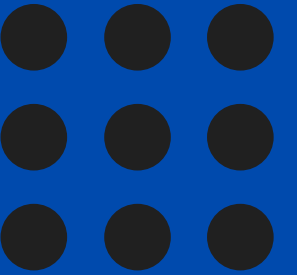


30 DE JANEIRO DE 2023

# INTRODUÇÃO A DATA SCIENCE: TRABALHO INDIVIDUAL I

**Aluna:** Luana Leite Rodrigues

**Repositório:** <https://github.com/LuanaLeite/trabalho-individual>



# Contexto

## A EMPRESA

Vende produtos de emagrecimento via canais digitais. Conta com uma audiência de mais de 6M de inscritos no Youtube e, aproximadamente, 980mil seguidores no Instagram. Além da produção de conteúdo orgânico e utilização de estratégias de lançamento de produto, a empresa também investe cerca de R\$800 mil reais em tráfego pago por mês para captação de leads.

# Contexto

## PROBLEMÁTICA

A empresa está no mercado desde 2019 e, recentemente, vem enfrentando novos desafios em relação ao aumento do custo de aquisição de clientes e redução de margem de lucro. Dado esse cenário, a grande pergunta feita pelo principal gestor da empresa, demandante da análise, foi ***\*Quem são os nossos melhores clientes? O que eles compraram?\****.

# INFORMAÇÕES

## DADOS FORNECIDOS

Para responder a pergunta feita pelo gestor da empresa, foi disponibilizada uma base de vendas histórica de janeiro de 2020 até novembro de 2022. A base é constituída do histórico de 5 plataformas de vendas (hotmart, guru-info, guru-sup, upnid e nutlog). Para organização e manipulação dos dados, tudo foi concatenado em excel, com tratativas específicas para deixar as bases no mesmo formato e padronizar seus campos.

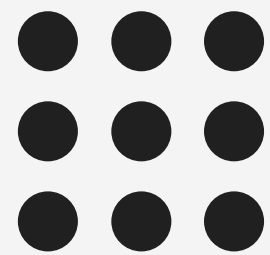
## DICIONÁRIO DE VARIÁVEIS

Para melhor compreensão dos dados fornecidos, foi desenvolvido um dicionário de variáveis que está disponível no link a seguir: [clique aqui para acessar](#).

# COLETA E PRÉ-PROCESSAMENTO

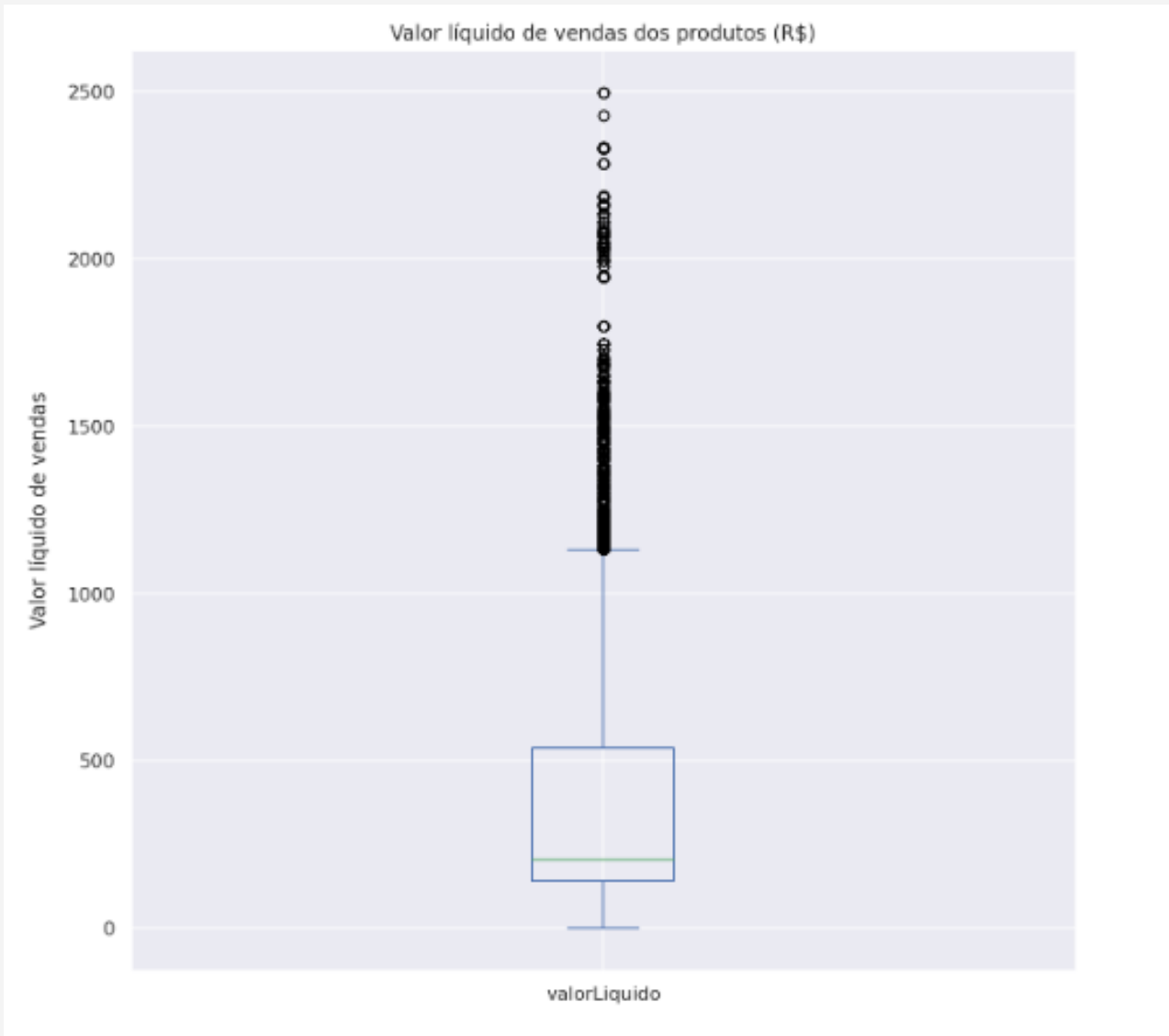
A empresa forneceu a base histórica das plataformas de vendas citadas anteriormente em Excel. Em domínio da base, foi necessário realizar ajustes de padronização, tratamento de formatação de variáveis como textos, números e datas. Após as alterações, foi gerado um documento em CSV. Devido ao tamanho do arquivo não foi possível salvar no repositório do Git, por isso, o arquivo com os dados foi salvo Google Drive da aluna.

# Análise exploratória



nomeProduto	valorLiquido	qtdProduto
Colágeno Tipo 2   30 dias	8638.455	117
Colágeno Tipo 2   Kit 60 dias	22902.825	165
Colágeno Tipo 2   Kit 90 dias	45473.045	223
Detox Sistêmico	604286.490	833
Ebook Detox de 15 Dias	282797.930	1612
...	...	...
Ômega Vita - Kit com 7 Potes	2946.200	5
Ômega Vita - Kit com 7 Potes (Frete grátis)	137001.420	218
Ômega Vita - Kit com 8 Potes	163056.170	333
Ômega Vita - Pote Unitário	10922.670	88
Ômega Vita - Pote Unitário (Frete grátis)	93085.280	739

188 rows x 2 columns



# Próximos passos

## COLETAR MAIS DADOS

Apesar do longo histórico da base de vendas fornecida pela empresa, para responder a pergunta chave levantada pelo gestor será importante ter acesso a base de dados não somente de vendas, mas também de leads gerados e valor investido em tráfego para mapear com mais precisão o perfil dos clientes, seu comportamento de compra e identificar possíveis alavancas para o negócio.