

RELATÓRIO BELLABEAT

Arthur Braz Santos

Fevereiro 2023

Contents

Introdução.	2
Tarefa de negócios.	2
Ferramentas.	3
Preparar	4
Integridade.	4
Processar	5
Tem alguma tendência no uso do dispositivo?	6
Quais são os horários de maior intensidade de exercicios, existe uma rotina?	7
Como os usuários dormem?	12
Conclusão.	18

Introdução.

O presente relatório tem como objetivo destrinchar a análise os dados de uso de dispositivos inteligentes para obter informações sobre como os consumidores usam dispositivos inteligentes que não são da Bellabeat.

Tarefa de negócios.

Identificar o comportamento de cliente que utilizam dispositivos inteligentes que não são Bellabeat, e aplicar esse conhecimento para melhorar o aplicativo da Bellbeat.

Aplicativo Bellabeat: O aplicativo Bellabeat fornece aos usuários dados de saúde relacionados à sua atividade, sono, estresse, ciclo menstrual e hábitos de atenção plena. Esses dados podem ajudar os usuários a entender melhor seus hábitos atuais e tomar decisões saudáveis. O aplicativo Bellabeat se conecta à sua linha de produtos inteligentes de bem-estar.

As perguntas que guiarão a análise são:

- **Como os usuários dormem?**
- **Quais são os horários de maior intensidade de exercícios, existe uma rotina?**
- **Tem alguma tendência no uso do dispositivo?**

Ferramentas.

O projeto foi totalmente desenvolvido na linguagem python através do Google Colab.

As bibliotecas utilizadas no projeto:

- pandas
- datetime
- matplotlib.pyplot
- seaborn

Preparar.

O conjunto de dados foi obtido através do Kaggle e contém um rastreador de condicionamento físico pessoal de trinta usuários do Fitbit, uma empresa estadunidense de eletrônicos e fitness. Trinta usuários elegíveis do Fitbit consentiram com o envio de dados pessoais do rastreador, incluindo os resultados a cada minuto de atividade física, frequência cardíaca e monitoramento do sono. São abrangidas informações sobre atividades diárias, passos e frequência cardíaca que podem ser usadas para explorar os hábitos dos usuários.

Os data frames utilizados são:

- `dailyActivity_merged.csv`
- `sleepDay_merged.csv`
- `hourlyIntensities_merged.csv`

A única tabela no formato longo é a `dailyActivity_merged.csv`

Fonte de dados: <https://www.kaggle.com/datasets/arashnic/fitbit>

Meta dados:

<https://www.fitabase.com/media/1930/fitbasedatadictionary102320.pdf>

Integridade.

O processo de integridade dos dados consiste em verificar se os dados estão completos, duplicados ou faltando. Para verificação criei uma função que verifica a quantidade de usuários, a quantidade de dados duplicados e a quantidade de dados faltantes.

A tabela `sleep_day` possui 24 usuários, como ela é vital para análise e possui uma quantidade relativamente próxima do total não irei descartá-la. Os outros data frames estão em conformidade. Existem 3 valores duplicados na tabela `sleep_day`, estes valores serão excluídos.

Processar.

Para a expanssão dos valores da coluna que marcar o dia e a hora nos data frames `sleep_day` e `hour_intensities` foi utlizada uma função que a partir desta coluna foram criadas três nova colunas, uma que marca a data, outra o dia e por último a hora, com objetivo de detalhar melhor os dados.

E para o data frame `day_activity` foram criadas apenas duas coluna, data e dia_semana.

Foi criado um novo data frame com uma junção entre `sleep_day` e `day_activity` chamado de `merge_sleep_activity`.

A coluna que horiginalmente marcava a data e a hora foi excluída.

Tem alguma tendência no uso do dispositivo?

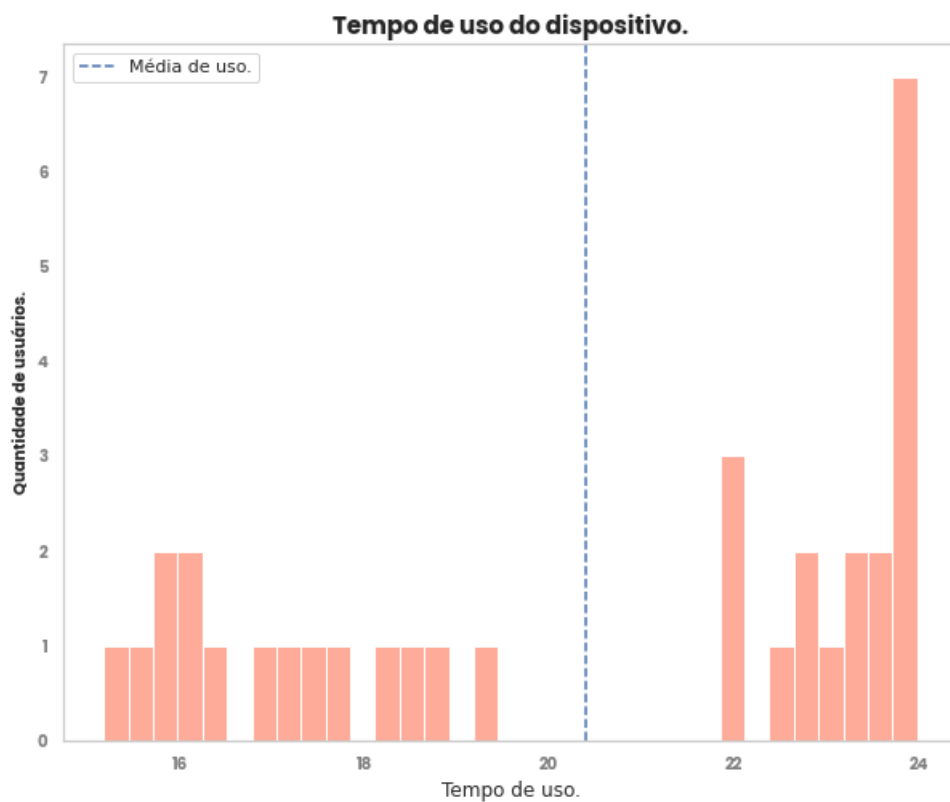


Figure 1: Histograma do tempo de uso de dispositivos inteligentes.

A média amostral do tempo de uso é de aproximadamente 20 horas e 15 usuários utilizam o dispositivo menos do que a média.

Quais são os horários de maior intensidade de exercicios, existe uma rotina?

A fitbit mensura a intensidade em uma escala de 0 a 3, sendo 0 o estado sedentário e 3 o estado muito ativo:

Valor	Estado
0	Sedentario
1	Leve
2	Moderado
3	Muito ativo

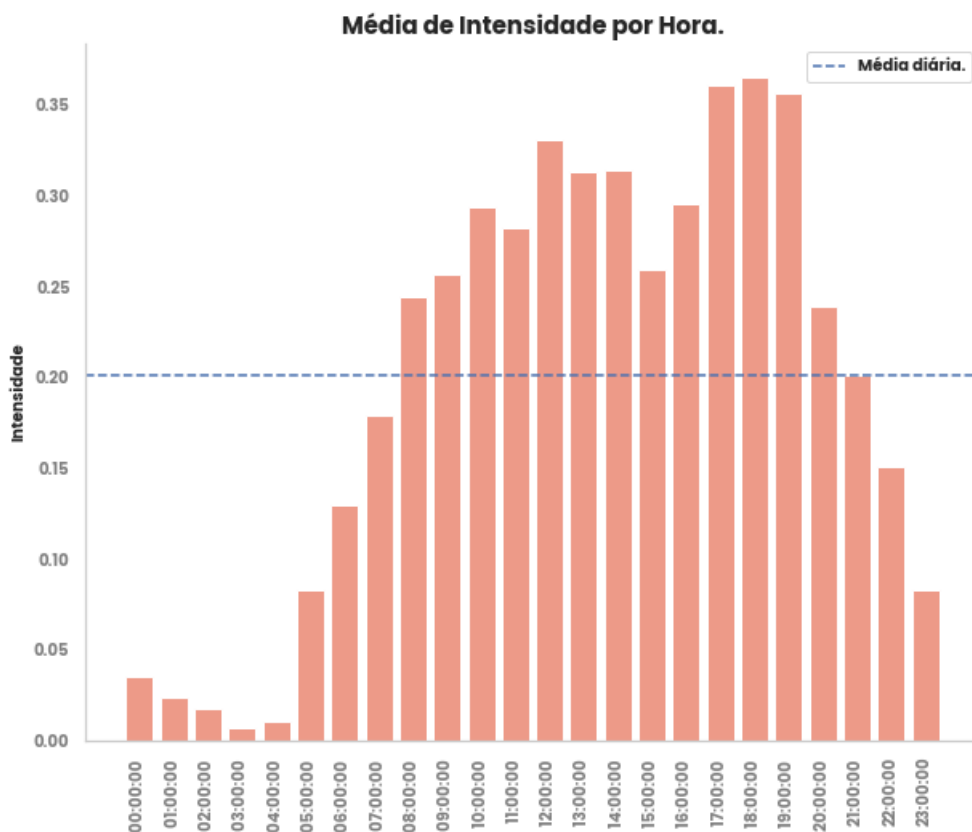


Figure 2: Média de intensidade por hora do dia.

Existe um pico de intensidade às 5 horas da manhã, possivelmente é nesse horário que os usuários estão acordando. Uma queda às 15 horas, possivelmente quando os usuários vão almoçar. E uma queda às 20 horas, possivelmente quando eles vão dormir.

O período ativo é das 5 às 19, possivelmente é o momento da prática de atividades físicas. Vale ressaltar, que esse período ativo, em média, não apresenta um nível de estado muito ativo, estando na escala como uma atividade entre sedentária e leve.

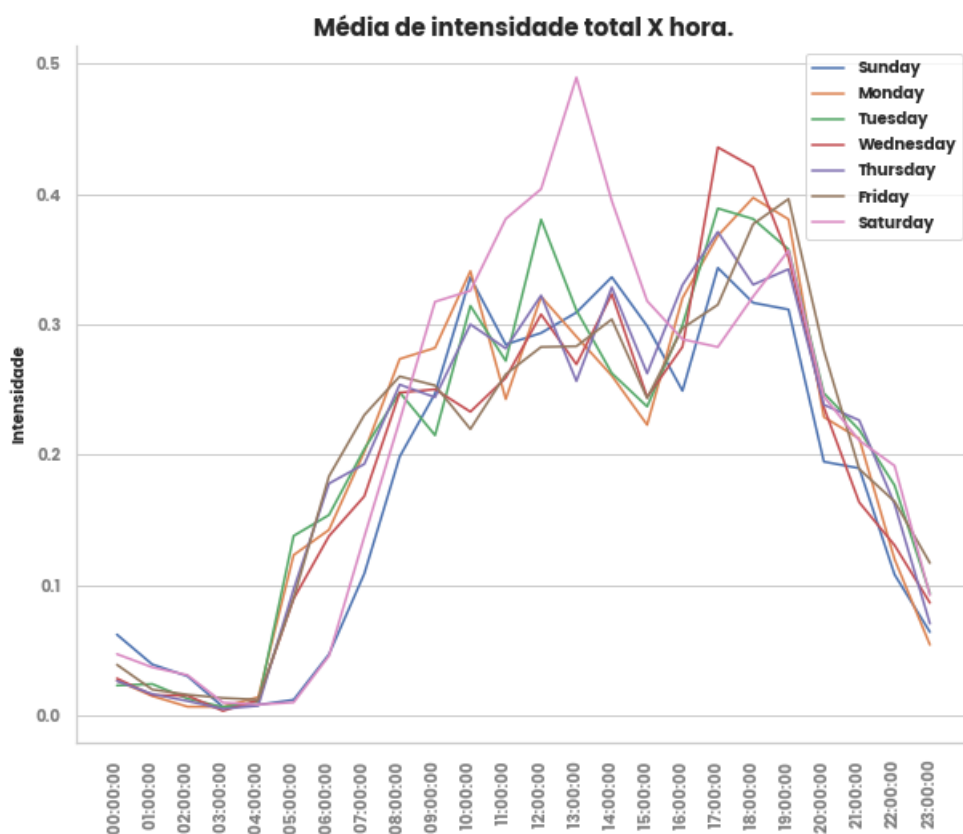


Figure 3: Intensidade média dos dias da semana por horas dos dia.

Existe um padrão para o comportamento da intensidade média, exceto para o sábado que apresenta um pico de intensidade às 13 horas enquanto que nos outros dias o pico ocorre entre às 17 e 19 horas.

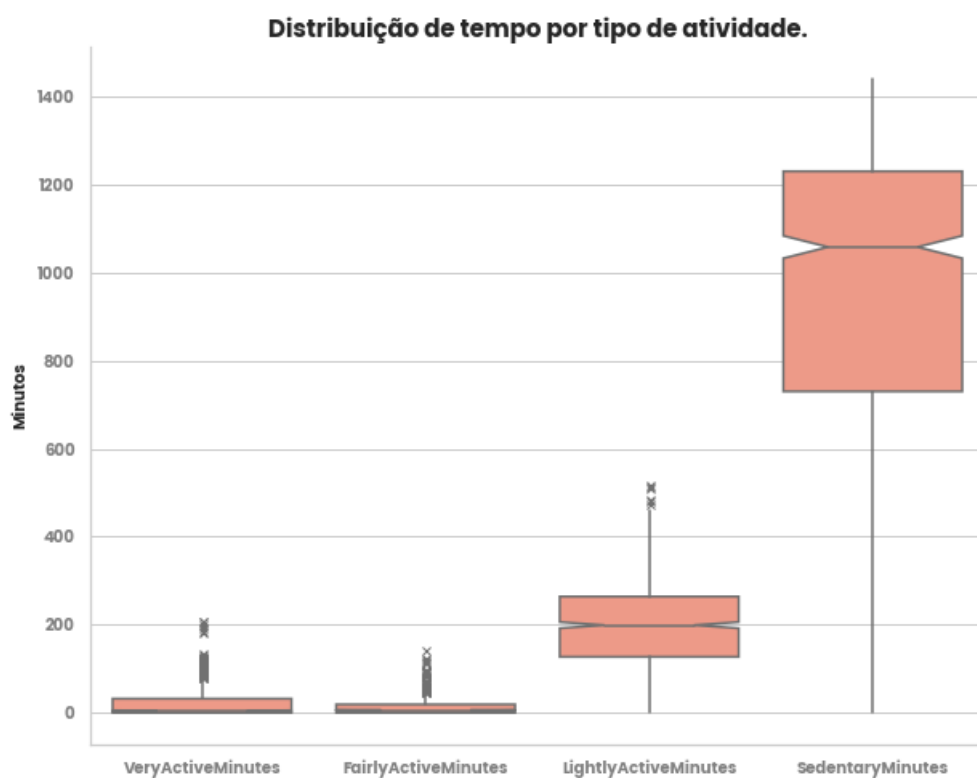


Figure 4: Distribuição do tipo nível de atividade dos usuários.

A maioria dos usuários passam a maior parte do seu tempo com atividades sedentárias.

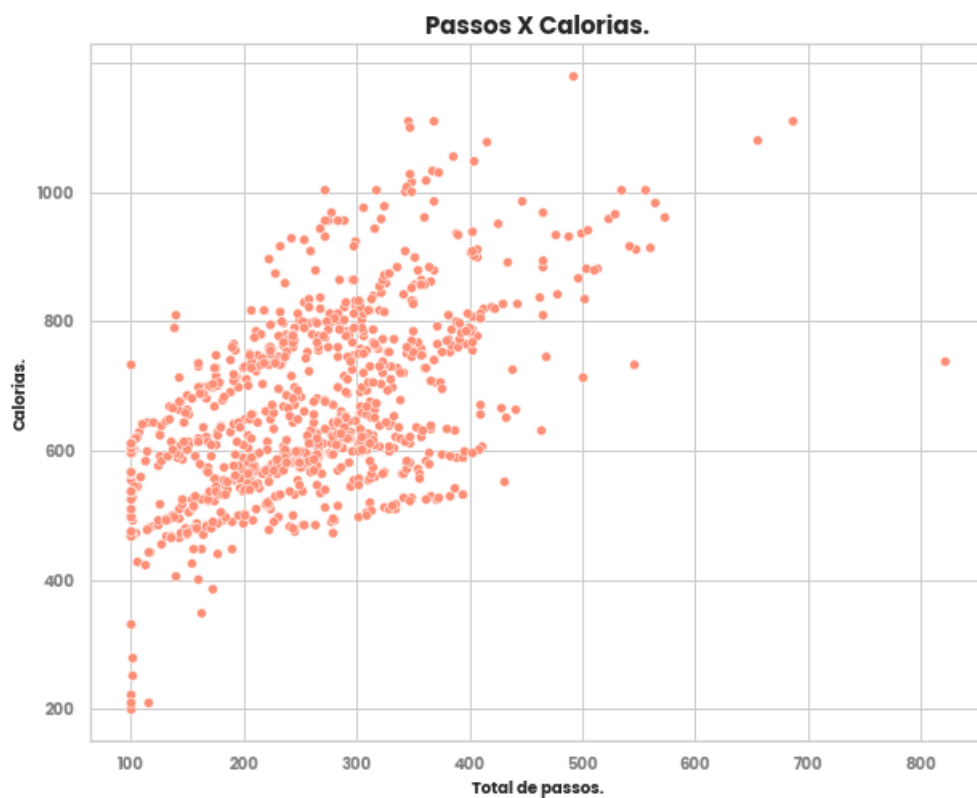


Figure 5: Gráfico de dispersão dos passos vs calorias.

Existe uma correlação moderada em relação ao tempo dormido e o total de passos por dia, o que significa que a quantidade de passos influencia parcialmente na quantidade de sono.

Como os usuários dormem?

Segundo a National Sleep Foundation, um instituto de pesquisa sem fins lucrativos dos Estados Unidos, uma pessoa entre 18 e 64 anos deve dormir entre 7 a 9 horas por dia.

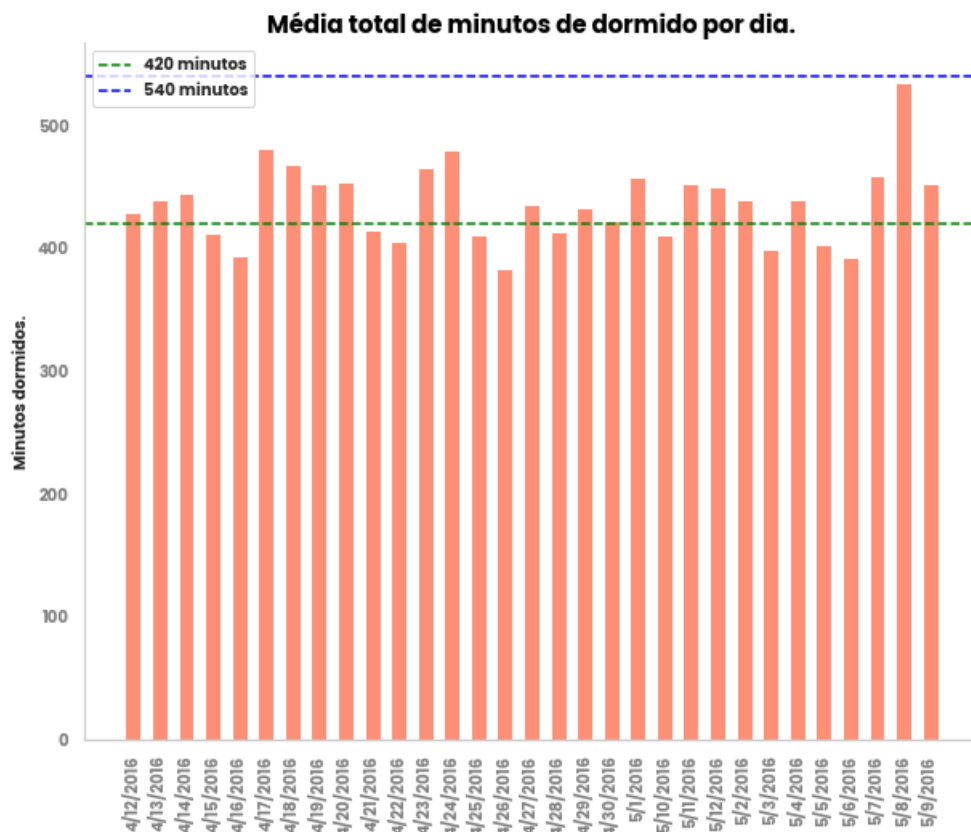


Figure 6: Gráfico de barras da média de tempo dormido em cada dia do período analisado.

Em média, pelo menos metade do mês os usuários não dormem o mínimo necessário.

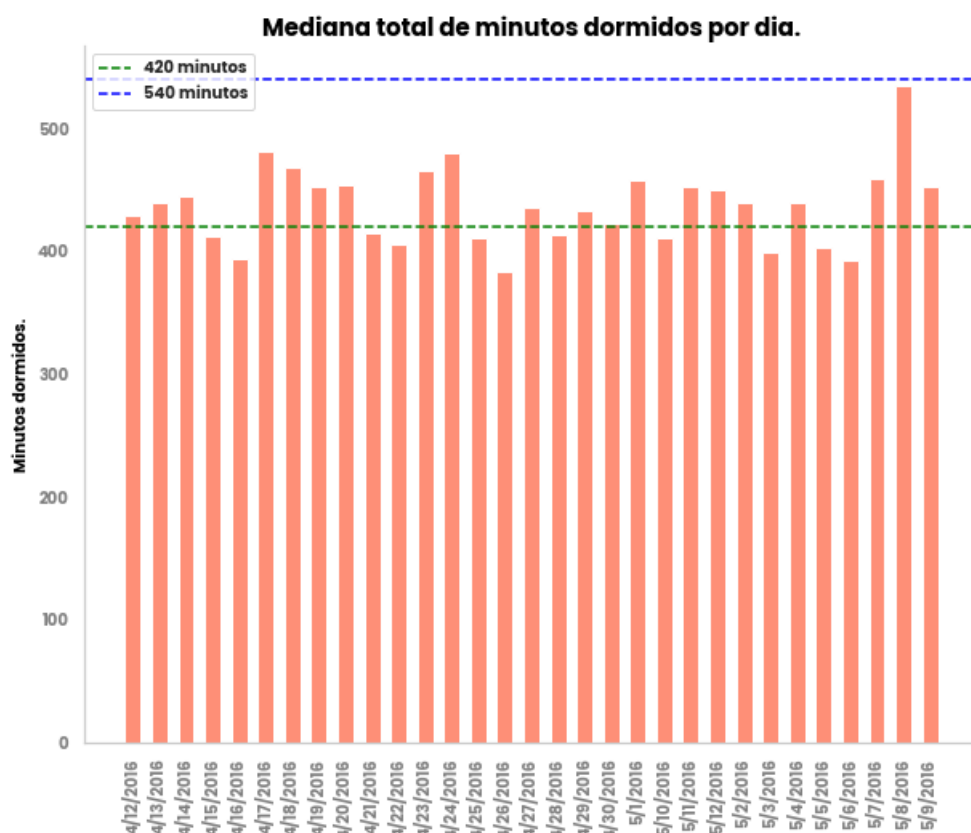


Figure 7: Gráfico de barras da mediana de tempo dormido em cada dia do período analisado.

Em mediana, 11 os usuários não dormem o mínimo necessário.

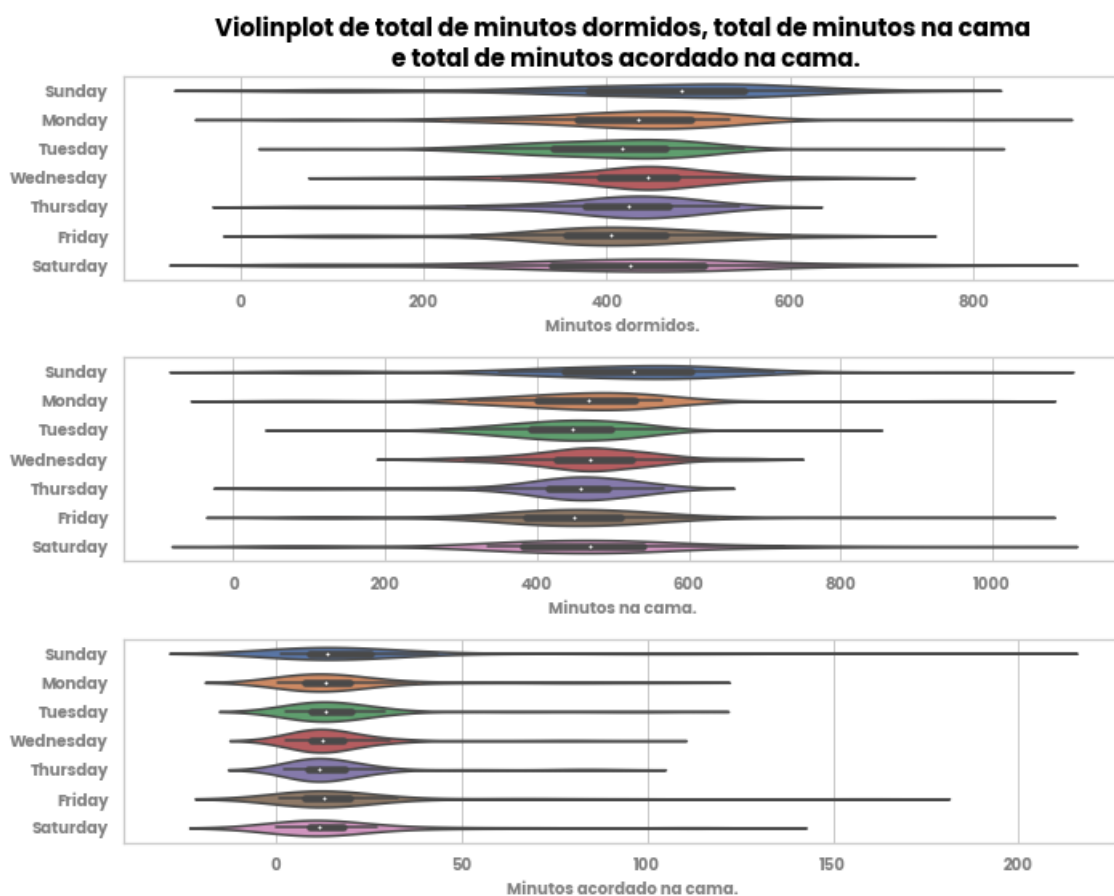


Figure 8: Gráficos das distribuições que os usuários passam dormindo em cada dia da semana, o tempo em que os usuários passam na cama e o tempo em que os usuários passam acordados na cama.

Domingo é dia em que os participantes mais dormem e sexta-feira é o dia que menos dormem.

Percebe-se uma redução drástica de minutos de sono de domingo para segunda, uma redução que se estende terça. No período de quarta a sexta-feira acontece um fenômeno parecido com o de domingo a terça-feira, na quarta-feira acontece um aumento relativamente pequeno de minutos dormidos em relação ao dia anterior, o período do primeiro ao terceiro quartil se posicionam de forma mais enxuta e num intervalo de minutos maior, porém existe uma redução de sono nos próximos dois dias quando no sábado existe um aumento repentino de minutos dormidos.

Isso pode significar uma quebra de rotina do sono, pois a cada três dias, o primeiro dia de sono tem uma qualidade maior do que os próximos dois dias, até que se repete o ciclo, exceto para o período de sábado para domingo.

Vale ressaltar que os participantes ficam em mediana 25,5 minutos por dia acordados na cama (foi utilizada a mediana, pois existem muito outliers nesta feature), sendo que, domingo é o dia em que os participantes mais ficam na cama acordados.

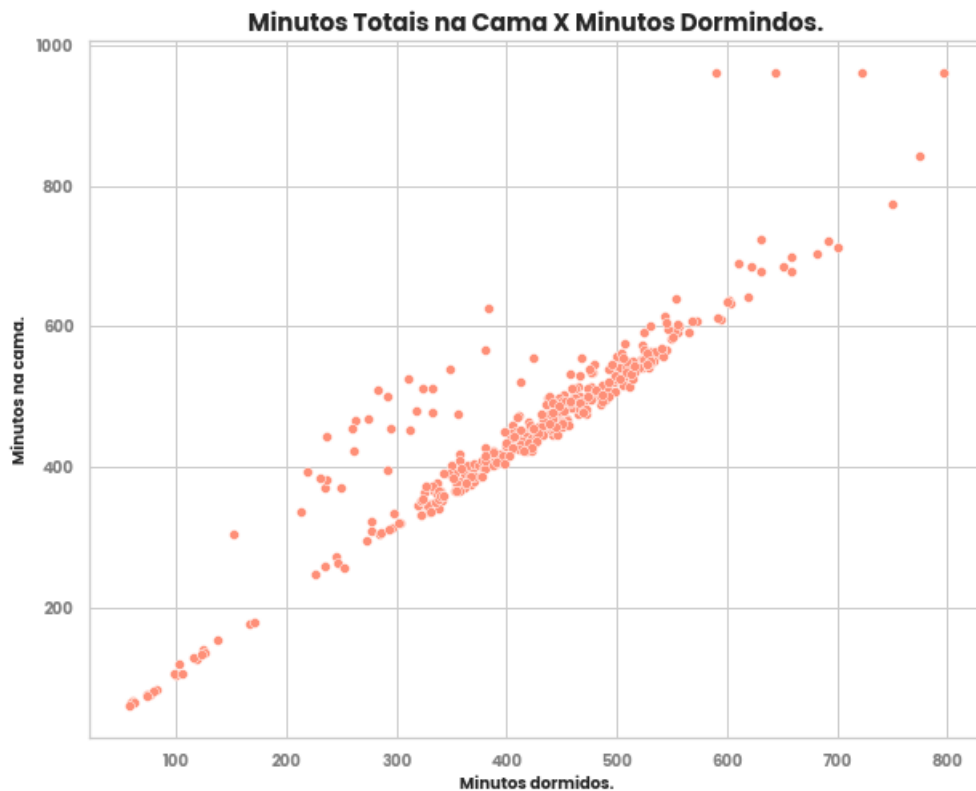


Figure 9: Gráfico de dispersão entre o tempo que os usuários passam na cama e o tempo que eles passam dormindo

Existe uma correlação de 0.93 entre o tempo em que os usuários passam na cama e o tempo que eles estão adormecidos, o que faz muito sentido. Porém existem alguns pontos no gráfico de dispersão que indicam usuários passando um bom tempo na acordados na cama , seria interessante saber as causas desse fenômeno.

- Celular?
- Streamig?
- Insônia?

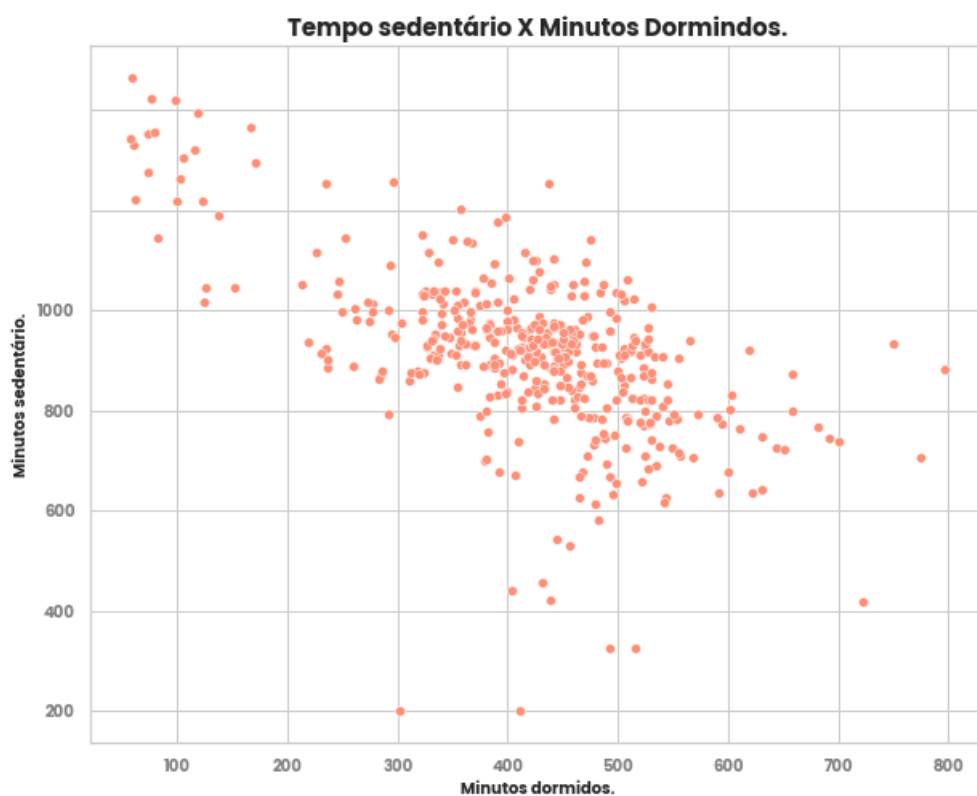


Figure 10: Gráfico de distribuição entre o tempo sedentário dos usuários e o tempo de sono.

Existe uma correlação negativa moderada de -0.6 entre o tempo sedentário no dia e quantidade de sono, ou seja o maior tempo sedentário no dia influencia parcialmente numa menor quantidade de tempo adormecido.

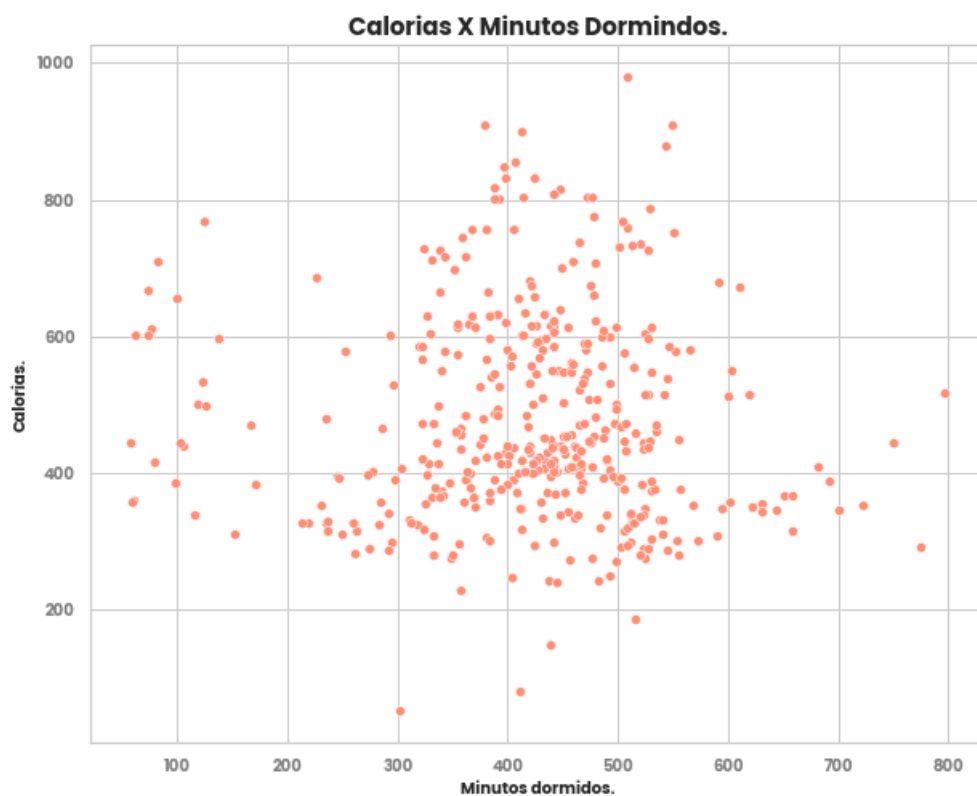


Figure 11: Gráfico de distribuição entre as calorias gastas dos usuários e a quantidade de sono.

Não existe correlação entre a quantidade de tempo adormecido e e o gasto calórico.

Conclusão

Proposta:

Como a maioria dos usuários não dormem o tempo necessário e nem tem uma constancia, o aplicativo da Bellabeat tem que ter a funcionalidade de performance do sono semanal, informando se o usuário está dormindo o tempo necessário, bem com um comparativo com os dias anteriores. Seria interessante nessa funcionalidade mostrar se essa performance de sono segue uma rotina, como dormir nos mesmo horários, acordar nos mesmos horários e mostrar o tempo dormido.

Como o aplicativo se conecta com outros dispositivos da marca, é de interesse da empresa investir em funcionalidade de alerta para hora de dormir, bem com orientações para uma melhor qualidade de sono, visando reduzir o tempo que se passa acordado na cama.

Criação de notificações sonoras para a prática de alguma atividade física, principalmente nos momentos de maior tempo de momentos sedentários.

Criação metas para os usuários quee incentivam a praticar mais exercícios físico. Nessa funcionalidade poderia mostrar quantos passos ou quilômetros foi percorrido no dia e comparar com o dia anterior e propondo aos usuários darem alguns passos a mais.

Criação de notificações, que ressaltam a importancia do uso do dispositivo para aqueles usuários que não ficam o tempo todo com o dispositivo.