**Abstract**

According to the American Society for the Prevention of Cruelty to Animals (ASPCA), every year, about 6.5 million animals are adopted from animal shelters, and about 1.5 million animals are euthanized. A better understanding of the factors contributing to shelter outcomes may allow humane organizations to save more animals. This project will explore which animal-related factors affect adoptions and euthanizations for cats and determine which of these factors have the greatest effect on shelter outcomes

While this research is exploratory, some significant association between adoption/euthanization and animal species, breed, color, sex, and age is expected as these factors influence people’s decision to adopt an animal.

The Austin Animal Center Outcomes dataset from October 1,2013 to October 22,2019 will be analyzed. Only cats who were either adopted or euthanized will be included in the analysis. Outcomes will be coded in two binary variables, *adopted* and *euthanized* (0 = no, 1 = yes). A logistic regression model will be fit in R using Hosmer’s purposeful selection criteria to find the most relevant factors.

If successful, this research will help inform shelters of current trends in adoptions and euthanizations so they may take appropriate action.

**Introduction**

According to the According to the American Society for the Prevention of Cruelty to Animals (ASPCA), every year, about 6.5 million animals are adopted from animal shelters, and about 1.5 million animals are euthanized. 3.2 million of the animals adopted and 860,000 of those euthanized are cats.

Adding to the body of research on animal shelter outcomes for cats may better inform shelters and other humane organizations of adoption and euthanization trends. With this information, they may be able to formulate better plans and improve outcomes. For this purpose, this study seeks to answer two questions:

1. What cat characteristics influence whether a shelter cat is adopted vs. euthanized?
2. What is the relationship between these influential factors and adoption/euthanization?

These questions are addressed through the application of a logistic regression model. By refining which factors are included the model, the cat-related characteristics that affect feline shelter outcomes and the nature of their relationship to these outcomes can be determined.

As this is an exploratory study, there is no initial hypothesis. However, there are some trends based common knowledge and opinions that one might expect to encounter in the data. First, people tend to be more eager to adopt kittens or young cats. Older cats are less likely to be adopted. Second, due to superstition and perhaps an inclination toward colorful cats, potential pet owners are often less likely to adopt black cats. Finally, cats who are spayed or neutered already may have a better chance of being adopted due to potential owners’ reproductive concerns.

**Methods**

The Austin Animal Center Outcomes Data from October 1, 2013 to October 22, 2019 was downloaded from their website and used in this study. This dataset, updated daily, is made publicly available by the city of Austin, Texas and provided by the Austin Animal Center. The data include outcome and demographic information for each animal that passes through the Austin Animal Center.

The original dataset had 109,915 cases. For this analysis, only cases of cats who were adopted or euthanized were used. Cases for other animal and outcome types were excluded. After applying the inclusion criteria, the dataset used for the study was comprised of 19,826 cats, 18,365 who were adopted (92.6%) and 1461 (7.4%) who were euthanized.

The outcome variable for this study was *adopted*: whether or not the cat was adopted (1 = yes, 0 = no). Due to the binary categorical nature of the outcome variable, a logistic regression model was applied to determine which cat demographic variables were related to adoption as well as the magnitude and direction of the relationships present in the data. An examination of how the data fit the assumptions of a logistic regression model follows in the Results section.

The following predictor variables were initially included in the model:

* Animal age at outcome – animal age at outcome in days
* Breed – individual variables for four different breeds: domestic shorthair, domestic longhair, snowshoe, American shorthair
* Mixed breed – whether or not the cat was a mixed breed (1 = yes, 0 = no)
* Color – individual variables for the # most common colors: black, orange, brown, white, gray, multi
* Mixed color – whether or not the cat was a mixed color
* Gender – whether the cat was female or not (1 = female, 0 = male)
* Intactness – whether or not the cat was intact (1 = intact, 0 = not intact (spayed or neutered))
* Has name – whether or not the cat had been given a name (1 = yes, 0 = no)

Except for animal age at outcome, all other predictors were binary categorical indicator variables.

The entire analysis was performed in RStudio. Hosmer’s Purposeful Selection method was used for variable selection and refinement of the logistic regression model. Goodness of fit was assessed using the Hosmer-Lemeshow goodness-of-fit test.

**Results**

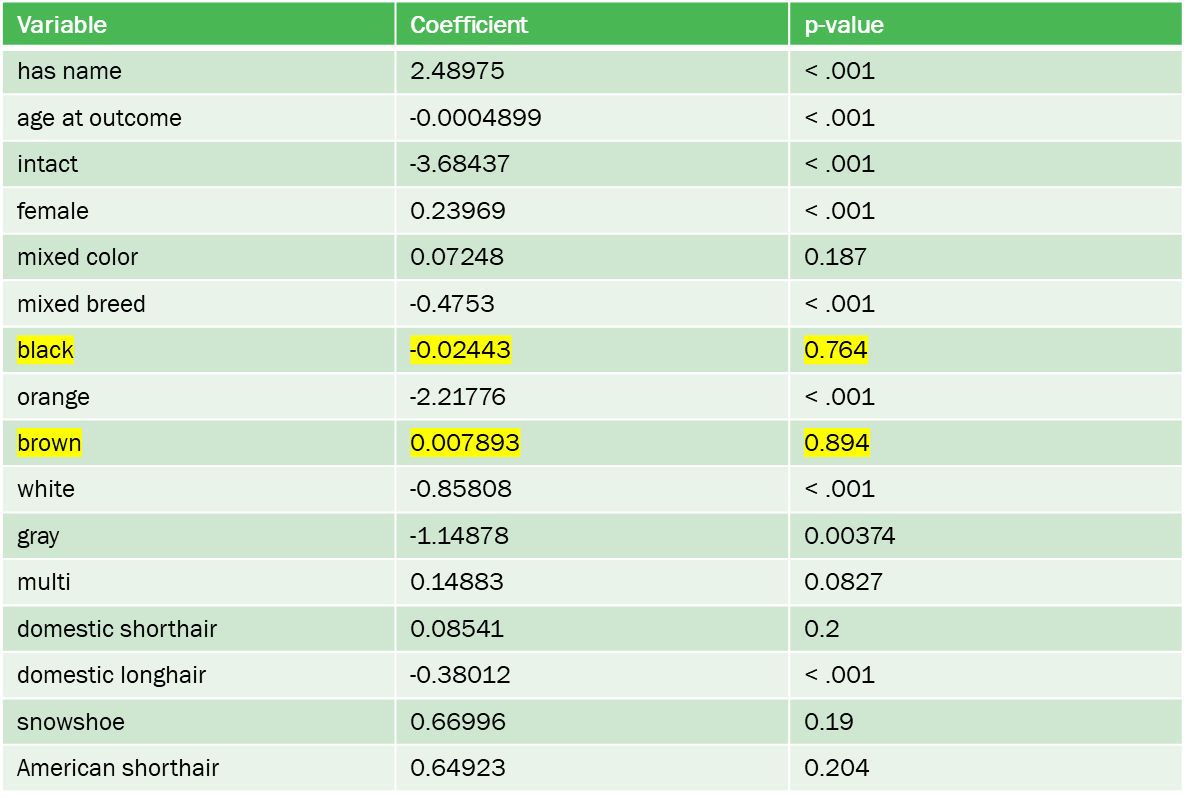
Each step of the logistic regression model fitting is detailed in this section, followed by interpretation of the final results. In addition, the assumptions necessary for logistic regression are examined. Selected output is included. In the interest of legibility, full model outputs are included only when interpretation of the coefficients was relevant. Otherwise, individual coefficient and LRT p-values are reported within the narrative. For more details, please refer to the RMarkdown document included in the submission.

*Variable Selection*

Step 1: Univariable logistic regression

Univariable logistic regressions were run with each predictor and the outcome variable. All variables with p > 0.25 were eliminated during this step. Only black (p = 0.764) and brown (p = 0.894) were non-significant at this level, so all other variables were included in the next step.

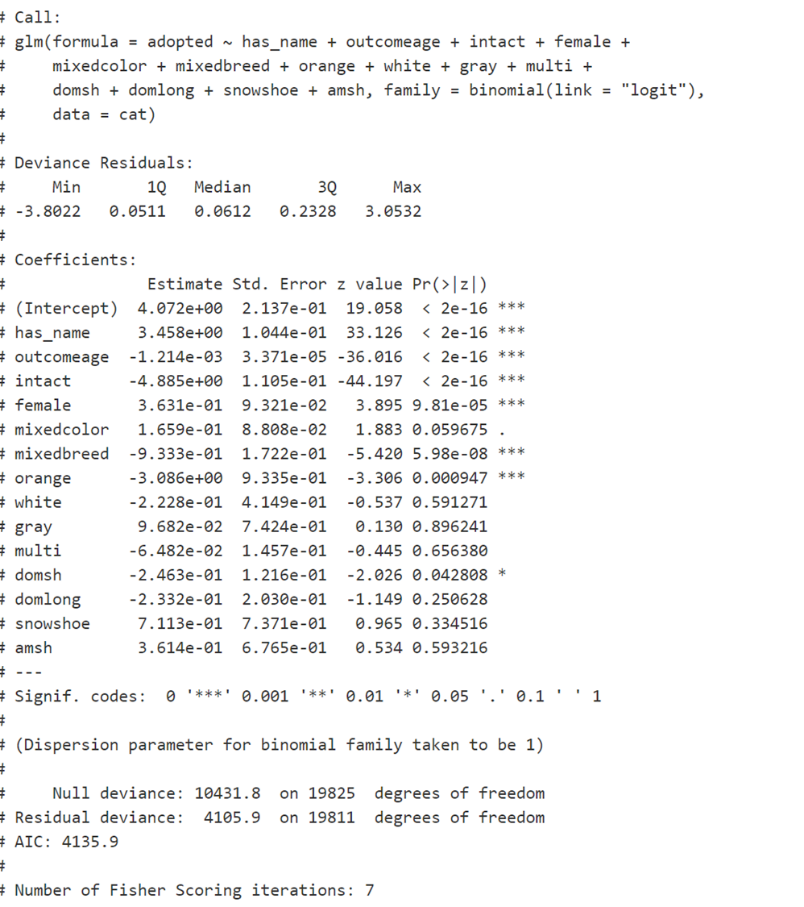
*Table 1: Coefficients and p-values for univariable logistic regressions*



Step 2: Full model

All significant variables from Step 1 were added to the full logistic regression model. Variables with p > 0.05 were considered for removal. At this step, mixed color, white, gray, multicolored, domestic longhair, snowshoe, American shorthair were removed.

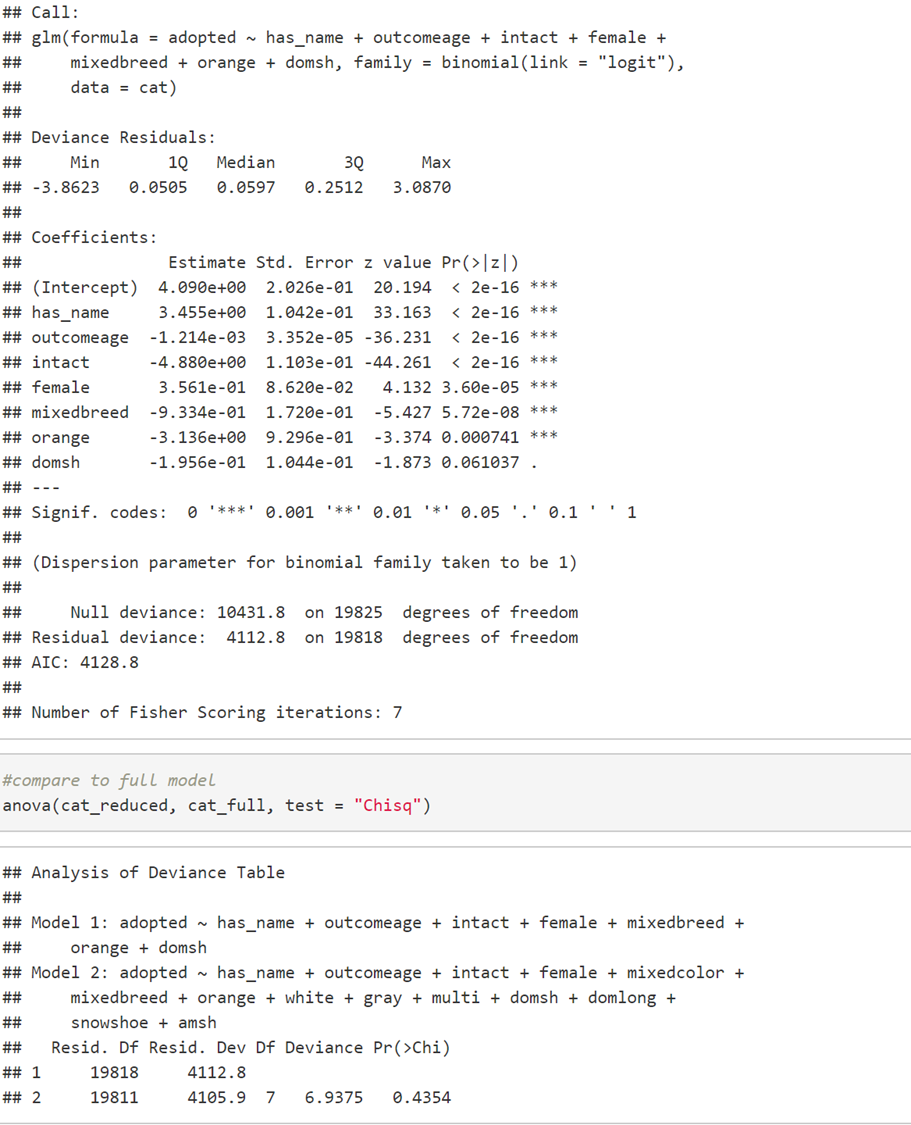
*Figure 1: Results of logistic regression with full model*



Step 2: Reduced model

A reduced model was fit with the variables that remained from the previous step. The likelihood ratio test (LRT) p-value was used to compare the fit of the reduced model to that of the full model. A non-significant at alpha = 0.05 (p = 0.4354) indicated that the reduced model was a better fit, so the process was continued with the reduced model.

*Figure 2: Results of reduced model logistic regression and LRT*

**

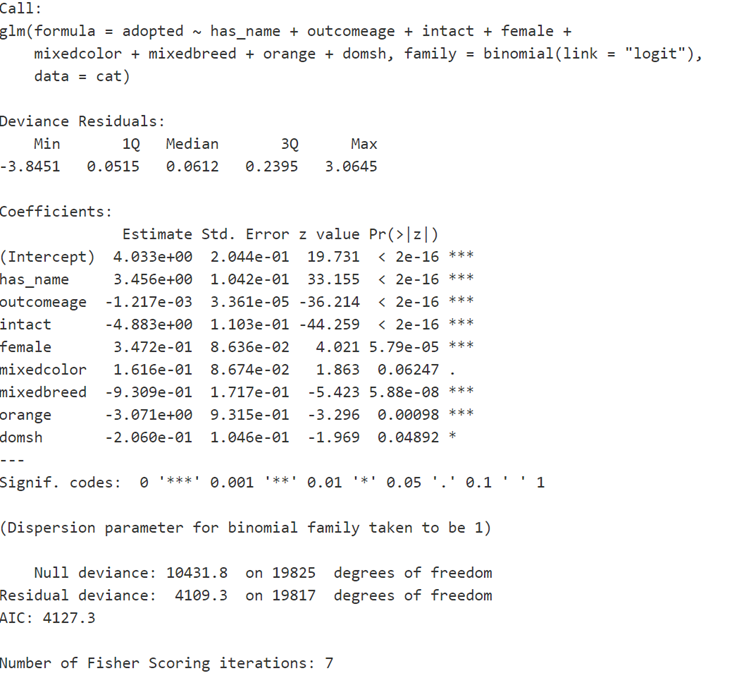
Step 3: Check for 20% coefficient change

The coefficients presented in Figure 2 above were then examined to determine whether any of them had changed more than 20% with the removal of the variables from the full model. These coefficients were visually inspected, and the change in coefficient was calculated for values that appeared as if they could have changed 20% or more. Upon inspection, domsh (domestic shorthair) had changed by 21%, indicating that at least one of the variables removed in the full model must be added back to the model.

Step 2: Add back mixed color

Mixed color was added back to the model as it was the non-significant variable closest to significance in the full model (p = 0.059675). Upon examination, compared to the full model, domsh changed only 16%, and all other variables showed minimal change. A LRT comparing the reduced model + mixed color to the full model was non-significant (p = 0.7502), indicating that the reduced model + mixed color fit the data better than the full model. Thus, the process continued with the reduced model + mixed color.

*Figure 3: Results of reduced + mixed color logistic regression model*

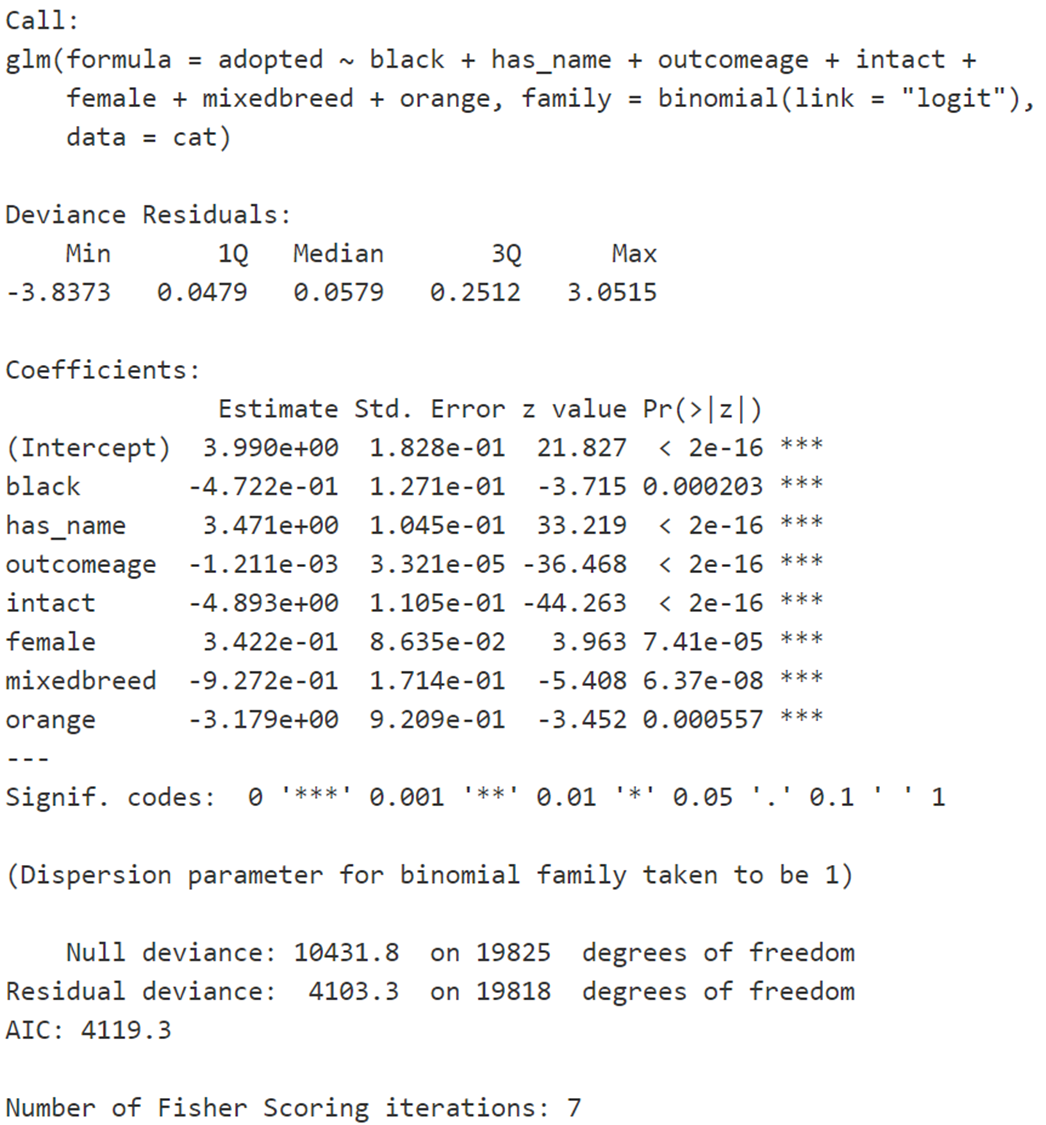


Step 4: Add back variables removed in Step 1 to the reduced + mixed color model

Black and brown, which were removed in Step 1, were added back individually to the reduced + mixed color model to determine if either of them affected adoption in the presence of the reduced + mixed color model predictors. Both black (p = 0.002) and brown (p = 0.03) were individually significant. When both variables were added back to the model together, black (p = 0.02) but not brown (p = 0.05) was significant. Thus, only black was added to the reduced + mixed color model.

However, adding black resulted in domsh (p = 0.075) and mixed color (p = 0.45) being non-significant, so domsh and mixed color were removed from the model. The LRT comparing the model with black added and domsh and mixed color removed to the reduced + mixed color + black model was non-significant (p = 0.164), indicating a better model fit. This new model, illustrated in Figure 4 below, became the preliminary main effects model.

*Figure 4: Preliminary main effects model*

**

Step 5: Check linearity of age at outcome with logit of adopted

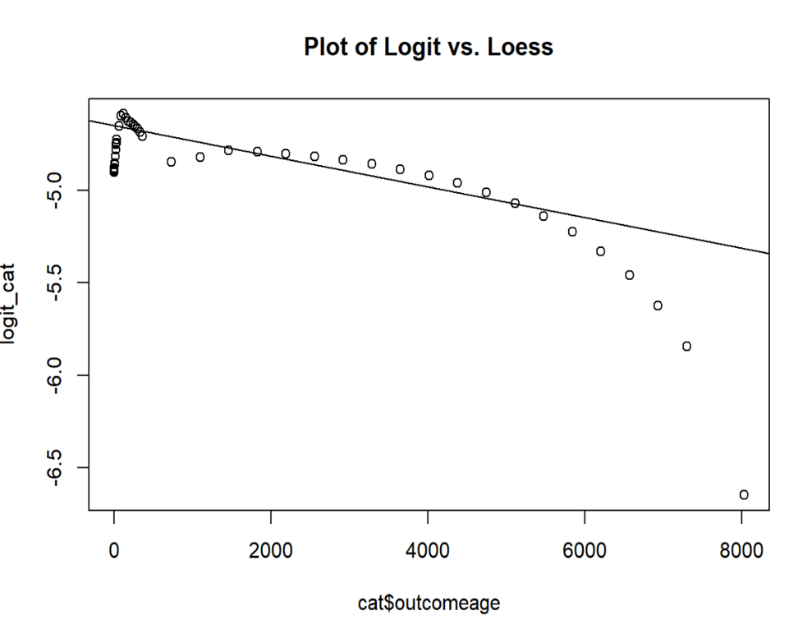
A logit vs. loess plot was used to examine the logistic regression assumption of linearity of continuous predictor variables (age at outcome) with the logit of the outcome variable (adopted).

Figure 5 illustrates that the relationship does not appear to be linear. There is a sharp rise at young ages followed by short, sharp drop, gradual decline, and then another more rapid drop. To handle the non-linearity, age at outcome was reparametrized into three indicator categories with

kittens 1 year or younger as an implied reference group. Category cut-offs were estimated based on visual cutoffs in the graph and logical ages based on cat adoption.

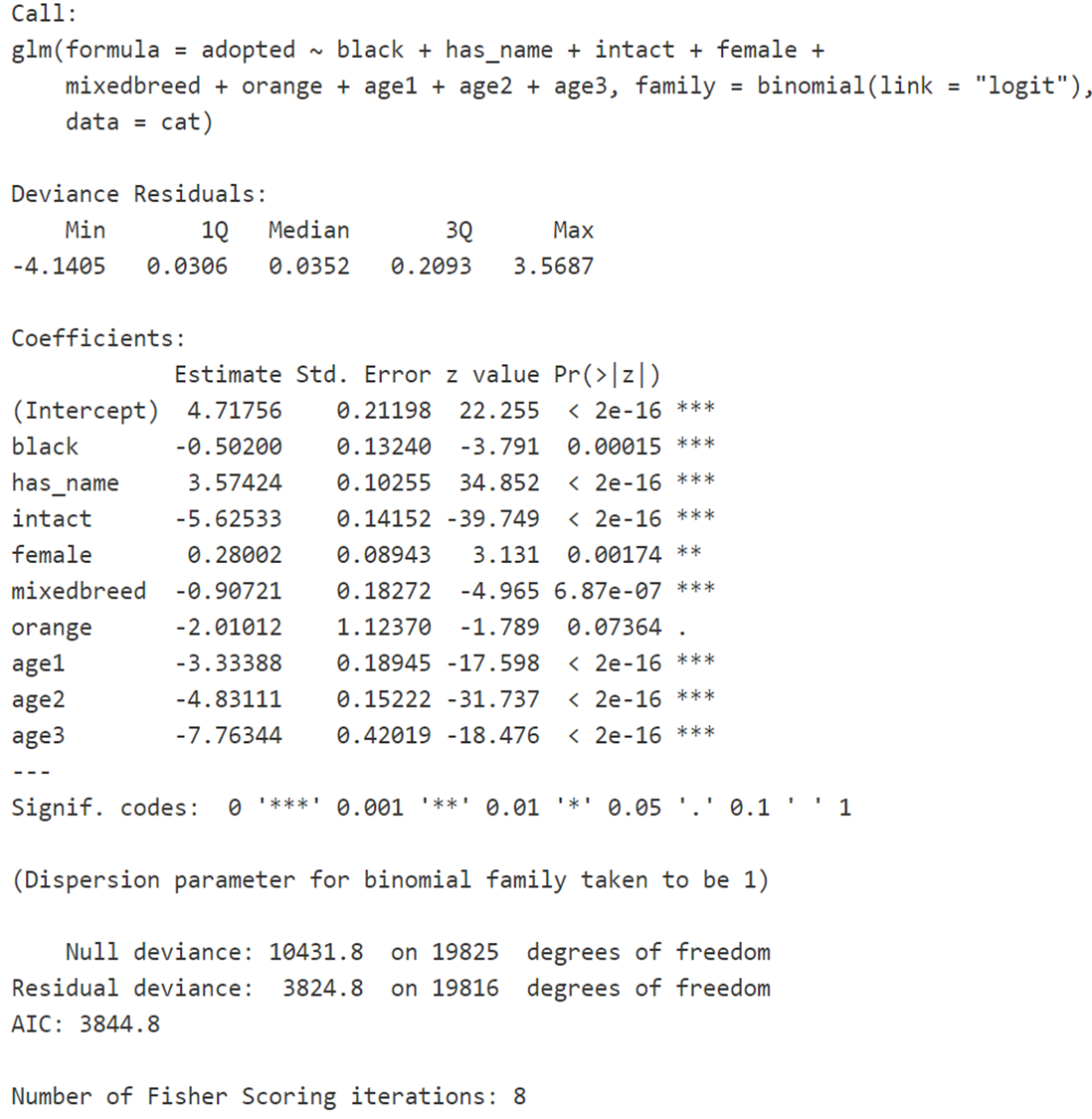
* age1 – > 365 days (1 year) and <= 913 days (2.5 years)
* age2 - > 913 days (2.5 years) and < 5900 days (16 years)
* age3 - > 5900 days (16 years)

*Figure 5: Logit vs. loess plot for linearity of age at outcome with logit of adopted*

**

The fit of the categorical age categories was then assessed by adding each individually to the model, removing the old age at outcome variable. All three of the age categories were significant (p < .001 for all). Then, all the age categories were added to the model together. All three age categories remained significant (p <.001 for all). As shown in Figure 6, variation was present in the coefficients for each category, reflecting the non-linearity of age at outcome.

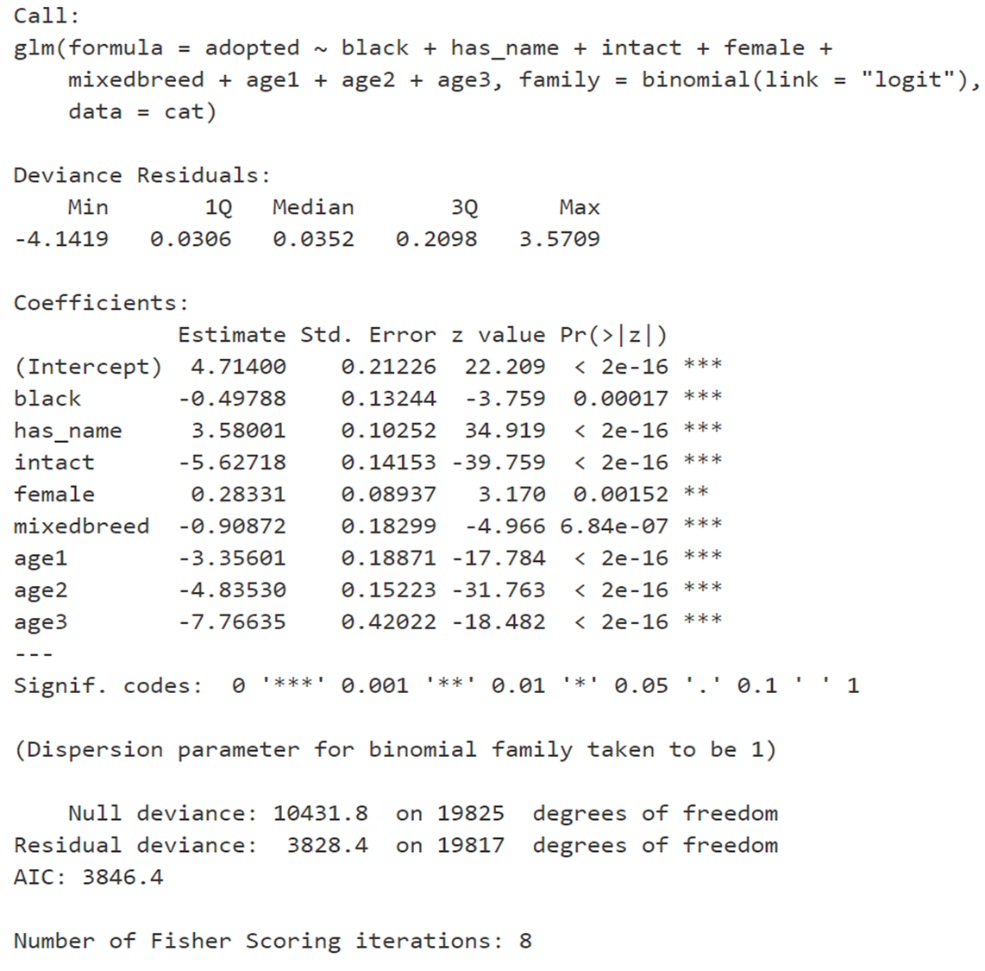
*Figure 6: Preliminary main effects logistic regression model with all age categories*

**

After adding the age categories to the model, orange was no longer significant (p = 0.07) and had to be removed. The LRT comparing the model without orange to the model with orange was barely non-significant (p = 0.059), indicating a slightly better fit without orange.

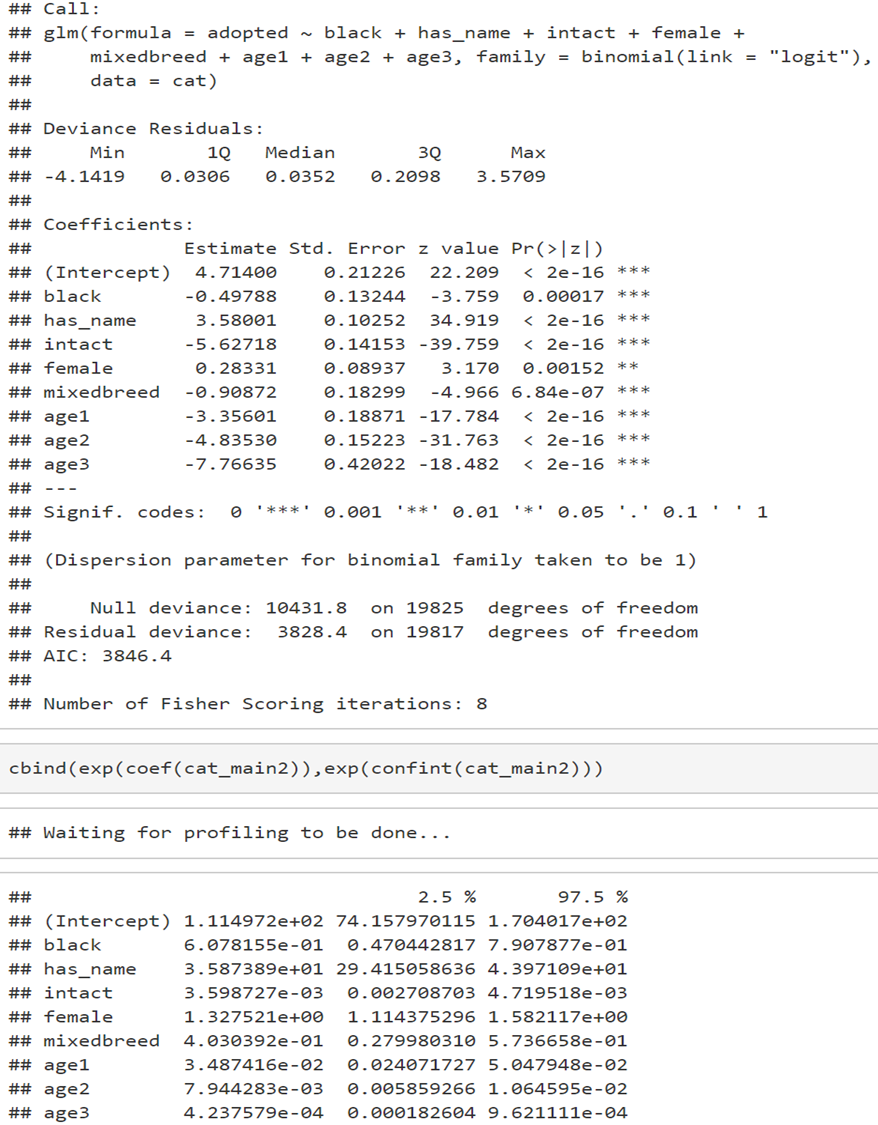
Thus, the final main effects model shown in Figure 7 was achieved.

*Figure 7: Final main effects model*



*Interpretation of Main Effects Model*

*Figure 8: Final main effects model with odds ratios and confidence intervals*

**

The final main effects model with odds ratios and confidence intervals obtained by exponentiating the coefficients and confidence intervals from the original logit output is shown above in Figure 8.

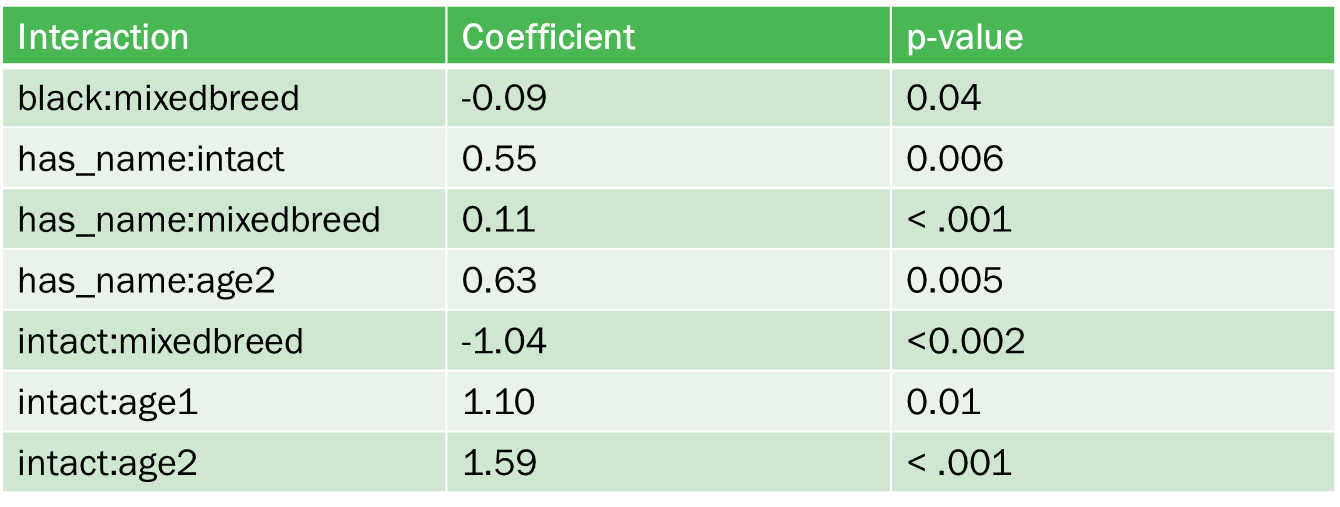
For each main effect interpretation, all other variables are held constant. Black cats have 0.6 of the odds of being adopted compared to non-black cats (OR = 0.60, CI = [0.47, 0.79]). Cats that are named have 3.59 times the odds of cats without names of being adopted (OR = 3.59, CI = [29.4, 44.0]). Cats who are not spayed/neutered have 0.004 of the odds of cats who are spayed/neutered of being adopted (OR = 0.004, CI = [0.0027, 0.0047]). Female cats have 1.33 times the odds of male cats of being adopted (OR = 1.33, CI = [1.11, 1.58]). Mixed breed cats have 0.40 of the odds of purebred cats of being adopted (OR = 0.40, CI = [0.28, 0.57]). Cats between 1 year old and 2.5 years old have 0.03 of the odds of cats 1 year old or younger of being adopted (OR = 0.03, CI = [0.02, 0.05]). Cats between 2.5 years old and 16 years old have 0.008 of the odds of cats 1 year old or younger of being adopted (OR = 0.008, CI = [0.006, 0.011]). Cats older than 16 years old have 0.0004 of the odds of cats 1 year old or younger of being adopted (OR = 0.0004, CI = [0.0002, 0.0009]).

*Interactions*

Step 6: Examine interactions

Interactions between all predictor variables in the finalized main effects model were computed and added to the main effects model individually. Table 2 illustrates which of these interactions were significant at the individual level.

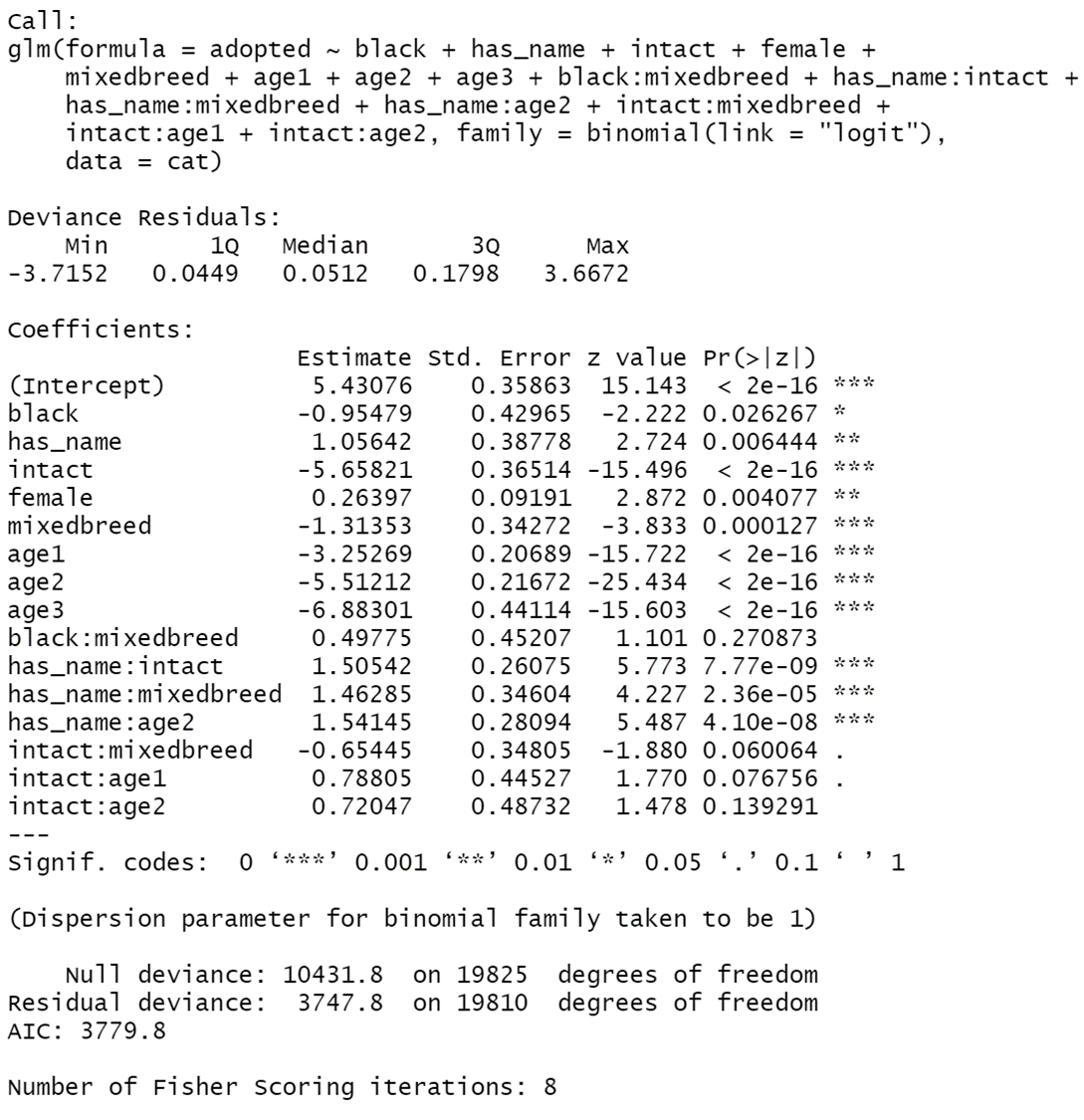
*Table 2: Significant individual interactions at p < 0.05*

**

The interactions in Table 2 were then added to the main effects model together. As shown in Figure 9, black:mixedbreed, intact:mixedbreed, intact:age1, and intact:age2 were non-significant at

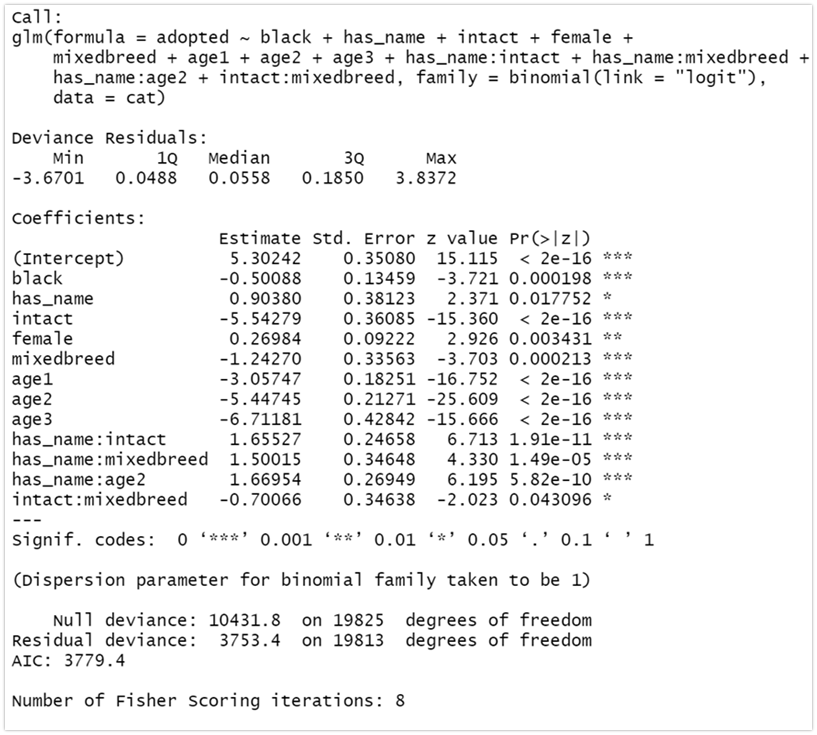
p < 0.05 and were eliminated from the model.

*Figure 9: Main effects model with significant interactions added*



The reduced interaction model was then inspected. The AIC of the reduced interaction model was higher than that of the original interaction model (reduced AIC = 3781.5, original AIC = 3779.8), suggesting at least one interaction should be added back to the model. As intact:mixedbreed was closest to significant in the previous step (p = 0.06), this interaction was reintroduced into the model. The resulting LRT between the model with intact:mixedbreed vs. the model without intact:mixed breed was significant (p = 0.04), indicating that the model with the variable re-added was a better fit. The LRT between the reduced model with intact:mixed breed and the full interaction model was non-significant (p = 0.13), so the reduced model with intact mixed breed became the final model. The results of this model are given in Figure 10.

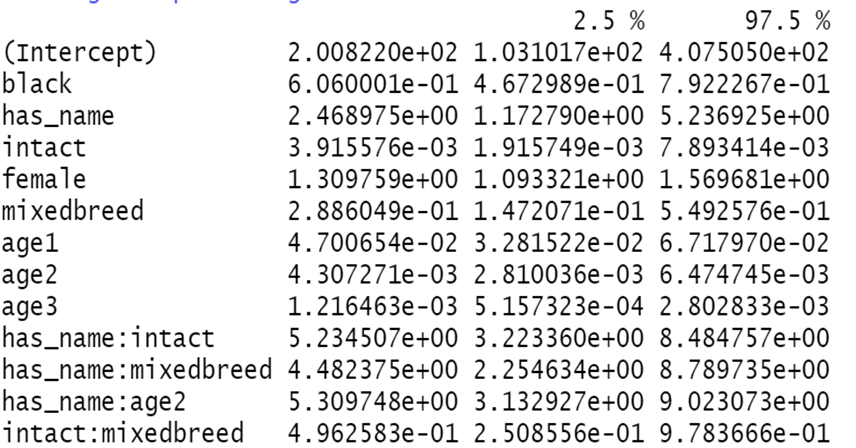
*Figure 10: Final logistic regression model*

**

*Interpretation of Final Model with Interactions*

In Figure 11, the odds ratios and coefficients for the final model are shown.

*Figure 11: Odds ratios and coefficients for the final model including interactions*

**

The interaction terms were added to the final main effects model mostly for model fit and because of the presence of the interaction step in Hosmer’s Purposeful Selection criteria. Because interactions terms are complicated to interpret, particularly in the presence of their corresponding main effects, it seems more appropriate simply to note which factor interactions should be considered in the context of cat adoption vs. euthanization.

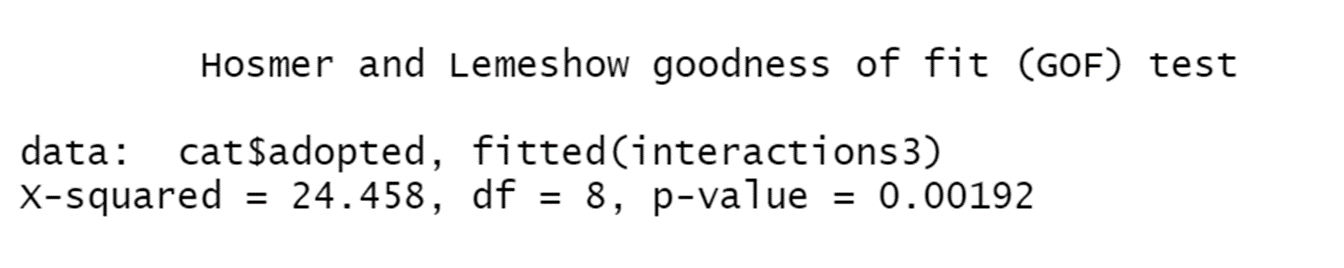
For cats, having a name impacted the effects on adoption of being intact, being a mixed breed, and being between the ages of 2.5 and 16 years. Being intact also impacted the effects of being a mixed breed.

*Goodness of fit*

Step 7: Checking for goodness of fit

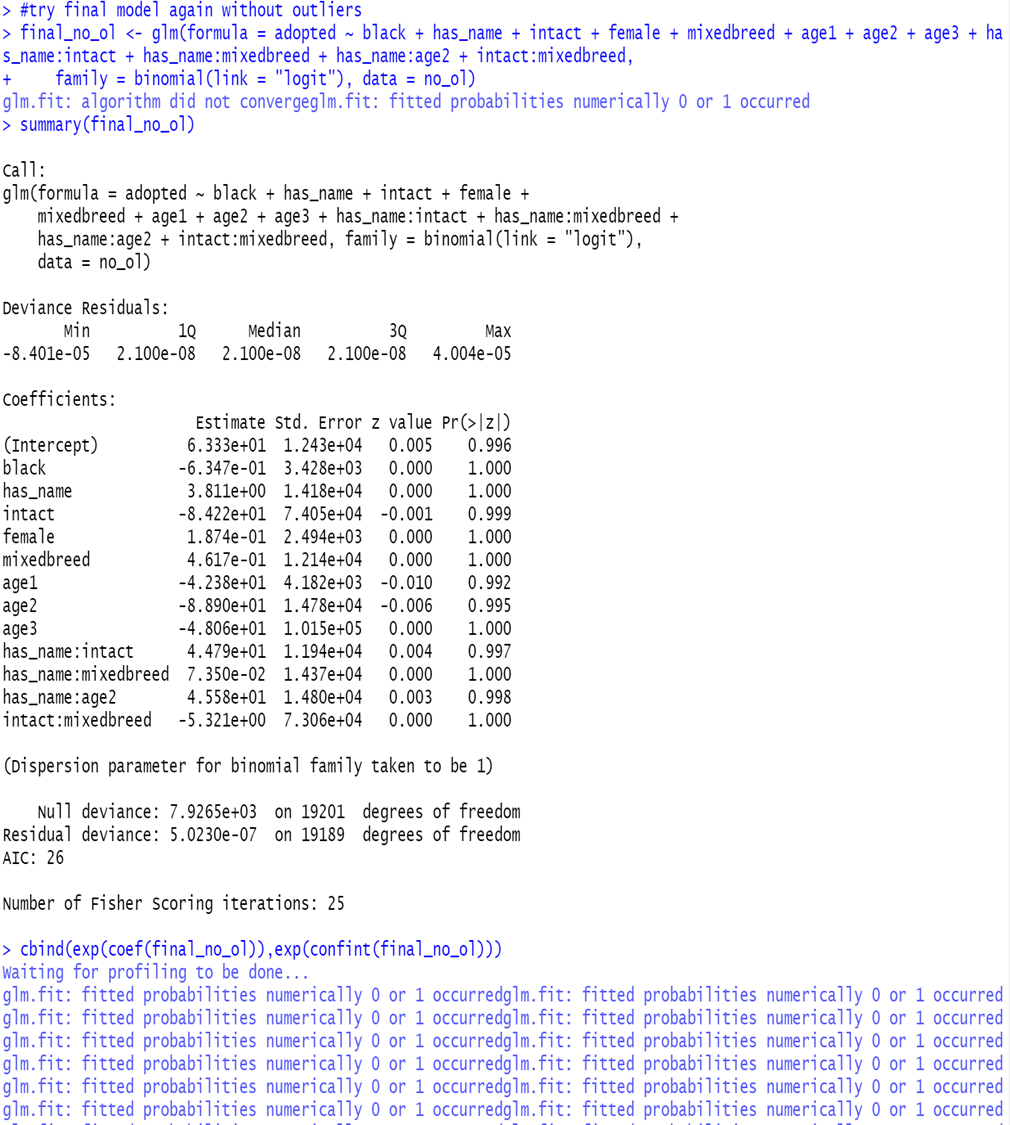
Hosmer and Lemeshow’s goodness of fit test was applied to the final model to check how well the model captures the data. The results are shown in Figure 12. A very significant p-value of 0.00192 indicates that the model is a poor fit.

*Figure 12: Hosmer and Lemeshow goodness of fit test*

**

Following the goodness-of-fit test, influential values were examined to see if they played a role in the model lack of fit. Cook’s distances for each case were calculated from the final model. Cases with a Cook’s distance greater than 4/n, where n was the total number of cases in the data, were marked as having high leverage. There were 624 high leverage points in the dataset. With this large number of high leverage points and with the large size of the data set, it seemed like it would be difficult to identify individual outliers. To try and create a better fitting model, all the high leverage points were removed, and the final model with interactions was re-run on the data with the intent of running another Hosmer-Lemeshow goodness of fit test on the new model. However, the attempt failed because a logistic model could not be fit. The error messages displayed are shown in Figure 13.

*Figure 13: Attempt at fitting logistic regression to data without high leverage points*



The errors depicted above typically indicate perfect separation, suggesting that some combination of predictor variables resulted in perfect prediction of adoption or euthanization. Given the relatively low percentage of euthanizations in the dataset, it is possible that eliminating the high leverage points removed all the euthanizations. However, the cause of the lack of fit is still unclear.

*Assumptions*

The logistic regression assumptions were examined to ensure that a logistic regression model was appropriate for the data.

1. The outcome variable must be binary.

Indeed, adopted, the outcome variable, is a binary categorical variable (1 = adopted, 0 = not adopted).

1. Independent observations

Independence can be reasonably assumed because most of the cats are likely not related, so for the most parts, the characteristics and outcome of one cat should not affect that of another cat.

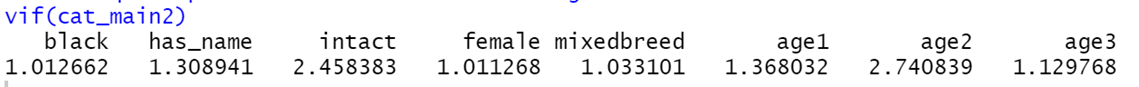
1. There must be a linear relationship between continuous predictor variables and the logit of the outcome.

The only continuous predictor in the model was age. The non-linear relationship between adopted and age of outcome was accounted for by the reparameterization of age into three categorical variables (see Step 5 for details).

1. No multicollinearity should be present among the predictor variables.

The variance inflation factors (VIF) were examined for all variables in the final main effects model. The VIFs are shown in Figure 14 . A VIF over 5 is considered evidence of multicollinearity. Here, the highest VIF is 2.74, so there is no evidence of multicollinearity between the variables.

*Figure 14: Variance inflation factors for main effects model predictor variables*



**Discussion**

The results of this study indicate that factors that matter for cat adoption are sex, being black vs. another color, age, intactness, being purebred or not, and having a name. It is doubtful that having a name is a causal factor in adoptions as whether or not names are assigned to a cat may be dependent on how long the cat has been in the shelter and the manner in which the cat ended up at a shelter (e.g., a cat who was surrendered by their owner might already have a name).

Based on the odds ratios from the full model, a cat’s odds of being adopted are increased if they are female, not black, 1 year old or younger, spayed/neutered, named, and purebred. In terms of interactions, having a name interacts with the effects on adoption of being a mixed breed, being between the ages of 2.5 and 16 years old, and being intact. Being intact also interacts with the effects on adoption of being a mixed breed.

The identification in this study of age, sex, intactness, and blackness as factors that influence cat adoption is consistent with the findings of previous studies. A study by Lepper, Cass, and Hart in 2002 found that the major factors related to feline adoption were “age, sex, coat color, and reason for relinquishment” (p. 1). Alberthsen et al. (2013) also found a lower rate of euthanasia in kittens than in adult cats as well as in cats who had been spayed or neutered.

However, the literature also highlights one of the limitations of this study. There are many factors that affect cat adoption that were not examined in this study or were not adequately addressed in the data. Both Lepper, Cass, and Hart (2002) and Alberthsen et al. (2013) identified factors outside the scope of the current study. Lepper, Cass, and Hart pointed to reason for relinquishment, while Alberthsen et al. (2013) noted that cats that had been surrendered by their owners were less likely to be adopted than stray cats. In addition, other studies have examined the effects of other factors such as adoption price (Zito et al., 2015) and condition of the animal on shelter intake (Hawes, Carrigan, & Morris, 2018). There are many factors not captured in the current data set, particularly adopter-related factors or cat personality characteristics, that could be even more predictive of shelter outcomes than the variables examined in this study. It is possible that the lack of such factors in this analysis contributed to the lack of fit in the data.

In addition, a more accurate model may have been fit if having/not having a name was not included in the analysis. As mentioned previously, this variable may be an effect rather than a predictor of shelter outcomes, so it may be more reasonable to exclude it.

**Conclusion**

These results contribute to the body of literature on factors influencing outcomes for cats in animal shelters. In particular, they examine the odds of adoption based on various cat characteristics.

Animal organizations can use this information to become more aware of trends in animal adoption and direct their resources and efforts towards cats who may otherwise be at a disadvantage. Examples of potential action might include increasing efforts to spay/neuter cats at the shelter/feral cats in the community or increasing advertising for cats who don’t fit the ideal demographic (i.e., for older, black, male, mixed-breed cats). Future directions of study could include

including time of outcome in the analysis, finding a better classification system for age, breed, and color, repeating the analysis on data from other shelters to see if the results are reproducible, and measuring and including human-related or cat personality factors as potential predictors of shelter outcomes. It is hoped that research like this will help find as many cats as possible a happy home.

**References**

Alberthsen, C., Rand, J. S., Bennett, P. C., Paterson, M., Lawrie, M., & Morton, J. M. (2013). Cat

admissions to RSPCA shelters in Queensland, Australia:   
Description of cats and risk factors for euthanasia after entry. Australian Veterinary Journal, 91(1-2), 35-42.

American Society for the Prevention of Cruelty to Animals. (2019). Retrieved from

<https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics>

Austin Animal Center Outcomes. (2019). Retrieved from

<https://data.austintexas.gov/Health-and-Community-Services/Austin-Animal-Center-Outcomes/9t4d-g238>

Hawes, S., Kerrigan, J., & Morris, K. (2018). Factors informing outcomes for older cats and dogs in

animal shelters. Animals, 8(3), 36.

Lepper, M., Kass, P. H., & Hart, L. A. (2002). Prediction of adoption versus euthanasia among dogs

and cats in a California animal shelter. Journal of Applied Animal Welfare Science, 5(1), 29-

42.

Zito, S., Paterson, M., Vankan, D., Morton, J., Bennett, P., & Phillips, C. (2015). Determinants of cat

choice and outcomes for adult cats and kittens adopted from an Australian animal shelter.

Animals, 5(2), 276-314.