

UNIVERSIDADE PRESBITERIANA MACKENZIE
FACULDADE DE COMPUTAÇÃO E INFORMÁTICA

André Dalle Vedove Canassa
Kaymmi Antunes Costa Silva
Matheus Gomes
Valdo Alvim da Rocha Junior

Estudo sobre Doações de Alimentos

SÃO PAULO
2023

André Dalle Vedove Canassa – 23021292

Kaymmi Antunes Costa Silva – 23018690

Matheus Gomes – 22516220

Valdo Alvim da Rocha Junior – 23007796

**Análise das Doações de Alimentos
Realizadas por Estabelecimentos
Comerciais no Brasil**

Trabalho final da disciplina de Projeto Aplicado I apresentado à Faculdade de Computação e Informática da Universidade Presbiteriana Mackenzie, como parte dos requisitos obrigatórios de conclusão de curso.

Docente: Prof. Everton Knihs.

SÃO PAULO

2023

RESUMO

Neste trabalho, abordamos o problema do desperdício de alimentos no Brasil, que impacta tanto o meio ambiente quanto a segurança alimentar. Utilizamos um dataset obtido do Kaggle, contendo informações detalhadas sobre doações de alimentos por estabelecimentos comerciais. Através de análises exploratórias, identificamos tendências temporais, estatísticas-chave e correlações entre variáveis. Detectamos a presença de outliers e analisamos categorias categóricas relevantes, como tipos de alimentos e principais fornecedores.

O objetivo principal é identificar oportunidades para melhorar a eficiência das doações de alimentos e reduzir a fome no Brasil. Utilizamos o python em nossa abordagem metodológica, que incluiu KDD e mineração de dados.

Sumário

1	INTRODUÇÃO	3
1.1	Contexto do estudo.....	3
1.2	Descrição da origem.....	3
1.3	Descrição do dataset.....	3
2	DESENVOLVIMENTO E ANÁLISE DE DADOS	4
2.1	Ferramentas de modelagem.....	4
2.2	Objetivos.....	4
2.3	Metodologia.....	4
2.4	Resultados esperados.....	5
2.5	Cronograma.....	5
3	APRESENTAÇÃO/ANÁLISE EXPLORATÓRIA DE DADOS.....	6
3.1	Proposta analítica.....	6
3.2	Análise Exploratória de Dados.....	6
3.3	Descrição Estatística dos dados.....	6
3.4	Quantidade.....	6
3.5	Valor calórico.....	7
3.6	Dias restantes.....	7
3.7	Análise temporal das doações.....	8
3.8	Correlação entre as variáveis.....	9
3.9	Análise de Outliers.....	9
3.10	Análise de categorias.....	9
4	REFERENCIAS	11
5	GLOSSÁRIO	12

1 INTRODUÇÃO

1.1 Contexto do estudo

O desperdício de alimentos é um problema global que afeta tanto o meio ambiente quanto a segurança alimentar. No Brasil, estima-se que cerca de 40 mil toneladas de alimentos sejam desperdiçadas por dia, o que equivale a 15 bilhões de reais por ano. Uma forma de combater esse problema é incentivar as doações de alimentos por parte dos estabelecimentos comerciais que possuem excedentes ou produtos próximos ao vencimento. Essas doações podem beneficiar pessoas em situação de vulnerabilidade social, além de reduzir o impacto ambiental da produção e descarte de alimentos.

1.2 Descrição da origem

O dataset "Registros de Doações de Alimentos" foi obtido a partir do site Kaggle, uma plataforma online que disponibiliza conjuntos de dados públicos para fins de análise e aprendizado de máquina. O dataset foi criado pelo usuário buzzlightyear64, que coletou e organizou os dados a partir de fontes confiáveis, como o Programa Mesa Brasil Sesc, o Banco de Alimentos e o Instituto Stop Hunger..

1.3 Descrição do dataset

O dataset contém oito (8) colunas com informações detalhadas sobre doações de alimentos realizadas por diferentes estabelecimentos comerciais. O conjunto de dados abrange informações como o nome do estabelecimento comercial, o tipo de alimento doado, a data da doação, a data de validade do alimento, a quantidade doada, o valor calórico do alimento, o tipo de alimento (perecível ou não perecível) e os dias restantes até a data de validade.

Esses registros refletem uma variedade de estabelecimentos, incluindo supermercados conhecidos, como Walmart e Carrefour, e restaurantes renomados, como Fogo de Chão, D.O.M., Maní e Mocotó. Os alimentos doados também apresentam uma ampla gama de tipos, desde itens básicos da

despensa, como arroz, feijão e macarrão, até pratos preparados, como frango assado, bife à parmegiana e pizzas.

Cada registro contém informações essenciais para acompanhar e gerenciar as doações de alimentos. A data da doação permite rastrear quando o alimento foi doado, enquanto a data de validade fornece informações importantes sobre a vida útil do produto. A quantidade doada indica a quantidade específica de cada alimento, permitindo uma melhor compreensão do volume de doações realizadas. Além disso, o valor calórico do alimento pode ser útil para avaliar o conteúdo nutricional das doações.

Uma informação relevante presente, no dataset é o tipo de alimento, que é classificado como perecível ou não perecível. Essa classificação é essencial para identificar alimentos que possuem uma vida útil limitada e precisam ser consumidos rapidamente, garantindo que não sejam desperdiçados. Os dias restantes até a data de validade fornecem uma medida clara da urgência em consumir determinados alimentos antes que eles expirem.

Entretanto, encontramos um problema na coluna de data da doação, onde constatamos que há datas no futuro o que prejudica o rastreamento de quando o alimento foi doado.

2 DESENVOLVIMENTO E ANÁLISE DOS DADOS

2.1 Ferramenta de modelagem

Estaremos trabalhando com modelos em Python e R com apoio do Excel e Power Bi em algumas demonstrações.

2.2 Objetos

Data Set de Doações
Ferramentas de análises

2.3 Objetivos

Identificar oportunidades de aumentar a eficiência e eficácia das doações de alimentos realizadas por estabelecimentos privados para a redução da fome no Brasil

2.4 Metodologia

Estaremos realizando os procedimentos de KDD e mineração de dados. Com estas ações, procurando modelos estatísticos com calculos para propor cenários de melhoria no panorama.

2.5 Resultados esperados

Cenários de melhor contribuição da redução da fome e de sustentabilidade no aproveitamento de produtos antes do vencimento.

2.6 Cronograma

DATA		Item
08/08/2023	Terça-Feira	ES 19:30 as 20:30 1/2 - AULA 1
12/08/2023	Sábado	Criação de Grupo Whats´up / Registro do grupo no fórum
14/08/2023	Segunda-Feira	Realizada Reunião de Equipe nº1
15/08 A 18/08		Pesquisas, estudos, avaliações para propor possibilidades de Assunto do PA
18/08/2023	Sexta-Feira	Reunião de Equipe
21/08/2023	Segunda-Feira	Definir Assunto do PA
22/08/2023	Terça-Feira	ES 19:30 as 20:30 2/2 - AULA 1
26/08/2023	Sábado	DEFINIÇÃO DO ASSUNTO DE PA
27/08/2023	Domingo	ELABORAÇÃO DO DOCUMENTO
28/08/2023	Segunda-Feira	ENTREGA 1
29/08 A 04/09		ANALISE EXPLORATÓRIA
05/09/2023	Terça-Feira	ES 19:30 as 20:30 1/2 - AULA 2
06/09 A 12/09		ELABORAÇÃO DA PROPOSTA ANALÍTICA
19/09/2023	Terça-Feira	ES 19:30 as 20:30 2/2 - AULA 2
20/09 A 24/09		CONSIDERAÇÕES FINAIS PARA A ENTREGA 2
25/09/2023	Segunda-Feira	ENTREGA 2
10/10/2023	Terça-Feira	ES 19:30 as 20:30 1/2 - AULA 3
16/10 A 30/10		PRODUÇÕES DO CONTEÚDO DE STORYTELLING DE APRESENTAÇÃO/CONSIDERAÇÕES
26/10/2023	Quinta-Feira	ES 19:30 as 20:30 1/2 - AULA 3
30/10/2023	Segunda-Feira	ENTREGA 3

3 APRESENTAÇÃO/ANÁLISE EXPLORATÓRIA DE DADOS

Visando o acesso completo ao estudo realizado, foi disponibilizado no repositório o dataset nomeado: Dataset_Registro_de_doacoes_de_alimentos e o arquivo: PA.ipynb que possui toda a estrutura utilizada para a exploração de dados. Abaixo faremos uma breve síntese dos estudos iniciados:

3.1 Proposta analítica

Nesta seção, vamos apresentar a proposta analítica, que se baseará na análise dos dados fornecidos. O objetivo principal é entender o comportamento das doações de alimentos ao longo do tempo, identificar tendências, padrões e possíveis insights que possam ser úteis para a tomada de decisões e o planejamento de futuras doações.

3.2 Análise Exploratória de Dados

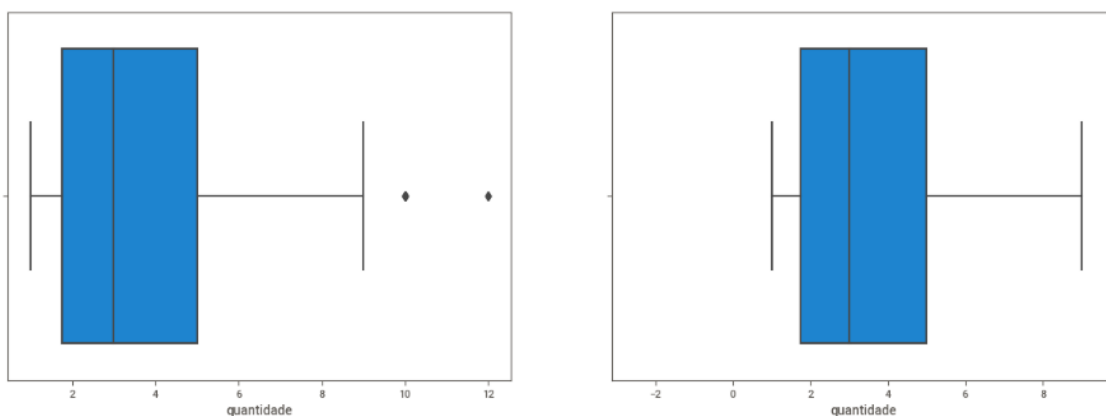
A análise exploratória de dados é uma etapa crucial para compreender o conjunto de dados e extrair informações relevantes. Neste estudo, utilizamos diversas ferramentas de análise de dados, incluindo visualizações gráficas e estatísticas descritivas.

3.3 Descrição Estatística dos Dados

Primeiramente, realizamos uma análise descritiva dos dados para obter uma visão geral das variáveis envolvidas nas doações de alimentos. Abaixo, apresentamos algumas estatísticas-chave:

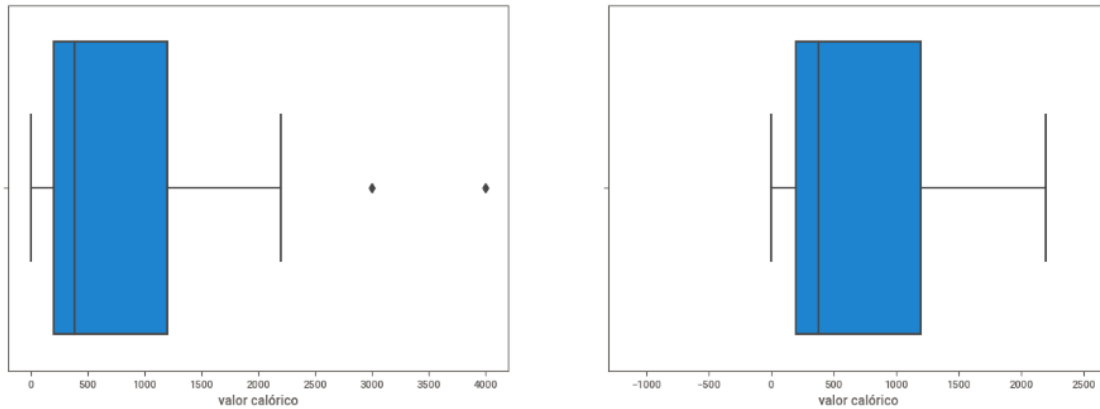
3.4 Quantidade

A quantidade média de alimentos doados é de aproximadamente 3.87 unidades, variando de 1 a 12 unidades. A mediana é de 3 unidades.



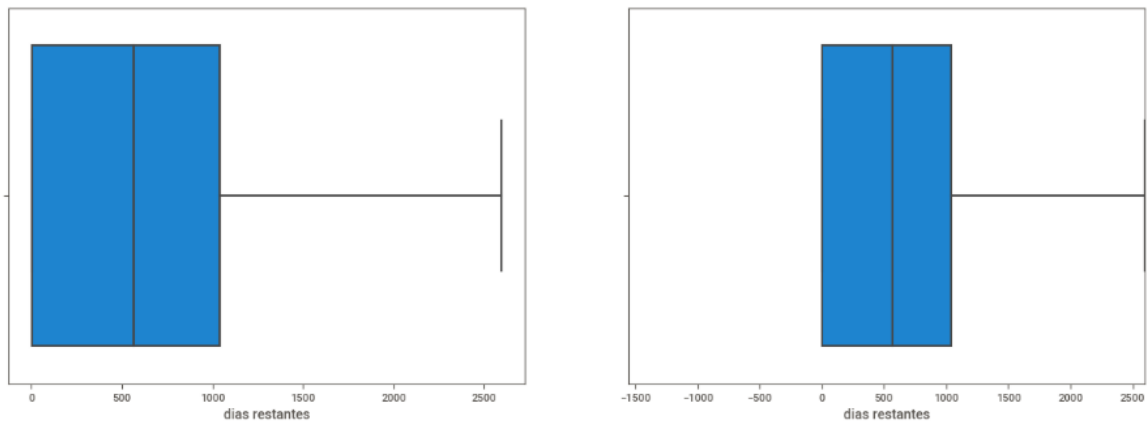
3.5 Valor Calórico

O valor calórico médio dos alimentos doados é de cerca de 738.42 calorias, variando de 0 a 4000 calorias. A mediana é de 380 calorias.



3.6 Dias restantes

O prazo médio de validade dos alimentos doados é de aproximadamente 658.37 dias, variando de 3 a 2596 dias. A mediana é de 567.50 dias.

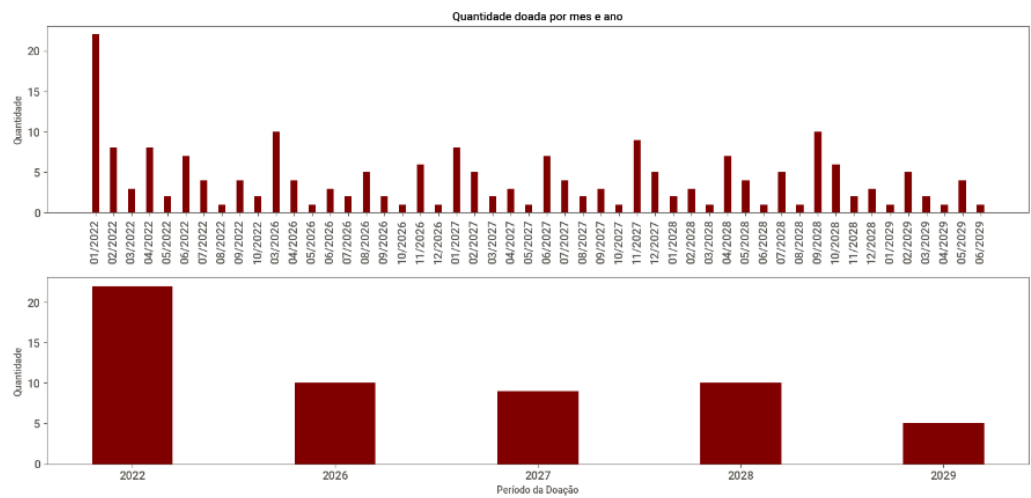


3.7 Análise temporal das doações

Para entender como as doações de alimentos evoluíram ao longo do tempo, analisamos os dados relacionados às datas de doação e datas de validade.

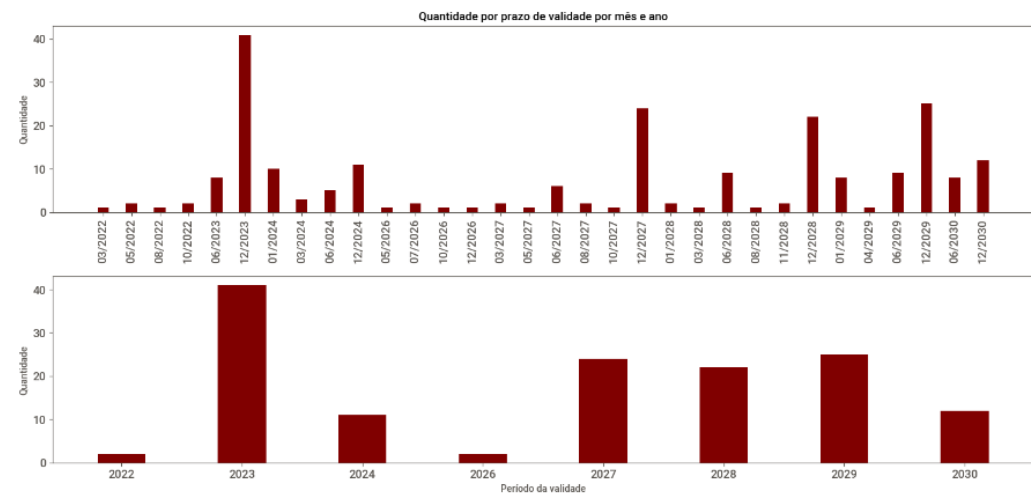
a. Quantidade doada por mês e ano

- Observamos que a quantidade de alimentos doados tende a variar ao longo dos meses e anos. Picos de doações podem ser identificados em alguns períodos, enquanto outros têm menos doações. Isso pode indicar sazonalidade nas doações ou eventos específicos que impulsionaram as doações.



b. Quantidade por prazo de validade por mês e ano:

- Examinamos também a relação entre a quantidade de alimentos doados e o prazo de validade. Novamente, notamos variações significativas na quantidade de alimentos doados em relação à validade. Isso pode sugerir que as doações estão correlacionadas com a proximidade da data de validade dos alimentos.



3.8 Correlações entre as variáveis

Também calculamos as correlações entre as variáveis quantitativas, como quantidade, valor calórico e dias restantes. A matriz de correlação nos ajuda a entender se existe alguma relação linear entre essas variáveis.

Quantidade vs. Valor Calórico: A correlação entre quantidade e valor calórico parece ser baixa, indicando que a quantidade de alimentos doados não está fortemente relacionada ao seu valor calórico.

Quantidade vs. Dias Restantes: A correlação entre quantidade e dias restantes também é baixa, sugerindo que a quantidade doada não depende fortemente do prazo de validade.

3.9 Análise de Outliers

Realizamos uma análise de outliers para identificar valores extremos nas variáveis quantitativas. Para isso, utilizamos diagramas de caixa (boxplots) que nos permitem visualizar a distribuição dos dados e identificar possíveis outliers.

- Identificamos outliers nas variáveis de quantidade, valor calórico e dias restantes. Esses valores atípicos podem indicar casos excepcionais de doações.

3.10 Análise de Categorias

Por fim, analisamos as categorias categóricas presentes nos dados, como o tipo de alimento, o comércio (fornecedor) e o próprio alimento doado.

Tipo de Alimento: Observamos a distribuição dos tipos de alimentos doados para entender quais categorias são mais comuns nas doações.

Maiores fornecedores: Identificamos os principais fornecedores (comércio) que mais contribuem com doações.

Maiores Alimentos: Identificamos os alimentos individuais que são mais frequentemente doados.

Essa análise nos permite entender as preferências de doação e as parcerias mais significativas com fornecedores específicos.

4 REFERÊNCIAS

Acesso ao dataset: <https://www.kaggle.com/datasets/buzzlightyear64/registros-doacoes-alimentos?resource=download>

Acesso ao projeto no Github: <https://github.com/Equipe-de-projetos-I-CD-Mackenzie>

Contato do integrante: André Dalle Vedove Canassa
TIA: 23021292 E-mail: andredvcanassa@gmail.com

Contato do integrante: Kaymmi Antunes Costa Silva
TIA: 23018690 E-mail: kaymmicosta@hotmail.com

Contato do integrante: Matheus Gomes
TIA: 22516220 E-mail: matheus.vinicius97@hotmail.com

Contato do integrante: Valdo Alvim da Rocha Junior
TIA: 23007796 E-mail: valdo.rocha@yahoo.com.br

6 GLOSSÁRIO

Dataset: Um dataset, em tradução livre, pode ser chamado de conjunto de dados. É uma coleção estruturada de informações, geralmente organizada em tabelas, planilhas ou outros formatos, que contém dados relevantes para análise, pesquisa ou processamento.

Outliers: são pontos de dados que se destacam devido a sua diferença significativa em relação aos outros dados e podem ter um impacto substancial nas análises estatísticas e de dados, se não forem adequadamente identificados e tratados.