

Tarefa 2 - Estatística

André Filipe da Silva Fernandes - 19100515

Outubro 2022

1

Os histogramas a seguir foram obtidos filtrando os dados pelas marcas "SPATEN" e "HEINEKEN" respectivamente e entre embalagens "C16 - LATA - 350 ml", desconsiderando todos os valores nulos. Para facilidade de leitura os gráficos compartilham a mesma proporção em seus dois eixos e a amplitude do eixo de preços foi fixada entre R\$ 0.00 e R\$ 10.00. Além disso cada classe possui amplitude de R\$ 0.50.

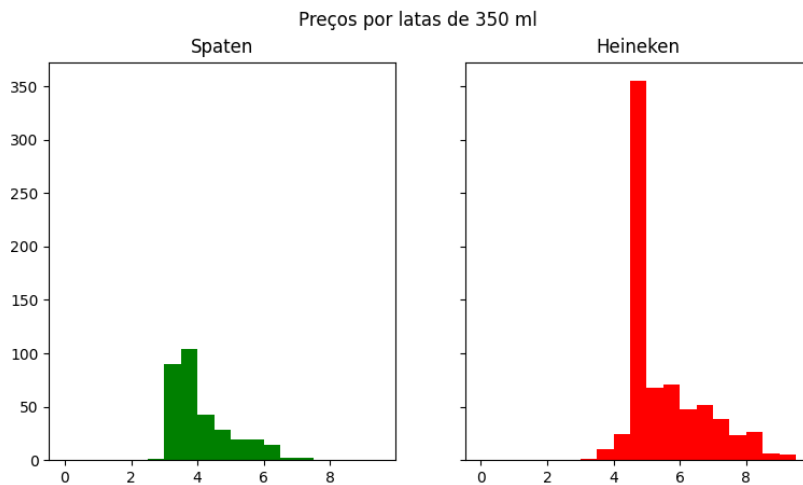


Figure 1: Histograma de preço das cervejas

2

O histograma da cerveja Spaten é uma distribuição assimétrica com uma tendência à esquerda, e a maior parte dos valores estão concentrados entre R\$ 3.00 e R\$ 3.99. Progressivamente, ao observar valores mais altos, a quantidade de estabelecimentos que oferece o produto diminui.

No caso da cerveja Heineken algo diferente acontece. Em geral os preços parecem quase seguir uma distribuição normal com a moda em torno de 5 reais, mas não é este o caso pois uma quantidade muito grande de estabelecimentos oferece o produto com valores entre R\$ 4.50 e R\$ 4.99. Acredito que para aproximar este comportamento seja necessário usar utilizar função mais complexa diferente das comumente utilizadas.

Além destas observações, nota-se pelo gráfico que a cerveja Heineken é vendida em mais lugares do que a Spaten. Também é possível observar que em geral a cerveja Heineken custa um pouco mais caro.

3

3.1 Estatísticas Spaten

A tabela a seguir apresenta as principais estatísticas descritivas obtidas sobre o preço da cerveja Spaten em latinhas de 350 ml.

Estatística	Valor
Média	4.199
Mediana:	3.99
Moda:	3.49
Variância:	0.6646
Desvio padrão:	0.8152
Erro padrão:	0.04543
Coeficiente de variação:	19.41
Assimetria:	0.009901
Curtose:	0.0
Quartis 1:	3.49
Quartis 3:	4.5
Mínimo Relevante:	2.79
Máximo Relevante:	6.0

3.2 Estatísticas Heineken

A tabela a seguir apresenta as principais estatísticas descritivas obtidas sobre o preço da cerveja Heineken em latinhas de 350 ml.

Estatística	Valor
Média	5.538
Mediana:	4.99
Moda:	4.79
Variância:	1.105
Desvio padrão:	1.051
Erro padrão:	0.03896
Coefficiente de variação:	18.98
Assimetria:	0.6694
Curtose:	0.0
Quartis 1:	4.79
Quartis 3:	6.0
Mínimo Relevante:	3.15
Máximo Relevante:	7.8

4

Observando o box-plot é possível chegar à mesma conclusão, de que a cerveja Heineken em geral é um pouco mais cara, mas pode ser muito mais cara, mesmo se ignorarmos valores que são muito discrepantes. Isso é possível de perceber pois este tipo de gráfico destaca claramente o que é e o que não é exceção de acordo com um critério heurístico.

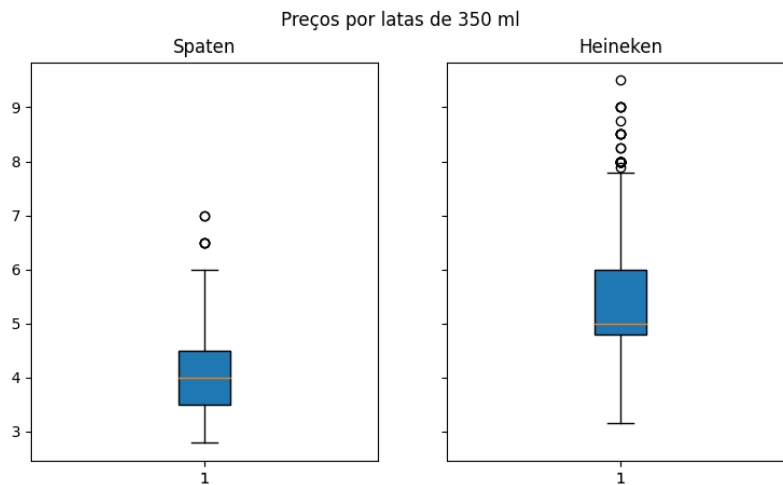


Figure 2: Box-plot do preço das cervejas Spaten

5

Em ambos os plots existem alguns valores discrepantes de acordo com a regra analítica, ou seja, através do seguinte critério:

$$irq = Q3 - Q1 \quad (1)$$

$$outlier = \{x \mid (x < Q1 - 1.5 * irq) \text{ and } (x > Q3 + 1.5 * irq)\} \quad (2)$$

Entre os preços obtidos da cerveja Spaten todos os valores discrepantes foram extremos, são eles: R\$ 6.49, R\$ 6.50 e R\$ 7.00. No caso da cerveja Heineken foram encontrados mais valores discrepantes extremos, são eles: R\$ 7.90, R\$ 7.99, R\$ 8.0, R\$ 8.25, R\$ 8.50, R\$ 8.75, R\$ 9.00 e R\$ 9.50.

Do ponto de vista do consumidor estes são casos em que as bebidas estão custando um valor muito caro, pois em ambas as marcas todos os outliers encontrados são preços muito acima média e nunca abaixo.