



Instituto Politécnico de Setúbal

Escola Superior de Tecnologia do Barreiro

**Trabalho de Grupo Projeto de Introdução à
Estatística**

Licenciatura em Bioinformática

**Estudo estatístico sobre
Ranunculus sp., *Geranium sp.*,
*Crepis sp.***

Junho de 2022

Grupo 07

João Almeida (202100068)

Pedro Pacheco (202100957)

Hélder Marques (202100959)

Guilherme Parreira (202100206)

Índice

1	Introdução	1
2	Metodologia	1
2.1	Conceitos utilizados	1
2.2	População alvo e variáveis	1
2.3	Análise estatística	1
3	Caraterização da amostra	2
3.1	Análise Descritiva Univariada e Bivariada	2
3.1.1	Análise do comprimento das pétalas por espécie(Bivariada)	2
3.1.2	Análise da largura das pétalas por espécie (Bivariada) . .	3
3.1.3	Análise da Humidade das pétalas por cor (Bivariada) . .	4
3.1.4	Humidade por espécie (Bivariada)	5
3.1.5	Gráfico circular para a variável cor (Univariadas)	6
3.1.6	Gráfico circular para a variável espécie (Univariadas) . . .	7
3.1.7	Gráfico circular para a variável Humidade (Univariadas) .	8
3.1.8	Gráfico para a variável comprimento de petala (Univariadas)	9
3.1.9	Gráfico para a variável largura da petala (Univariadas) .	10

1 Introdução

Neste projeto vamos realizar uma análise descritiva univariada e bivariada sobre 3 (três) espécies de flor: *Ranunculus sp.*, *Geranium sp.*, *Crepis sp.*, o seu número, comprimento e largura das pétalas, cor (amarelo e roxo) e humidade. Vamos abordar o comprimento e a largura das pétalas por espécie e a humidade por cor da pétala e por espécie. Vamos também analisar o tipo de espécie em relação a [ESCOLHA] e a correlação entre [ESCOLHA1] e [ESCOLHA2]. Em todas estas análises incluiremos tabelas e gráficos, assim como também todos os cálculos realizados para fundamentar as mesmas.

2 Metodologia

2.1 Conceitos utilizados

Desenvolvemos este projeto com recurso à análise estatística. Análise estatística permite recolher, explorar, organizar e apresentar grandes quantidades de dados. Desenvolvemos estas atividades através de testes de hipóteses, análise exploratória e modelação de dados. Utilizámos uma população alvo que é um conjunto de indivíduos sobre os quais é efetuada uma avaliação e/ou uma análise estatística.

2.2 População alvo e variáveis

A nossa população de estudo é constituída por 150 (cento e cinquenta) amostras das seguintes espécies *Ranunculus sp.*, *Geranium sp.*, *Crepis sp.* Para avaliar a nossa população alvo utilizámos as seguintes variáveis: Tipo de espécie (*Ranunculus sp.*, *Geranium sp.*, *Crepis sp.*) ; Cor (Amarelo e Roxo); Nº pétalas; Comprimento das Pétalas; Largura das Pétalas.

2.3 Análise estatística

Para esta população utilizámos análise univariada e bivariada. A análise univariada analisa a distribuição e dispersão dos dados, inclui a análise dos valores da média, moda, mediana e quartis. A análise bivariada permite observar como duas variáveis se comportam na persença uma da outra, ou seja a sua associação, correlação e dependência.

No nosso projeto utilizámos todas as medições estudadas da estatística univariada, ou seja, média, mediana, moda e quartis. Em relação à análise bivariada utilizámos a associação de Cramer e a correlação de Pearson. Utilizámos a associação de Cramer porque

3 Caraterização da amostra

3.1 Análise Descritiva Univariada e Bivariada

3.1.1 Análise do comprimento das pétalas por espécie(Bivariada)

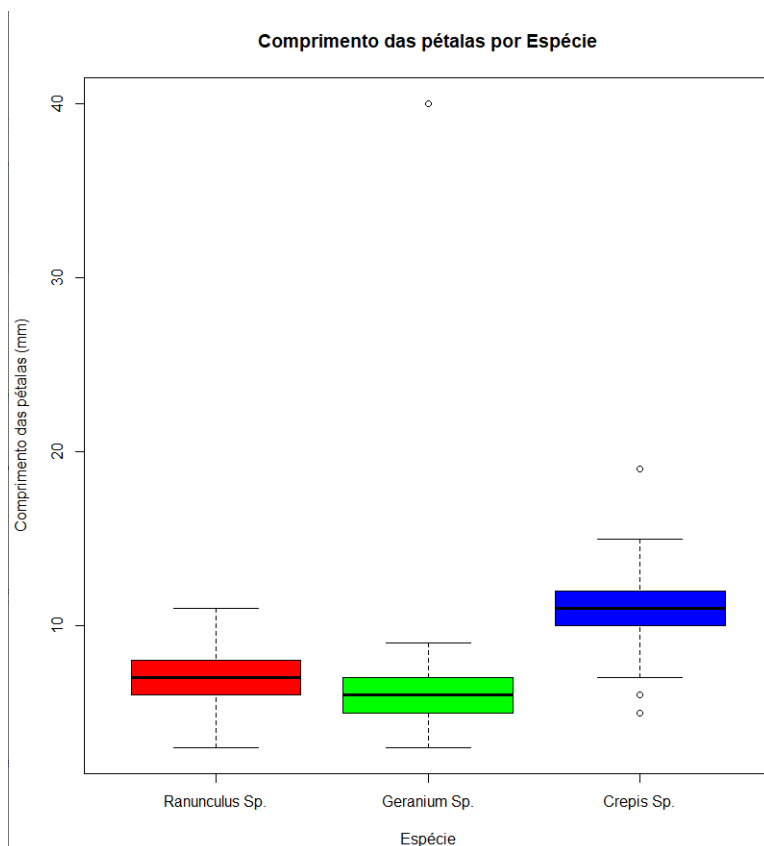


Figura 1: Análise do comprimento das pétalas por espécie.

Como podemos observar neste grafico o comprimento das pétalas (por mm) tem uma pequena divergência de espécie para espécie. Sendo que a espécie *Crepis Sp* apresenta em média o maior comprimento de pétalas entre as tres (3) espécies, e também podemos observar que a espécie *Geranium sp* tem em média o menor comprimento de pétala entre as tres (3) espécies.

Ranunculus sp: minimo de 3, primeiro quartil de 6, mediana de 7, média de 7.1, terceiro quartil 8,máximo11, variância(2,33),desvio-padrão(1,53),coef-dispersão(0.213)

Geranium sp: minimo(3),quartil1(5),mediana(6),média(6.6),quartil3(7),máximo(40),variência(25,59), desvio-padrão(5,06),coef-disperção(0.765)

Crepis Sp: minimo(3),quartil1(10),mediana(11),média(10,620),quartil3(12),máximo(19), variência(6,77),desvio-padrão(2,6),coef-disperção(0.245)

3.1.2 Análise da largura das pétalas por espécie (Bivariada)

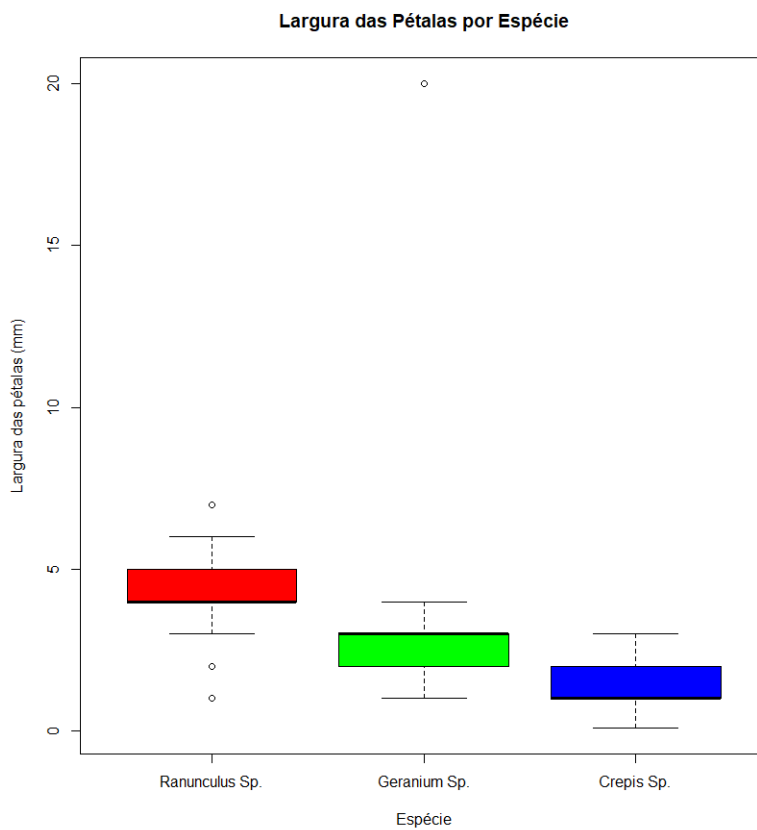


Figura 2: Análise do comprimento das pétalas por cor.

Com a análise deste gráfico podemos compreender que em média a espécie *Ranunculus sp* apresenta uma maior largura de pétalas(mm) entre as tres espécies enquanto que *Crepis Sp* apresenta em média a menor largura de pétalas.

3.1.3 Análise da Humidade das pétalas por cor (Bivariada)

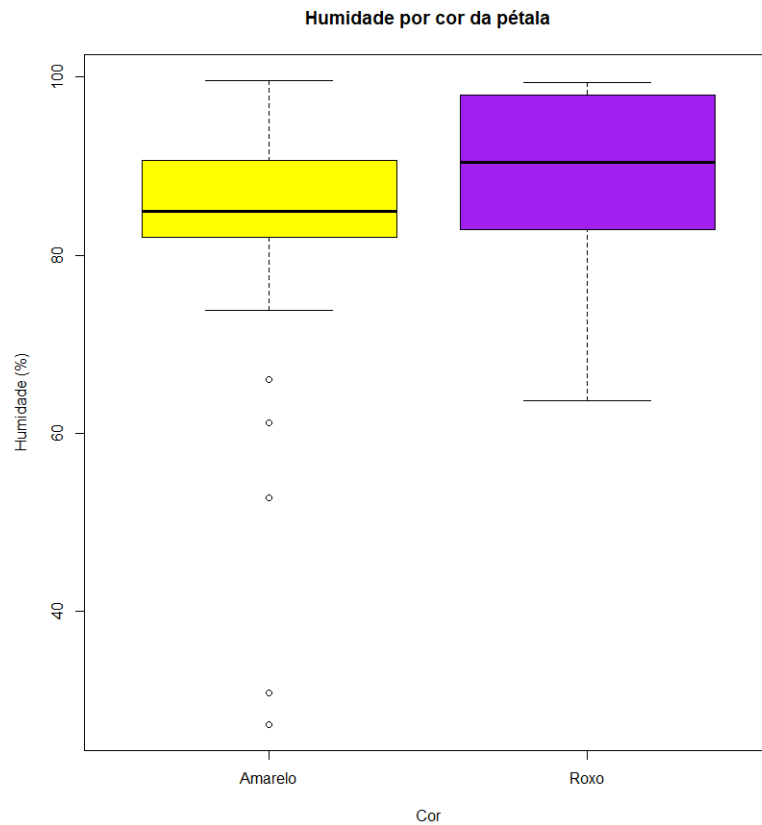


Figura 3: Análise da Humidade das pétalas por cor.

Com este gráfico podemos inferir que a percentagem de humidade nas pétalas de cor roxa é em média maior que as pétalas de cor amarela e também podemos observar a existência de outliers nas pétalas de cor amarela ou seja existem pétalas de cor amarela com uma baixa percentagem de humidade.

3.1.4 Humidade por espécie (Bivariada)

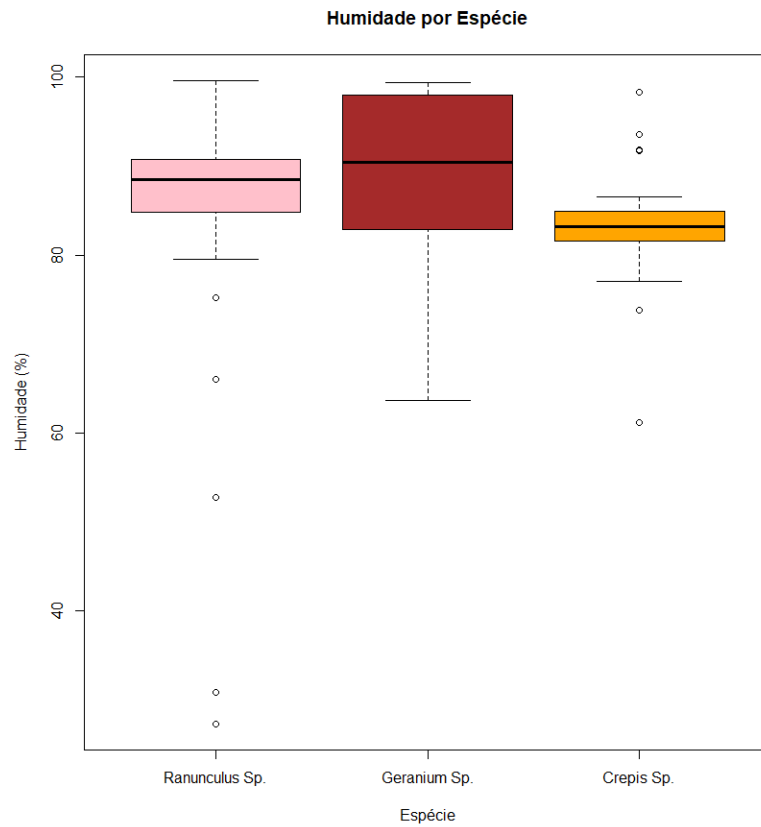


Figura 4: Humidade por espécie

Neste grafico podemos inferir que a espécie *Geranium sp* apresenta em média uma maior percentagem de humidade por espécie. Podemos também observar que a espécie *Crepis Sp* apresenta outliers, ou seja, existem petalas desta espécie com maior e com menor percentagem de humidade. Mas a espécie *Ranunculus Sp* também apresenta outliers com percentagem de humidade inferior.

3.1.5 Gráfico circular para a variável cor (Univariadas)

Gráfico circular para a variável 'Cor'

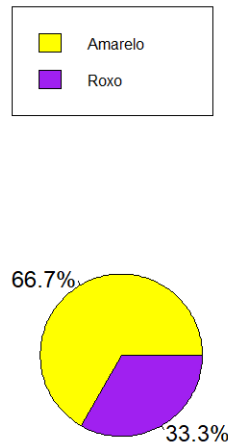


Figura 5: Gráfico circular para a variável cor

Com a análise deste gráfico podemos concluir que a moda é a cor amarela.

3.1.6 Gráfico circular para a variável espécie (Univariadas)

Gráfico circular para a variável 'Espécie'

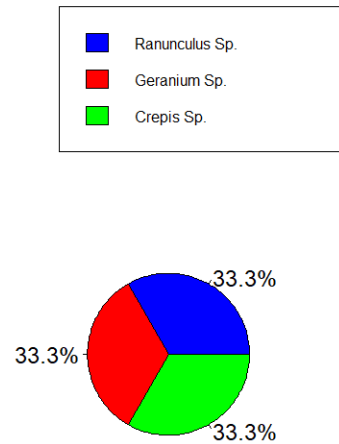


Figura 6: Gráfico circular para a variável espécie

De acordo com a análise do gráfico concluímos que a variável espécie não tem moda pois tem a mesma quantidade de cada espécie.

3.1.7 Gráfico circular para a variável Humidade (Univariadas)

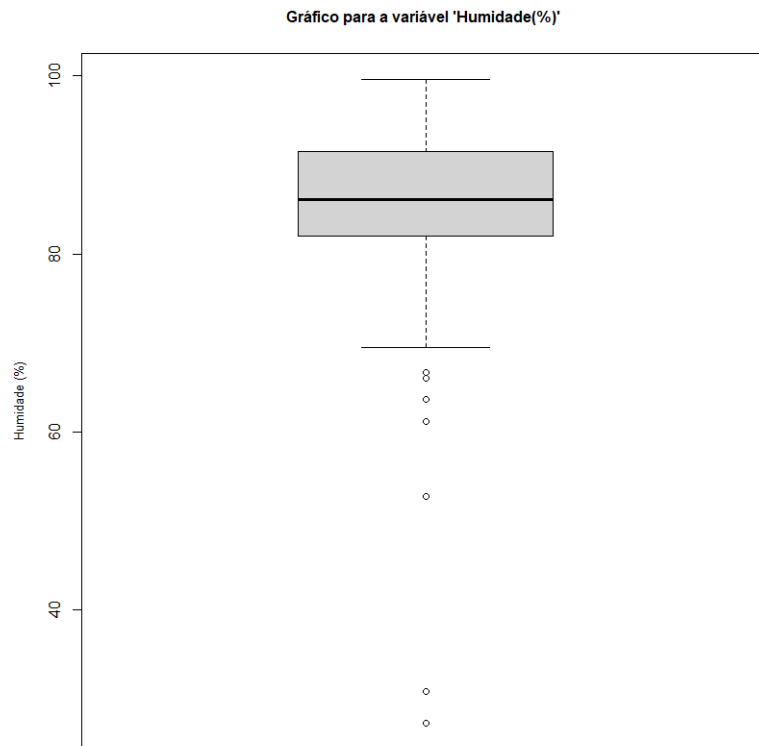


Figura 7: Gráfico circular para a variável Humidade

Neste gráfico podemos observar um valor mínimo de 27.27 e máximo de 99.6, primeiro quartil de 82.01 e terceiro quartil de 91.45, mediana de 86.11, média de 86.01, variância de 108.18 e desvio padrão de 10.40

3.1.8 Gráfico para a variável comprimento de petala (Univariadas)

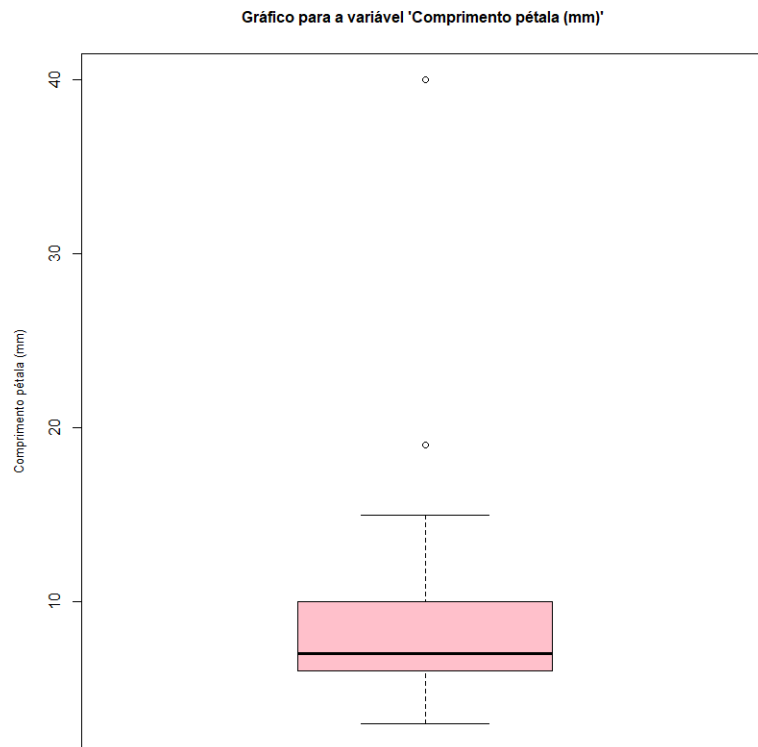


Figura 8: Gráfico circular para a variável Comprimento

Neste gráfico podemos observar um valor mínimo de 3.0 e máximo de 40.0, primeiro quartil de 6.0 e terceiro quartil de 10.0, mediana de 7.0, média de 8.13, variância de 14.5 e desvio padrão de 3.81

3.1.9 Gráfico para a variável largura da petala (Univariadas)

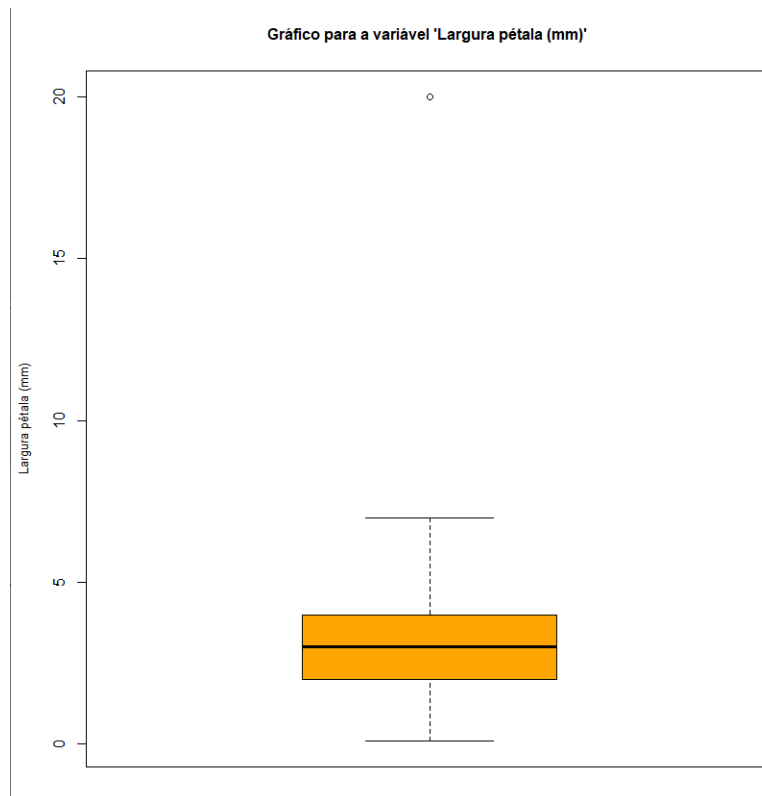


Figura 9: Gráfico circular para a variável Largura

Neste gráfico podemos observar um valor mínimo de 0.1 e máximo de 20.0, primeiro quartil de 2.0 e terceiro quartil de 4.0, mediana de 3.0, média de 2.96, variância de 4.09 e desvio padrão de 2.02