

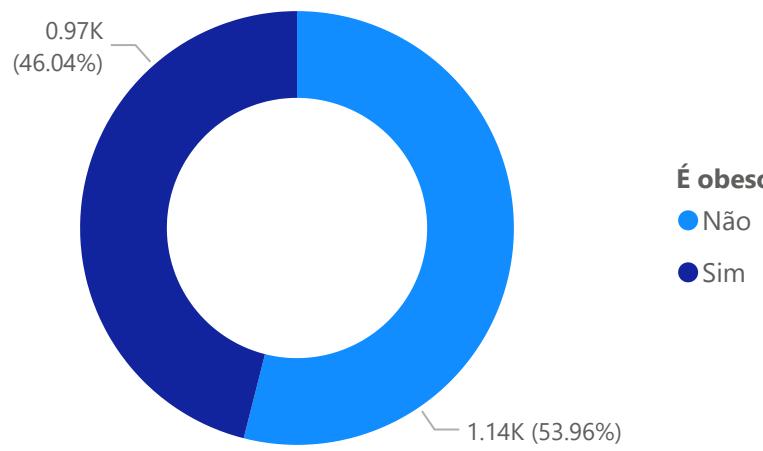
Obesidade e seus fatores contribuintes

Victoria Costi Menezes | RM361292

Os dados foram fornecidos pelos clientes, e descrevem o *status* de **obesidade** do paciente e vários fatores biológicos e cotidianos que podem afetar a saúde e peso.

A separação dos dados em grupos para análise foi de uma simples pergunta binária: se o paciente é obeso, ou não.

1. Obesidade



A separação dos dados em grupos para análise foi de uma simples pergunta binária: se o paciente é obeso, ou não.

Temos 46% sim, e 54% não (fig. 1). Como a divisão é aproximadamente 50/50, podemos observar e comparar números brutos entre esses grupos sem convertê-los em proporção.

Ignoramos os diferentes tipos de obesidade, pois a base apresenta um desbalanço no quesito de gênero: III é exclusiva a mulheres, e II a homens. A divisão iria apenas atrapalhar os insights que buscamos.

A incidência de obesidade entre homens e mulheres, nesse dataset, é aproximadamente a mesma (fig. 4).

É obeso

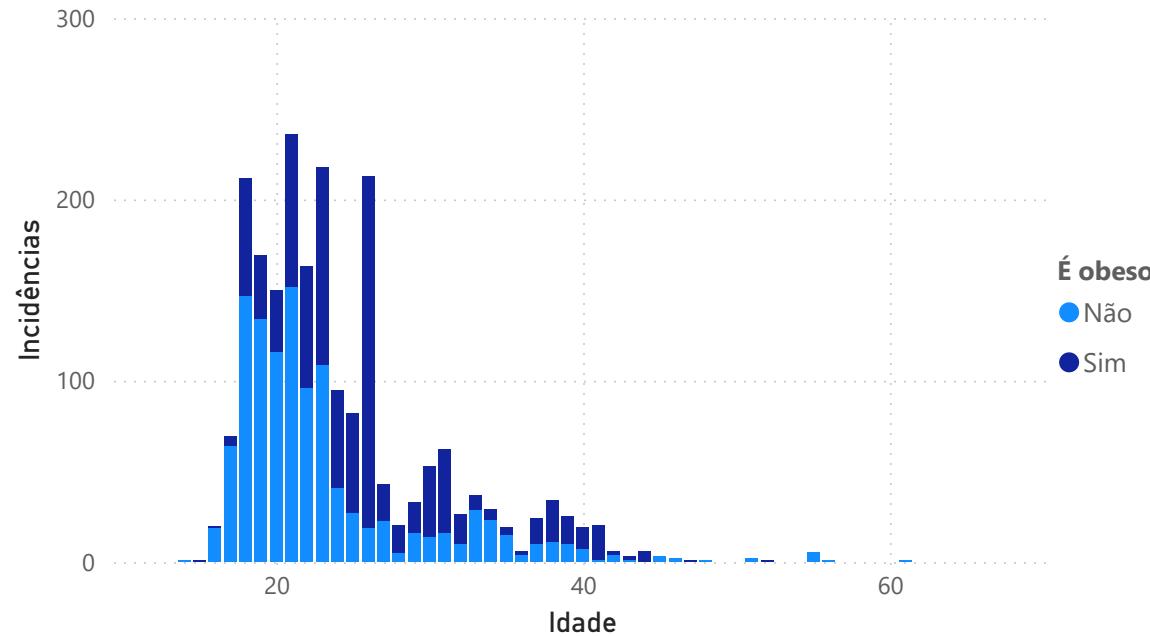
▼ Gênero

- Homem
- Mulher

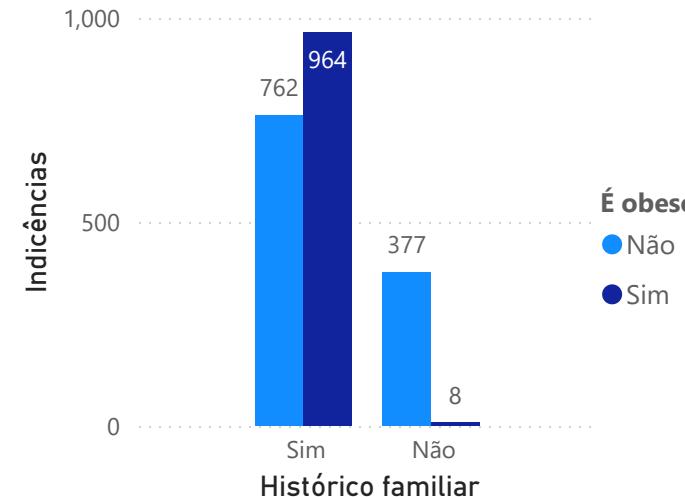
▼ 2.111K Respondentes

A maioria dos casos de obesidade se acumulam na faixa de 18 a 30 anos, com um enorme pico aos 26 anos de idade (fig. 2). Esse pico é composto primariamente por mulheres.

2. Faixa de peso por idade

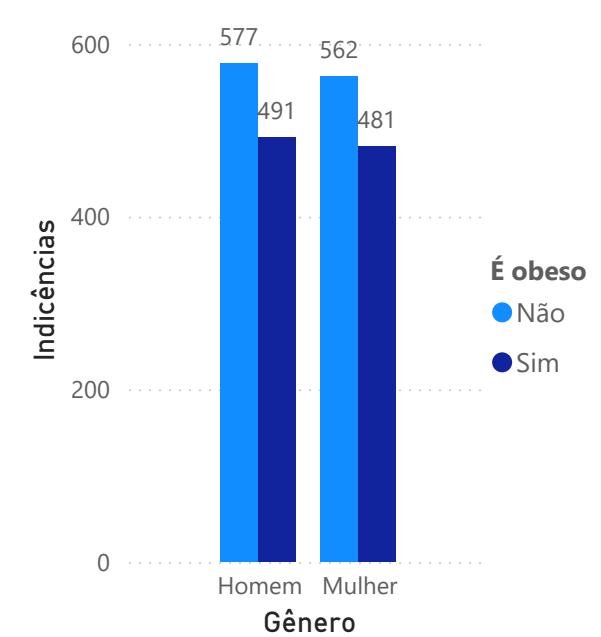


3. Faixa de peso por Histórico familiar

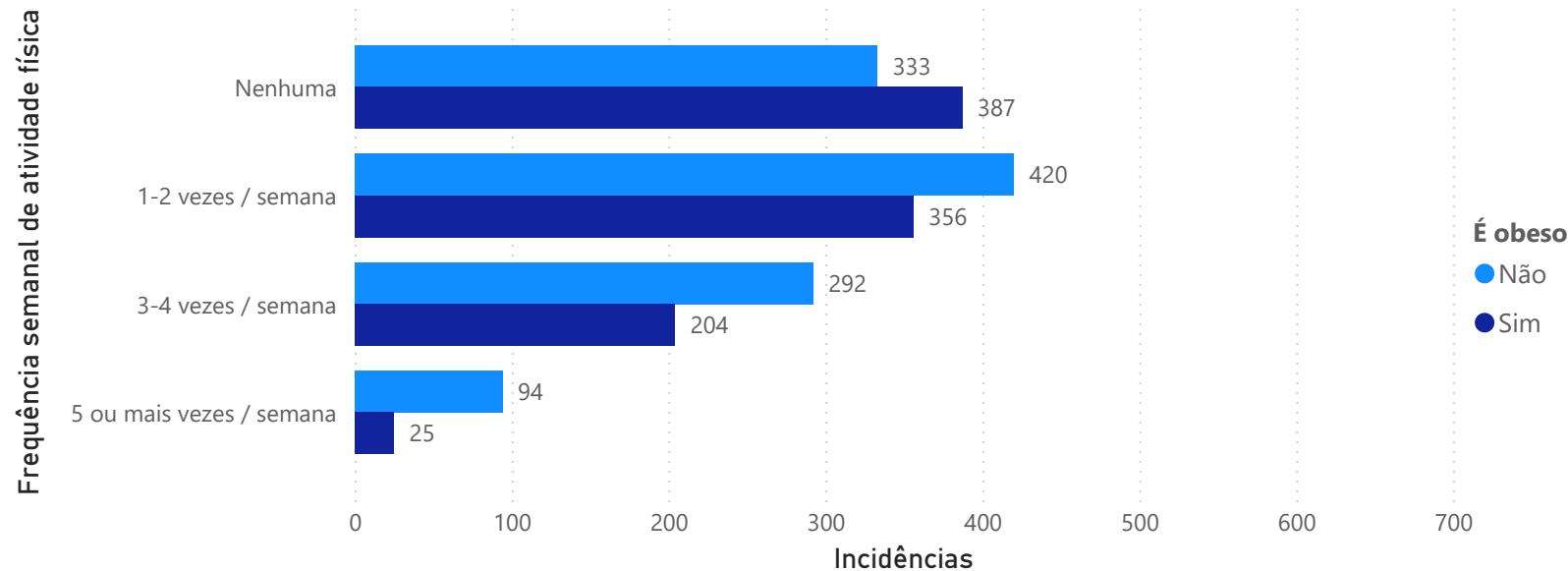


Quase todos os respondentes que apresentam obesidade, têm histórico familiar (fig. 3).

4. Faixa de peso por gênero



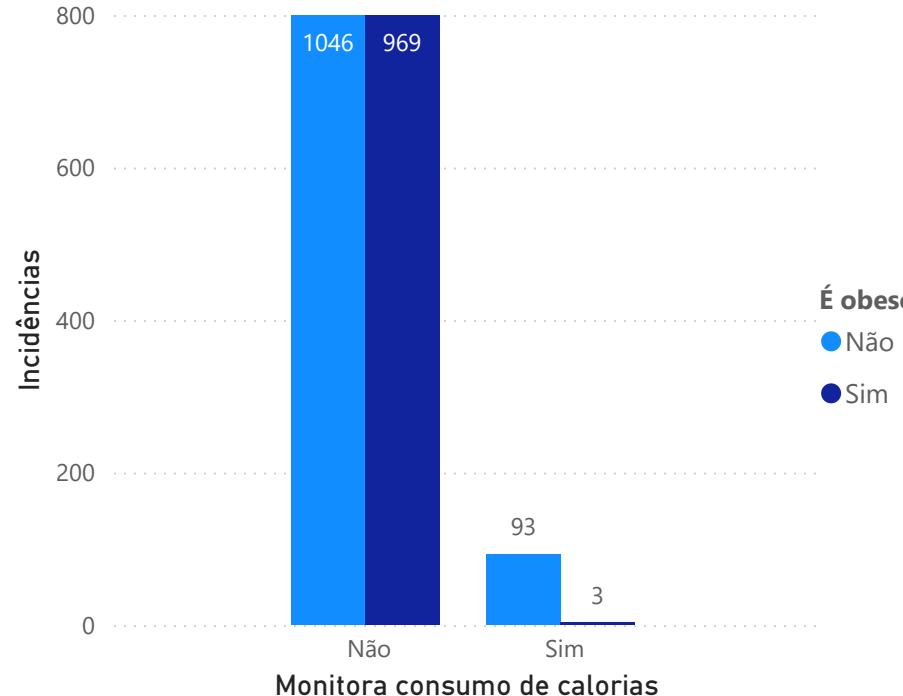
1. Frequência de atividade física por faixa de peso



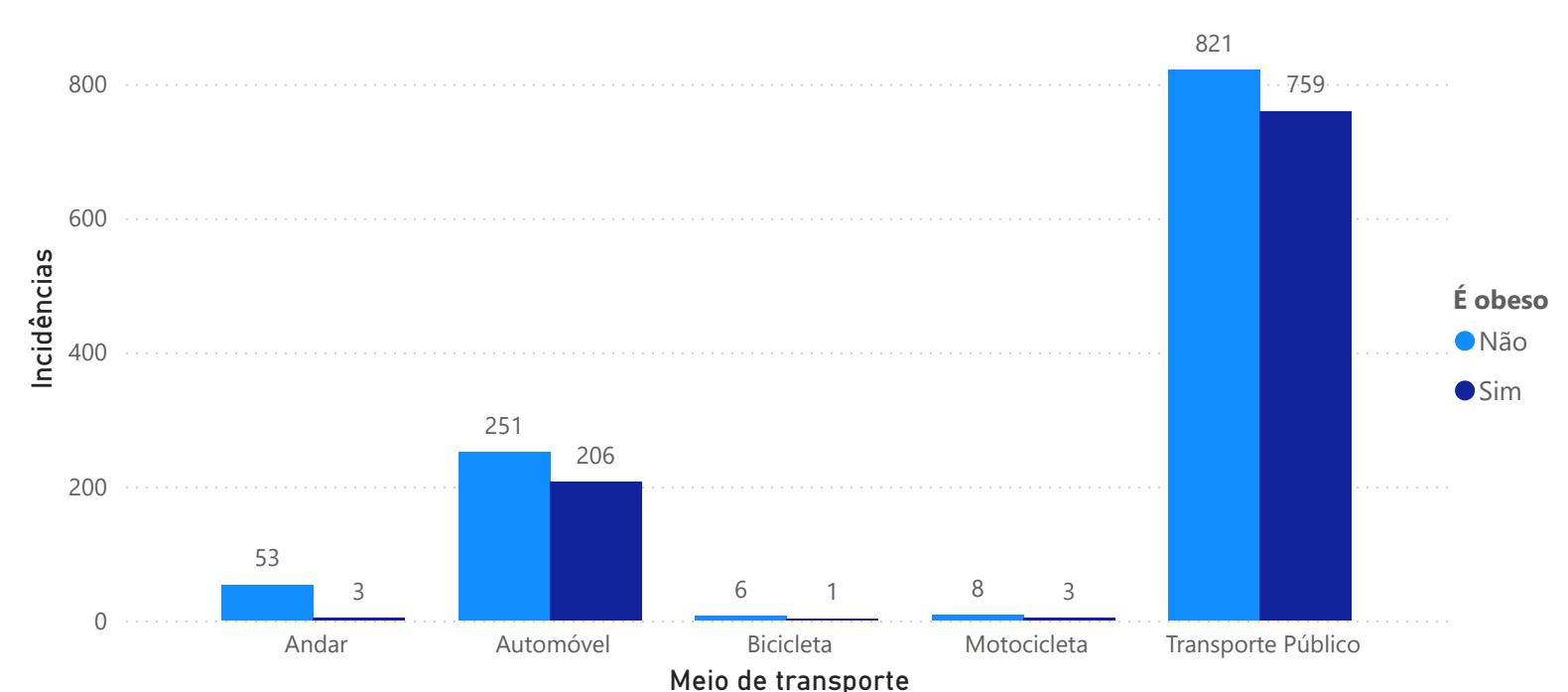
Observando os hábitos dos respondentes, podemos afirmar que:

- Pessoas com obesidade tendem a se exercitar com menor frequência que pessoas não-obesas (fig. 1)
- Aqueles que monitoram calorias raramente são obesos (fig. 2)
- Aqueles que andam como meio primário de transporte raramente são obesos (fig. 3)
 - Bicicleta e Motocicleta obtiveram poucos respondentes, portanto não iremos fazer afirmações sobre essas opções.

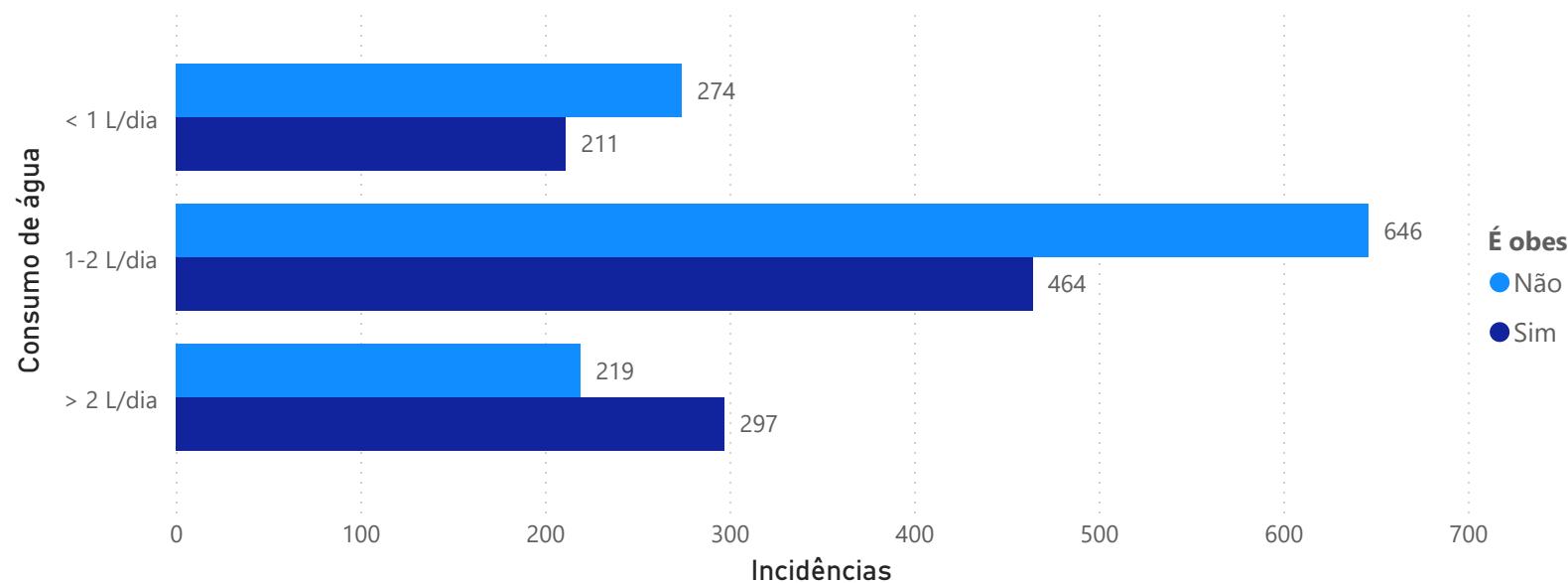
2. Monitoração de calorias por faixa de peso



3. Meio de transporte por faixa de peso



1. Consumo de água por faixa de peso

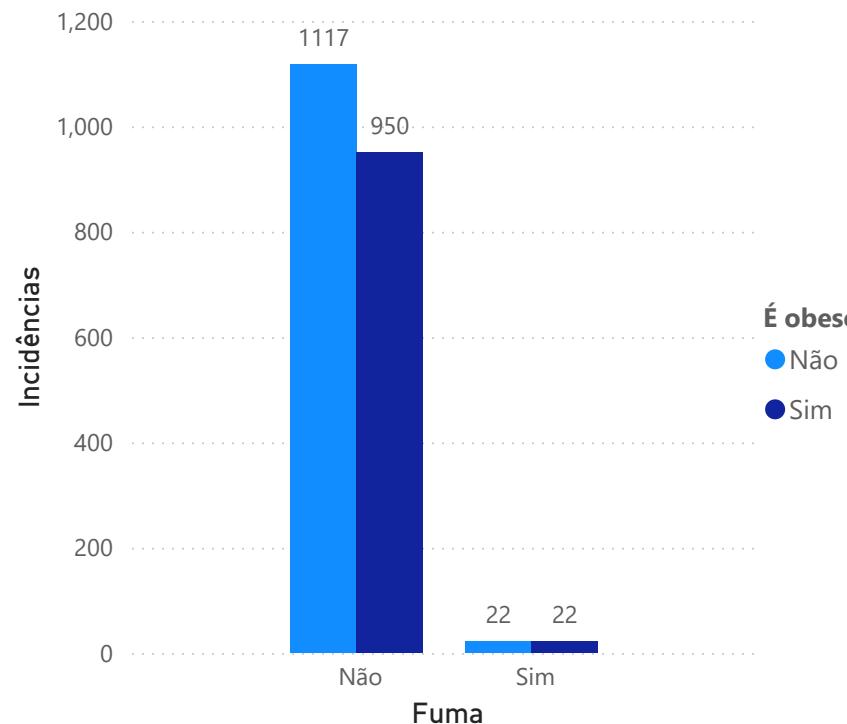


- Pessoas obesas parecem consumir mais água que pessoas não-obesas (fig. 1)

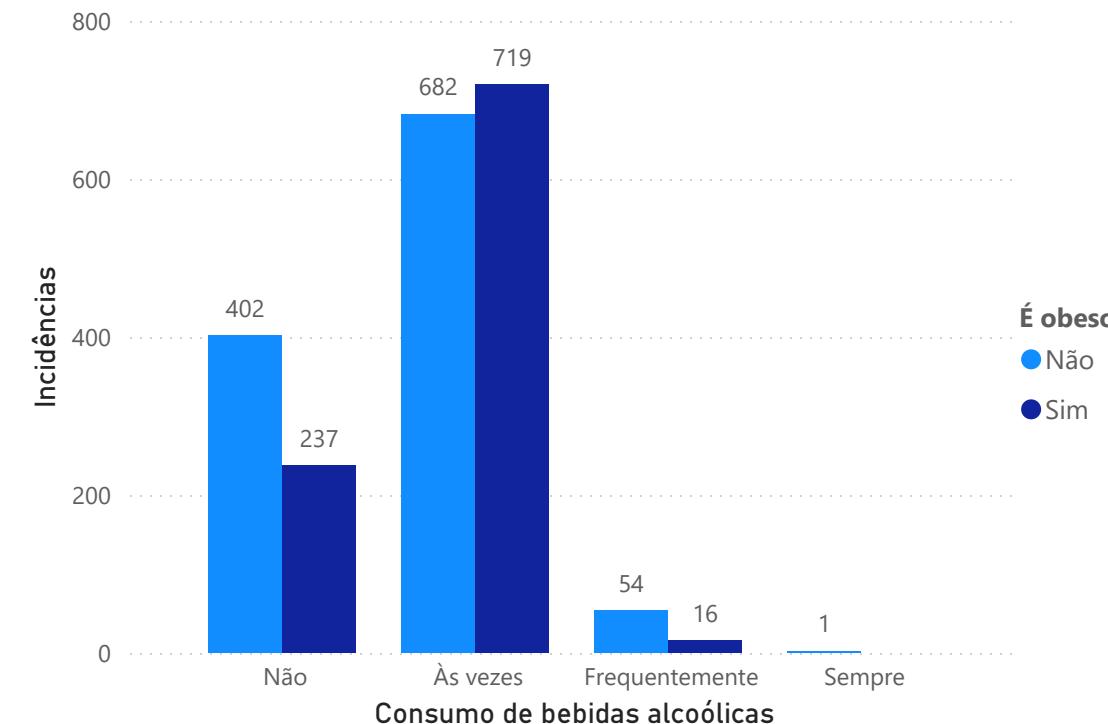
- O hábito de fumar não parece diretamente se correlacionar à obesidade (fig. 2)

- Aqueles que não consomem álcool são obesos com menor frequência, mas o consumo de álcool entre obesos e não obesos é similar (fig. 3)

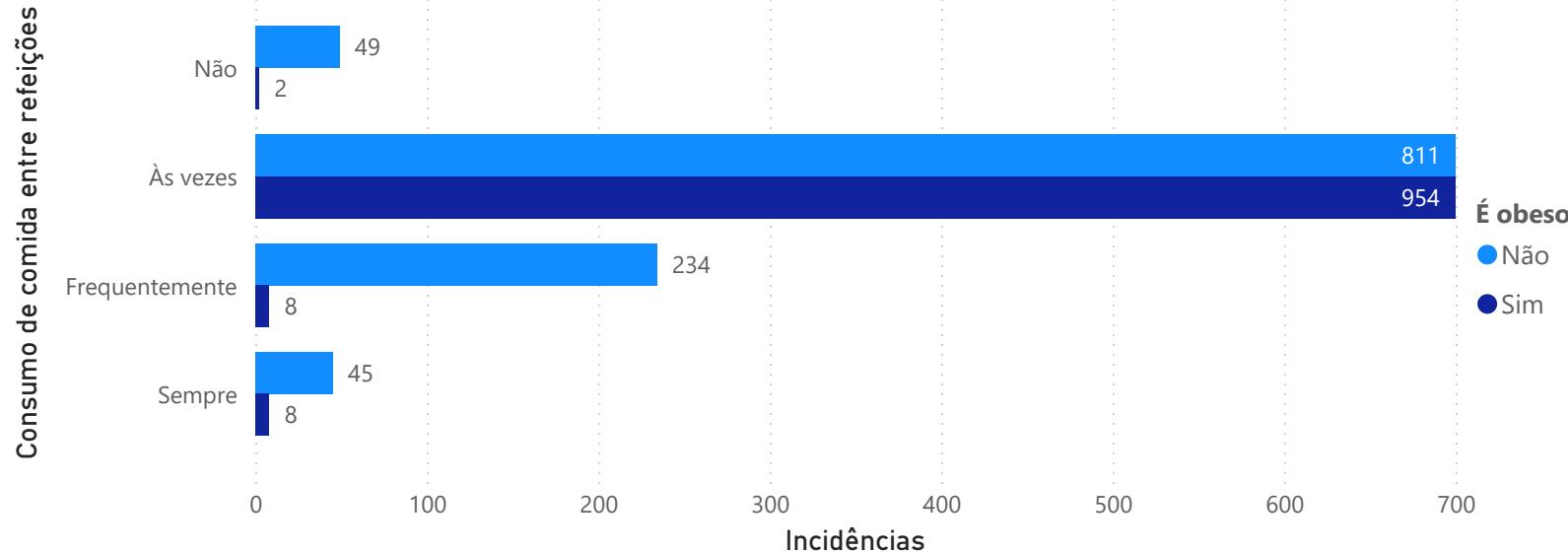
2. Hábito de fumar por faixa de peso



3. Consumo de bebidas alcóolicas por faixa de peso

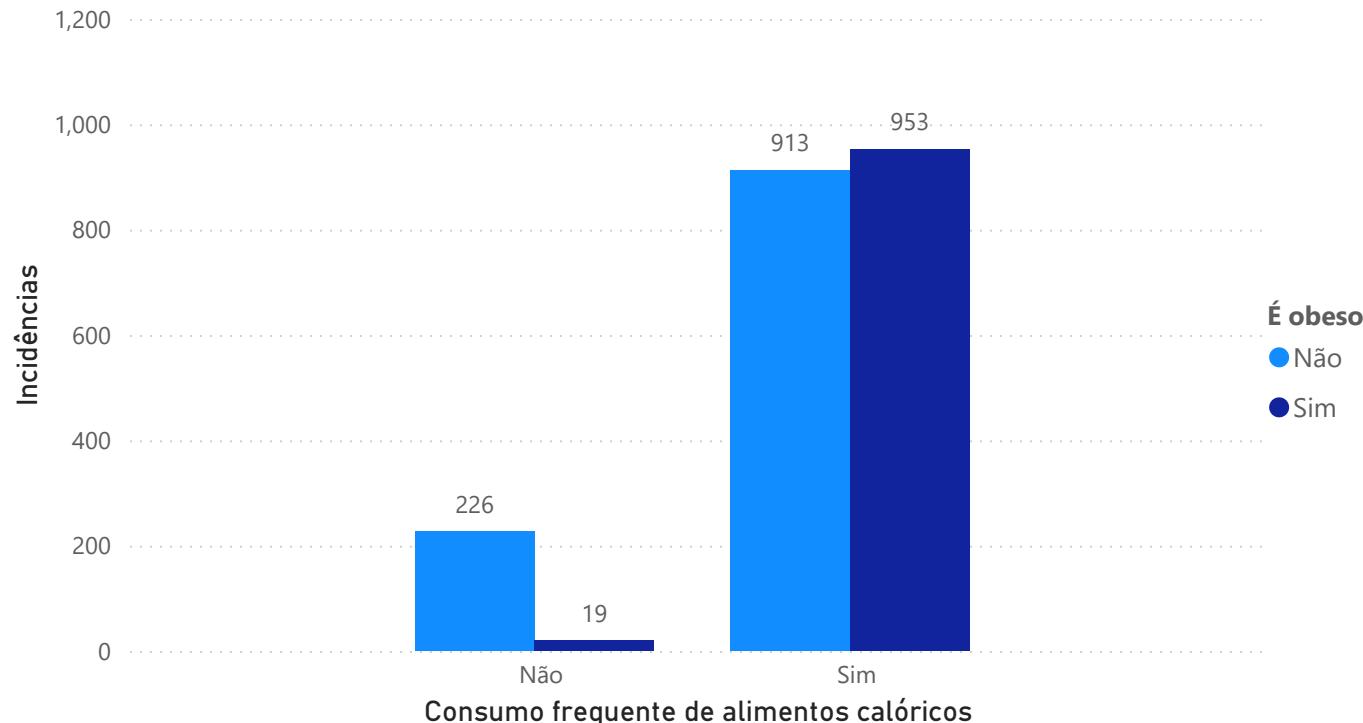


1. Consumo de comida entre refeições por faixa de peso

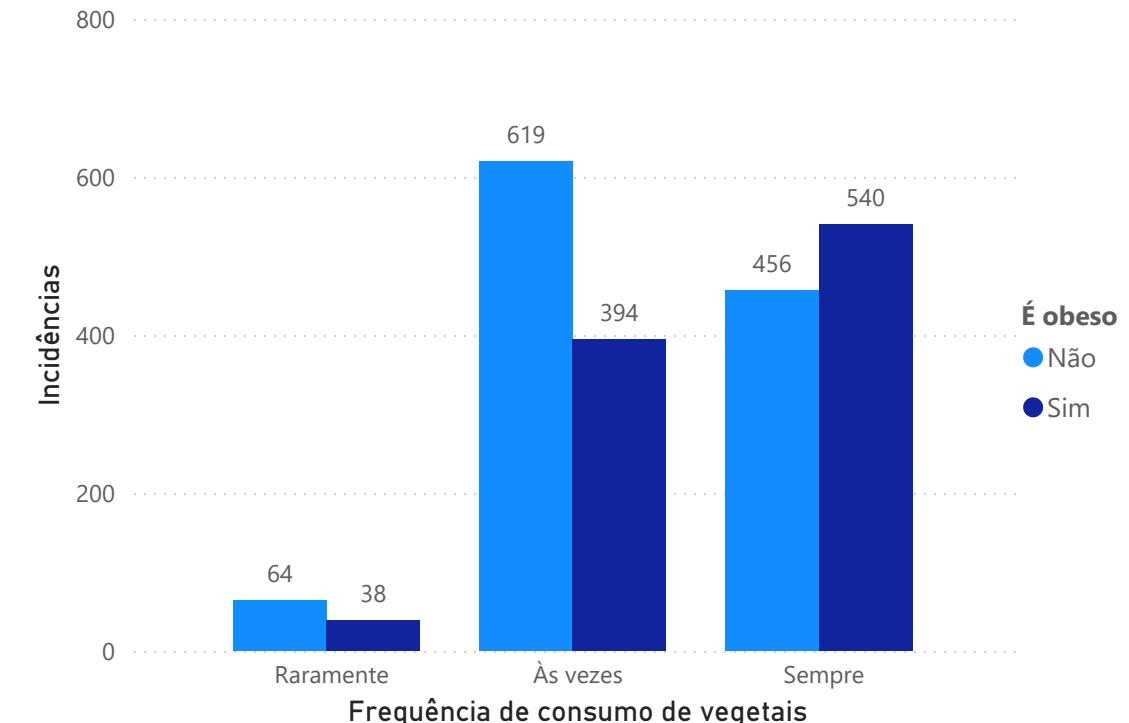


- Quase todos os obesos neste dataset se **alimentam entre refeições** “às vezes”; não-obesos apresentam uma maior variedade de respostas (fig. 1)

2. Consumo de alimentos calóricos por faixa de peso



3. Consumo de vegetais por faixa de peso



- Aqueles que **não consomem alimentos calóricos**, raramente são obesos (fig. 2)
- Obesos, nesse dataset, **consomem vegetais** com maior frequência que não-obesos (fig. 3)

Obesidade e seus fatores contribuintes

Conclusão

O fator que mais diretamente contribui para a incidência de obesidade é o **histórico familiar**. Com base nos nossos insights, os seguintes fatores foram selecionados para a construção do modelo:

- Idade
- Altura
- Histórico familiar
- Consumo de alimentos calóricos
- Alimentação entre refeições
- Monitoramento de calorias
- Atividade física
- Meio de transporte

O tratamento dos dados foi feito com uma aplicação de uma pipeline para padronizar as respostas em dados numéricos quando possível, e em dados booleanos categóricos em outros casos. A seleção do modelo foi feita a partir da uma simples avaliação de acurácia, e incidência de falsos positivos e falsos negativos.

O modelo escolhido foi um de KNearestNeighbors, com uma **acurácia de 80%**.