PENNSTATE

Tel: (814) 863-0281
Fax: (814) 863-7114

Statistical Consulting Center
Department of Statistics
Eberly College of Science

The Pennsylvania State University
323 Thomas Building
University Park, Pa 16802

Date 2/25/2019

From: *Jonathan Fivelsdal (jxf404@psu.edu),* Graduate Student Consultant

To: *Louisville Metro Animal Services*

Re: *Analysis of LMAS Shelter Data*

RESEARCH FIELD: Animal Welfare and Behavior

PROJECT TITLE: Research on the Adoption and Shelter Characteristics of Domestic Shorthair and Domestic Mediumhair Cats from the Louisville Metro Services Shelter

1.0 - PROJECT DESCRIPTION

Many people across the U.S. have pet cats. According to the American Pet Products Association (APPA), in 2016 there were about 94.2 million pet cats in the U.S.[1]. The ASPCA says that out of about 6.5 million companion animals that enter shelters in the U.S. every year, 3.2 million are cats[2]. Two common types of cats are the domestic shorthair cats (DSH) and domestic mediumhair cats (DMH). The Pet Premium website says that DSH cats possess a medium amount of kid friendliness, affection level and activity level[3]. In contrast, Pet Premium says that DMH cats tend to have low amounts of kid friendliness, affection level and activity level[4]. Both DSH cats and DMH cats have no specific breeding pattern and so the personality traits vary widely among DSH cats and personality traits also vary widely among DMH cats. Louisville

[1] Springer, J. (2017). The 2017-2018 APPA National Pet Owners Survey Debut Trusted Data for Smart Business Decisions. https://americanpetproducts.org/Uploads/MemServices/GPE2017_NPOS_Seminar.pdf
[2] "Pet Statistics." ASPCA, The American Society for the Prevention of Cruelty to Animals, Accessed 11 January 2019, https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics
[3] "Domestic Shorthair." PetPremium, PetPremium, Inc., Accessed 12 January 2019, https://www.petpremium.com/cat-breeds/domestic-shorthair/
[4] "Domestic Mediumhair." PetPremium, PetPremium, Inc., Accessed 12 January 2019, https://www.petpremium.com/cat-breeds/domestic-medium-hair/

Metro Animal Services (LMAS) provides data on the animals in their shelter on the Louisville Open Data website. Using data from LMAS, this report will consider factors related to the adoption of DSH and DMH cats. Data from 2017 on DSH and DMH cats will be used and the data can be found at this link https://data.louisvilleky.gov/dataset/animal-service-intake-and-outcome. The analysis was conducted in R[5]. Understanding how different factors are related to adopt could help a shelter to implement programs that could potentially boost adoption.

1.1 – RESEARCH AND STATISTICAL QUESTIONS

Initial analysis will address the questions 1.) How does the adoption of stray cats differ from that of non-stray cats? and 2.) How does the length of time spent in a shelter (length of stay) and age affect adoption? The remaining questions will be considered which are 3.) Is the proportion of DSH cats that are adopted from a shelter different from the proportion of DMH cats that are adopted from a shelter? and 4.) What factors are significant with respect to DSH and DMH cats being adopted?

1.2 - VARIABLES OF INTEREST

The analysis is based on data from the Louisville Metro Animal Services (LMAS) which was accessed from the Louisville Open Data website on January 10th 2019. The webpage where the data was downloaded is https://data.louisvilleky.gov/dataset/animal-service-intake-and-outcome . Data on DSH cats and DMH cats that were admitted to the shelter in 2017 was analyzed. Cats that didn't have an intake date, outcome date or had an outcome date that was recorded as being earlier than the intake date were excluded from the analysis. Cats over the age of 25 years old were excluded in the analysis (only one cat was over 25 years old). The number of cats that are included in the dataset

[5] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

being analyzed is 3275. Out of the 3275 cats considered, 2,945 cats are DSH cats and the remaining 330 cats are DMH cats. Analysis of shelter data can identify what factors affect adoption. One study suggests that "understanding factors that affect an adopter's choice of a cat is not only important for shelter management strategies but can also benefit the welfare of shelter cats, yet, few studies have researched this topic"[6]. Earlier research done on the adoption of cats from animal shelters provides guidance as to some of the factors that affect adoption.

The case study of the Guelph Humane Society noted that previous research suggests that age was the most important factor that adopters considered when choosing to adopt a cat[7]. The original data set for LMAS contains a birthdate and the intake date (the date when an animal enters the shelter) and from those variables an age variable is created which indicates the age of the cat at intake. After creating the age variable, it was condensed into three categorical levels of kitten, young adult and older adult and this variable is used in the analysis of the LMAS data. For the categorical age variable, cats that are less than one year old are considered to be kittens, cats that are at least one year old but less than seven years old are classified as young adult and cats that are seven years of age or older are classified as older adult[8]. In the LMAS data the intake type variable corresponds to the source/entry type variable referred to in the Guelph Humane Society case study. Research indicates that potential adopters tend to have preconceived notions about stray cats that negatively affects their chances of being adopted[9]. A new variable called stray condenses the levels of the

---

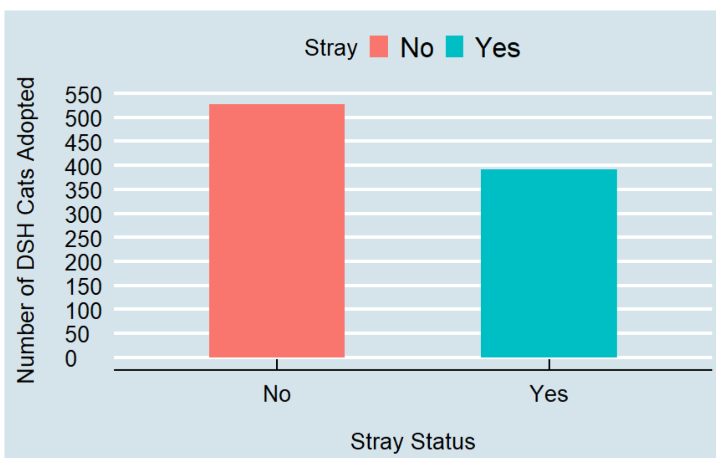
[6] Janke et.al. Risk factors affecting length of stay of cats in an animal shelter: A case study at the Guelph Humane Society, 2011-2016. Prev Vet Med. 2017 Dec 1;148: 44-48. doi:10. 1016/j.prevetmed.2017.10.007. Epub 2017 Oct 16. PubMed PMID: 29157373.
[7] Ibid.
[8] Downing, Robin. "Feeding Your Young Adult Cat." *VCA Hospitals*, Accessed 15 January 2019, vcahospitals.com/know-your-pet/feeding-your-young-adult-cat.
[9] Dybdall, Kathryn & Strasser, Rosemary. (2014). Is There a Bias Against Stray Cats in Shelters? People's Perception of Shelter Cats and How It Influences Adoption Time. Anthrozoos A Multidisciplinary Journal of The Interactions of People & Animals. 28. 603-614. 10.2752/089279314X14072268688087.

intake type variable down to 2 levels which are stray and non-stray. The length of stay (LOS) of a shelter cat is also of interest. For this report, LOS was calculated by taking the intake date and subtracting it from the outcome date and rounding the result to the nearest day. A study was conducted in which the average time for owner surrendered cats and stray cats to develop an illness after entering a shelter was measured to be 11 and 14 days respectively[10]. A variable called weeks which represents whether or not a cat stays in an animal shelter for more than 2 weeks and this will be the LOS variable used in the data analysis.

2.0 - EXPLORATORY DATA ANALYSIS FOR DOMESTIC SHORTHAIR CATS (DSH)

The next two sections will present exploratory data analysis (EDA) for DSH and DMH cats. As mentioned earlier, whether a cat is a stray or not a stray cat can affect whether or not a cat is adopted from a shelter. The following graph provides information on the number of DSH cats that were adopted by stray group:

Figure 2.1: Number of DSH Cats Adopted by Each Stray Category



[10] Dybdall, Kathryn & Strasser, Rosemary & Katz, Tanja. (2006). Behavioral differences between owner surrender and stray domestic cats after entering an animal shelter. Applied Animal Behaviour Science. 104. 10.1016/j.applanim.2006.05.002.

Figure 2.1 indicates that about 525 non-stray DSH cats and about 400 young adult stray DSH cats were adopted. This plot suggests that is a negative relationship between stray status and adoption. Now the other variables of interest which are age and weeks will be considered. The plot provided on the next page provides information on the number of domestic shorthair cats that were adopted by age group:

Figure 2.2: Number of DSH Cats Adopted by Each Age Category

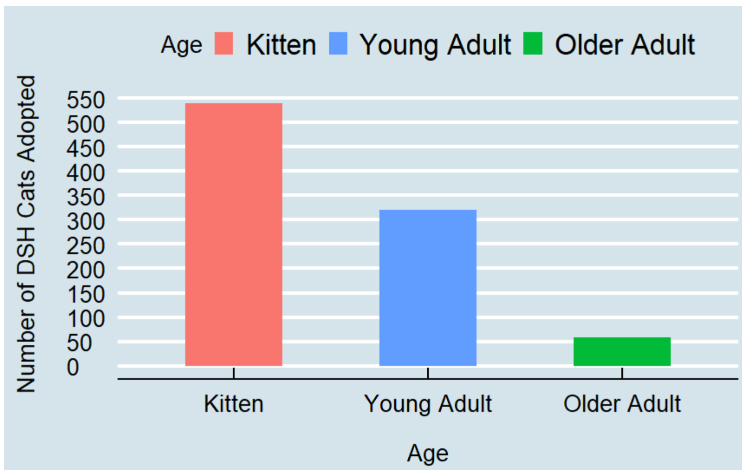


Figure 2.2 shows that a little less than 550 kittens, about 325 young adult cats and a little more than 50 older adult cats were adopted among the DSH cats in the shelter. This plot suggests that is a negative relationship between age and adoption. The next plot considers the number of DSH cats that were adopted by week category (DSH cats that stayed in the shelter for two weeks or less and those that stayed in the shelter for more than two weeks).

Figure 2.3: Number of DSH Cats Adopted by Length of Stay Category (Weeks)

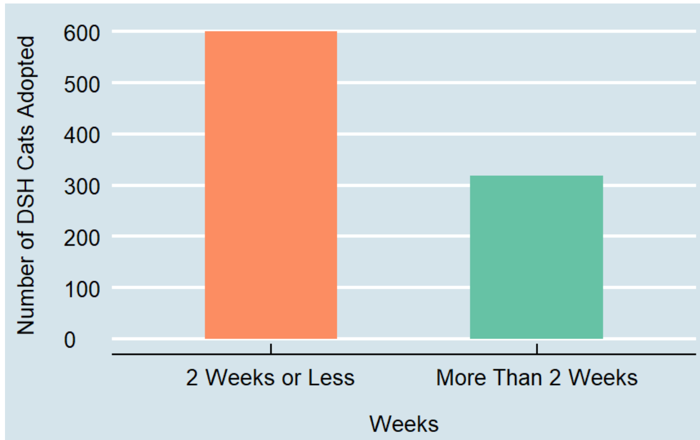


Figure 2.3 shows that 600 DSH cats that spent two weeks or less in the shelter were adopted while a little more than 300 DSH cats that spent more than 2 weeks in the shelter were adopted. Therefore, it appears that spending more than 2 weeks in a shelter has a negative impact on adoption.

3.0 - EXPLORATORY DATA ANALYSIS FOR DOMESTIC MEDIUMHAIR CATS (DMH)

Just as was done in the exploratory data analysis section for DSH cats, we will first visualize the relationship between stray status and adoption for DMH cats. The following is a fourfold plot which visualizes the number and proportion of stray and non-stray cats that have been adopted and stray and non-stray cats that have not been adopted:

Figure 3.1: Number of DMH Cats Adopted by Stray Category

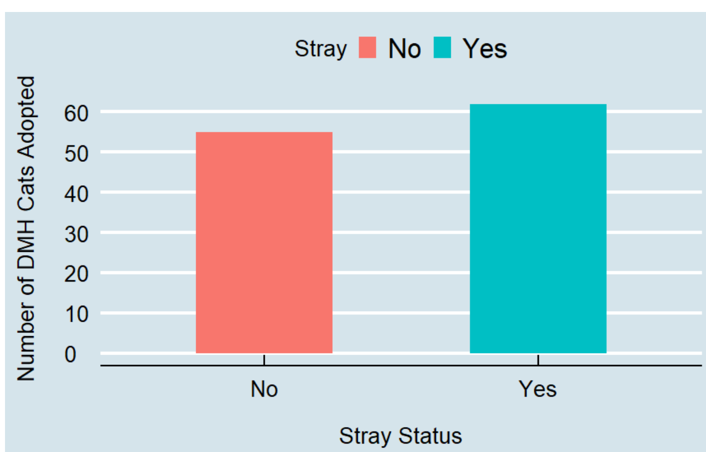


Figure 3.1 indicates that 55 non-stray DMH cats and a little more than 60 young adult stray DMH cats were adopted. The plot suggests that there might be a positive relationship between stray status and adoption, but the number of adopted stray and non-stray DMH cats are almost equal. When statistical models are considered later, the direction of the relationship between the stray status of DMH cats and the adoption of DMH cats may change when more factors such as the age of the cat and how long the cat has stayed in a shelter (length of stay) are taken into consideration.

Now we will investigate the relationship between the adoption of DMH cats and age. The following display provides information on the number of domestic shorthair cats that were adopted by age group:

Figure 3.2: Number of DMH Cats Adopted by Age Category

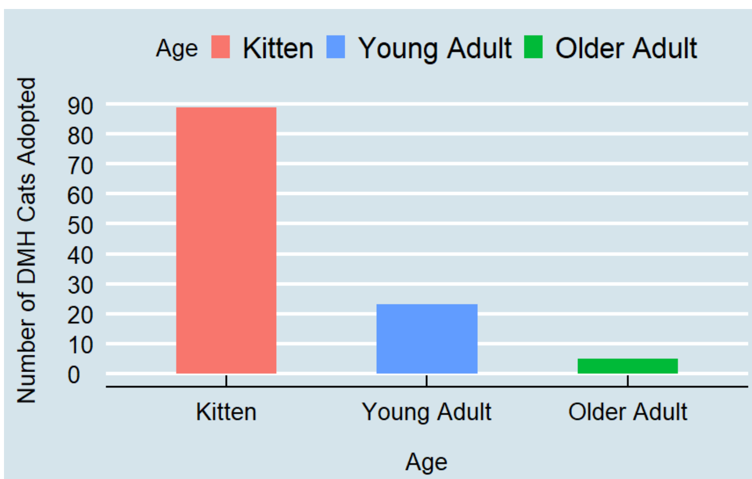


Figure 3.2 shows that about 90 kittens, a little more than 20 young adult cats and about 5 older adult cats were adopted among the DMH cats in the shelter. This plot suggests that there is a negative relationship between age and adoption. The next plot considers the number of domestic mediumhair cats that were adopted by week category (DMH cats that stayed in the shelter for two weeks or less and those that stayed in the shelter for more than two weeks).

Figure 3.3: Number of DMH Cats Adopted by Length of Stay Category (Weeks)

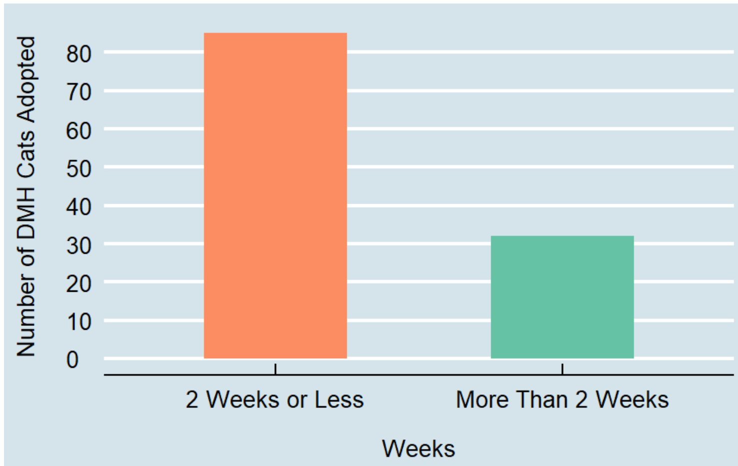


Figure 3.3 shows that more than 80 DMH cats (85 DMH cats) that spent two weeks or less in the shelter were adopted while only a little more than 30 DMH cats that spent more than 2 weeks in the shelter were adopted. This is similar to the case with DSH cats in which more cats that spent 2 weeks or less were adopted than cats that spent more than 2 weeks in the shelter and thus it appears that for DMH cats that spending more than 2 weeks in the shelter has a negative effect on adoption.

4.0 –STATISTICAL ANALYSIS

From the 2945 domestic shorthair cats admitted to the Louisville Metro Animal Services (LMAS) shelter, 919 of these cats were adopted. Out of the 330 domestic mediumhair cats that were admitted to LMAS in 2017, 117 of these cats were adopted. Below is a table that represents the aforementioned information along with estimated proportions for cats adopted by each cat type:

Table 4.1: Number and Proportion of DSH and DMH Cats Adopted

	Domestic Shorthair Cat	Domestic Mediumhair Cat
Sample Size	2945	330
Number Adopted	919	117
Estimated Proportion	919/2945 = 0.3121	117/330 = 0.3545

The following table provides the percentage of cats that were adopted and those that were not adopted for each cat type (domestic shorthair cats and domestic mediumhair cats):

Table 4.2: Percentage of DSH and DMH Cats Adopted and Not Adopted

	Adopted	Not Adopted
Domestic Shorthair	31.21%	68.79%
Domestic Mediumhair	35.45%	64.55%

Representing the adoption percentages as proportions, we have that 0.3121 of the DSH were adopted and 0.3545 DMH cats were adopted. From the sample of cats under study, the estimated percentages (and equivalently the proportions) of cats adopted appear to be similar, but we will use a formal statistical test to determine if there is a significant difference between the proportion of DSH cats that are adopted and the proportion of DMH cats among the population of domestic shorthair and domestic mediumhair cats that enter the shelter. A difference of proportions test was performed and the p-value resulting from the test is 0.116. A p-value that is less than 0.05 indicates a significant difference of proportions and since in this case the p-value resulting from the test is greater than 0.05, it is determined that there is no significant difference in the proportion between the proportion of DSH cats that are adopted versus the proportion of DMH cats that are adopted. The difference in proportion test also produces a confidence interval which provides information on the likely range in which the proportions differ. The confidence interval produced by the difference in proportions test is (-0.097,0.012). Since the confidence interval for the differences in proportions contains zero, this indicates that there is not a significant difference between the proportion of DSH cats that are adopted and the proportion of DMH cats that are adopted from the shelter.

Logistic regression will be used to investigate factors that affect adoption. Logistic regression is used in cases where the outcome is a binary or yes/no type of response. Adoption is a yes/no type of outcome (was a cat adopted or not from a shelter?) and so logistic regression is an appropriate method to consider when examining what factors affect whether or not a cat is adopted. The article "Understanding logistic regression analysis in clinical reports: An introduction" provides a detailed example demonstrating how to interpret the results of a logistic regression[11]. A logistic regression model will be used to examine factors affecting adoption for DSH and DMH cats. Details related to how the final models were chosen can be found in the appendix. The following are the results of the logistic regression model for DSH cats:

Table 4.3: Logistic Regression Model for Domestic Shorthair Cats

	Results
Age (Age = Older Adult)	-1.699***
	(-2.587, -0.812)
Age (Age = Young Adult)	-0.924***
	(-1.370, -0.478)
Stray (Stray = Yes)	-2.331***
	(-2.760, -1.902)
Weeks (Weeks = More Than 2 Weeks)	-3.038***
	(-3.633, -2.443)
Constant	3.744***
	(3.030, 4.458)
Notes:	***Significant at the 1 percent level.
	**Significant at the 5 percent level.
	*Significant at the 10 percent level.

The first column of Table 4.3 provides the names of the factors in the logistic model, second column of the table contains the coefficient estimates for each factor in the logistic regression and underneath the estimates contains a range in parentheses which provides a 95% confidence interval for the coefficient estimates. From the logistic regression results table, notice that all of the factors of interest in the model are statistically significant at the 1% level (all factors are also significant at a 5% level of significance which is the typical cutoff for determining statistical significance). The

[11] P Anderson, Richard & Jin, Ruyun & Grunkemeier, Gary. (2003). Understanding logistic regression analysis in clinical reports: An introduction. The Annals of thoracic surgery. 75. 753-7. 10.1016/S0003-4975(02)04683-0.

coefficients in logistic regression represent the logarithm of the odds and these coefficients will be mathematically transformed (the natural base e will be raised to the value of the relevant logistic model coefficient in order to obtain the odds ratio) and conclusions will be made in terms of the reciprocal of the odds ratios based on the model coefficients. The table on the following page presents the reciprocal of the odds ratios for factors in the DSH cat regression model:

Table 4.4: Reciprocal of Odds Ratio for DSH Logistic Regression Model Factors

	OR.T	Lower	Upper
Kitten vs. Older Adult	5.47	2.32	13.82
Kitten vs. Younger Adult	2.52	1.62	3.97
Non-Stray vs. Stray	10.29	6.75	15.93
At Most 2 Weeks vs. More Than 2 Weeks	20.87	11.71	38.61

The column labeled "OR.T" presents the reciprocal of the estimated odds ratios that would result from transforming the coefficients of the regression model. The columns labeled "Lower" and "Upper" present the lower and upper bounds of the 95% confidence interval for the odds ratios presented in the table. Note that an odds ratio greater than 1 indicates a positive relationship, an odds ratio less than 1 indicates a negative relationship and an odds ratio of 1 indicates that there is no relationship between levels of a particular factor (statistical independence). Table 4.4 indicates that we are 95% confident that the odds of a DSH kitten being adopted are between about 2.3 to 13.8 times higher than that of an older adult DSH cat, the odds of a DSH kitten being adopted are between about 1.6 to 4 times higher than that of a young adult DSH cat, the odds of a non-stray DSH cat are between about 6.8 to 16 times higher than that of a stray DSH cat and the odds for a DSH cat being adopted that stays 2 weeks or less in a shelter are between about 11.7 to 38.6 times higher than that of a DSH cat that stays more than 2 weeks in a shelter. In summary, note that for DSH cats, kittens have better odds of being adopted than adults (both young adults and older adults), non-stray cats have better odds of being adopted than stray cats and cats that have stayed in the shelter for two weeks or less have better odds of being adopted than cats that have stayed in the shelter for more than 2 weeks. The same methods that were used to analyze the factors related to the adoption of DSH cats will now be applied to analyzing the factors related to the adoption of DMH cats. The following are the results of the logistic regression model for DMH cats:

Table 4.5: Logistic Regression Model for Domestic Mediumhair Cats

	Results
Age (Age = Older Adult)	-2.115^{***}
	(-3.317, -0.913)
Age (Age = Young Adult)	-1.326^{***}
	(-1.977, -0.676)
Stray (Stray = Yes)	-1.823^{***}
	(-2.410, -1.236)
Weeks (Weeks = More Than 2 Weeks)	-2.640^{***}
	(-3.475, -1.804)
Constant	3.441^{***}
	(2.447, 4.435)
Notes:	***Significant at the 1 percent level.
	**Significant at the 5 percent level.
	*Significant at the 10 percent level.

All of the terms in the regression model are significant at the 5 percent level. As was done with the model for DSH cats, the results of the regression model will be interpreted in terms of the reciprocal of the odds ratios. The following table presents the reciprocal of the odds ratios for factors in the DMH cat regression model and the corresponding 95% confidence interval:

Table 4.6: Reciprocal of Odds Ratio for DMH Logistic Regression Model Factors

	OR.T	Lower	Upper
Kitten vs. Older Adult	8.29	2.69	30.79
Kitten vs. Younger Adult	3.77	2.01	7.41
Non-Stray vs. Stray	6.19	3.48	11.27
At Most 2 Weeks vs. More Than 2 Weeks	14.01	6.29	33.75

Table 4.6 indicates that we are 95% confident that the odds of a DMH kitten being adopted are between about 2.7 to 30.8 times higher than that of an older adult DMH cat, the odds of a DMH

kitten being adopted are between about 2 to 7.4 times higher than that of a young adult DMH cat, the odds of a non-stray DMH cat are between about 3.5 to 11.3 times higher than that of a stray DMH cat and the odds for a DMH cat being adopted that stays 2 weeks or less in a shelter are between about 6.3 to 33.8 times higher than that of a DMH cat that stays more than 2 weeks in a shelter. In summary, the relationships that were obtained for DSH cats are the same as those for DMH cats in that kittens have better odds of being adopted than adults (both young adults and older adults), non-stray cats have better odds of being adopted than stray cats and cats that have stayed in the shelter for two weeks or less have better odds of being adopted than cats that have stayed in the shelter for more than 2 weeks.

5.0 – RECOMMENDATIONS

There is a tendency for older cats to be adopted less than younger cats. As was explained in the analysis section, kittens have greater odds of being adopted than young adult cats and kittens also have greater odds of being adopted than older adult cats. A study that analyzed the adoption of dogs and cats from a California shelter suggested that shelter staff could "discuss with people the advantages of adopting a mature dog rather than a puppy" in order to increase adoption among older dogs[12]. Similarly, shelter staff could speak to potential adopters about positive aspects of adopting an older cat or young adult cat which could potentially increase the adoption of adult cats.

In addition to older cats, stray cats also have difficulty being adopted. As was demonstrated in the analysis section stray cats have lower odds of being adopted versus non-stray cats. A study by Dybdall and Strasser suggests that animal shelters could minimize bias against stray cats among potential adopters by "working with cats to increase desired social behaviors or add toys in the

[12] Lepper, Merry & Kass, Philip & Hart, Lynette. (2002). Prediction of Adoption Versus Euthanasia Among Dogs and Cats in a California Animal Shelter. Journal of applied animal welfare science: JAAWS. 5. 29-42. 10.1207/S15327604JAWS0501_3.

kennels of stray cats to improve adopter perception"[13]. Cats that stay in a shelter for extended periods of time face difficulty in terms of adoption as well. From the analysis section, it was established that cats that stayed longer than 2 weeks had lower odds of being adopted versus cats that stayed only 2 weeks or less in the shelter. Different means of promotion such as highlighting an animal with colorful collars and tags, moving animals to an offsite adoption location, displaying overlooked cats in the shelter lobby, updating photos and descriptions of the cats seen online and promotions done on traditional media and social media are ways that could potentially increase the adoption of cats that have stayed in a shelter for an extended period of time[14]. Implementation of these strategies can potentially reduce euthanasia and increase the adoption of cats that may typically be overlooked by potential adopters.

6.0 – CONSIDERATIONS

A possible limitation of this study involves the precision of the odds ratios derived from the logistic regression models. In Table 4.4, the 95% confidence interval for the odds ratio for kitten vs. older adult ranges from 2.32 to 13.82 and the interval corresponding to the odds ratio comparing the categories of the weeks variable ranges from 11.71 to 38.61. Similarly, in Table 4.6, the 95% confidence interval for the odds ratio for kitten vs. older adult ranges from 2.69 to 30.79 and the interval corresponding to the odds ratio comparing the categories of the weeks variable ranges from 6.29 to 33.75. These wide confidence intervals could be caused by a limited sample size and possibly by factors that significantly affect adoption that were not included in the respective regression models. Although there is a lack of precision among some of the odds ratios, the logistic

---

[13] Dybdall, Kathryn & Strasser, Rosemary. (2014). Is There a Bias Against Stray Cats in Shelters? People's Perception of Shelter Cats and How It Influences Adoption Time.

[14] Reider, Linda. "Overlooked No More." *Animal Sheltering Online by The Humane Society of the United States*, 2 Sept. 2016, www.animalsheltering.org/magazine/articles/overlooked-no-more.

regression models provide insights as to how factors such as age, stray status and length of stay impact the adoption of DSH and DMH cats.

Thank you for providing me with the opportunity to work with you on this project.

APPENDIX A: Assessing Quality of the DSH Logistic Regression Model

The following table presents the output for the model that describes factors that affect the adoption of DSH cats before adjusting for overdispersion:

Table A.1: DSH Logistic Model Before Adjusting for Overdispersion

```
Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)        3.7440     0.1779  21.051  < 2e-16 ***
AgeOlder Adult    -1.6993     0.2212  -7.683 1.55e-14 ***
AgeYoung Adult    -0.9238     0.1111  -8.315  < 2e-16 ***
StrayYes          -2.3313     0.1069 -21.800  < 2e-16 ***
Week              -3.0381     0.1482 -20.495  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1168.788  on 23  degrees of freedom
Residual deviance:   52.737  on 19  degrees of freedom
AIC: 153.08
```

With the output from R presented in Table A.1, we can assess issues pertaining to how well the regression model fits the data. For a logistic regression model that fits well, the residual deviance should be approximately equal to the residual degrees of freedom (or equivalently, the result of dividing the residual deviance and the residual degrees of freedom should be approximately equal to 1). The residual deviance is about 52.7 and the residual degrees of freedom is 19 which when the residual deviance is divided by the residual degrees of freedom, the result is about 2.8 which is not close to 1 and suggests the presence of overdispersion in the model. Overdispersion refers to when there is more variability in the data than the regression model assumes. Above the line in Table A.1 which provides the null deviance and null degrees of freedom, it states that the dispersion parameter is taken to be 1. To adjust for overdispersion, the model is fit again but this time the parameter will not be set to 1 (a dispersion parameter of greater than one will be considered in the refit model). The figure on the following page presents the output for the model that describes factors that affect the adoption of DSH cats in which overdispersion is accounted for:

Table A.2: DSH Logistic Model After Adjusting for Overdispersion

```
Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)       3.7440     0.3642  10.280 3.36e-09 ***
AgeOlder Adult   -1.6993     0.4529  -3.752 0.001350 ** 
AgeYoung Adult   -0.9238     0.2275  -4.061 0.000667 ***
StrayYes         -2.3313     0.2190 -10.646 1.90e-09 ***
Week             -3.0381     0.3035 -10.009 5.19e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasibinomial family taken to be 4.19297)
```

Table A.2 shows the results of the regression model after adjusting for overdispersion and notice that the coefficient estimates for each variable in the model are the same as the estimates in table A.1 when the model was fit without adjusting for dispersion. Table A.2 displays that the dispersion parameter for the model is taken to about 4.192 now instead of 1. Although the parameter estimates have not changed in the model adjusting for overdispersion versus the unadjusted model, the standard error and p-values change. For each term in the model, the standard error increases and the p-values increase after adjusting for overdispersion. In some cases, the increase in p-value could result in a factor in the regression model being statistically significant in the model that is not adjusted for overdispersion but then the factor becomes statistically insignificant in the model after adjusting for overdispersion (the p-value increases such that it exceeds the 0.05 level). Table A.2, shows that although the p-values for each term in the model increase over the p-values presented for each corresponding factor in table A.1, all the factors in the model are still highly statistically significant even after adjusting for overdispersion. In addition to the standard error, the residuals from the logistic regression can also be adjusted for in the case of overdispersion. To adjust the residuals, the residuals from the model not adjusted for dispersion are divided by the square root of the dispersion parameter. The next figure on the following page displays the Pearson residuals for the model that adjusts for overdispersion:

Figure A.1: Residuals for the DSH Regression Model After Adjusting for Overdispersion

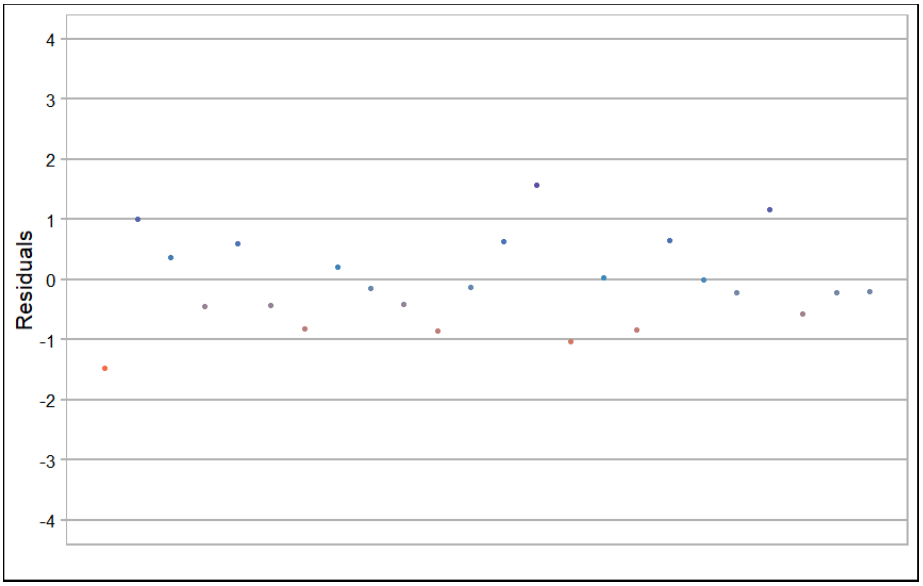


Figure A.1 shows that the residuals are between -2 and 2, form a horizontal band and there is no obvious curvature pattern among the residuals and thus provides evidence for the model possessing a good fit to the data. After adjusting for overdispersion, a chi-square goodness of fit test (GOF) was conducted and the results are presented below:

Figure A.2: Chi-Square GOF Test for DSH Model After Adjusting for Overdispersion

```
X2 = 19, df = 19, P(>X2) = 0.4568361
```

Figure A.2 displays that the observed chi-square statistic for the goodness of fit test has a value of 19 with 19 degrees of freedom and a p-value of about 0.46. The null hypothesis of a GOF test is that the model fits and so a p-value of less than 0.05 suggests that the model suffers from a lack of fit. Since the p-value is greater than 0.05 and this suggests that the model reasonably fits the data. Before adjusting for overdispersion, the chi-square statistic for goodness of fit had a value of 79.7 with 19 degrees of freedom and a p-value of $2.12 * 10^{-9}$ which is less than 0.05 which suggests a model that does not fit the data well. Adjusting for overdispersion helped improve the fit of the model.

APPENDIX B: Assessing Quality of the DMH Logistic Regression Model

The following table presents the output for the model that describes factors that affect the adoption of DMH cats before adjusting for overdispersion:

Table B.1: DMH Logistic Model Before Adjusting for Overdispersion

```
Coefficients:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)       3.4411     0.5072   6.784 1.17e-11 ***
AgeOlder Adult   -2.1154     0.6133  -3.449 0.000562 ***
AgeYoung Adult   -1.3262     0.3319  -3.996 6.44e-05 ***
StrayYes         -1.8229     0.2995  -6.087 1.15e-09 ***
Week             -2.6396     0.4263  -6.191 5.97e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 115.56  on 19  degrees of freedom
Residual deviance:  13.36  on 15  degrees of freedom
AIC: 62.435
```

With the output from R presented in figure B.1, we can assess issues pertaining to how well the regression model fits the data. The residual deviance is 13.36 and the residual degrees of freedom is 15. Dividing the residual deviance by the residual degrees of freedom results in a value of 0.8907 which is close to 1. The residual deviance and residual degrees of freedom are close in value and this provides evidence for a model that fits the data well. Unlike the logistic regression model for DSH cats, it appears unnecessary to adjust the logistic regression for DMH cats for overdispersion. The following are the results of a GOF test that was performed on the DMH logistic regression model:

Figure B.1: Chi-Square GOF Test for DSH Model After Adjusting for Overdispersion

```
X2 = 12.5476, df = 15, P(>X2) = 0.6371981
```

Figure B.1 displays that the observed chi-square statistic for the goodness of fit test has a value of 12.55 with 15 degrees of freedom and a p-value of about 0.64. Since the p-value is greater than 0.05, this suggests that there is not enough evidence to suggest that there is a lack of fit in the

DMH model. The following figure displays the Pearson residuals for the DMH logistic regression model:

Figure B.2: Residuals for the DMH Regression Model

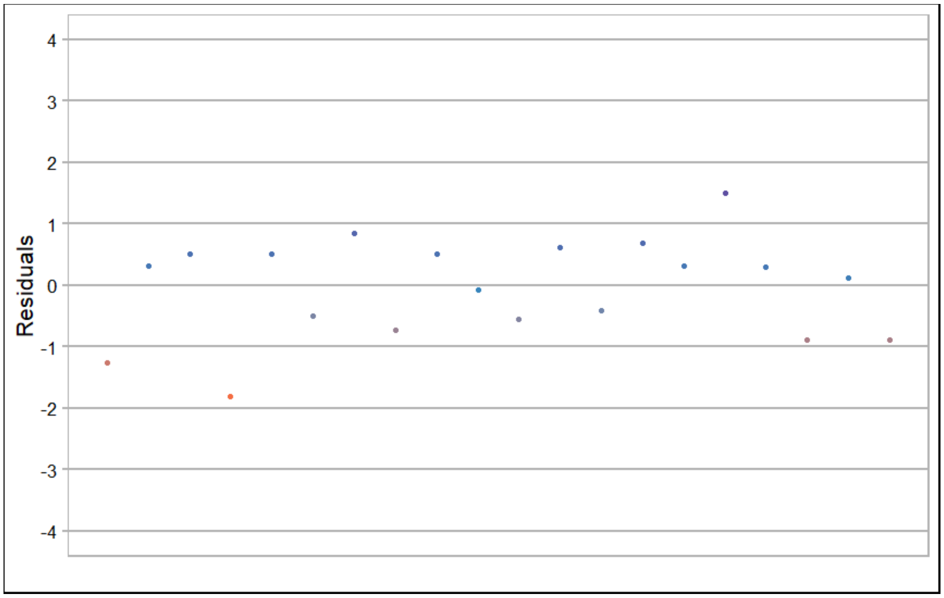


Figure B.2 shows that the residuals are between -2 and 2, form a horizontal band and there is no obvious curvature pattern among the residuals and thus provides evidence for the model possessing a good fit to the data.

APPENDIX C: Model Selection for the DSH and DMH Logistic Regression Models

At the beginning of the report it was mentioned that previous research guided the selection of relevant variables to consider when examining factors that affect the adoption of DSH and DMH cats. Potential models for cats that considered possible interactions between the stray factor and other variables (age and weeks) and a model with no interaction terms were compared in model selection processes. In addition to previous research, model selection by means of comparing the Bayesian information criterion (BIC) values associated with potential candidate models was implemented. Regression models that fit the data better have lower BIC values. The official website for the statistical package JMP provides more information on BIC[15] The following table presents the BIC values for the candidate models that were considered to describe factors that affect the adoption of DSH cats:

Table C.1: BIC Values for Candidate DSH Models

	BIC
Age + Stray + Week	158.98
Age + Stray + Week + Stray*Age	159.43
Age + Stray + Week + Stray*Week	153.44
Age + Stray + Week + Stray*Age + Stray*Week	156.14

Each model that was compared contains the age, stray and week factors. The three models that were compared are 1.) main effects model (Age + Stray + Week model that contains no interaction terms) 2.) the model that contains an interaction term between stray and age 3.) the model that contains an interaction term between stray and week and 4.) the model that contains both the stray-age interaction and the stray-week interaction terms. Table C.1 indicates that the model with the lowest BIC model is the model that contains the stray and week interaction term (this model has a BIC value of 153.44). Although the BIC values computed would suggest that the model which contains

[15] "Likelihood, AICc, and BIC." *Case Study Library | JMP*, 2 Jan. 2019, www.jmp.com/support/help/14-2/likelihood-aicc-and-bic.shtml.

the stray-week interaction term is the best model, there was evidence of overdispersion for the DSH cat models and previous research indicates that model selection criteria such as AIC and BIC perform poorly in cases where there is overdispersion[16]. A criteria value that can be used for model selection when there is evidence of overdispersion is the corrected quasi Akaike information criterion (QAICc) which provides a model selection value like AIC and BIC are used when there is no overdispersion[17]. Similar to BIC, lower values of QAICc indicate a better fitting model. The following table presents the QAICc values for the candidate models that were considered to describe factors that affect the adoption of DSH cats when adjusting for overdispersion:

Table C.2: QAICc Values for Candidate DSH Model

	QAICc
Age + Stray + Week	65.78
Age + Stray + Week + Stray*Age	72.42
Age + Stray + Week + Stray*Week	66.86
Age + Stray + Week + Stray*Age + Stray*Week	75.47

Table C.2 indicates that the model with the lowest QAICc model the main effects model and this was the final model which was chosen (this model has a QAICc value of 65.78).

Table C.3: BIC Values for Candidate DMH Models

	BIC
Age + Stray + Week	67.41
Age + Stray + Week + Stray*Age	69.99
Age + Stray + Week + Stray*Week	68.71
Age + Stray + Week + Stray*Age + Stray*Week	71.68

Table C.3 indicates that the model with the lowest BIC model is the main effects model (this model has a BIC value of 67.41). Similar to the model chosen for DSH cats, the model chosen for DMH

[16] M. Fitzmaurice, Garrett. (1997). Model selection with overdispersed data. Journal of the Royal Statistical Society: Series D (The Statistician). 46. 81 - 91. 10.1111/1467-9884.00061.
[17] Kim, Hyun-Joo & Cavanaugh, Joseph & A. Dallas, Tad & Foré, Stephanie. (2013). Model selection criteria for overdispersed data and their application to the characterization of a host-parasite relationship. Environmental and Ecological Statistics. 21. 10.1007/s10651-013-0257-0.

cats is the most parsimonious model that was chosen out of the candidate models that were under consideration.