**Relatório de Análise de Dados: Insights para a Estratégia Bellabeat**

**1. Introdução**

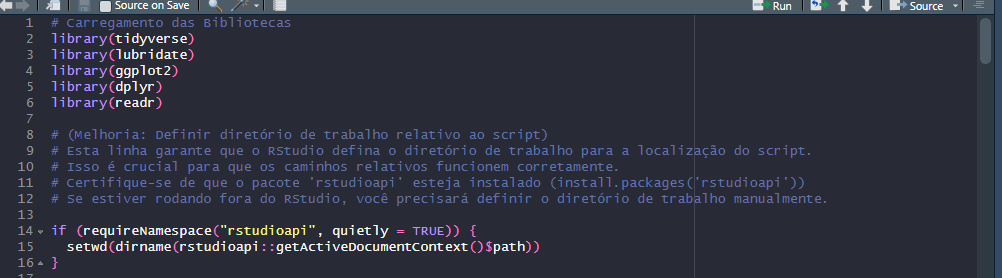
Este relatório técnico detalha o processo de análise de dados de dispositivos inteligentes não-Bellabeat, com o objetivo de extrair insights valiosos para aprimorar a estratégia de marketing da Bellabeat, focando especificamente no produto Bellabeat Leaf. A análise abrange padrões de atividade física, sono e peso, buscando identificar tendências de comportamento do consumidor que possam ser aplicadas para otimizar o engajamento dos usuários e o desenvolvimento de produtos.

**2. Visão Geral do Código e Metodologia**

O projeto foi desenvolvido em **R**, utilizando o ambiente **RStudio** e pacotes do **tidyverse** para manipulação e visualização de dados. A metodologia seguiu as etapas de carregamento, limpeza, transformação, fusão de dados, análise descritiva, segmentação de usuários e geração de visualizações.

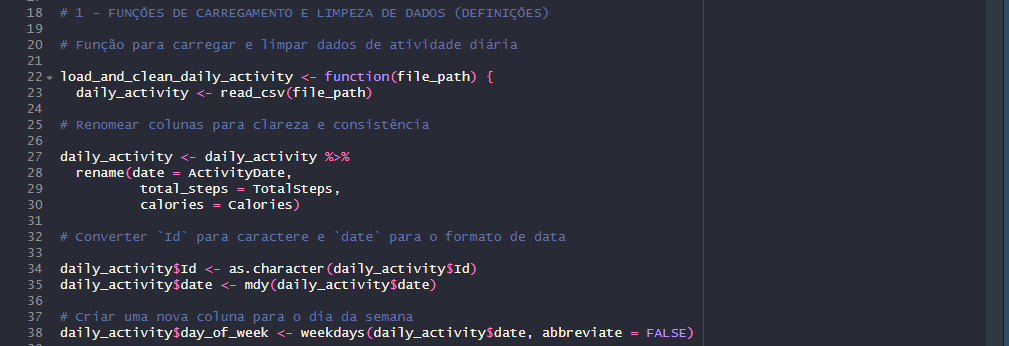
**2.1. Carregamento de Bibliotecas e Configuração do Ambiente**

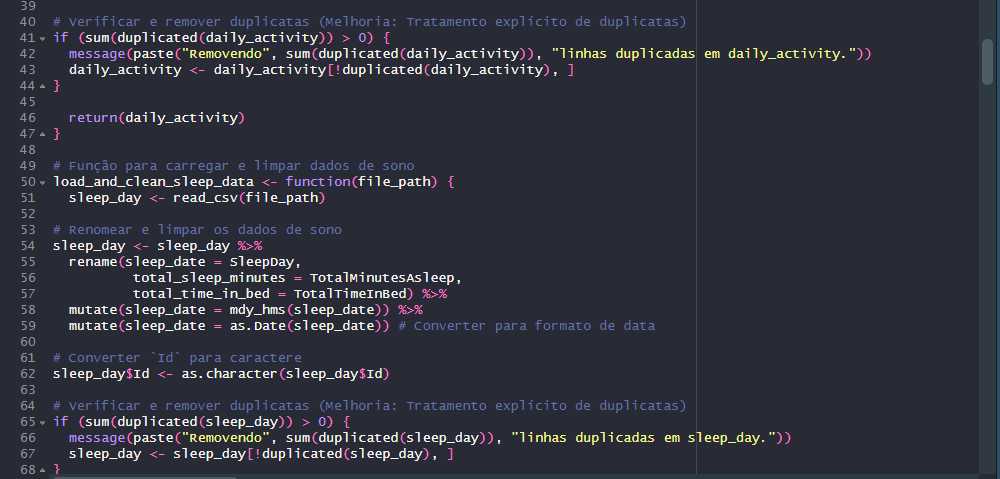
Esta seção inicial carrega as bibliotecas necessárias para a análise e configura o diretório de trabalho para garantir que todos os arquivos sejam localizados corretamente.

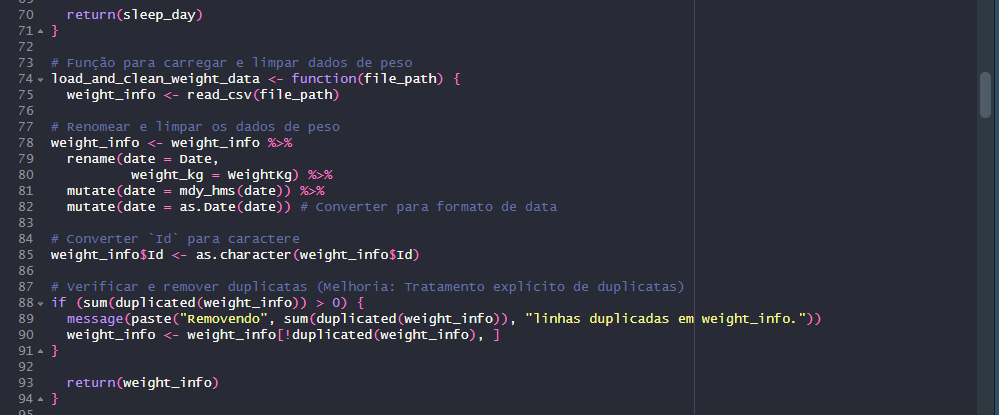


**2.2. Funções de Carregamento e Limpeza de Dados**

Foram criadas funções modularizadas para carregar e realizar a limpeza inicial de cada dataset (atividade diária, sono e peso). Isso inclui renomear colunas, converter tipos de dados (especialmente datas) e remover duplicatas, garantindo a consistência e qualidade dos dados antes da fusão.

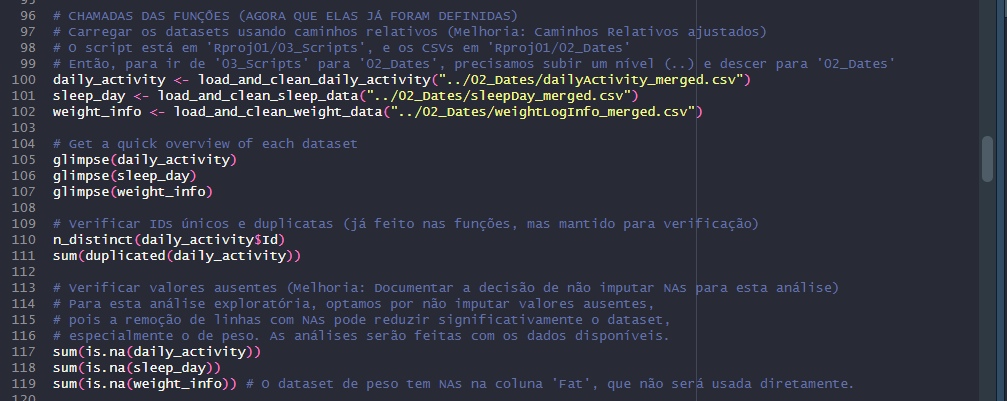


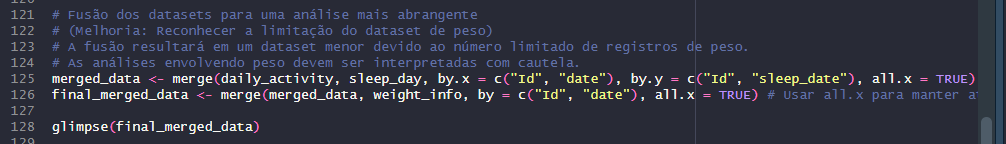




**2.3. Carregamento e Fusão dos Datasets**

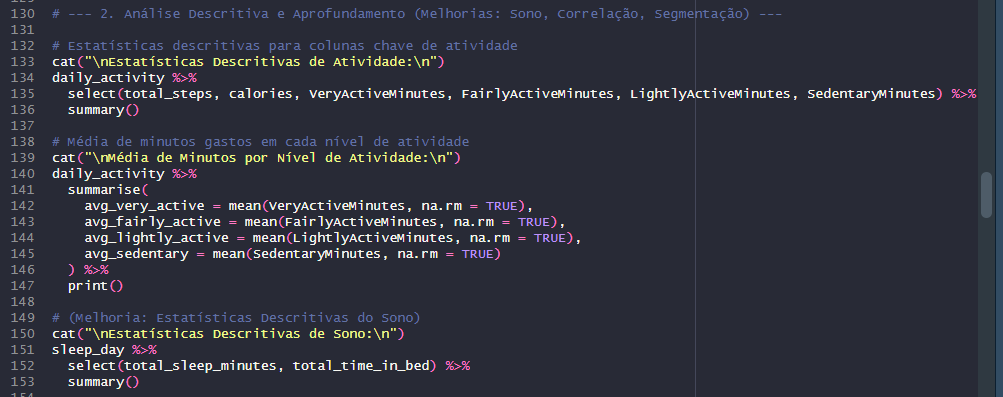
Após a definição das funções, os datasets são carregados utilizando caminhos relativos ajustados à estrutura do projeto e, em seguida, mesclados. A fusão é realizada com (*all.x = TRUE*) para garantir que todas as entradas de atividade sejam mantidas, mesmo que não haja dados correspondentes de sono ou peso para todas as datas/IDs. É importante notar que o dataset de peso possui uma amostra limitada, o que exige cautela na interpretação das análises que o envolvem.

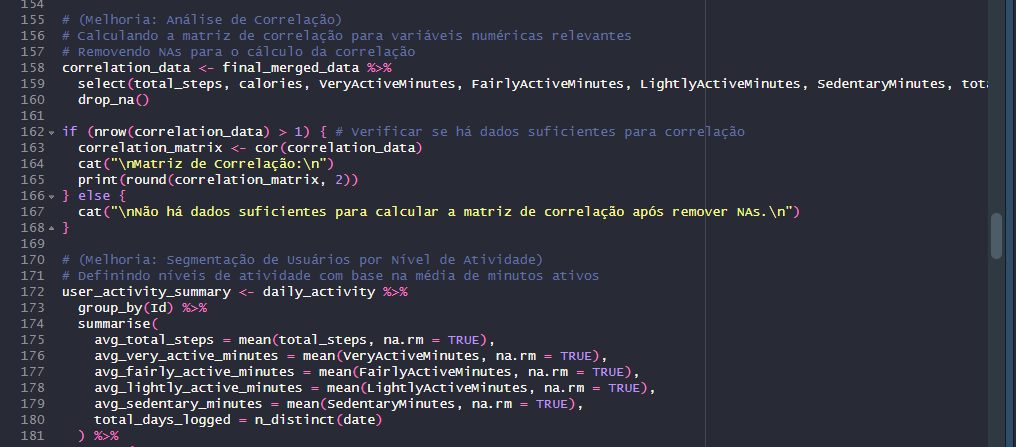


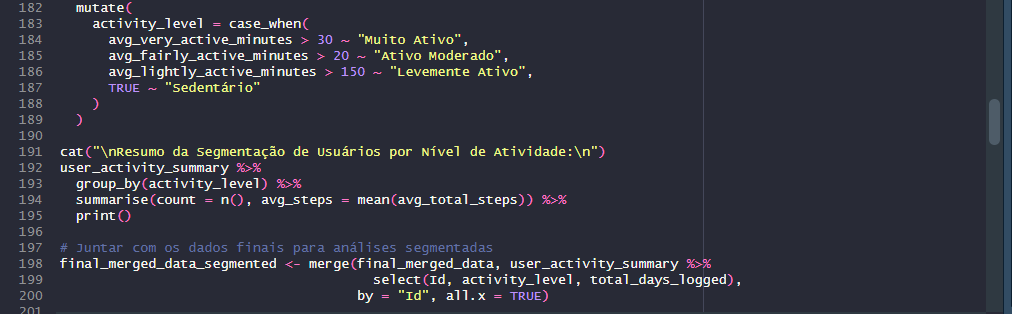


**2.4. Análise Descritiva e Segmentação de Usuários**

Esta etapa foca na obtenção de estatísticas descritivas para as variáveis-chave de atividade e sono, além de calcular uma matriz de correlação para identificar relações entre as variáveis. Uma segmentação de usuários baseada em seus níveis de atividade média também é realizada para permitir análises mais direcionadas.

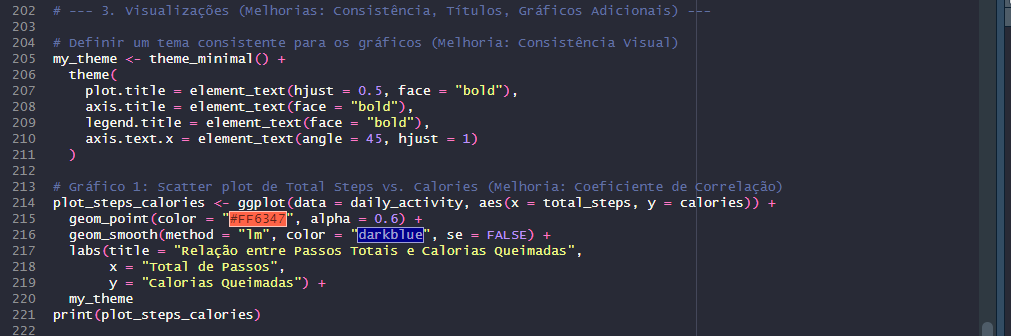


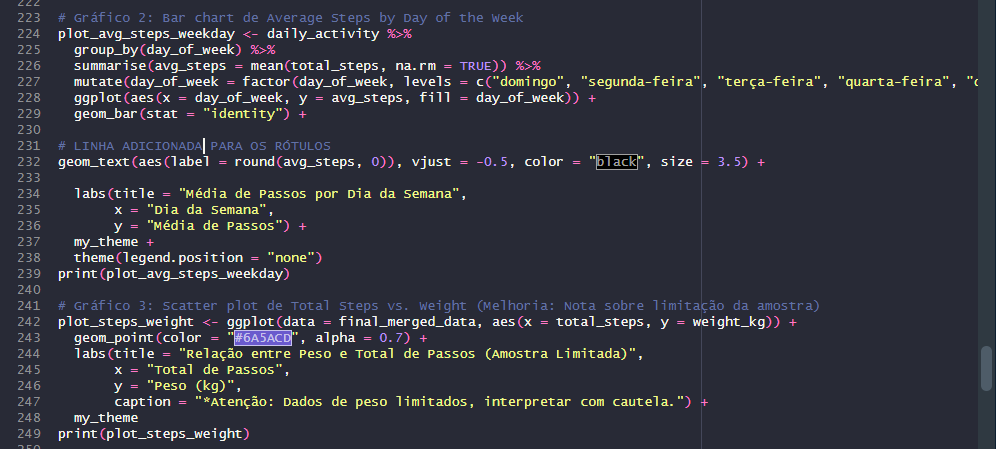


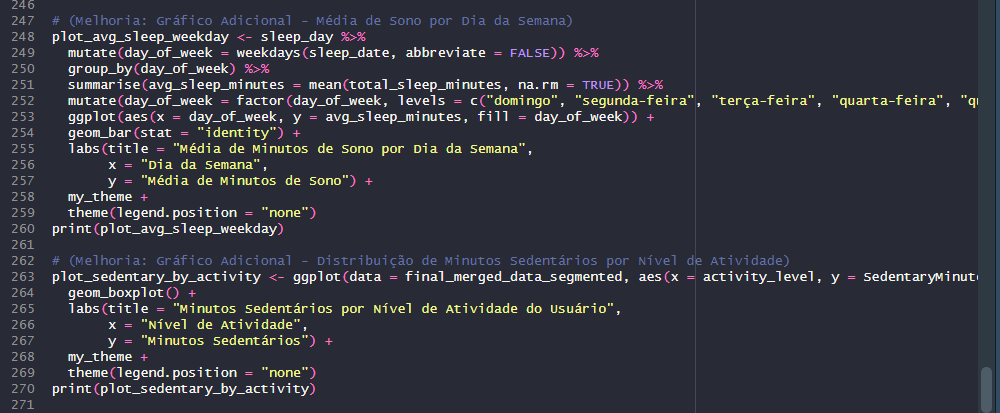


**2.5. Visualizações**

Esta seção gera os gráficos que ilustram as principais tendências e relações nos dados. Um tema visual consistente é aplicado para garantir a clareza e profissionalismo das visualizações.







**3. Tabela de Funções Comuns (Glossário)**

Esta tabela resume as principais funções e operadores utilizados no código R, fornecendo uma referência rápida para sua compreensão.

|  |  |
| --- | --- |
| Atribuição/Função | Descrição |
| library() | Carrega pacotes (bibliotecas) do R para uso no script. |
| setwd() | Define o diretório de trabalho atual do R. |
| dirname() | Retorna o caminho do diretório de um arquivo. |
| rstudioapi::getActiveDocumentContext()$path | Retorna o caminho completo do arquivo R atualmente aberto no RStudio. |
| read\_csv() | Lê arquivos CSV, parte do pacote readr (do tidyverse). |
| %>% (pipe) | Operador que "passa" o resultado da expressão à esquerda como o primeiro argumento da função à direita, tornando o código mais legível. |
| rename() | Renomeia colunas de um data frame. |
| mutate() | Adiciona novas colunas ou modifica colunas existentes em um data frame. |
| mdy() / mdy\_hms() | Funções do lubridate para converter strings em formato de data (mês/dia/ano) ou data e hora (mês/dia/ano hora:minuto:segundo). |
| as.Date() | Converte um objeto para o formato de data. |
| weekdays() | Extrai o dia da semana de uma data. |
| as.character() | Converte um objeto para o formato de caractere (texto). |
| sum(duplicated()) | Conta o número de linhas duplicadas em um data frame. |
| sum(is.na()) | Conta o número de valores ausentes (NA) em um data frame. |
| merge() | Combina dois data frames com base em colunas comuns. |
| glimpse() | Fornece uma visão rápida da estrutura de um data frame. |
| n\_distinct() | Conta o número de valores únicos em uma coluna. |
| summary() | Gera um resumo estatístico para as colunas de um data frame. |
| mean() | Calcula a média de um vetor numérico. |
| na.rm = TRUE | Argumento comum em funções estatísticas para remover valores NA antes do cálculo. |
| cat() | Imprime texto no console. |
| cor() | Calcula a matriz de correlação entre variáveis numéricas. |
| drop\_na() | Remove linhas que contêm valores NA. |
| group\_by() | Agrupa um data frame por uma ou mais variáveis para realizar operações em grupo. |
| summarise() | Cria um novo data frame com resumos de dados agrupados. |
| case\_when() | Função condicional útil para criar novas variáveis com base em múltiplas condições. |
| ggplot() | Inicia a criação de um gráfico no ggplot2. |
| aes() | Mapeia variáveis do data frame para atributos estéticos do gráfico (e.g., x, y, cor, tamanho). |
| geom\_point() | Adiciona uma camada de pontos (gráfico de dispersão). |
| geom\_smooth() | Adiciona uma linha de suavização (e.g., regressão linear) a um gráfico. |
| geom\_bar() | Adiciona uma camada de barras (gráfico de barras). |
| geom\_boxplot() | Adiciona uma camada de box plots. |
| labs() | Adiciona ou modifica rótulos e títulos do gráfico. |
| theme() | Personaliza elementos não-dados do gráfico (e.g., títulos, eixos, legendas). |
| print() | Exibe um objeto (como um gráfico) no console ou painel de plots. |

**4. Resultados e Discussão: Insights para a Bellabeat**

Com base na análise dos dados, foram identificadas tendências significativas no uso de dispositivos inteligentes que podem ser diretamente aplicadas à estratégia de marketing da **Bellabeat**, com foco no produto **Leaf.**

**4.1. Principais Tendências Identificadas**

1.Correlação Positiva entre Passos e Calorias Queimadas: Uma relação clara onde o aumento da atividade (passos) leva a um maior gasto calórico. Esta é uma tendência fundamental para motivar usuários a se manterem ativos.

Gráfico, Gráfico de dispersão

O conteúdo gerado por IA pode estar incorreto.

Gráfico 1- Relação entre passos totais e calorias queimadas

2.Padrões de Atividade Variam ao Longo da Semana: A atividade dos usuários não é constante, apresentando flutuações com dias de maior e menor engajamento. Isso sugere oportunidades para intervenções direcionadas.

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

Gráfico 2 - Média de Passos por Dia da Semana

3.Importância do Sono para o Bem-Estar Geral: Os dados de sono revelam padrões diversos, e a qualidade do sono é um fator crítico que pode influenciar os níveis de energia e atividade do dia seguinte.

Gráfico, Gráfico de barras

O conteúdo gerado por IA pode estar incorreto.

Gráfico 3- Média de Minutos de Sono por Dia da Semana

4.Diferentes Perfis de Usuários: A segmentação dos usuários por nível de atividade (Sedentários, Levemente Ativos, Ativos Moderados, Muito Ativos) demonstra que existem comportamentos e necessidades distintas, exigindo abordagens personalizadas.

Gráfico, Gráfico de caixa estreita

O conteúdo gerado por IA pode estar incorreto.

Gráfico 4 - Minutos Sedentários por Nível de Atividade do Usuário

Obs.: Este gráfico ilustra a distribuição dos minutos sedentários diários para cada perfil de usuário. A análise revela que todos os grupos, incluindo os "Muito Ativos", passam uma quantidade significativa de tempo parados (mediana acima de 800 minutos). O grupo "Sedentário" apresenta não só a maior mediana, mas também a maior variabilidade no comportamento. Isso indica que o hábito sedentário é um desafio presente em todos os segmentos de usuários.

**4.2. Aplicação dos Insights ao Bellabeat Leaf**

O Bellabeat Leaf, como um monitor de atividade, sono e estresse, está perfeitamente posicionado para capitalizar sobre essas tendências:

•Motivação por Resultados: Usuárias do Leaf podem ser motivadas ao verem a quantificação direta da queima calórica associada aos seus passos, incentivando a manutenção ou aumento da atividade.

•Incentivo em Dias de Baixa Atividade: O Leaf pode enviar lembretes e sugestões de atividades leves em dias onde o engajamento tende a ser menor, ajudando as usuárias a manterem a consistência.

•Otimização do Sono: Ao monitorar o sono, o Leaf permite que as usuárias compreendam seus padrões e recebam dicas personalizadas para melhorar a qualidade do descanso, impactando positivamente sua energia diária.

•Experiência Personalizada: A segmentação de usuários permite que o aplicativo Bellabeat e o Leaf ofereçam recomendações e metas adaptadas ao perfil de atividade de cada usuária, tornando a experiência mais relevante e eficaz.

**4.3. Recomendações para a Estratégia de Marketing da Bellabeat**

1.Campanhas Focadas em Benefícios Tangíveis: Destacar como o Leaf quantifica o esforço (passos, calorias) e os benefícios diretos para a saúde e bem-estar. Criar desafios de "passos diários" com recompensas virtuais no aplicativo.

2.Marketing Contextualizado por Dia da Semana: Desenvolver campanhas específicas para fins de semana, promovendo atividades leves e relaxantes. Enviar lembretes personalizados para incentivar a atividade em dias de menor engajamento.

3.Ênfase na Qualidade do Sono: Campanhas que eduquem sobre a importância do sono e como o Leaf auxilia no monitoramento e melhoria dos hábitos de sono, oferecendo dicas de higiene do sono e meditações guiadas.

4.Segmentação de Mensagens: Adaptar a comunicação de marketing aos diferentes perfis de usuários. Para usuárias mais sedentárias, focar nos benefícios da atividade gradual; para as mais ativas, em otimização de desempenho e recuperação.

**5. Conclusão**

Este projeto demonstra o poder da análise de dados de dispositivos inteligentes para gerar insights acionáveis. As tendências identificadas fornecem uma base sólida para a **Bellabeat** refinar sua estratégia de marketing, personalizar a experiência do usuário do **Leaf** e, em última instância, fortalecer sua posição no mercado de tecnologia de bem-estar. A implementação dessas recomendações pode levar a um maior engajamento do usuário e a um crescimento sustentável para a empresa.

**Lucas Gonzaga Pereira | lucasgpec@gmail.com**

|  |  |  |
| --- | --- | --- |
| [Linkedin](https://www.linkedin.com/in/lucasgpec/) | [GitHub](https://github.com/lucasgpec) | [Kaggle](https://www.kaggle.com/lucasgonzagapereira) |

[Ícone

O conteúdo gerado por IA pode estar incorreto.](https://www.kaggle.com/lucasgonzagapereira)[Forma

O conteúdo gerado por IA pode estar incorreto.](https://github.com/lucasgpec)[Ícone

O conteúdo gerado por IA pode estar incorreto.](https://www.linkedin.com/in/lucasgpec/)