

T1 - Trabalho sobre Estatísticas Descritivas
Probabilidade e Estatística para Bioinformática – 2º sem/2023

O conjunto de dados *lixo* contém a informação registrada para uma amostra de pesos (em libras) do lixo descartado por uma residência em uma semana, em um total de 62 residências de um bairro da Califórnia. Faça um relatório respondendo às perguntas abaixo.

Variáveis do estudo:

- *Size* – Tamanho da residência (1 – até 60m², 2 - de 60 a 100m², 3 – mais de 100m²)
- *Metal* – Peso dos metais descartados.
- *Paper* - Peso dos artigos de papel descartados.
- *Plastic* – Peso dos artigos de plástico descartados.

a) Construa o gráfico adequado para a variável qualitativa. Comente o gráfico.

Para as variáveis *Metal*, *Paper* e *Plastic*:

- b) Calcule a média, a mediana e a moda. Qual destas medidas melhor representa a tendência central dos dados em cada caso? Por que?
- c) Calcule o primeiro e o terceiro quartis. Explique o significado destes números.
- d) Calcule o percentil 80.
- e) Calcule a variância e o desvio-padrão.
- f) Qual das três variáveis é mais homogênea? Por que?

Análise cruzada

- g) Faça box-plots para a variável *Metal* separada por *Size*. Comente o gráfico comparando a variável por *Size*.
- h) Faça box-plots para a variável *Paper* separada por *Size*. Comente o gráfico comparando a variável por *Size*.
- i) Faça box-plots para a variável *Plastic* separada por *Size*. Comente o gráfico comparando a variável por *Size*.
- j) Faça gráficos de dispersão com as variáveis *Metal* x *Paper*, *Metal* x *Plastic* e *Paper* x *Plastic*. Comente os gráficos.