB Prophet, um procedimento criado pela equipe de ciência de dados do Facebook para análise de séries temporais.

Dois fatores foram determinantes para a escolha de tal modelo: a facilidade de uso e os métodos estatísticos utilizados pelo modelo. A simplicidade de utilização da ferramenta foi um fator de grande importância pois dispensa a necessidade de criação e configuração manual de modelos estatísticos complexos.

Além da simplicidade, a gama satisfatória de modelos estatísticos utilizados pelo FB Prophet compreende o Auto ARIMA, ETS, SNaive e o Tbats, os quais o FB Prophet atribui automaticamente para o conjunto de dados fornecidos.

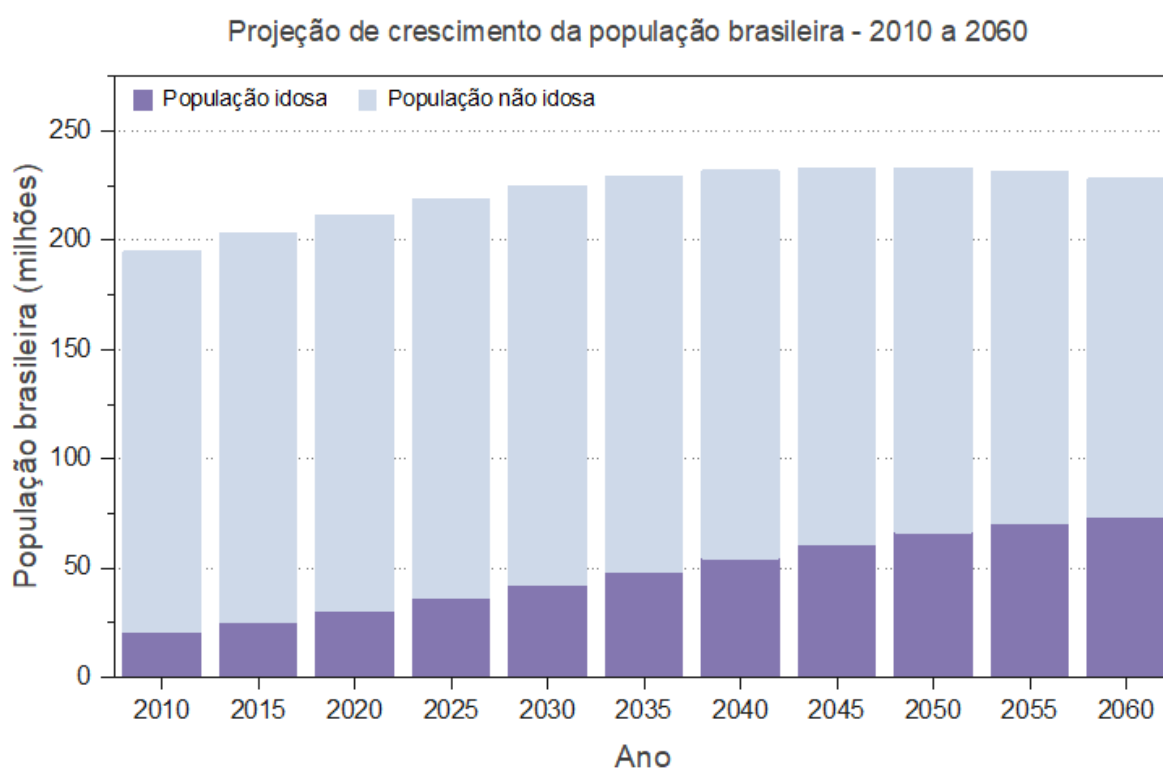
4.5. Avaliação

A avaliação dos resultados é apresentada com mais detalhes na seção de Resultados e Discussão.

5. RESULTADOS E DISCUSSÃO

O IBGE disponibiliza a projeção do crescimento populacional até o ano de 2060. A análise destes dados por quinquênios é apresentada na **Figura 2**. Observa-se que a população idosa, com idade igual ou superior a 60 anos, está crescendo em uma taxa maior do que a população não idosa.

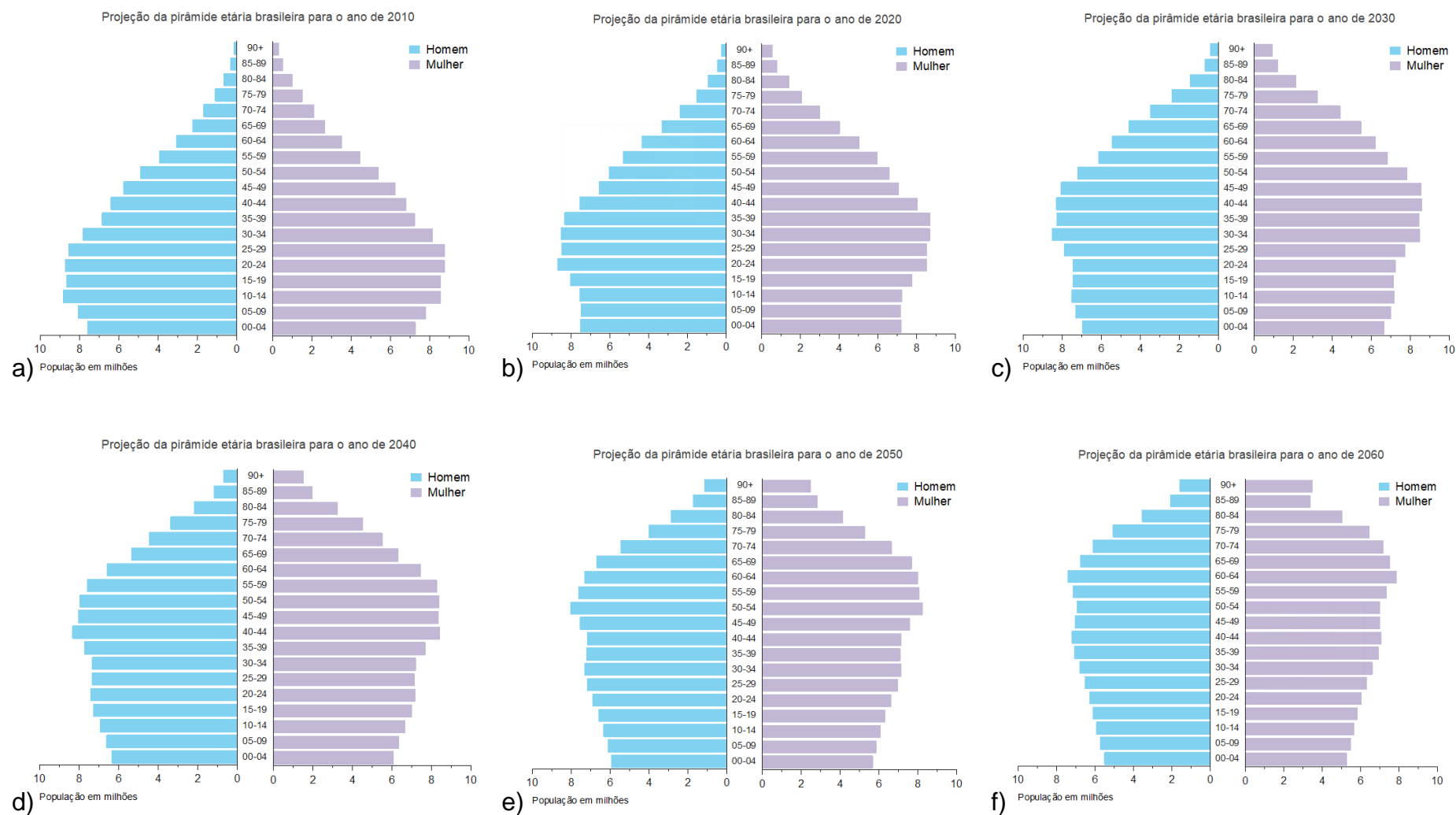
Figura 2: Gráfico representativo da projeção de crescimento da população brasileira no período de 2010 a 2060.



Fonte: Os autores, 2020. Dados obtidos do IBGE (2018)

As projeções indicam que em 2050 iniciará o decréscimo do total da população. No entanto, a população de idosos continuará em crescimento. Uma das formas de visualizar essa alteração da distribuição da população ao longo dos anos consiste na construção de uma pirâmide etária (**Figura 3**), no qual é possível observar o contínuo estreitamento da base da pirâmide devido ao declínio da fecundidade, e alargamento de outras faixas em virtude do aumento da expectativa de vida.

Figura 3: Projeções da pirâmide etária brasileira.



Legenda: Da esquerda para a direita, a partir da primeira linha, projeções para o ano de (a) 2010, (b) 2020, (c) 2030, (d) 2040, (e) 2050 e (f) 2060. Fonte: Os autores, 2020. Dados obtidos do IBGE (2018)

Na **Tabela 2** são apresentados de forma numérica os dados que compõem a **Figura 2**. Cabe ressaltar que os cálculos de crescimento apresentados são relativos ao período de 5 anos. Tomando como exemplo o ano de 2025, a população estimada é de 219 milhões de pessoas, representando um crescimento de 3,43% em relação à 2020. Neste período, a população idosa aumentará 19,5%, passando de 30,2 milhões em 2020 para 36,1 milhões em 2025, representando 16,47% da população.

Tabela 2: Dados do crescimento populacional, com dados referentes à população total e de idosos

Ano	População total (em milhões)	População idosa (em milhões)	Quantidade de idosos (%)	Crescimento da população (%)	Crescimento do nº de idosos (%)
2010	194,9	20,87	10,71	--	--
2015	203,5	25,04	12,31	4,41	19,98
2020	211,8	30,20	14,26	4,07	20,60
2025	219,0	36,08	16,47	3,43	19,50
2030	224,9	42,12	18,73	2,67	16,74
2035	229,2	47,97	20,93	1,91	13,88
2040	231,9	54,40	23,46	1,20	13,41
2045	233,1	60,88	26,11	0,53	11,90
2050	232,9	66,27	28,45	-0,09	8,85
2055	231,3	70,29	30,39	-0,70	6,07
2060	228,3	73,46	32,18	-1,30	4,52

.Fonte: Os autores (2020). Dados obtidos do IBGE (2018)

Em função dos dados apresentados e tendo como parte do problema inicial a compreensão da queda entre idosos, um minucioso estudo quantitativo e qualitativo dos dados do Sistema Único de Saúde foi realizado. Essa etapa foi realizada de duas formas, uma através da análise do banco de dados disponibilizado pela organizadora do evento e outra através do acesso direto às tabelas disponibilizadas no endereço eletrônico do DATASUS – TABNET.

A análise do banco de dados permitiu extrair informações significativas, como a relação dos exames mais realizados, o tempo médio de internação e o valor total empenhado. Para tal, foram realizadas previamente as etapas de limpeza e

formatação dos dados e integração dos dados provenientes de diversas fontes, como o SIGTAP, que forneceu os custos relativos a cada procedimento. Observa-se que há dados faltantes em alguns campos, indicando que há falhas no preenchimento das informações, podendo resultar em subnotificação de casos e dificultar as análises e ações a serem tomadas.

Uma descrição mais detalhada destes dados são apresentadas nos *notebooks* (arquivo *ipynb*), disponibilizados no Github. Dado a quantidade de informações disponíveis, as análises foram agrupadas por Estado.

O link de acesso é: <https://github.com/Datanoninho>

O acesso ao TABNET permitiu extrair informações mais específicas mediante o uso dos filtros disponibilizados nas páginas. Tendo como foco a ocorrência de queda em idosos, a primeira abordagem consistiu em uma verificação global do número total de internações, óbitos, o custo e média de dias de internações entre o período de 2008 a 2020, relativo a causas externas.

O link de acesso é: https://github.com/Datanoninho/Quedas_de_idosos

As quedas, categorizadas conforme o CID W00 a W19, correspondem a cerca de 50% do número total de internações de idosos por causas externas, com quantidades superiores a 100 mil internações por ano e gastos na ordem de 200 milhões por ano. Destas, cerca de 20% das quedas ocorreram no mesmo nível por escorregão, tropeção ou passos em falsos, 23% em queda em mesmo nível e outros 43% receberam classificação genérica (sem especificação). Evidencia-se que a queda dos idosos está diretamente associada a atividades rotineiras e não por realizar uma atividade classificada como de risco. Este tipo de queda, que ocorre sem motivo aparente, é característico de quadros de fraqueza muscular, osteoporose e sarcopenia.

Desmembrando estes dados, na **Tabela 3** é mostrado o quantitativo e custo associado a cada CID. Para referência, serão apresentados os dados do ano 2018. Dados relativos aos outros anos, bem como referentes à população não idosa, estão consolidados e disponibilizados online.

O link de acesso é: https://github.com/Datanoninho/Quedas_de_idosos

Tabela 3: Dados do quantitativo e custo total de ‘internações de idosos em 2018 em virtude da queda.

Cod.	Descrição	Número de ocorrências	Percentual em relação ao total (%)	Valor total (R\$)
W00	Queda no mesmo nível envolvendo gelo e neve	789	0,63%	914.207,26
W01	Queda no mesmo nível por escorregão, tropeção ou passos em falsos [traspés]	23399	18,67%	38.770.867,11
W02	Queda envolvendo patins de rodas ou para gelo, esqui ou pranchas de rodas	269	0,21%	395.781,32
W03	Outras quedas no mesmo nível por colisão com ou empurrão por outra pessoa	1787	1,43%	2.158.929,69
W04	Queda, enquanto estava sendo carregado ou apoiado por outra(s) pessoa(s)	470	0,37%	750.581,30
W05	Queda envolvendo uma cadeira de rodas	87	0,07%	147.232,42
W06	Queda de um leito	915	0,73%	1.355.766,73
W07	Queda de uma cadeira	290	0,23%	467.345,53
W08	Queda de outro tipo de mobília	327	0,26%	622.933,96
W09	Queda envolvendo equipamento de "playground"	101	0,08%	114.025,60
W10	Queda em ou de escadas ou degraus	3604	2,88%	6.320.802,22
W11	Queda em ou de escadas de mão	548	0,44%	582.219,44
W12	Queda em ou de um andaime	372	0,30%	522.755,16
W13	Queda de ou para fora de edifícios ou outras estruturas	806	0,64%	1.434.768,67
W14	Queda de árvore	251	0,20%	547.750,96
W15	Queda de penhasco	23	0,02%	26.050,56
W16	Mergulho ou pulo na água causando outro traumatismo que não afogamento ou submersão	75	0,06%	97.970,75
W17	Outras quedas de um nível a outro	7012	5,59%	10.678.287,57
W18	Outras quedas no mesmo nível	28032	22,36%	47.181.927,62
W19	Queda sem especificação	56194	44,83%	85.321.433,82
	Total	125351	100,00%	198.411.637,69

Fonte: Os autores, 2020.

As quedas podem desencadear uma série de complicações à saúde do idoso, como traumatismo craniano, fratura de fêmur, sepse, pneumonia, confusão mental, depressão, redução da mobilidade, redução da sociabilidade e óbito. Considerando o crescimento da população idosa, faz-se necessário reforçar a adoção de ações preventivas capazes de minimizar as quedas e minimizar o impacto das consequências na saúde dos idosos. Cabe ressaltar que, a osteoporose e a sarcopenia possuem uma prevalência de aproximadamente 10% da população idosa e são fatores relacionados à ocorrência de quedas.

Por análise preditiva de série temporal, baseada em *machine learning* e utilizando a biblioteca Prophet, foram estimadas a quantidade e custo total das quedas de idosos até 2021. (**Figura 4**)

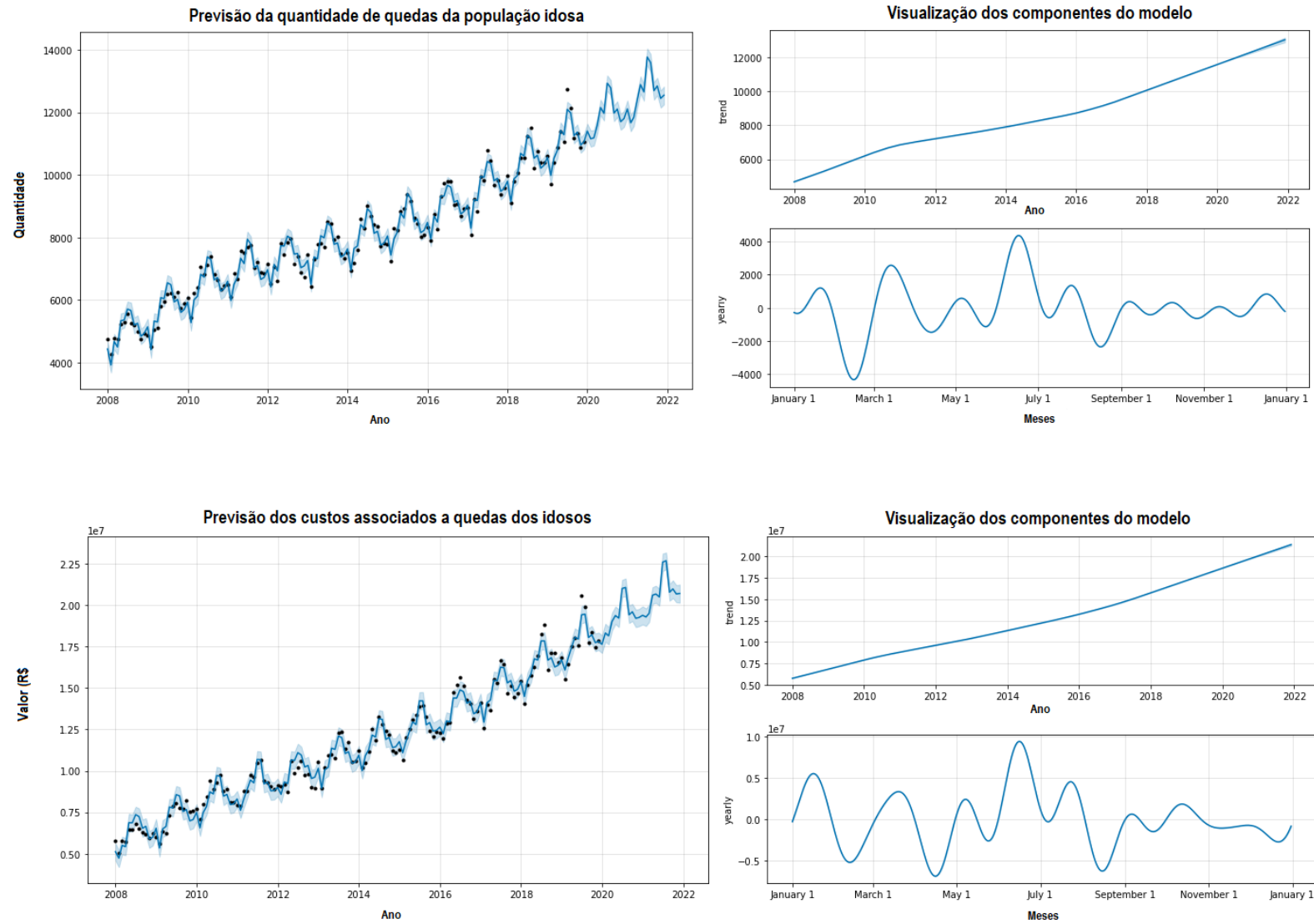
A análise dos dados permite inferir que irão ocorrer cerca de 140 mil eventos de quedas de idosos ao longo do ano 2021, um valor médio de 12,7 mil quedas por mês. Também foi possível quantificar os possíveis custos das internações para o SUS, com valor anual de aproximadamente 229 milhões de reais e média de 20,8 milhões de reais por mês.

A análise dos componentes preditivos do modelo, *trend* (tendência) e *yearly* (sazonalidade anual) permite visualizar a tendência de crescimento do número de quedas e, conseqüentemente, os custos associados. É possível ainda notar a sazonalidade dessa série temporal, na qual eventos de queda ocorrem em maior quantidade durante os meses de julho a setembro, período de inverno, que afeta diretamente a população idosa.

De forma complementar, também foram realizadas análises preditivas de outras causas de internação que estão correlacionadas à queda de idosos, dentre elas a fratura do colo do fêmur e fratura de outros ossos dos membros. Os resultados estão disponíveis no Github do projeto.

O link de acesso é: https://github.com/Datanoninho/Quedas_de_idosos

Figura 4: Resultado da análise preditiva do quantitativo e custos de quedas da população idosa.



Fonte: Os autores, 2020..

As quedas são classificadas como reduzíveis por ações adequadas de promoção à saúde, vinculadas a ações de atenção à saúde. A busca por alternativas que minimizem o impacto do envelhecimento, seja na promoção de maior qualidade de vida ou na redução de custos associados à enfermidades características dessa fase da vida devem ser intensificadas.

A associação de fatores como alimentação inadequada, falta de acesso a condições básicas de sobrevivência e avanço da idade podem comprometer o futuro da população idosa. É de interesse do Estado minimizar os custos e consequências da queda de idosos, dado o elevado impacto econômico e social que isso ocasiona.

6. *INSIGHTS*

A contratação de serviços ou aquisição de produtos pela Administração Pública requer um processo administrativo obrigatório, a licitação. No Brasil existem diversos portais para tratar do tema, porém a busca por editais tende a ser morosa. E em alguns casos editais acabam por serem dados como desertos, por simples falta de conhecimento.

Para apresentarmos a solução proposta, antes devemos explicar (mesmo de forma simplificada) duas importantes definições que compõem a organização por trás dessa resolução. A primeira delas, o *web crawler*, também conhecido como rastreador web ou *bot*, é um algoritmo usado para analisar o código de um website em busca de informações, e depois usá-las para gerar insights ou classificar os dados encontrados. O Processamento de Linguagem Natural (PLN) é um campo da ciência da computação, inteligência artificial e computação linguística que tem como foco a interação entre computadores e linguagem (natural) humana.

Na solução apresentada, o *crawler* terá como função a busca automática por informações sobre os processos de licitação, sem necessidade da intervenção humana, e as aplicações de PLN tem como função analisar, filtrar e classificar as informações relevantes.

Os dados consolidados serão apresentados em um portal voltado para *Business Intelligence*, construído com recursos de Inteligência Artificial.

Link de acesso: <https://xd.adobe.com/view/0497a8e5-54d3-48f7-a5f8-b48039ebd403-51c4>

Na página inicial, identificada pela aba Hoje, **Figura 5**, será possível acompanhar as licitações do dia, com informações atualizadas automaticamente, bem como os responsáveis por cada processo.

Na aba Previsões, **Figura 6**, é apresentado uma estimativa de possíveis licitações por cidade, mas que ainda não foram anunciadas. Outra informação relevante apresentada nessa aba é a estimativas de lucro futuro pela análise do histórico de compras.

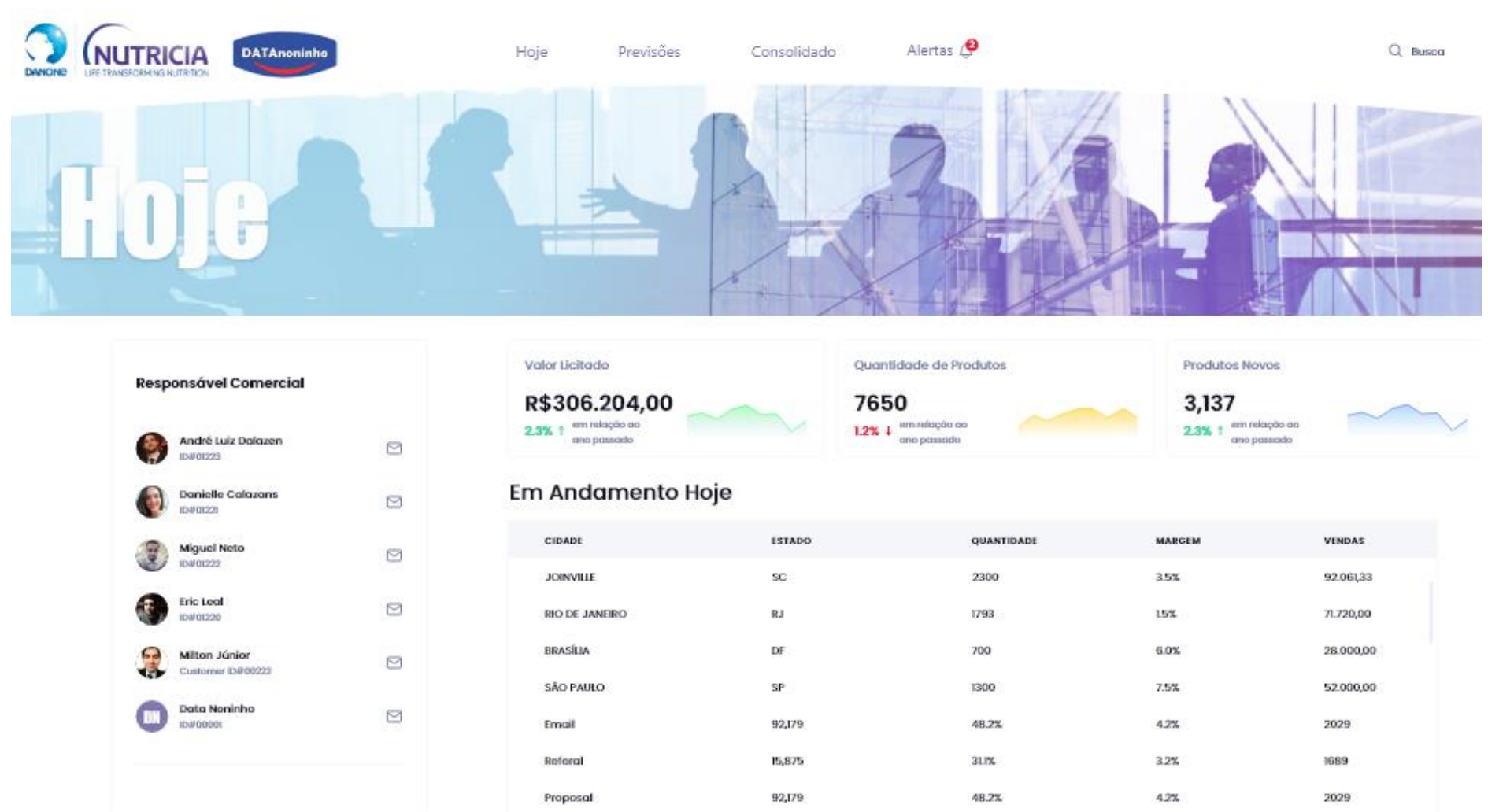
Na parte denominada Consolidado, **Figura 7**, há informações sobre o histórico de licitações, dados de venda por produtos e o quantitativo de licitações ganhas, com informações sobre o crescimento e valor médio por cliente.

O item Alerta, **Figura 8**, permite avaliar quais licitações possivelmente estão atrasadas, calcular se o provisionamento de estoque atende à demanda, e qual a estimativa de valor ainda não licitado.

Também será possível realizar buscas por Cidades, **Figura 9**, com informações precisas das necessidades individuais e possibilidade de atuação.

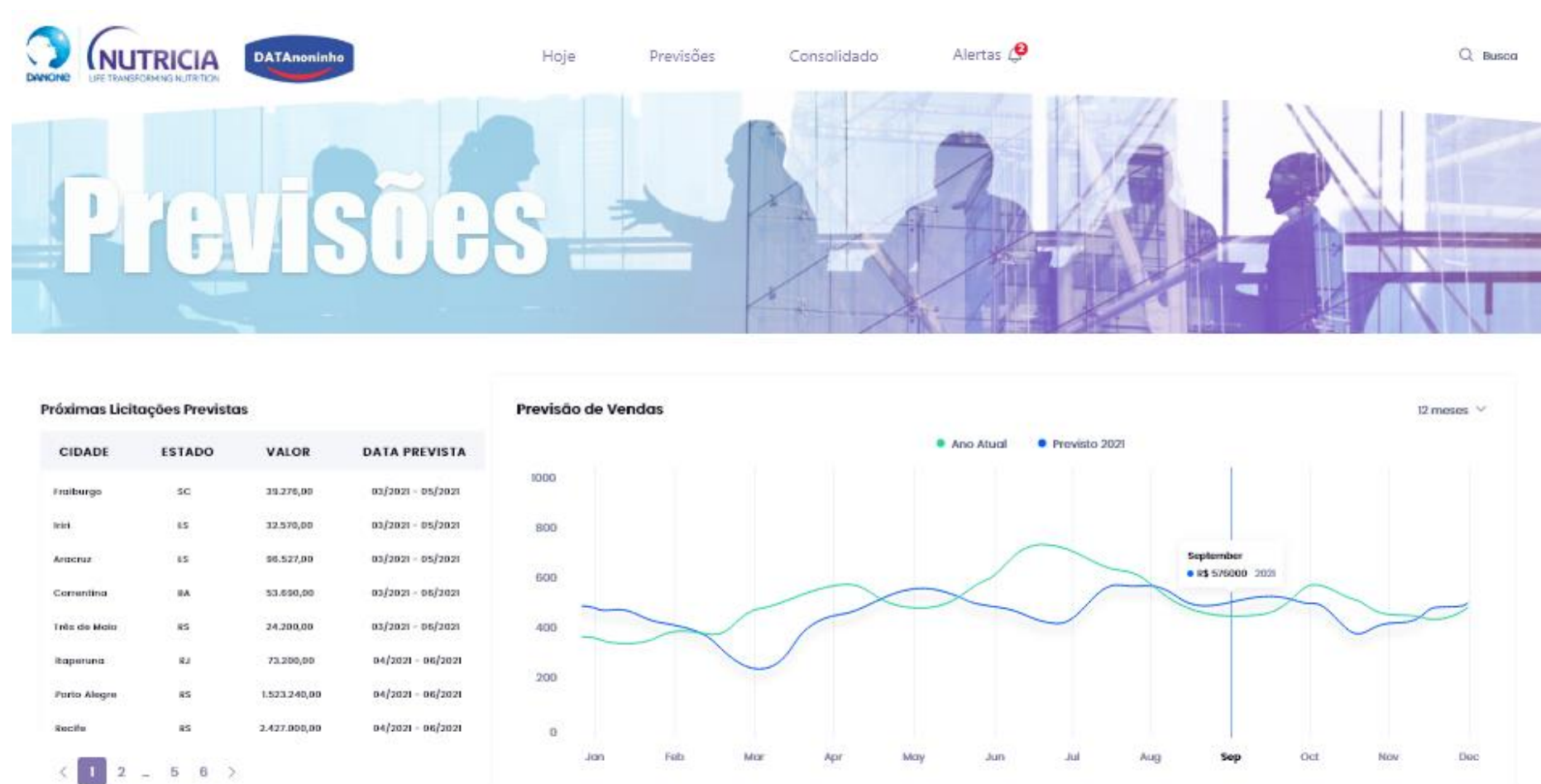
Com isso, a ideia é que se possa organizar toda essa inteligência de negócios em um único local, facilitando a pesquisa e o acesso as informações abrangentes dos processos licitatórios, assim como também trazer uma atualização dos mesmos existentes e os ganhos proporcionais vinculados a eles.

Figura 5: Tela de licitações disponíveis no dia



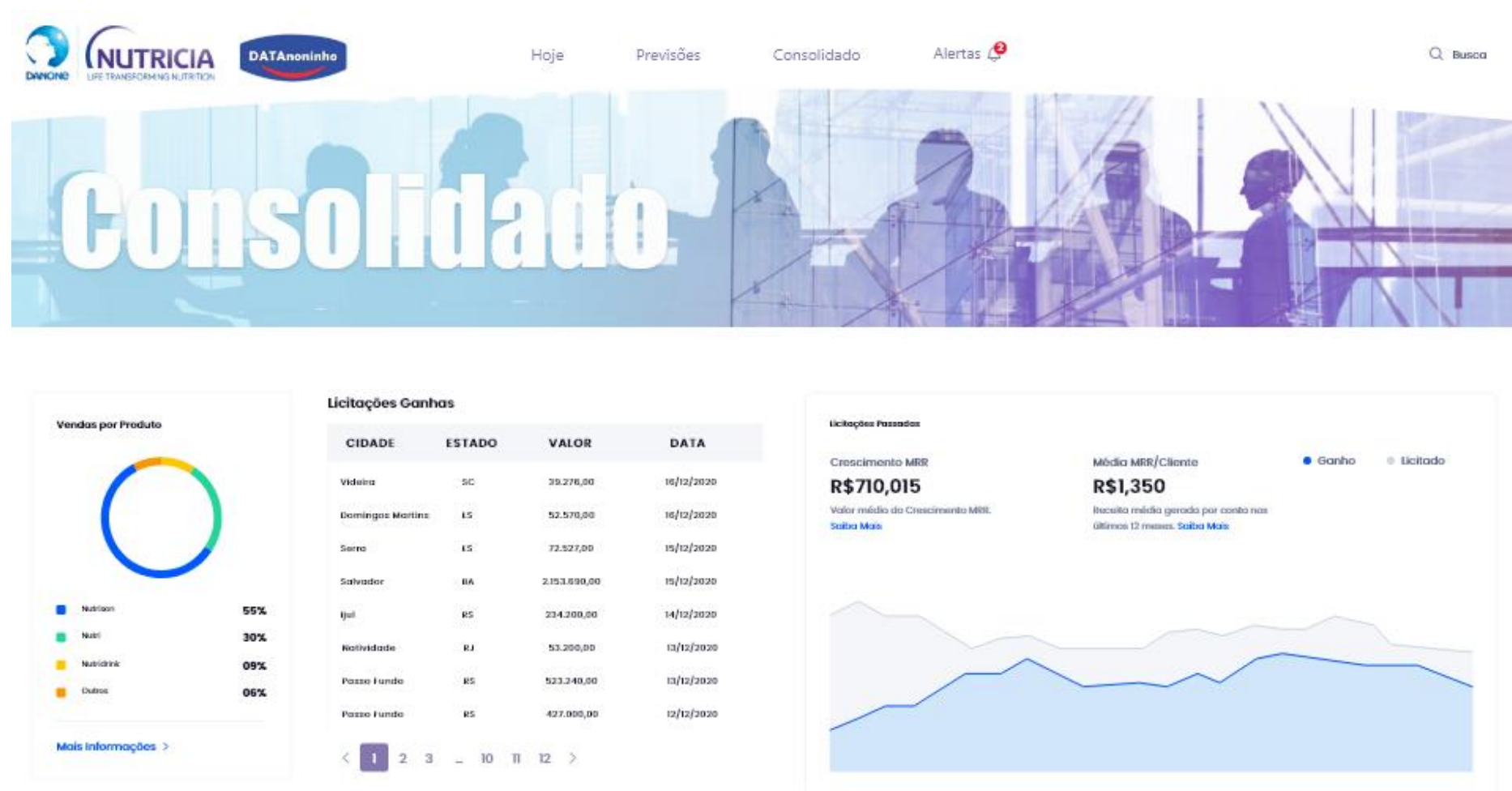
Fonte: Os autores, 2020.

Figura 6: Tela de previsões de licitações.



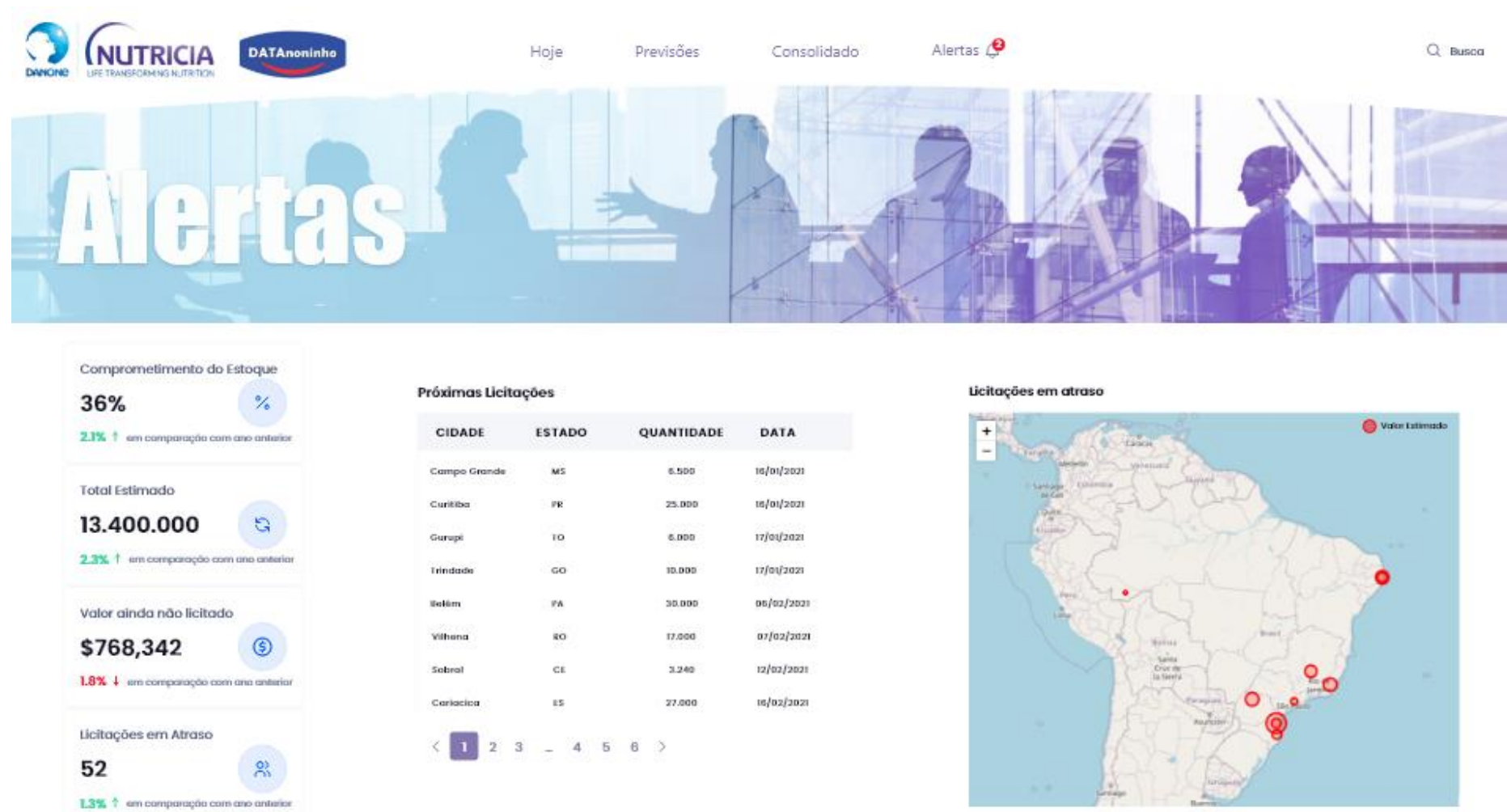
Fonte: Os autores, 2020.

Figura 7: Tela de licitações ocorridas.



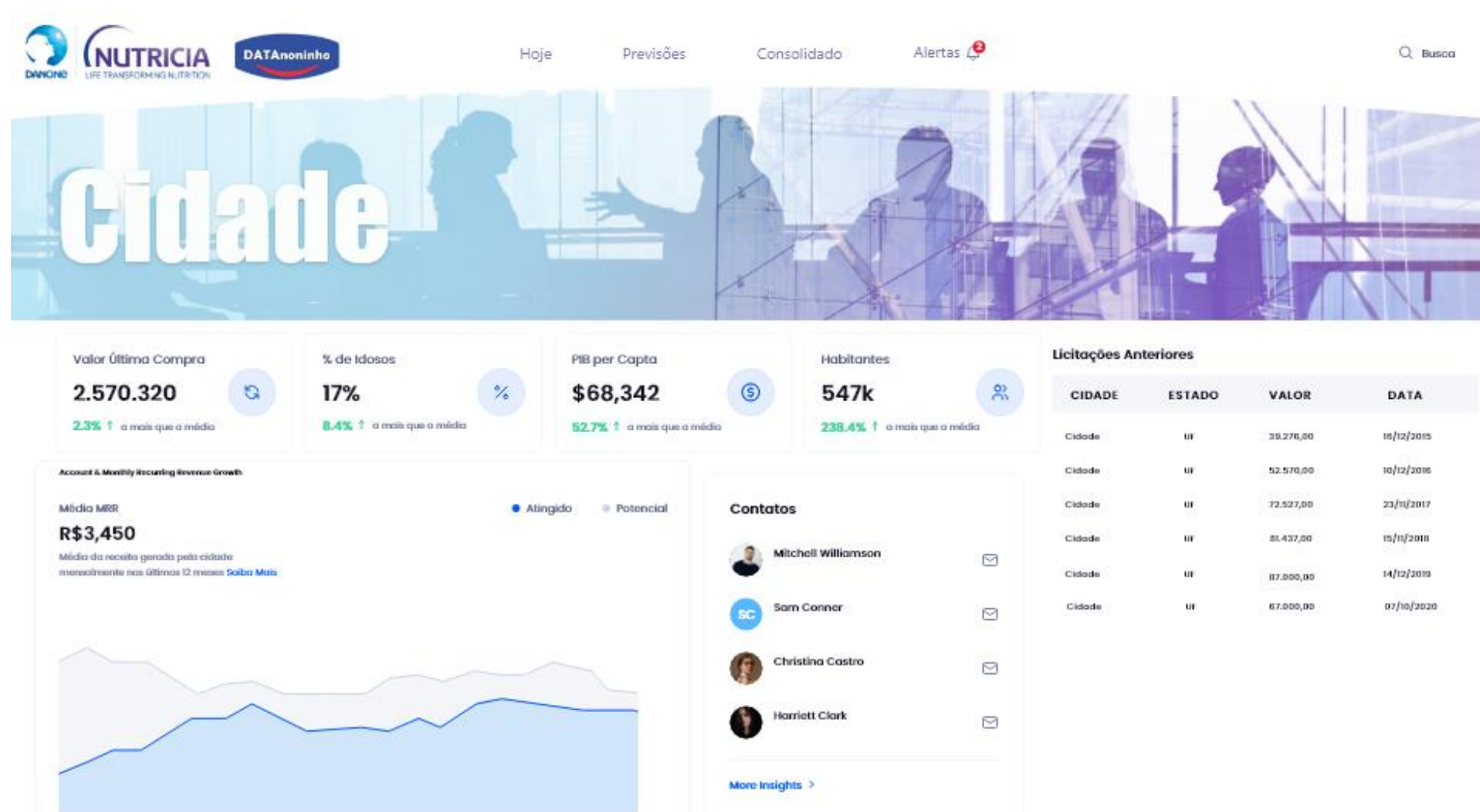
Fonte: Os autores, 2020.

Figura 8: Tela de alerta de licitações e disponibilidade de estoque para atender a demanda.



Fonte: Os autores, 2020.

Figura 9: Tela de licitações previstas por cidade.



Fonte: Os autores, 2020

7. CONCLUSÕES

A ampliação do número de idosos no Brasil e os tratamentos disponíveis tendem a ser de alto custo, em termos tanto de diagnóstico como de tratamento. O montante dos gastos com saúde dependerá essencialmente o tipo de doença e suas possíveis consequências, que podem ser reduzidas por investimentos na prevenção e no tratamento das pessoas ao longo da vida, não apenas na idade avançada

As quedas correspondem a 50% dos motivos de internações por causas externas da população idosa. Como consequência, pode-se citar a redução ou perda da autonomia, da independência, da qualidade de vida da pessoa idosa e aumento de danos sociais, financeiros, emocionais e mentais, levando à hospitalização, institucionalização e aumento da morbidade e mortalidade.

Intervenções para prevenir e tratar a sarcopenia têm demonstrado ser eficazes no aumento da massa muscular, força física e desempenho. A suplementação nutricional associada à atividade física evidencia a competência na prevenção e no tratamento da sarcopenia e osteoporose. Dentro dos suplementos nutricionais, destacam-se os mais utilizados, sendo eles: proteínas, aminoácidos e vitamina D.

Uma das formas de garantir uma boa nutrição e efetivamente prevenir osteoporose e sarcopenia é através de ações do sistema de saúde. A disponibilização de um portal para acompanhamento e previsão de processos licitatórios poderá aumentar a eficiência dos processos das empresas proponentes.

8. REFERÊNCIAS

ALEXANDRE, T. S. et al. **Prevalência e fatores associados à sarcopenia, dinapenia e sarcodinapenia em idosos residentes no Município de São Paulo - Estudo SABE**. Revista Brasileira de Epidemiologia. São Paulo, v. 21, supl. 2, 180009, 2019. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-790X2018000300406&lng=en&nrm=iso>. Acesso em 12 de dezembro de 2020.

ANS - Agência Nacional de Saúde Suplementar. **Sistema de informações de Beneficiários**. 2020. Disponível em: <<http://ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor/sala-de-situacao>>. Acesso em 12 de dezembro de 2020.

ANVISA - Agência Nacional de Vigilância Sanitária. Resolução RDC nº 269, de 22 de setembro de 2005. **Regulamento Técnico sobre a Ingestão Diária Recomendada (IDR) de proteína, vitaminas e minerais**. Disponível em: <<https://coffito.gov.br/nsite/wp-content/uploads/2016/08/resolucao-rdc-n-269-2005-ingestao-diaria-recomendada-idr-de-protenas-vitaminas-e-minerais.pdf>>. Acesso em 10 de dezembro de 2020.

BRASIL. **Estatuto do idoso: Lei Federal nº 10.741**, de 01 de outubro de 2003. Brasília, DF: Secretaria Especial dos Direitos Humanos, 2004. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2003/l10.741.htm#art118>. Acesso em 12 de dezembro de 2020.

CELICH, K.L.S. et al. **Fatores que predispõem às quedas em idosos**. Revista Brasileira de Ciências do Envelhecimento Humano. Passo Fundo, v. 7, n. 3, p. 419-426, 2010. Disponível em: <<http://www.seer.upf.br/index.php/rbceh/article/view/776>>. Acesso em 11 de dezembro de 2020.

DISTRITO. **Estratégia da Danone Nutricia foca em novos negócios**. Inovação. 18/09/2020. Disponível em <<https://distrito.me/danone-nutricia-novos-negocios/>>. Acesso em 11 de dezembro de 2020.

GUSMÃO, C.M.G e FILHO, S.V. **Análise de quedas na população do estado de Pernambuco: mapeamento em base de dados do DATASUS**. Revista Brasileira de Inovação Tecnológica em Saúde, v. 8, n. 1 (2018). Disponível em: <https://www.researchgate.net/publication/328945792_ANALISE_DE_QUEDAS_NA_POPULACAO_DO_ESTADO_DE_PERNAMBUCO/fulltext/5becc65da6fdcc3a8dd6e059/ANALISE-DE-QUEDAS-NA-POPULACAO-DO-ESTADO-DE-PERNAMBUCO.pdf>. Acesso em 14 de dezembro de 2020.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Projeções da População do Brasil e Unidades da Federação por sexo e idade: 2010-2060**. 2018. Disponível

em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html?=&t=resultados>>. Acesso em 10 de dezembro de 2020.

____ - Instituto Brasileiro de Geografia e Estatística. **Estimativas da população**. Tabelas. Estimativas de população enviadas ao TCU. 2020a. Disponível em: <<https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?=&t=o-que-e>>. Acesso em 13 de dezembro de 2020.

____ - Instituto Brasileiro de Geografia e Estatística. **Pesquisa de Orçamentos Familiares**, 2017-2018. Tabelas - Avaliação Nutricional da Disponibilidade Domiciliar de Alimentos no Brasil. Disponível em <<https://www.ibge.gov.br/estatisticas/sociais/protecao-social/24786-pesquisa-de-orcamentos-familiares-2.html?edicao=29142&t=sobre>>. Acesso em 12 de dezembro de 2020.

____ - Instituto Brasileiro de Geografia e Estatística. SIDRA - Sistema IBGE de Recuperação Automática. **Estimativas da população. Tabela 6579 - População residente estimada**. 2020b. Disponível em: <<https://sidra.ibge.gov.br/tabela/6579>>. Acesso em 10 de dezembro de 2020.

KRAUSE, M.M.; MAHAN, L. K.; ESCOTT-STUMP, S. **Alimentos Nutrição e Dietoterapia**. Rio de Janeiro: Elsevier, 12^a ed., 2011.

MALAFARINA, V. *et al.* **Effectiveness of nutritional supplementation on muscle mass in treatment of sarcopenia in old age: a systematic review**. J. Am. Med. Dir. Assoc. v. 14, n. 1, p. 10-17, 2013. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S1525861012002459>>. Acesso em 11 de dezembro de 2020.

NASCIMENTO, I.M e MOREIRA, P.A. **Efeito da suplementação nutricional no idoso com sarcopenia: uma revisão sistemática**. Cientefico. V. 19, N. 40, Fortaleza, jul./dez. 2019. Disponível em: <<https://cientefico.emnuvens.com.br/cientefico/article/download/613/415>>. Acesso em 12 de dezembro de 2020.

PETERS, B.S.E.; MARTINI, L. A.. **Nutritional aspects of the prevention and treatment of osteoporosis**. Arquivos Brasileiros de Endocrinologia & Metabologia, São Paulo, v. 54, n. 2, p. 179-185, 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0004-27302010000200014&lng=en&nrm=iso>. Acesso em: 12 de dezembro de 2020.

ROSA, T. S. M. et al. **Perfil epidemiológico de idosos que foram a óbito por queda no Rio Grande do Sul**. Revista Brasileira de Geriatria e Gerontologia, 18(1), 59-69. Rio de Janeiro. 2015. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1809-98232015000100059&lng=en&nrm=iso>. Acesso em: 14 de dezembro de 2020.

SILVA, M.H.F et al. **Pesquisa dos fatores de risco para quedas na população idosa de uma unidade básica do município de Itaúna – MG**. Revista Médica de Minas Gerais, 2018. Disponível em: <<http://rmmg.org/exportar-pdf/2362/e1938.pdf>> Acesso em 13 de dezembro de 2020.

WINGERTER DG et al. **Mortalidade por queda em idosos: uma revisão integrativa**. Revista Ciência Plural. v 6, n 1, UFRN, 2020. p. 119-136. Disponível em:<<https://periodicos.ufrn.br/rcp/article/view/18366/12532>>. Acesso em: 14 de dezembro de 2020.

YEUNG, S.S.Y et al. **Sarcopenia and its association with falls and fractures in older adults: A systematic review and meta-analysis**. Journal of Cachexia, Sarcopenia and Muscle. [Volume10, Issue3](https://onlinelibrary.wiley.com/doi/full/10.1002/jcsm.12411), p. 485-500. 2019. Disponível em <https://onlinelibrary.wiley.com/doi/full/10.1002/jcsm.12411> Acesso em 11 de dezembro de 2020.