



**UNIVERSIDADE
FEDERAL DO CARIRI**

**UNIVERSIDADE FEDERAL DO CARIRI
CENTRO DE CIÊNCIAS E TECNOLOGIA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**1ª Avaliação
Rosilda Benício de Souza**

Wanderson Faustino Patricio - MAT: 2022005052

1 Variável Qualitativa

- a) Nesse item estudaremos uma variável qualitativa nominal, a quantidade de mortes decorridas da covid-19 ao redor do mundo.

Para fazer o tratamento dos dados precisamos das informações a respeito da quantidade de pessoas que perderam a vida em uma amostra de países. Para esse estudo foram 30 países de diferentes continentes.

- b) Apresentando a tabela:

Mortes por covid por país		
País	frequência	frequência (%)
USA	1,132,935	21.64
India	530,740	10.14
France	164,233	3.14
Germany	165,711	3.16
Brazil	697,074	13.31
Japan	68,399	1.31
S. Korea	33,486	0.64
Italy	186,833	3.57
UK	204,171	3.90
Russia	395,108	7.55
Turkey	101,492	1.94
Spain	118,434	2.26
Vietnam	43,186	0.82
Australia	18,615	0.36
Argentina	130,421	2.49
Taiwan	16,356	0.31
Netherlands	22,989	0.44
Iran	144,749	2.76
Mexico	332,198	6.34
Indonesia	160,817	3.07
Poland	118,736	2.27
Colombia	142,486	2.72
Austria	21,689	0.41
Greece	35,630	0.68
Portugal	26,022	0.50
Ukraine	111,020	2.12
Chile	63,812	1.22
Malaysia	36,942	0.71
Israel	12,193	0.23
DPRK	74	0.00
Total	5,236,551	100.00

Disponível em <<https://www.kaggle.com/datasets/themrityunjaypathak/covid-cases-and-deaths-worldwide?resource=download>> Acesso em 03 de março de 2023

c) Apresentando os dados em um gráfico de colunas:

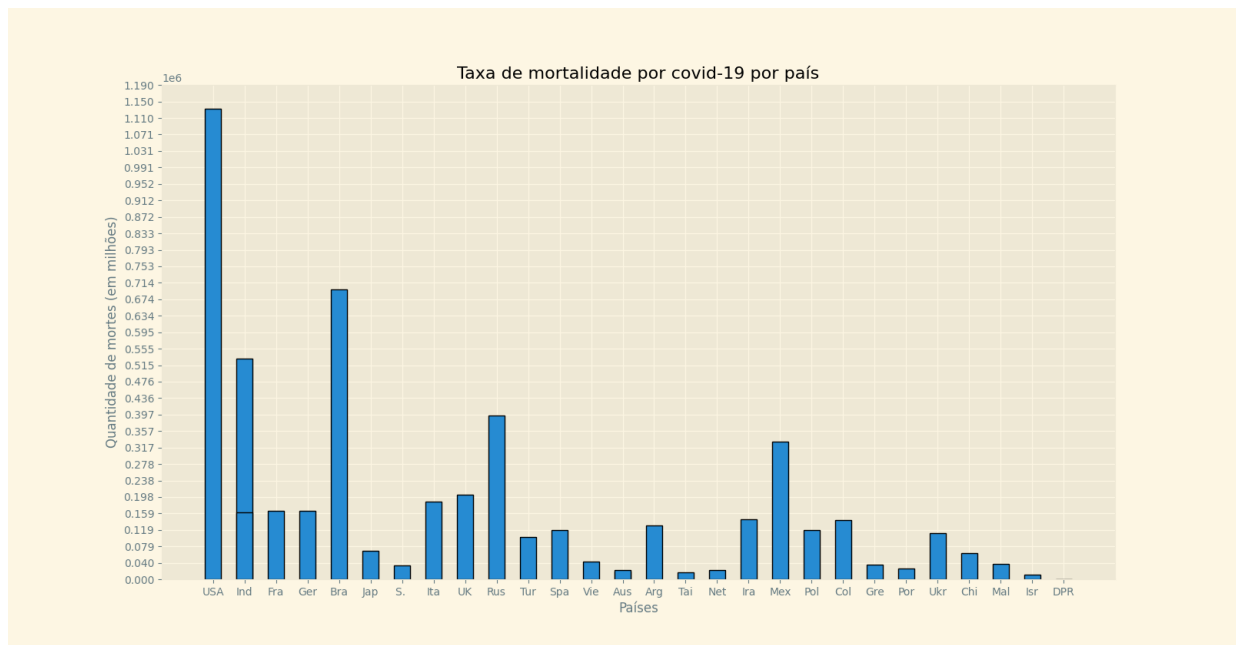


Figura 1: Gráfico de colunas da taxa de mortalidade

d) Analisando o gráfico e a tabela podemos retirar as seguintes conclusões:

- Os Estados Unidos da América foi o país com a maior quantidade de mortes, e a Coreia do Norte (DPRK) foi o país com menos mortes.
- Países mais bem desenvolvidos tecnologicamente (como Japão, França e Alemanha) tiveram uma quantidade de mortes menor em relação aos outros países.
- A quantidade de mortes média dos países foi de 174552 mortes, porém o desvio padrão desse conjunto de dados foi de 241137.22, o que indica uma alta variabilidade nos dados, pois o coeficiente de variação é 138.15%.
- Ao olharmos para a amplitude dos dados (1132861 mortes) vemos uma alta variação. Para se ter um exemplo disso podemos comparar a maior quantidade de mortes (USA) com a menor (DPRK), visto que a quantidade de mortes nos Estados Unidos é 1531 vezes maior que a quantidade na Coreia.

2 Variável Quantitativa contínua

- a) Para a nossa variável quantitativa contínua utilizaremos o preço do videogame Playstation 5 no site do Mercado Livre.

Ao olhar os preços (em R\$) dos videogames no site encontramos a seguinte lista de dados:

4499.00, 4599.00, 4799.00, 4464.00, 4399.00, 4770.00, 4647.00, 5343.00, 5099.00, 5600.00, 4637.00, 4199.00, 5199.00, 4647.00, 4299.00, 5000.00, 5199.00, 5699.00, 4449.00, 4732.00, 4903.00, 4899.00, 4990.00, 5599.00, 5099.00, 5499.00, 4697.00, 5499.00, 5197.00, 4890.00, 5318.00, 4299.00, 4748.00, 7799.00, 4697.00, 4799.00

Colocando os dados em rol temos:

4199.00, 4299.00, 4299.00, 4399.00, 4449.00, 4464.00, 4499.00, 4599.00, 4637.00, 4647.00, 4647.00, 4697.00, 4697.00, 4732.00, 4748.00, 4770.00, 4799.00, 4799.00, 4890.00, 4899.00, 4903.00, 4990.00, 5000.00, 5099.00, 5099.00, 5197.00, 5199.00, 5199.00, 5318.00, 5343.00, 5499.00, 5499.00, 5599.00, 5600.00, 5699.00, 7799.00

- b) Apresentando a tabela de frequências temos:

**Distribuição do preço do videogame Playstation 5
no site do Mercado Livre**

Preço (R\$)	f	f(%)	F	F(%)	$x_m(R\$)$
4000.00 ┤ 4800.00	18	50.00	18	50.00	4400.00
4800.00 ┤ 5600.00	15	41.67	33	91.67	5200.00
5600.00 ┤ 6400.00	2	5.56	35	97.23	6000.00
6400.00 ┤ 7200.00	0	-	35	97.23	6800.00
7200.00 ┤ 8000.00	1	2.77	36	100.00	7600.00
Total	36	100.00	-	-	-

Disponível em <[https://lista.mercadolivre.com.br/ps5#D\[A:playstation%205\]>](https://lista.mercadolivre.com.br/ps5#D[A:playstation%205]>)

Acesso em 03 de março de 2023

- c) Calculando as medidas de posição e dispersão:

- i) Média

$$\bar{x} = \frac{\sum_{i=1}^n f_i \cdot x_{m_i}}{\sum_{i=1}^n f_i} = \frac{18 \cdot 4400 + 15 \cdot 5200 + 2 \cdot 6000 + 0 \cdot 6800 + 1 \cdot 7600}{36}$$

$$\bar{x} = R\$ 4911,11$$

- ii) Mediana

Como Exatamente 50% dos preços estão na primeira classe e 50% acima dela, consideraremos que a classe mediana é a primeira (não irá fazer diferença se utilizarmos a primeira ou a segunda).

$$med = l_i + \frac{\left(\frac{\sum_{i=1}^n f_i}{2} - F_{ant} \right)}{f} \cdot h = 4000 + \frac{\frac{36}{2} - 0}{18} \cdot 800$$

$$med = R\$ 4800,00$$

- iii) Moda

$$mo = l_i + \frac{f_{max} - f_{ant}}{2f_{max} - (f_{ant} + f_{post})} \cdot h = 4000 + \frac{18 - 0}{2 \cdot 18 - (0 + 15)} \cdot 800$$

$$mo = R\$ 4685,71$$

iv) Coeficiente de variação

* Variância

$$s^2 = \frac{\sum_{i=1}^n f_i(x_{m_i} - \bar{x})^2}{n - 1}$$

$$s^2 = 444444,44 (R\$)^2$$

* Desvio Padrão

$$s = \sqrt{s^2} \rightarrow \boxed{s = R\$ 666,67}$$

* Coeficiente de variação

$$cv = \frac{s}{\bar{x}} \cdot 100\%$$

$$\boxed{cv = 13,57\%}$$

d) Apresentando os gráficos

i) Histograma

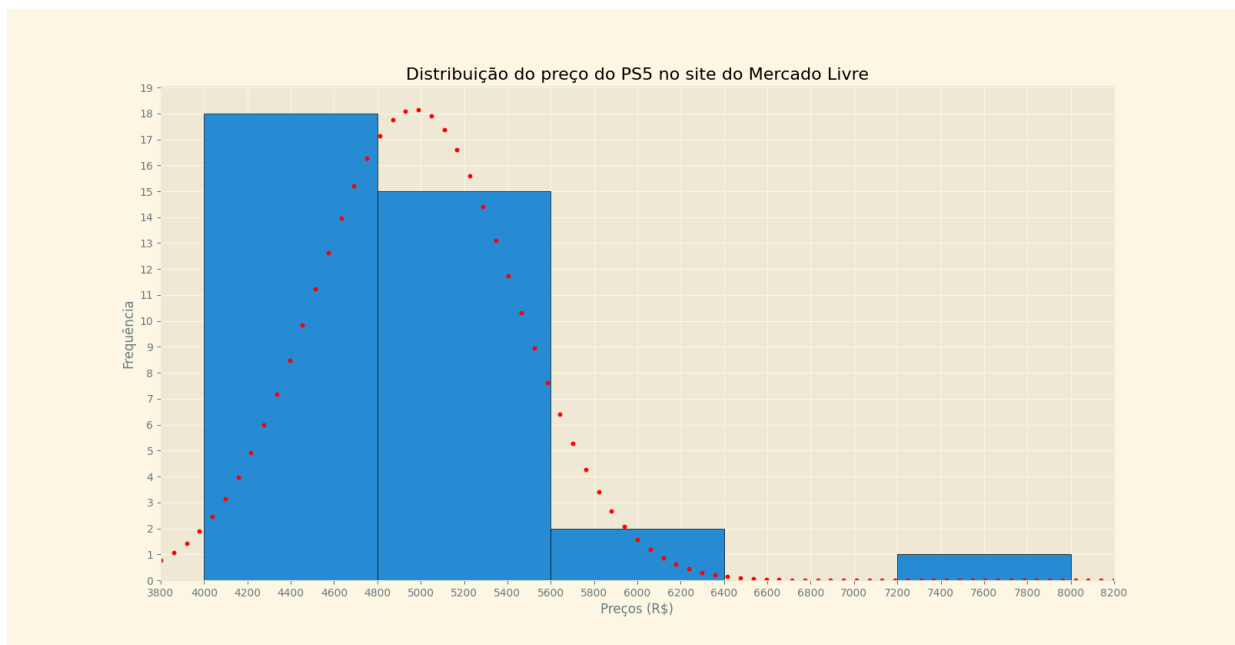


Figura 2: Histograma de distribuição de preços (a linha pontilhada vermelha representa a distribuição normalizada)

ii) Boxplot

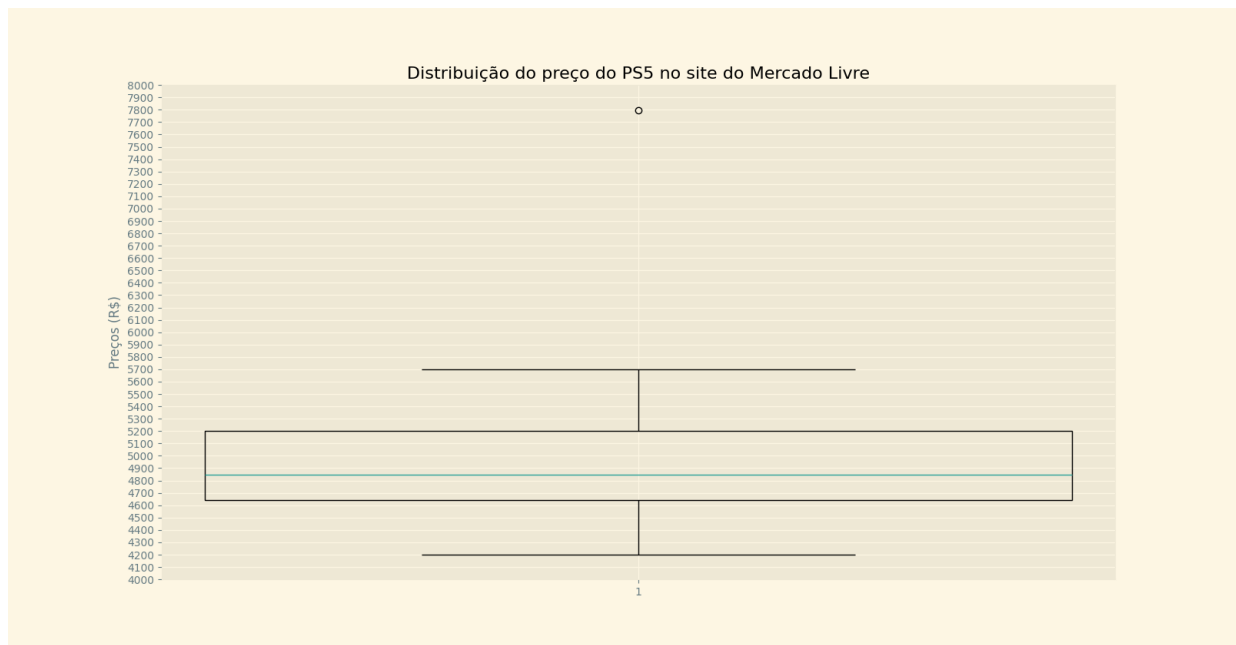


Figura 3: Gráfico Boxplot

e) Analisando os dados apresentados pelo gráfico:

- A distribuição de preços tem uma distribuição assimétrica positiva $\left(a_s = \frac{\bar{x} - mo}{s} \approx 34\%\right)$.
- Não há muita variabilidade nos preços, visto que o coeficiente de variação é de 13,57%.
- A amplitude de preços é R\$ 3600,00.
- A concentração dos preços está muito próximo dos R\$ 4800,00, e o preço de R\$ 7799,00 está muito fora do range, logo, é um preço outlier.
- O preço médio é aproximadamente 3,7 vezes maior que o salário mínimo brasileiro.
- Como visto no gráfico boxplot a mediana (Q_2) está mais perto de Q_1 do que de Q_3 , o que demonstra a assimetria positiva.