Data Dictionary
for Prepared Pets Adoption Dataset

updated 23 Mar 2023

Introduction

This write-up documents the data dictionary of the prepared Pets Adoption Datasets from PetFinder.my. The prepared dataset is extracted and summarized from **kaggle** data source: https://www.kaggle.com/c/petfinder-adoption-prediction/data

Data Dictionary

1. pets_prepared.csv

ID	Column Name	Description
1	PetID	Unique hash ID of pet profile
2	AdoptionSpeed	Categorical speed of adoption. Lower is faster. This is the value to predict. The value is determined by how quickly, if at all, a pet is adopted. The values are determined in the following way: 0 - Pet was adopted on the same day as it was listed. 1 - Pet was adopted between 1 and 7 days (1st week) after being listed. 2 - Pet was adopted between 8 and 30 days (1st month) after being listed. 3 - Pet was adopted between 31 and 90 days (2nd & 3rd month) after being listed. 4 - No adoption after 100 days of being listed. (There are no pets in this dataset that waited between 90 and 100 days).
3	Type	Type of animal (1 = Dog, 2 = Cat)
4	Name	Name of pet (Empty if not named)
5	Age	Age of pet when listed, in months
6	Breed1	Primary breed of pet (Refer to BreedLabels reference "breed_labels.csv")

ID	Column Name	Description
7	Breed2	Secondary breed of pet, if pet is of mixed breed (Refer to BreedLabels reference "breed_labels.csv")
8	Gender	Gender of pet (1 = Male, 2 = Female, 3 = Mixed, if profile represents group of pets)
9	Color1	Color 1 of pet (Refer to ColorLabels reference "color_labels.csv")
10	Color2	Color 2 of pet (Refer to ColorLabels reference "color_labels.csv")
11	Color3	Color 3 of pet (Refer to ColorLabels reference "color_labels.csv")
12	MaturitySize	Size at maturity (1 = Small, 2 = Medium, 3 = Large, 4 = Extra Large, 0 = Not Specified)
13	FurLength	Fur length (1 = Short, 2 = Medium, 3 = Long, 0 = Not Specified)
14	Vaccinated	Pet has been vaccinated (1 = Yes, 2 = No, 3 = Not Sure)
15	Dewormed	Pet has been dewormed (1 = Yes, 2 = No, 3 = Not Sure)
16	Sterilized	Pet has been spayed / neutered (1 = Yes, 2 = No, 3 = Not Sure)
17	Health	Health Condition (1 = Healthy, 2 = Minor Injury, 3 = Serious Injury, 0 = Not Specified)
18	Quantity	Number of pets represented in profile
19	Fee	Adoption fee (0 = Free)
20	State	State location in Malaysia (Refer to StateLabels reference "state_labels.csv")
21	RescuerID	Unique hash ID of rescuer
22	VideoAmt	Total uploaded videos for this pet
23	PhotoAmt	Total uploaded photos for this pet
24	Description	Profile write-up for this pet. The primary language used is English, with some in Malay or Chinese.
25	AgeBins	Age binning ([0, 2), 2, [3, 6), [6, 12), [12, 24), [24, 60), [60,255])
26	FeeBins	Fee binning (0, [1, 100), [100, 200), [200,3000])
27	BreedBins	Breed_labels grouping (307,266,265,299,264,292, Others)
28	StateBins	State_labels grouping (41326,41401,41327,41336,41330,41332, Others)
29	VideoAmtBins	VideoAmt grouping (0, 1, >=2)
30	PhotoAmtBins	PhotoAmt grouping (0, 1, 2, 3, 4,5,>=6)

ID	Column Name	Description
31	QuantityBins	Quantity grouping (1, 2, 3, 4,5,>=6)
32	BreedPure	Derived variable from Breed1 / Breed2 The pet is purebred or mixed breed (1/0)
33	ColorAmt	Derived variable from Color1/ Color1 /Color1 The amount of color (1,2,3)
34	NameorNO	Derived variable from Name The pet has name or not (1/0)
35	Adopted	Derived variable from Adoptionspeed The pet can be adopted in 100 day or not (1/0)
36	TypeName	The name of Type (1 = Dog, 2 = Cat)
37	GenderName	The name of Gender (1 = Male, 2 = Female, 3 = Mixed, if profile represents group of pets)
38	MaturitySizeName	The name of MaturitySize (1 = Small, 2 = Medium, 3 = Large, 4 = Extra Large, 0 = Not Specified)
39	FurLengthName	The name of FurLength (1 = Short, 2 = Medium, 3 = Long, 0 = Not Specified)
40	HealthName	The name of Health (1 = Healthy, 2 = Minor Injury, 3 = Serious Injury, 0 = Not Specified)
41	VaccinatedName	The name of Vaccinated (1 = Yes, 2 = No, 3 = Not Sure)
42	DewormedName	The name of Dewormed (1 = Yes, 2 = No, 3 = Not Sure)
43	SterilizedName	The name of Sterilized (1 = Yes, 2 = No, 3 = Not Sure)
44	Breedname	The name of Breed1 (Refer to BreedLabels reference "breed_labels.csv")
45	BreedBinname	The name of BreedBin (Refer to BreedLabels reference "breed_labels.csv")) (Group less frequent classes into "Others")
46	StateName	The name of State (Refer to StateLabels reference "state_labels.csv")
47	StateBinName	The name of StateBin (Refer to StateLabels reference "state_labels.csv") (Group less frequent classes into "Others")
48	ColorName	The name of Color (Refer to ColorLabels reference "color_labels.csv")

ID	Column Name	Description
49	AdoptedName	The name of Adopted (Y/N)

2. breed_labels.csv

ID	Column Name	Description
1	BreedID	Unique ID of pet breed (1-306)
2	Type	Type of animal (1 = Dog, 2 = Cat)
3	BreedName	Name of pet Breed

3. color_labels.csv

ID	Column Name	Description
1	ColorID	Unique ID of pet color (1-7)
2	ColorName	Name of pet color

4. state_labels.csv

ID	Column Name	Description
1	StateID	Unique ID of State in Malaysia
2	StateName	Name of state in Malaysia

@end