

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EM GESTÃO E TECNOLOGIA
DEPARTAMENTO DE COMPUTAÇÃO

Bruno dos Santos Ferraz | 796085
Marcus Vinícius Caruso Leite | 792200
Vinicius Carvalho Venturini | 793250

Processamento Massivo de Dados
Orientador(a): Prof. Sahudy Montenegro González

A acessibilidade da alimentação no mundo

Fase Intermediária I

Grupo 13

Sorocaba
17 de junho de 2025

Índice

1. Resumo do tema	3
2. Tecnologias utilizadas	4
3. Objetivo	4

1. Resumo do tema

O projeto tem como tema a acessibilidade da alimentação no mundo, utilizando dados públicos disponíveis na plataforma *ourworldindata*, especificamente sobre preços e disponibilidade de alimentos. Para isso, o Apache Spark será utilizado para o processamento e cruzamento de dados em JSON, utilizando atributos em comum para cruzar os dados como: país, ano, continente e etc. O banco de dados NoSQL MongoDB será utilizado para armazenar os resultados desses cruzamentos, para facilitar a análise desses dados por meio de consultas *Map/Reduce* e agregações (Figura 1). O objetivo é explorar aspectos relevantes de acessibilidade alimentar globalmente, identificando padrões e disparidades entre regiões.

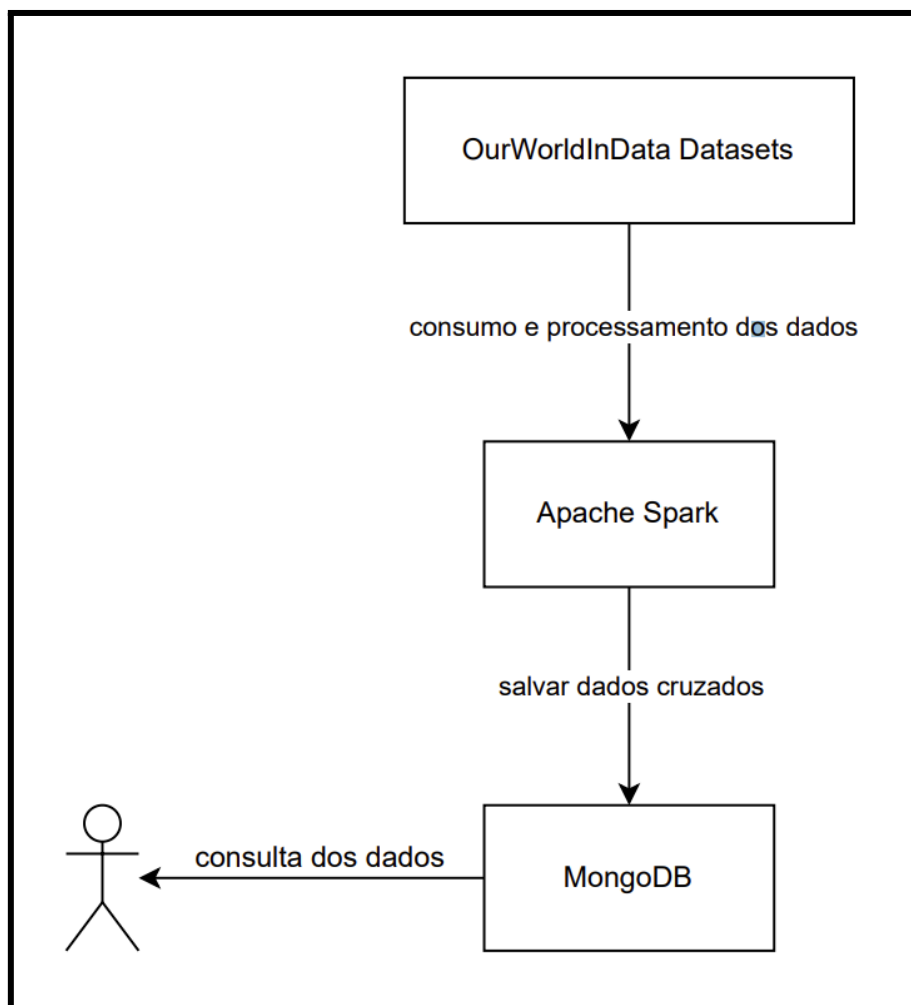


Figura 1: Fluxograma do funcionamento

2. Tecnologias utilizadas

Spark: Utilizado para o cruzamento de dados de diversas fontes de dados (JSON), visto que consegue lidar com um grande conjunto de dados.

Mongo: Utilizado para salvar e consultar os dados cruzados utilizando consultas map/reduce e agregações.

3. Objetivo

Cruzar dados sobre alimentação ao redor do mundo, analisando sua acessibilidade e outras informações relevantes por meio de consultas (Tabela 1).

Título da consulta	Descrição	Dataset 1	Dataset 2
Retorne a porcentagem de quantos não conseguem sustentar uma dieta saudável, incluindo a porcentagem do preço superfaturado do continente de 2017 até 2022	A ideia é mostrar um comparativo da porcentagem da população por país que consegue manter uma dieta saudável e o quanto isso é caro em questão de inflação.	https://ourworldindata.org/grapher/share-healthy-diet-unaffordable?time=latest	https://ourworldindata.org/grapher/share-of-countries-recording-high-food-prices?tab=table
Dado um tipo de comida (vegetais, carne vermelha, frutas e etc.) retorne o gasto diário em dólar por país comparado dividido pelo custo diário em dólar de uma dieta saudável em 2021.	Dividindo o gasto diário em uma comida específica pelo gasto diário em uma dieta saudável, temos a porcentagem do comparativo do gasto que uma pessoa faz com comidas específicas com a dieta saudável diariamente naquele país.	https://ourworldindata.org/grapher/cost-foods-healthy-diet?tab=table	https://ourworldindata.org/grapher/cost-healthy-diet?time=2022
Dado um ano X, retorne quanto uma pessoa gastou com comida naquele ano e inclua o gasto feito em tabaco e álcool daquele mesmo ano por país.		https://ourworldindata.org/grapher/food-expenditure-per-person-per-year	https://ourworldindata.org/grapher/share-expenditure-alcohol-tobacco

Evolução do custo de uma dieta saudável em relação à renda média per capita	Mostrar como o custo diário de uma dieta saudável evoluiu ao longo dos anos em comparação com a renda média per capita da população, para identificar se ficou mais acessível ou mais difícil manter uma dieta saudável.	https://ourworldindata.org/grapher/cost-healthy-diet	https://ourworldindata.org/grapher/gdp-per-capita-worldbank
Comparar custo da dieta saudável com a prevalência de insegurança alimentar	Comparar se um país com o custo menor de uma dieta saudável, possui também uma menor taxa de pessoas com insegurança alimentar.	https://ourworldindata.org/grapher/cost-healthy-diet	https://ourworldindata.org/grapher/share-of-population-with-moderate-or-severe-food-insecurity

Tabela 1: Dados a serem cruzados e explorados