Andrew Gatchalian, Davidson Rajasekar, Devin Xiang Tian, Kaidi Huang, Yuanhui Yao

BA 288 – Team 2 Data Dictionary

Final Project Data Dictionary #1

All deliverables can be found here:
https://drive.google.com/drive/folders/1Hk-9iLcJxFPdy31-NaT666KhD0rwRMDR?usp=sharing

Our project aims to leverage historical data from an animal shelter to develop a predictive model for animal outcomes. We will be utilizing data from the Austin Animal Center Shelter Outcomes dataset, as found on Kaggle. To assist with our analysis, we have **cleaned** the data for modeling and interpretability purposes. We dropped several redundant variables and opted to keep only a select few that we saw fit for our analysis. Additionally we feature engineered several variables (such as converting days of the week to a weekend binary and converting outcome hours to a categorical variable that describes the shelter's morning, afternoon, and closed hours). We have produced three CSV files for modeling and forecasting, and a separate file for EDA. The data dictionary describes the modeling CSV files:

cat_data_cleaned_updated.csv & **cat_data_cleaned_updated_continous_age.csv**

Variable	Variable Name	Data Type	Measurement Unit	Allowed Values	Description
Animal ID	animal_id	Qualitative	string	Alphanumeric characters	Unique ID to identify each cat
Outcome Age (years)	outcome_age_(years)	Quantitative	numeric	0 to 22	Age (in years of the cat upon which the outcome occurred) * NOTE this variable is only included in cat_data_cleaned_updated_continous_age.csv
Has Name	has_name	Qualitative	binary	0 - no, 1 - yes	Whether the cat has a name or not
Adoption Status ***(Dependent variable)***	is_adopted	Qualitative	binary	0 - not adopted, 1 - adopted	Whether the cat was adopted or not. Adoption means the animal was adopted to a home. Not adopted imply the cat was transferred, euthanized, or other.
Sex	sex_male	Qualitative	binary	0 - female, 1 - male	Sex of the cat

Spay/Neuter	spay_neuter	Qualitative	binary	0 - no, 1 - yes	If a cat has been spayed/neutered
CFA Approved	cfa_approved	Qualitative	binary	0 - False, 1 - True	Designates if this is a CFA approved breed
Cat/Kitten Status	is_kitten	Qualitative	binary	0 - kitten, 1 - cat	If it is a cat or kitten (cats of around 6 months of less)
Season	season_Fall season_Spring season_Summer season_Winter	Qualitative	categorical	(Fall, Spring, Summer, Winter) 0 - outcome did not occur during the corresponding season, 1 - outcome did occur	The season in which the outcome occurred
Weekend	is_weekend	Qualitative	binary	0 - no, 1 - yes	Whether the outcome occurred during the weekend (Saturday, Sunday) or not
Time of Day	time_of_day_afternoon time_of_day_closed time_of_day_morning	Qualitative	categorical	(Morning, afternoon, closed) 0 - the outcome did not occur during the corresponding time, 1 - outcome did occur	Time of day the outcome occurred. Time is split into 3 categories that reflect the actual operating hours of the Austin Animal Shelter. Morning: 11am-3pm, Afternoon: 3pm-7pm, Closed: all other hours.
Shorthair	is_shorthair	Qualitative	binary	0 - no, 1 - yes	Refers to the breed of the cat, whether it is a shorthair breed or not
Solid Pattern	is_solid_pattern	Qualitative	binary	0 - no, 1 - yes	Refers to the coat pattern of the cat, whether it is a solid coat pattern or not
Color	color_black color_white color_other	Qualitative	categorical	(Black, White, Other) 0 - the cat is not the corresponding color, 1 - the cat is the color	Refers to the color of the cat. Color is separated into 3 categories (Black, White, and Other)

Final Project Data Dictionary #2

Additionally, we have created a separate dataset for time series forecasting which captures the total number of adopted cats for each given month in our dataset:

cat_data_cleaned_updated_forecast.csv

Variable	Variable Name	Data Type	Measurement Unit	Allowed Values	Description
Date (month-year)	Date	Qualitative	string	MM/1/YYYY	Month
Number of cats adopted	number_of_adopted	Quantitative	numeric	-	Number of cats adopted at the shelter during the given month