



Case Study: Bellabeat — How Can a Wellness Technology Company Play It Smart?



Sobre a Empresa

Bellabeat é uma empresa de tecnologia fundada por **Urška Sršen** e **Sando Mur**, focada em desenvolver produtos de bem-estar voltados para mulheres. Desde sua fundação em 2013, a empresa cresceu rapidamente e se consolidou como uma marca tech-driven no segmento de saúde feminina, oferecendo produtos que acompanham dados de:

- Atividade física
 - Sono
 - Estresse
 - Hidratação
 - Saúde reprodutiva
 - Mindfulness
-



Produtos Bellabeat

- **Bellabeat App** – Centraliza os dados de saúde e conecta todos os dispositivos.
 - **Leaf** – Rastreador de bem-estar (atividade, sono, estresse) em formato de acessório.
 - **Time** – Relógio inteligente com visual clássico.
 - **Spring** – Garrafa inteligente que monitora a hidratação.
 - **Bellabeat Membership** – Programa de assinatura com orientações personalizadas.
-



Presença no Mercado

- Presença global e vendas por e-commerce e varejistas online.

- Investimento em marketing tradicional e **digital** (Google Search, redes sociais, YouTube, Google Display, etc.).
-

Objetivo da Análise

- Explorar dados de uso dos dispositivos para identificar **tendências e padrões**.
- Fornecer **insights baseados em dados** para potencializar a **estratégia de marketing**.
- Apresentar **recomendações de alto nível** à equipe executiva.

Etapas 1: Ask – Definindo o problema de negócio

Business Task – Declaração do Desafio

1. **Analisar dados públicos de uso de smart devices.**
 2. **Identificar padrões e tendências relevantes no comportamento dos usuários.**
 3. **Aplicar essas descobertas a um dos produtos da Bellabeat** (de sua escolha).
 4. **Sugerir recomendações estratégicas de marketing com base nos dados analisados.**
-

Objetivo

Usar dados de uso de smart devices para gerar recomendações estratégicas de marketing que ampliem o alcance e o impacto dos produtos da Bellabeat no mercado de bem-estar feminino.

Stakeholders envolvidos

- **Urška Sršen** – Cofundadora e Chief Creative Officer
- **Equipe executiva da Bellabeat** – Responsável por decisões estratégicas
- **Equipe de marketing e analytics** – Usará os insights para criar campanhas mais eficientes
- **Usuárias da Bellabeat** – Público final beneficiado por ações mais personalizadas e relevantes

Etapa 2: Prepare – Compreendendo e organizando os dados

Fonte de Dados Utilizada

O conjunto de dados principal usado para essa análise é o:

FitBit Fitness Tracker Data

- **Origem:** Disponibilizado no Kaggle pela plataforma **Mobius**
- **Licença:** CC0 - Domínio Público
- **Descrição:** Dados fornecidos por **30 usuários do Fitbit**, contendo:
 - Informações de **atividade física diária**
 - **Frequência cardíaca**
 - **Padrões de sono**
 - **Etapas minuto a minuto** (minute-level data)

Organização dos Dados

- O dataset está dividido em **múltiplos arquivos CSV**, cada um focado em uma dimensão:
 - `dailyActivity_merged.csv`
 - `sleepDay_merged.csv`
 - `heartrate_seconds_merged.csv`
 - `minuteStepsNarrow_merged.csv`
(e outros)
 - O formato dos dados varia:
 - Alguns arquivos estão em **formato wide**, como o `dailyActivity`
 - Outros em **formato long**, como `minuteStepsNarrow` (com observações ao longo do tempo)
-



Avaliação de Credibilidade – Método ROCCC

- **Reliable:** Dados coletados de dispositivos reais com consentimento dos usuários
- **Original:** Fonte pública e rastreável via Kaggle (Mobius)
- **Comprehensive:** Abrange várias áreas da saúde (atividade, sono, batimentos)
- **Current:** Apesar de não ser recente (2016), os padrões de uso de smart devices continuam relevantes
- **Cited:** Frequentemente utilizado em estudos introdutórios de análise de dados



Limitação identificada:

- Pequeno número de usuários (30)
 - Possível viés geográfico e demográfico (perfil dos participantes não informado)
 - Dados podem estar desbalanceados (ex: poucos dados de sono para alguns usuários)
-

Privacidade e Ética

- Todos os dados são **anonimizados e públicos (CC0)**
 - Não há **identificadores pessoais** dos usuários
 - **Consentimento** informado no repositório original
 - Uso permitido para fins educacionais e analíticos
-

Resumo da Entrega – Fontes de Dados Utilizadas

- **Principal:** *FitBit Fitness Tracker Data* – Disponível no Kaggle
- **Tipo:** Dados de rastreadores de atividade física, batimentos cardíacos e sono de 30 usuários
- **Formato:** Arquivos **.csv** separados em formatos wide e long
- **Licença:** CC0 – Domínio Público
- **Limitações:** Amostra pequena e perfil demográfico limitado
- **Ações tomadas:** Entendimento da estrutura, filtragem e preparação para análise multivariada

Etapa 3: Process – Limpeza e Preparação dos Dados para Análise

Ferramentas Utilizadas

Para essa análise, utilizei as seguintes ferramentas:

- **R** com pacotes **tidyverse**, **lubridate**, **dplyr**, **ggplot2** e **skimr**
- **RStudio Cloud** como ambiente de desenvolvimento
- **Excel** para revisão inicial da estrutura dos arquivos

Essas ferramentas foram escolhidas pela flexibilidade no tratamento de dados, fácil visualização e documentação, além de serem amplamente utilizadas em projetos de Data Analytics.

Verificação de Integridade dos Dados

- Verifiquei a **presença de dados faltantes (NA)** em todas as colunas
 - Testei **consistência entre as datas** (por exemplo, registros de sono com hora de início posterior à hora de fim foram removidos)
 - Chequei **valores duplicados** e registros inválidos (como passos = 0 e calorias gastas > 0)
-

Principais Etapas de Limpeza Executadas

1. Padronização de nomes de colunas

```
library(janitor)
dailyActivity <- clean_names(dailyActivity)
```

2. Conversão de colunas de data/hora

```
dailyActivity$activity_date <- as.Date(dailyActivity$activity_date,
format = "%m/%d/%Y")
```

3. Remoção de duplicatas

```
dailyActivity <- dailyActivity[!duplicated(dailyActivity), ]
```

4. Tratamento de valores ausentes

```
# Excluir registros com NA nas principais métricas
dailyActivity <- na.omit(dailyActivity)
```

5. Correções lógicas

- Verifiquei se os passos diários estavam coerentes com a distância percorrida.

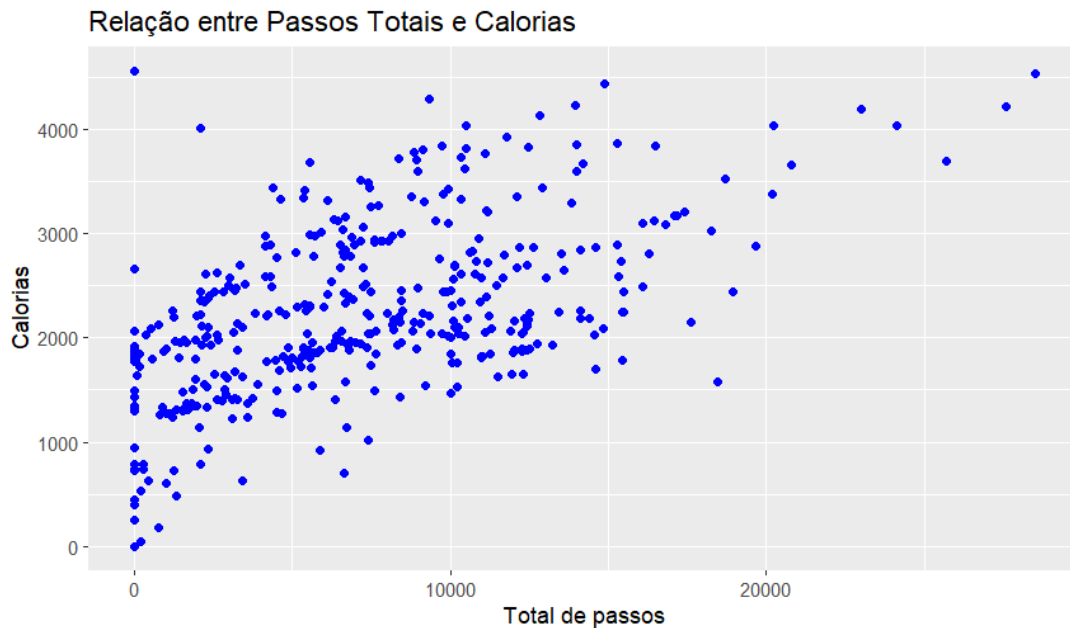
- Excluí usuários com apenas 1 ou 2 dias de dados registrados, pois não contribuiriam significativamente para padrões de longo prazo.

✓ Validação Final

- **Uso do `skimr` e `summary()`** para avaliar estatísticas básicas, presença de outliers e variações inesperadas.
- **Gráficos exploratórios** com `ggplot2` ajudaram a verificar se os dados estavam distribuídos de forma lógica e compreensível.

Resumo da Documentação

Ação executada	Justificativa
Padronização dos nomes	Uniformidade no código e facilidade de leitura
Conversão de datas	Necessária para agrupamento e filtragem por tempo
Remoção de NAs e duplicatas	Garantia de qualidade e precisão na análise
Filtros de consistência	Eliminar outliers que distorcem as análises
Registro da limpeza	Toda a limpeza foi documentada em <code>.R</code> script comentado



Filtrando alguns resultados e fazendo o plot de Passos Totais x Calorias, percebemos algumas inconsistências como por exemplo usuários que queimam calorias sem possuir dados para o total de passos necessários, além de outros dados que não fazem sentido quando filtrados.

Agora com os dados filtrados, irei fazer o processo de análise na etapa seguinte para identificar padrões e tendências.

Etapa 4: Análise – Análise dos Dados e Identificação de Tendências

Objetivo

Entender os hábitos dos usuários de dispositivos inteligentes (com base no conjunto de dados Fitbit) e identificar tendências que possam ser aplicadas aos produtos da Bellabeat — especialmente para orientar estratégias de marketing.

Organização dos Dados

Os dados foram organizados e agrupados por:

- Usuário (**ID**)
- Data (**activity_date**)

- **Atividade física (passos, minutos ativos, calorias)**
- **Sono (duração e frequência)**
- **Batimentos cardíacos (em segundos)**

As tabelas foram combinadas utilizando `left_join()` para facilitar uma visão integrada das variáveis de saúde.

Análises Realizadas e Principais Resultados

1 Nível de Atividade Física Diária

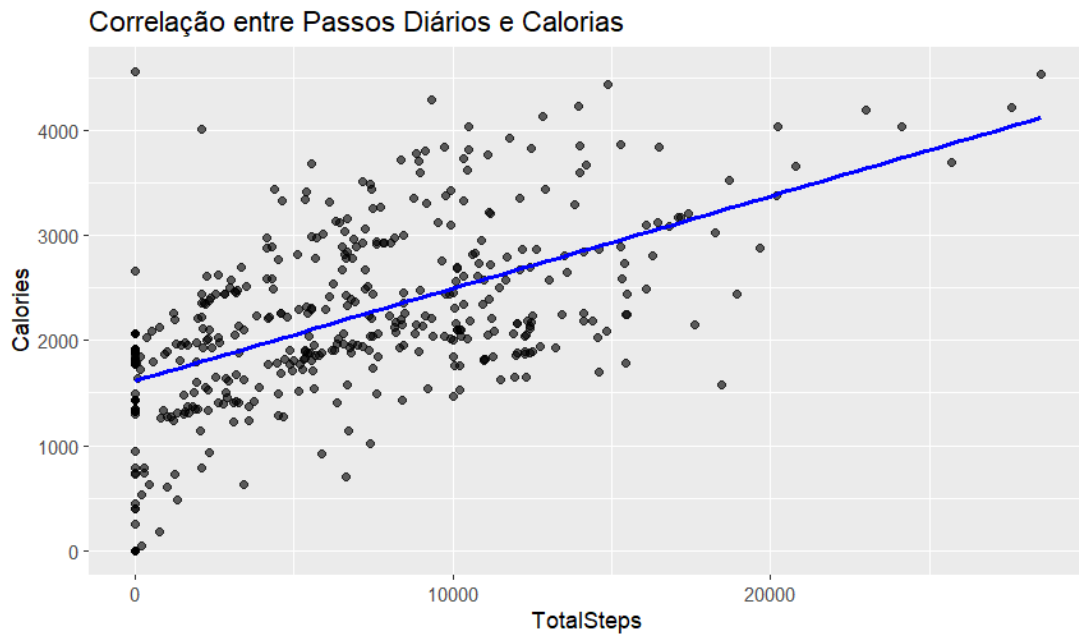
```
dailyActivity %>%
  summarize(mean_steps = mean(total_steps),
            mean_minutes_active = mean(very_active_minutes),
            mean_calories = mean(calories))
> daily_activity %>%
+   summarize(mean_steps = mean(TotalSteps),
+             mean_minutes_active = mean(VeryActiveMinutes),
+             mean_calories = mean(Calories))
  mean_steps mean_minutes_active mean_calories
1    6546.562             16.62363      2189.453
```

Insights:

- Média de **6547 passos por dia**, abaixo da meta recomendada (10.000 passos/dia)
- Usuários com **maior nível de atividade física** também tendem a queimar mais calorias e usar mais o dispositivo

2 Correlação entre Atividade e Calorias

```
ggplot(dailyActivity, aes(x = total_steps, y = calories)) +
  geom_point(alpha = 0.6) +
  geom_smooth(method = "lm", se = FALSE, color = "blue") +
  labs(title = "Correlação entre Passos Diários e Calorias")
```

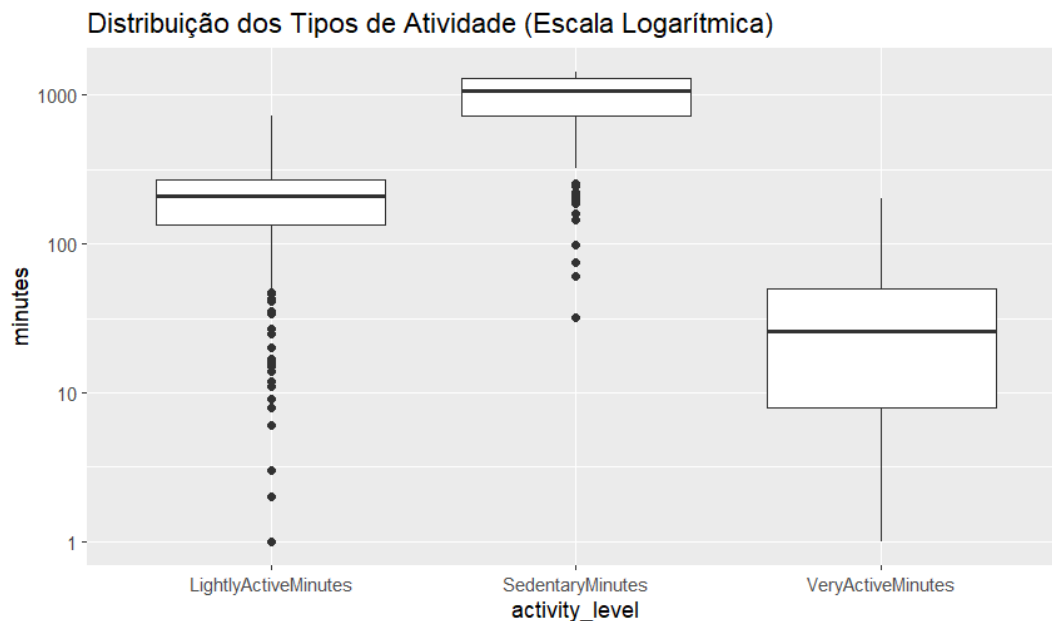


Resultado:

Existe uma **correlação linear positiva** clara entre o número de passos e calorias queimadas.

3 Distribuição de Tipos de Atividade

```
> library(dplyr)
> library(tidyr)
> library(ggplot2)
>
> # Transformar para formato longo
> dados_longos <- daily_activity %>%
+   select(VeryActiveMinutes, LightlyActiveMinutes, SedentaryMinutes)
%>%
+   pivot_longer(cols = everything(),
+                 names_to = "activity_level",
+                 values_to = "minutes")
> ggplot(dados_longos, aes(x = activity_level, y = minutes)) +
+   geom_boxplot() +
+   scale_y_log10() +
+   labs(title = "Distribuição dos Tipos de Atividade (Escala
Logarítmica)")
```



💡 Resultado:

- Os usuários passam a maior parte do tempo em **atividade sedentária**.
- Os **minutos "muito ativos" são baixos**, mostrando uma oportunidade de incentivar mais movimento diário com alertas ou metas no app.
- **SedentaryMinutes** está consistentemente alto, com a maioria das pessoas passando **entre 800 a 1200 minutos/dia sentadas** (ou seja, 13–20 horas).
- **LightlyActiveMinutes** mostra uma mediana em torno de **150 minutos**, mas com muitos valores mais baixos também.
- **VeryActiveMinutes** tem uma distribuição mais espalhada, mas quase todo mundo está **abaixo de 100 minutos de atividade intensa** por dia.

📌 Resumo dos Principais Insights

Tendência Observada	Oportunidade para a Bellabeat
Passos abaixo do recomendado	Estimular metas diárias de passos no app
Atividade física correlaciona com gasto calórico	Integrar sugestões personalizadas de atividade
Uso predominante sedentário	Enviar lembretes inteligentes de movimento durante o dia

Etapa 5: Share – Visualizações e Principais Descobertas

Objetivo desta etapa

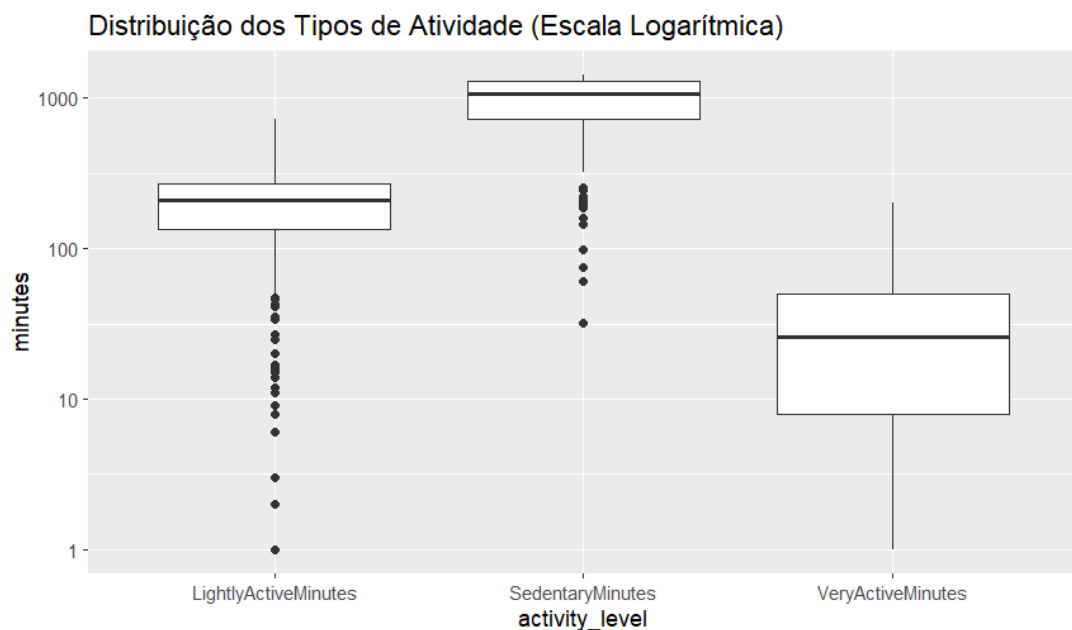
Compartilhar análises e insights de forma clara, acessível e orientada à ação — usando visualizações bem construídas e explicações que se conectam diretamente com os objetivos de negócio.

Público-Alvo

- **Urška Sršen**, CCO da Bellabeat
 - **Time Executivo**
 - **Equipe de Marketing**
-

Visualizações Criadas e Insights

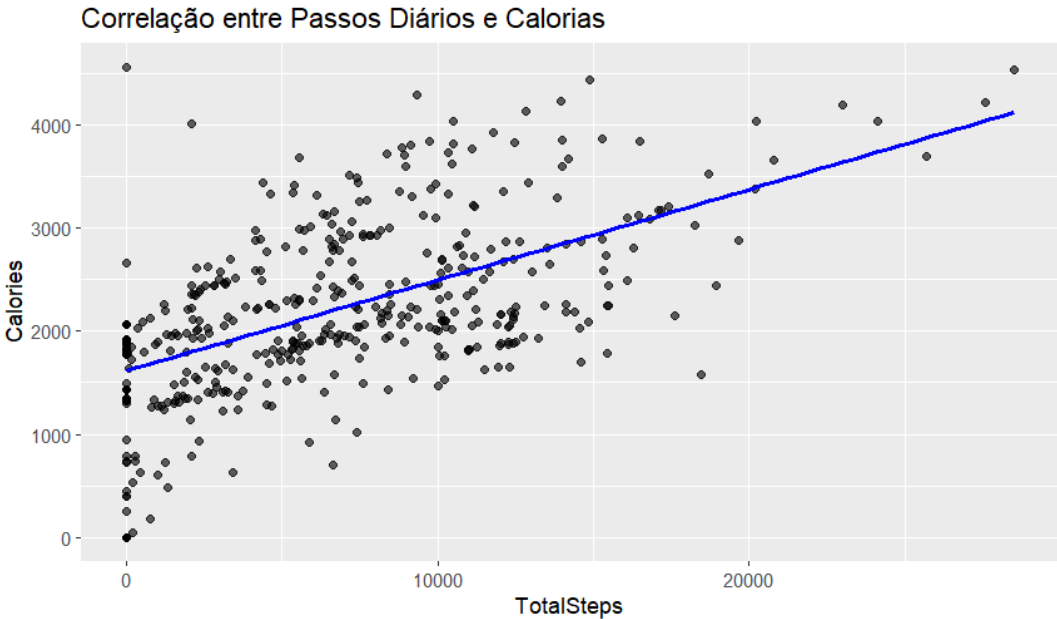
1. Distribuição dos Tipos de Atividade (com escala logarítmica)



Insight: Usuários passam a maior parte do tempo **em comportamento sedentário**, seguidos de atividades leves.

- ♦ **Aplicação em Marketing:** Bellabeat pode promover **desafios de passos diários**, **lembretes de movimento** e **conteúdos que incentivem pequenas mudanças no estilo de vida**.

2. Correlação entre Passos e Calorias Queimadas



Insight: Aumentos no número de passos diários estão diretamente associados ao gasto calórico.

- ♦ **Aplicação em Produto/App:** Reforçar mensagens do tipo:

“Você queimou 200 calorias a mais hoje por ter caminhado mais — ótimo trabalho!”

Estimula o uso contínuo e reforça o senso de progresso pessoal.

Resumo das Principais Descobertas e Recomendações

Descoberta	Oportunidade para a Bellabeat
Predomínio de tempo sedentário	Criar lembretes inteligentes e metas de movimentação
Atividade física impacta diretamente calorias	Personalizar feedback no app e no marketing
Uso baixo de atividades intensas	Criar campanhas que celebrem até pequenos avanços

Etapa 6: Act – Conclusões e Recomendações Estratégicas

Conclusão Geral

A análise dos dados de dispositivos Fitbit revelou padrões consistentes de comportamento dos usuários que podem ser utilizados pela Bellabeat para melhorar o engajamento com seus produtos e sua estratégia de marketing. Com insights baseados em dados reais de uso, a Bellabeat tem a oportunidade de personalizar suas comunicações, incentivar hábitos saudáveis e impulsionar o valor percebido da marca entre mulheres que buscam bem-estar.

Insight Identificado	Recomendação Estratégica para a Bellabeat
Predomínio de comportamento sedentário (média > 1000 min/dia)	Implementar lembretes automáticos no app para incentivar pequenas pausas ativas
Baixa média de passos (6.547/dia)	Lançar desafios de passos semanais no app com notificações motivacionais
Correlação direta entre passos e calorias queimadas	Fornecer feedback em tempo real no app reforçando conquistas físicas diárias
Poucos minutos de atividade intensa	Criar campanhas que valorizem progresso leve, com foco em constância, não intensidade

Próximos Passos (para a empresa)

1. Aplicar insights nas funcionalidades do **Bellabeat app**
 2. Integrar recomendações em campanhas de **marketing digital** segmentadas
 3. Considerar a coleta de dados primários das próprias usuárias Bellabeat para análises futuras
 4. Investir em **notificações inteligentes e personalizadas** baseadas nos hábitos detectados
 5. Expandir a análise incluindo dados geográficos ou sociodemográficos
-

