

Problema: destino de animais em abrigos

Aproximadamente 7,6 milhões de animais são recolhidos em abrigos. Muitos desses animais tem a sorte de voltar aos seus donos ou ainda serem adotados, mas não todos contam com essa sorte. Muitos animais esperam a vida inteira por uma família que possa adotá-lo. 2.7 milhões de cachorros e gatos sofrem eutanásia todo ano nos Estados Unidos. O Centro Animal de Austin divulgou dados sobre mais de 25.000 animais para a comunidade de Data Mining auxiliá-los a prever qual será o destino dos animais.

Para este exercício foram selecionados 30 registros: 10 de animais que foram adotados, 10 animais que sofreram eutanásia e 10 animais que retornaram para os donos. O conjunto de dados é apresentado na página 2. Desenvolva as seguintes atividades para fazer com que conjunto de dados fique mais apropriado para o processo de mineração:

1. Apresente a frequência e percentual dos possíveis valores dos atributos **AnimalType**, **SexuponOutcome** e **AgeuponOutcome**.

Atributo **AnimalType**

Valor	Cat	Dog
Frequência	7	23
%	23,33	76,67

Atributo **SexuponOutcome**

Valor	Intact Female	Intact Male	Neutered Male	Spayed Female	Unknown
Frequência	1	2	15	11	1
%	3,33	6,67	50	36,67	3,33

Atributo **AgeuponOutcome**

Valor	Frequência	%
1 year	7	23,33
10 months	1	3,33
11 years	1	3,33
12 years	2	6,67
2 months	1	3,33
2 weeks	1	3,33
2 years	5	16,67
3 months	1	3,33

Valor	Frequência	%
3 years	3	10
4 years	1	3,33
5 months	1	3,33
5 years	2	6,67
6 months	1	3,33
6 years	1	3,33
7 years	1	3,33
9 years	1	3,33

2. O atributo **AgeuponOutcome** descreve quantidades, mas é apresentado de forma nominal, misturando anos, meses e semanas. Transforme o atributo **AgeuponOutcome** para número de dias **DaysuponOutcome** (considere o ano com 365 dias e o mês com 30 dias). Apresente a distribuição de frequência do novo atributo em ordem crescente de valor.

Valor	14	60	90	150	180	300	365	730	1095	1460	1825	2190	2555	3285	4015	4380
Frequência	1	1	1	1	1	1	7	5	3	1	2	1	1	1	1	2
%	3,33	3,33	3,33	3,33	3,33	3,33	23,33	16,67	10	3,33	6,67	3,33	3,33	3,33	3,33	6,67

Transforme o atributo **DaysuponOutcome** em um atributo discreto utilizando as técnicas a seguir e apresente a nova distribuição de frequência:

- a. Intervalos de tamanho pré-definido. Intervalos: [0;3), [3;8), [8;12] anos.
- b. Quatro (4) intervalos de igual tamanho.

	[0;3)	[3;8)	[8;12]
Frequência	18	8	4
%	60	26,67	13,33

	[0;4)	[4;8)	[8;12)	[12;16)
Frequência	21	5	2	2
%	70	16,67	6,67	6,67

- c. Cinco (5) intervalos com o mesmo número de elementos.

	[0;1)	[1;2)	[2;3)	[3;6)	[6;13)
Frequência	6	7	5	6	6
%	20	23,33	16,67	20	20

3. Ainda em relação ao atributo **DaysuponOutcome** calcule: Média, Desvio padrão, Mediana, Valor mínimo Valor máximo.

Média	Desvio padrão	Mediana	Valor mínimo	Valor máximo
1206,63	1300,79	730	14	4380

a. Normalize o atributo **DaysuponOutcome** com as técnicas de normalização **linear no intervalo [0,1]** (f_1) e **desvio padrão** (f_2).

Valor	14	60	90	150	180	300	365	730	1095	1460	1825	2190	2555	3285	4015	4380
$f_1(x)$	0,00	0,01	0,02	0,03	0,04	0,07	0,08	0,16	0,25	0,33	0,41	0,50	0,58	0,75	0,92	1,00
$f_2(x)$	-0,92	-0,88	-0,86	-0,81	-0,79	-0,70	-0,65	-0,37	-0,09	0,19	0,48	0,76	1,04	1,60	2,16	2,44

	AnimalType	SexuponOutcome	AgeuponOutcome	OutcomeType
1	Dog	Neutered Male	2 years	Adoption
2	Dog	Spayed Female	5 months	Adoption
3	Dog	Spayed Female	1 year	Adoption
4	Dog	Spayed Female	2 years	Adoption
5	Dog	Neutered Male	4 years	Adoption
6	Cat	Neutered Male	3 months	Adoption
7	Dog	Spayed Female	2 years	Adoption
8	Dog	Neutered Male	10 months	Adoption
9	Dog	Spayed Female	6 months	Adoption
10	Dog	Spayed Female	5 years	Adoption
11	Cat	Spayed Female	1 year	Euthanasia
12	Cat	Spayed Female	3 years	Euthanasia
13	Dog	Intact Male	3 years	Euthanasia
14	Dog	Intact Male	6 years	Euthanasia
15	Dog	Spayed Female	11 years	Euthanasia
16	Cat	Intact Female	1 year	Euthanasia
17	Cat	Unknown	2 weeks	Euthanasia
18	Dog	Neutered Male	12 years	Euthanasia
19	Dog	Spayed Female	2 years	Euthanasia
20	Dog	Neutered Male	1 year	Euthanasia
21	Dog	Neutered Male	1 year	Return_to_owner
22	Dog	Neutered Male	2 years	Return_to_owner
23	Dog	Neutered Male	1 year	Return_to_owner
24	Dog	Neutered Male	1 year	Return_to_owner
25	Dog	Neutered Male	2 months	Return_to_owner
26	Cat	Neutered Male	7 years	Return_to_owner
27	Dog	Neutered Male	5 years	Return_to_owner
28	Dog	Neutered Male	3 years	Return_to_owner

29	Dog	Neutered Male	12 years	Return_to_owner
30	Cat	Spayed Female	9 years	Return_to_owner