US Animal Shelter Intakes and Adoptions Analysis

Mavis Wang
Data Visualization, Department of Data Analytics
San Jose State University
San Jose, US

Chujie Cai
Data Visualization, Department of Data Analytics
San Jose State University
San Jose, US

Zuojun Zheng
Data Visualization, Department of Data Analytics
San Jose State University
San Jose, US

***Abstract*— The overpopulation of homeless animals continually grows in the United States. Millions of dogs and cats entered animal shelters are euthanized every year because of overcrowding. According to the law, once an animal enters a shelter, the owner has three to ten days to reclaim the animal before being placed for adoption, sold, or euthanized [7]. Animal shelters are faced with challenges due to limited resources in space, staffing, and facilities [1]. Statistics and literature focus on reasoning and adoption prediction; however, there is little discussion of the actionable implementation to resolve the crowdedness problem. This paper analyzes a series of national dogs and cats data in 2020 to discover factors and trends linked to overcrowding and adoption subjectively and geographically. Through the explorative visualizations, we discovered that the total number of intakes in an animal organization does not necessarily reflect the crowdedness because of the uneven resource distribution among states and organizations. Our study also shows an increasing seasonal trend of the total intakes towards December. Among the adopted dogs and cats, we discover that age, breed, and coat are the top three features affecting the adoption rate. Further, there is a correlation between the total number of licensed breeders and total intakes; however, we do not have enough evidence to conclude such phenomenon as causation, especially when online dog shipping services become prevalent. We aim to provide actionable insights to help animal shelters reduce the crowdedness by (a) transporting animals to the optimal location and (b) increasing their adoption rate.**

Keywords— data visualization, shelter animal analysis

I. INTRODUCTION

According to SPCA [6], there are 6.5 million dogs and cats entering animal shelters every year, and approximately 1.5 million of them are euthanized. Local shelters and rescues face challenges to handle such an amount of animal intake due to limited resources and spaces and reduce the intake amount. Although recent statistics indicate a decline in euthanasia, some analysts suspect that shelter intake reduction may be one of the latent causes [4]. In this project, we investigated a series of pet data and statistics in the U.S. to understand and discover the latent causes of why the animal enters and leaves the shelter through an explorative visualization approach. There is literature discussing solutions of tackling overpopulation by reducing the stays or increasing adoption rate; however, those analyses mainly were based on the total number of intakes instead of the crowdedness rate because each state and organization has a different amount of resources. Our goal in this project is to bring awareness to the general public and provide actionable insights to help animal shelters resolve the crowdedness issue by analyzing seasonal and regional associations to help relocate the adoptable animals and optimize their adoption rate.

II. RELATED WORK

In an adoption speed prediction research [12], the author Soowhan concludes that young animals which are younger than 4 months old are adopted faster than adults, usually with a spike at 2 months old. Non-pure breeds are faster to be adopted, instead of pure breeds. Compared with female pets, male pets are easier to be adopted. Another finding in this research is that people have different preferences for pet's health conditions. Dewormed pets are easier to be adopted. However, people tend to prefer non-vaccinated and non-sterilized pets when adopting. What's more, healthy pets without information are rarely adopted, which indicates that it's always helpful to provide information on pet health conditions no matter if the pet is healthy or not. According to Racheal et al. [9] cats usually stay in shelter 4 days longer than dogs, which takes 18 days in average to be adopted. New-born puppies and kittens usually stay up to 70-80 days in shelter, with decreasing time of staying to 27 days for young pets from 1-6 months' old and about 14.8 days for adult pets from 1-3 years old. Bradley and Rajendran [3] found out that age and color are the main factors that influence pet adoption speed. According to the research of Protopopova in 2012 [11], breed, reason of intake would highly influence the adoption rate. Simona et al. (2015) suggest that except for the animal characteristics including breed, size, and age, shelter programs improving animal interaction with humans would significantly help in adoption.

III. METHODS

A. Data Collection

In order to analysis in various aspects, we collected cleaned data from multiple sources. Bellows are the cleaned dataset included in this project: Intakes and outcomes of Austin animal center [9], animal cares and control [14], 2019 household income in the U.S. [15], educational attainment rank [16].

For more detailed data regarding to adoption of 2020, we designed and implemented the python code to obtain data

from Petfinder API. For pets that were not adopted, we were able to collect all of them. There are approximately 16,000 records for dogs that were not adopted and 21,000 records. In contrast, there are over 370,000 adopted dogs and over 290,000 adopted cats in 2020. In order to balance the two datasets, we under sampled the adopted data using the random sampling approach. As a result, the adopted dataset has roughly the same number (16,000 dogs and 21,000 cats) of records as the adoptable dataset (pets were not adopted).

B. Data Cleaning

To perform visualization tasks in the later stage, it is necessary to clean the dataset first. Since the dataset contains attributes that are nested, we have to first flatten these fields using the Pandas' map function. For example, the 'breeds' field returned by the API is in the format of a Python dictionary. It contains primary breed, secondary breed, mixed and unknown fields. To convert these fields into meaningful columns, the flatten operation is required.

In addition, we have to determine what features are reasonable to keep for further analysis. Due to the image of pets is not the main focus of this project, features like descriptions and photos' hyperlinks are not considered. Moreover, columns such as colors tertiary that have more than 80% of missing value in both the adopted and adoptable datasets are eliminated. However, if a column has more than 80% of missing values in only one of the datasets (eg. Declawed attribute in cat dataset), we decided to keep the column and labeled the null values as "missing".

C. Data Preprocessing

To perform visualization tasks in the later stage, it is necessary to clean the dataset first. We dropped the attributes like descriptions and hyperlinks for photos, which are irrelevant to our analysis. Columns such as colors tertiary that have more than 80% of missing value in both datasets are also eliminated. However, if a column has more than 80% of missing values in only one of the datasets (eg. Declawed attribute in cat dataset), we decided to keep the column and marked the null values as missing.

D. Adoption Rate Visualization

Owing to our target audience may not have experience in data visualization, it is crucial to present the data in a comprehensible format. Hence, two different visualizations were designed, one for the analysis purpose, and the other for the comprehensive purpose. The former visualization does not require an additional calculation field, the latter one on the other hand requires an extra measure computed from the original dataset to represent the adoption rate. The logic is as follow (use dogs as an example).

$$Adoption\ Rate = \frac{\sum d_{adopted}}{\sum d_{adopted} + \sum d_{not\ adopted}}$$

Where $d_{adopted}$, $d_{not\ adopted}$ refers to dogs adopted, and dogs not adopted in the same category (eg. dogs of the same breed) respectively. To convert this equation into a calculation field, we need to use the combination of *iff* and *sum* function, which can be written as the following

```
sum(iif([Status] = "adopted", 1,0)) /
(sum(iif([Status] = "adopted", 1,0)) +
sum(iif([Status] = "adoptable", 1,0)))
```

IV. Results

A. Analysis 1: In-Out Reasons

1) Intake Reason Analysis

Our preliminary analysis emphasizes discovering reasons that dogs and cats enter and leave the shelter. We want to understand the factors and associations that are directly related to animal intake and adoption. (reference for the chart)

As we can see in the bubble chart (Fig. 1.), there are mainly three categories of intake seasons in three different colors: gray for social reasons including stray, transfer from other shelter, and born in shelter; orange for owners' financial reasons such as unable to afford, moving, and landlord issues; blue for other reasons related to owner such as incompatible with owner lifestyle, allergy, relationship split. Stray is the biggest reason that shelters hold a large number of intakes each year. The blue and orange bubbles represent the reasons related to owners' life changes such as relocating, relationship split, and unable to afford.

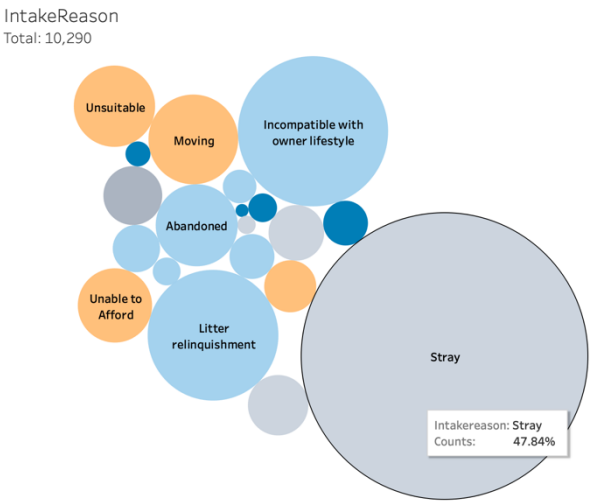

Fig. 1. Top intake reasons in animal shelter

2) Out Reason Analysis

The out-reasons are relatively simpler compared to various intake reasons. It's comforting to see that 56.11% of animals are adopted and 25.42% of animals are fostered in this no-kill shelter. 14.48% of animals are reclaimed by owners which indicates that there a number of owners would be willing to have their pets back after a certain problem is solved. We can see that only a small percentage (3.93%) of animals are transferred to other shelters.

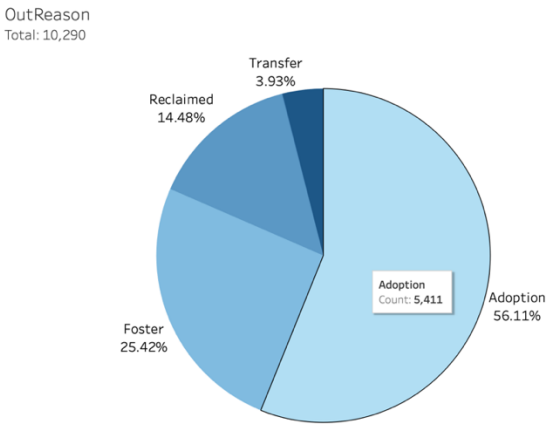

Fig. 2. Tops reasons that animal leave a shelter

B. *Analysis2: Crowdedness*

After understanding why dogs and cats enter and leave the shelter nationwide, our next challenge is defining crowdedness and identifying the most needed area to transport animals to optimal locations effectively. To achieve this, we need first to identify the state-level crowdedness and then drill down to the city level. The analysis involves three aspects of the intake data: total intakes, average intakes, and intake distribution.

1) *Total Intakes*

The total intake is the sum of all adoptable dogs and cats published by organizations in the United States. Fig. 3 maps the geographically distributed dog and cat intakes in the United States in 2020. There are over 36 thousand dogs, and cats enter animal organizations as of December 31, 2020. The highlighted states in blue indicate a higher volume of intakes compared to those in grey. Texas (3,857), California (3,144), Florida (2,924), and Pennsylvania (2,106) are the top four states of the most intakes, and there are a few clusters around the central and the east coast.

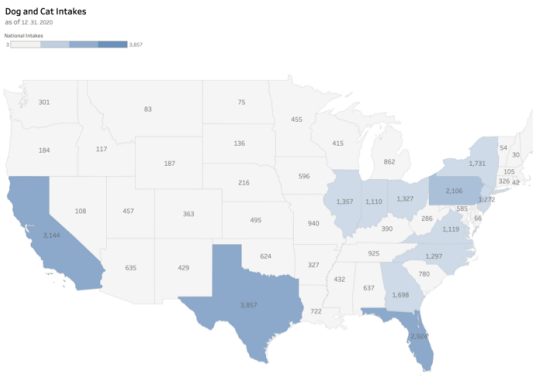

Fig. 3. National dog and cat intakes in 2020

It might be intuitive to associate the number directly with crowdedness; however, the population and resources in these states are also the highest compared to other states. Our data shows that the total number of organizations of the four states (1,742) is about one-third of the national organizations (5,730). Fig.4 manifests an interactive solution of using a tooltip to guide viewers in discovering macro and micro-level information of a state's intakes. We find a significant imbalance of resources among the states. In California, for instance, there are 599 animal organizations in total, while there are only 10 in South Dakota. Therefore, the total intakes might not be a suitable indicator of crowdedness.

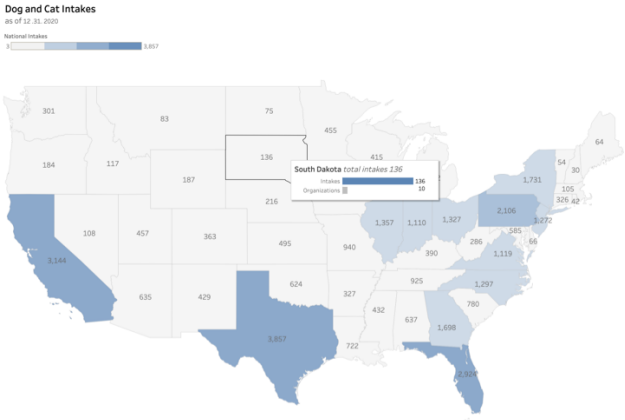

Fig. 4. Tooltip indicates the uneven distribution in different states

2) *Average Intakes*

To get a better measurement, we use the national average intake as our baseline to determine crowdedness. We first calculate the average intake per organization for each state and then divide it by the national average intake as our naive estimator. Fig. 5 shows a comparison between total national intake and average intake with the same color scale of measurement to represent crowdedness. The color-coded maps tell us a different story from our previous analysis. There are more blue shades across the entire country, and the average intake per organization in California is slightly below the national average (7 average intakes). The four large states are no longer the most crowded areas. Instead, South Dakoda (14 average intakes), Uta (13 average intakes), and New Mexico (13 average intakes) become our top three target areas.

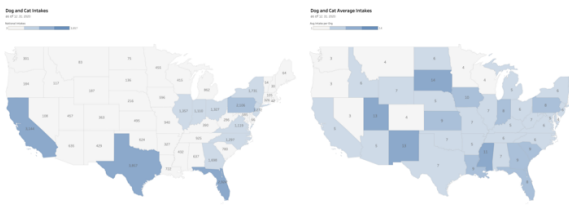

Fig. 5. Crwodedness estimator comparison by using total intake (left) and average intake (right)

The average intake may not be an unbiased estimator due to outliers, latent factors, and other limitations during data collection; however, it creates a simple and intuitive threshold for us to discover the questionable areas in the preparatory stage. Once we find the target states, we can drill down to the city-level to explore further in-depth. Fig. 6 shows the top 5 cities with the highest intake volumes in Utah. The tooltip reveals an unusual volume in Kanab and an imbalanced distribution among the top 5 intake cities. The

high volume of intake can be a reflection of the limited resources. Alternatively, it may simply suggest that many animals are registered under the same network in Kanab. Our next step is to drill down further by looking at the intake distribution of each city to find out the answers.

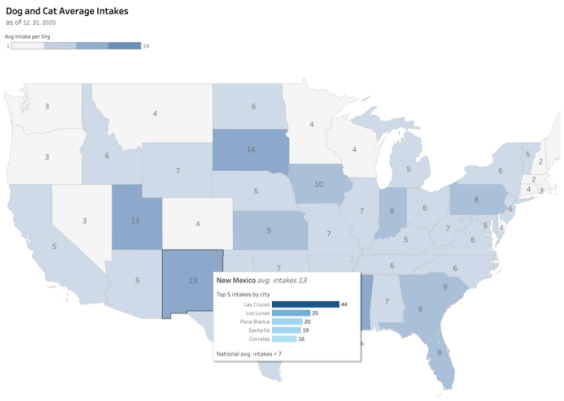

Fig. 6. With tooltips, users can drill down to citi-level for details

3) *Intake Distribution*

Fig. 7 provides more comprehensive information with the same color-coded system and an interactive map navigator to help the viewer quickly locate and identify the relatively overcrowded cities in each state. To avoid ambiguity, we also apply size alongside color as another attribute to emphasize the density of the crowdedness. The figure also marks the total number of intakes to organizations ratio of the most and the least compacted city in each state. From the top of the figure, we observe a high variance of average intakes ranging from 1 to 58 in Iowa compared to the state's average (10).

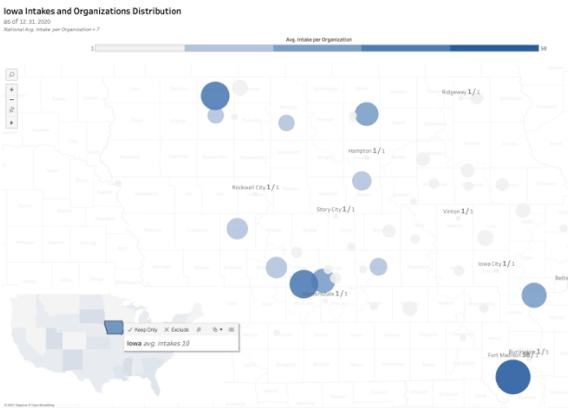

Fig. 7. National organization and intake distribution

Fig. 8 provides a complete picture of the resource and animal distribution in the city of De Soto by summarizing its average intake, the total amount of organization, intake, and the number of cats and dogs. The tooltip shows that there is only one organization in the city with 39 dog and cat intakes. The figure suggests that many shelters face crowdedness due to limited resources and imbalanced distribution.

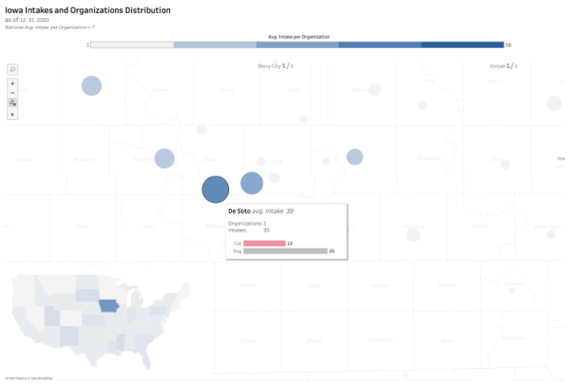

Fig. 8. The tooltip shows detailed dog and cat distribution in De Soto

C. *Analysis 3: Seasonal Trend*

Many shelters have struggled with the unexpected surge of intakes. In this analysis, we want to understand if there is a significant link or trend between seasons and the volume of intakes to help shelters and organizations prepare in advance.

1) *National Trend*

Fig. 9 shows the national dog and cat intakes over the twelve months in 2020. There is a significant upward trend toward the end of the year. In December, the number is surged by 41% to 8,014 than the previous month (5,681). It is almost eight times compared to the lowest volume in March (1,019). A similar trend applies to both cats and dogs.

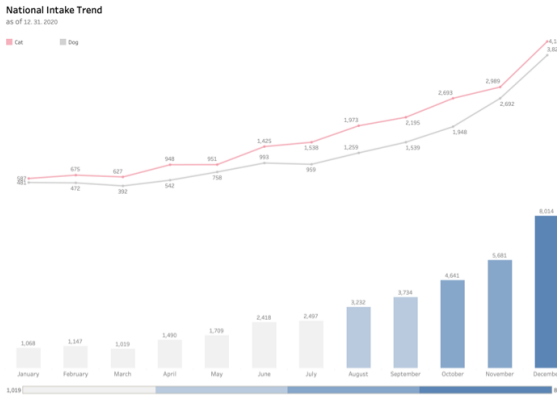

Fig. 9. Monthly dog and cat intakes in 2020

2) *Trend by State*

To verify if the seasonal trend is a norm, we further look into the monthly intake by states. Fig. 10 presents the intake distribution in Pennsylvania over the year 2020. We applied a third-degree polynomial trendline to signify the trend in contrast to the national one. The upward pattern appears towards the following half of the year and arrives at 421 in December. It is nearly 20% of the statewide intake (2,106) over the year. We also observe a 56% spike in April compared to March.

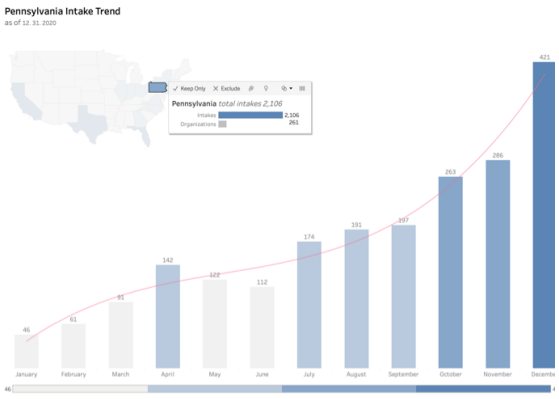

Fig. 10. Monthly intake trend by states

D. *Analysis 4: Essential Features of Dog and Cat*

To prioritize the rescue task, it is necessary to further understand what features affect the adoption rate. In this paper, we discovered the three most significant factors that influence the adoption rate the most: age, primary breeds and coat. As mentioned in the method section, these attributes are visualized in two approaches. One for comprehensive purposes and the other for analysis purposes. To ensure the chart is comprehensive, we simply showed the adoption rate on the bar charts for each of the three categories. To get more insight from each attribute, different charts were chosen when analyzing each of the three categories.

1) Age

To visualize the age percentage in each group(baby/young/adult/senior), the doughnut chart is utilized. It is the most suitable chart to represent the numerical proportion of age distribution. Through this visualization method, we are not only able to capture the dominant age group/groups for dogs/cats but to observe the difference between dogs and cats in terms of age perspective.

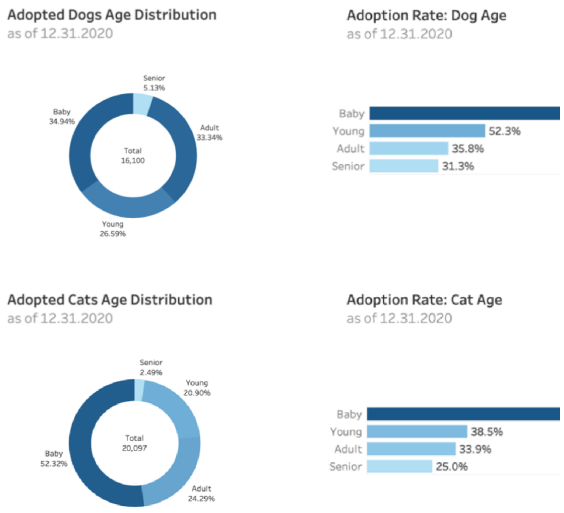

Fig. 11. 2020 adoped cats/dogs age distribution(left), adoption rate(right)

As the Fig. 11suggests, among 16,100 adopted dogs in 2020, the majority of them are infants and adults, but in terms of adoption rate, infants and young dogs have a higher chance to be adopted. For cats, most adopted cats are infants among 20,097 adopted cats in 2020 (adoption rate bar chart suggests the same). Hence, we conclude that infant dogs and infant cats have the highest opportunity to be adopted.

2) Primary breeds

For the graph with the hours of the day, we separated each location attributes to one color. We can easily see that at each hour, where is the most dangerous place on the street. If put this graph into real use, drivers will be notified to pay extra attention at certain location at certain time. A warning system can be built based on that.

Since there are many dog/cat breeds (219 for dogs, 59 for cats), we took only the top 10 most adopted breeds to perform analysis for each category. This is sufficient because the 11th most adopted breed has a relatively small population compared to the total sample size (2.05% for dogs, 0.67% for cats). Also, the secondary breed is not included in our analysis due to the percentage of missing values (over 80% are missing for both dog and cat datasets)

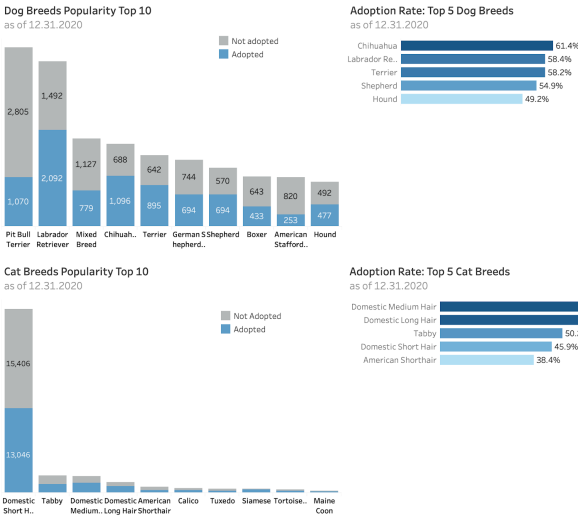

Fig. 12. 2020 adoped cats/dogs breeds popularity(left), adoption rate(right)

According to Fig. 12, certain dog breeds have a higher probability of being adopted (ie. adopted percentage is greater than adoptable percentage). For instance, Labrador Retriever, Chihuahua, Terrier, etc are breeds that have a higher adoption rate. As for the cat breeds, although the adoption rate for domestic medium/long haired cats results in higher adoption rate, it may not be representative due to sample size differences. As shown in Fig.XXX, there exists a dominant cat breed(domestic short hair) in our samples, it is almost 70% of the cat dataset. Therefore, the cat breeds may not be an efficient indicator for cat adoption rate since most of the shelter intakes are domestic short haired cats.

3) Coat

The coat feature is visualized using the connected dot graph. Our goal is to identify what type of coat infers the higher adoption rate. The connected dot graph enables the viewers to compare between a small number of classes (two classes: adoptable/adopted in our case) effectively. This

graph could provide the information intuitively even for people who are not familiar with data visualization.

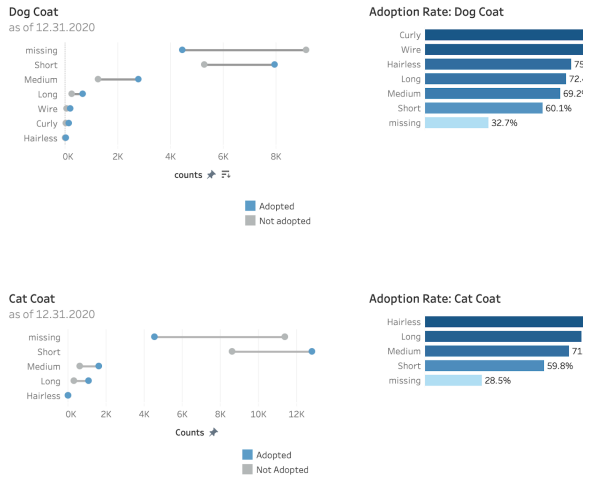

Fig. 13. 2020 adoped cats/dogs coat(left), adoption rate(right)

In Fig. 13, we can observe that all types of coat indicate the higher adoption rate except "missing" for both dogs and cats. Notice that only 32.7% of dogs and 28.5% of cats with coat attributes missing were adopted in 2020. On the dot graph, we could clearly observe that dogs and cats that are not adopted with missing coat attributes are doubled the size of the adopted ones. Specifically, adopted/not adopted ratios are 1 : 2.51 for cats, 1 : 2.06 for dogs. Hence, it is essential to have the coat information available for pets to increase the adoption rate.

E. Analysis5: Correlation

We find out that there is a significant difference among states in intakes and adoption numbers. To understand the factors which influence people abandoning and adopting animals, we carry out correlation analysis on the following factors: the number of licensed breeders, education level, and income level. In the following figures, every blue dot represents a certain state. Resolution is distributed according to the number of intake and adoption, which means that the darker blue dot is, the larger intake/adoption number is found in the according state.

1) Intakes and Breeders

As you can see in the Fig. 14., y-axis represents the number of intakes, and x-axis represents the number of licensed breeders. The trend line in the figure shows that intake number grows exponentially when the number of licensed breeders increases. For example, California has 493 licensed breeders, with 3,144 intakes of animals, while Nevada has 53 licensed breeders, with 108 animal intakes in shelter. Also, we can see that Ohio and Indiana are two outliers, with significant number of licensed breeders but rather low intake number, which we could analyses the reasons in the Discussion session.

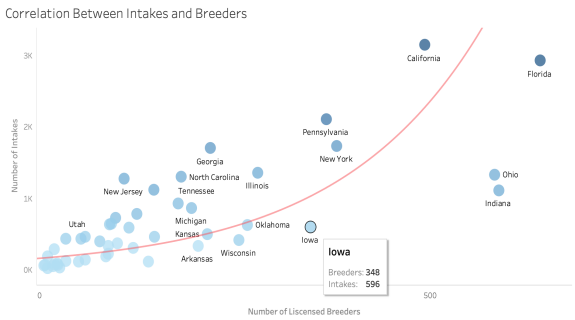

Fig. 14. Correlation between intakes and number of breeders

2) Adoption and Breeders

In Fig. 15., x-axis represents the number of licensed breeders and y-axis represents the adoption numbers. The trend line shows that the adoption number grows exponentially when the number of licensed breeders increases. For example, there are 681 licensed breeders in Taxes, with 2,368 animals adopted from shelter in 2020. Oregon has 88 licensed breeders with 448 animals adopted. As shown in the figure, there is also an outliner Missouri, with 1,127 licensed breeders and 686 animals adopted, would be discussed later.

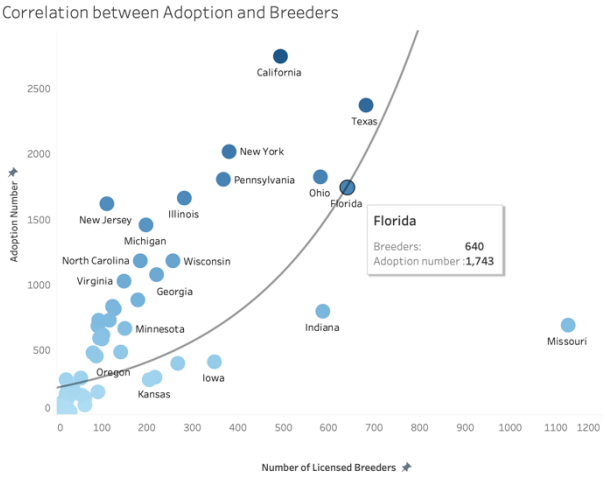

Fig. 15. The correlation between adoption and number of licensed bredders

3) Adoption and Education

In Fig. 16., x-axis represents the education score while y-axis represents the adoption number. We can see that there is not an obvious trend line to describe the relationship between adoption number and education level. However, the states with median education scores, such as California, Texas, and New York are found with higher adoption numbers. The states either with higher education scores, for example, Massachusetts (81.54), or the ones with lower education scores, for example, West Virginia (23.65), both have fewer animals adopted in 2020, with 480 and 271 respectively.

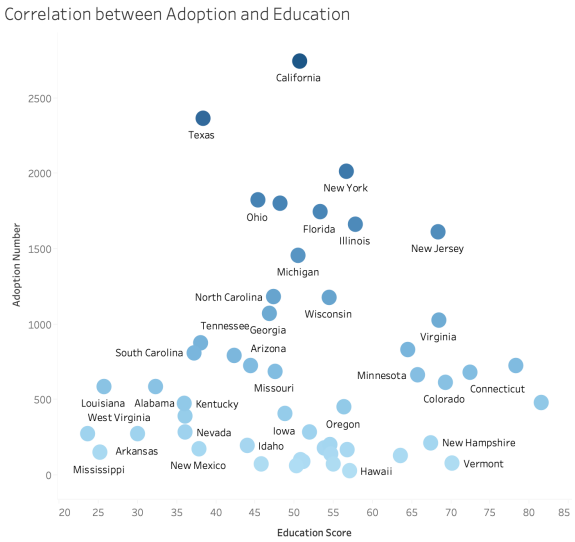

Fig. 16. The correlation between adoption and education level

4) Adoption and Income

In Fig. 17., x-axis represents median household income and y-axis represents adoption number. As shown in the figure, there is no significant correlation discovered between adoption and income levels. However, we can conclude from the figure that the states with higher income levels are more likely to have higher adoption rates. For example, California has the highest household income (80,440 dollars per year) and also adopts most animals (2,746) while Mississippi has the lowest household income (45,792 dollars per year) with 147 animals adopted in 2020.

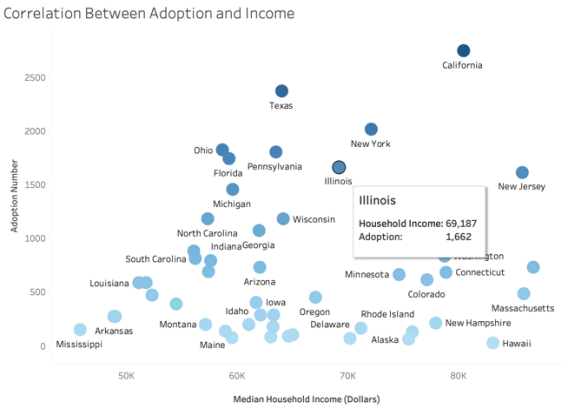

Fig. 17. The correlation between adoption and income level

V. Discussions

A. *In-out Reasons*

In intake reason analysis, we find out that almost half of the animals enter the shelter are stray animals. which could be solved by having stray animals spayed/neutered regularly. Decreasing the number of stray animals would help shelters significantly. The other intake reasons can mainly conclude to owners' abandoning for life changing reasons, for example, relationship split, and financial reasons such as landlord issues, unavailability of pet expenses. Therefore, it's important to have people fill a form stating their living conditions and marriage status before adoption. It would also help if shelters can carry out more orientation activities to educate people how to be a qualified owner before adopting. From the analysis, we can suggest that a qualified owner should have stable income, suitable accommodation for a pet, and good health conditions (not allergic to dog/cat). This information can help shelters to design background information check for potential adopters.

On the other hand, there are two findings in out-reason analysis that need attention. Not a small number of animals are reclaimed by the owners after a certain time. In this situation, shelters can tackle the problem of why a pet is abandoned (for example, biting or behavioral reasons) and therefore to make efforts to increase the possibility of the animal being reclaimed. Also, we notice that only a small portion of animals are transferred to other shelters. With the rapid development of transportation nowadays, shelters can consider transferring animals to nearby shelters to balance the resource.

B. *Crowdedness*

In the crowdedness analysis, the visualization tells an entirely different story between the total number of intakes and the average intakes. By adjusting the threshold from the total number of intakes to average intakes, we discover a high variance of organizations and intakes across the country. The total intake does not necessarily reflect overcrowding. The average intake provides a more intuitive and accurate measurement of crowdedness compared to the total intake. Our color-coded system guides the viewers to discover and identify the relatively crowded areas progressively. With the interactive user interface, the viewer can easily browse, search, and compare the up-to-date resources and intakes among the nearby cities to relocate the animals effectively.

C. *Seasonal Trend*

Our analysis suggests that there is a direct connection between season and intake nationwide. The intake number is constantly increased by month and accelerated during the half of the year. The peak wave mainly occurs in December, and we suspect that there might be an association with the holidays. Many experienced shelter workers suggest that most people surrender their pets during the holidays because they want to travel or have to move to a new place (Kress, 2021).

D. *Essential Features*

In this paper, we found three most significant features that influence the adoption rate: age, breed and coat. For both cats and dogs, the infants tend to have a higher adoption rate based on our analysis. In contrast, the breed's prosperity does not share the similar outcome. The certain dog breeds such as Chihuahua and Labrador have better opportunity to be adopted than other dog breeds. For cat breeds, even that the domestic medium hair and domestic long hair showed a high adoption rate, it is not representative due to the dominant breed: domestic short hair. Lastly, the coat feature showed one essential fact, which is that the pet with missing coat attribute has extremely low probability of being adopted.

E. Correlations

In correlation analysis, we find out that the intake and adoption numbers are exponentially related to the number of breeders. The correlation indicates that more resources should be allocated to shelters in states with more breeders. It would also helpful if shelters in 'more-breeder states' can transfer overloaded animals to the 'fewer-breeder states'. As we notice in the analysis there are some outliers such as Ohio, Indiana, and Missouri which have the largest number of licensed breeders but with low intake/adoption counts. According to the Humane Society of the United States [13], these states are proved to have most puppy mills where dogs are bred for commercial purposes, which catches social attention in recent years. Also, we notice that the states with either high or low education levels presents lower adoption rates, which gives us an interesting indication that people with medium education levels are more likely to be a potential adopter. Moreover, we find out the states with more median household income adopt more animals, which again suggests that it's important to consider income level in an adopter background information check. The correlations analysis would inspire animal shelters to better allocate resources and to find a stable, long-time adopter efficiently.

VI. Future Work

Based on our preliminary descriptive analysis, we understand the reasons, regions, seasonal trends, and the essential factors linked to adoption. Our goal aims to help animal shelters resolve the crowdedness problem by transporting the adoptable animals to the best locations with the highest chance of being adopted. Using the information, we found, we can find the maximum likelihood of adoption rate based on animal factors, season, region, and other possible features. We include two approaches in our future work: statistical models and machine learning algorithms. We first build statistical models to discover latent factors and formalize the association that connects to adoption and overpopulation. Next, we will use the machine learning tool to build a predictive model based on the associations to predict the most likelihood of adoption based on region and season for the adoptable animals. Moving forward, we plan to build an automated transportation network to help shelters balance the uneven distribution between animal intakes and resources across the nation. Further, we look forward to helping animals transport to the best location with the highest chance of being adopted.

VII. What Did We Learned

Throughout the project, we learned how to apply visualization principles, technical skill sets, and domain knowledge to build a meaningful product that conveys important messages and benefits our society.

A. Technical

1) Python

In the data collection and cleaning stage of the project, we learned how to write Python code to obtain data from the Petfinder API. During this process, we implemented our algorithm to random sampling the adopted cats and dogs data. Being able to design and implement algorithms for data collection is significant for the potential future extension of this project.

Furthermore, we successfully extracted useful information from returned JSON files and converted them into data frames with the support of the Python Data Analysis Library: Pandas. A substantial amount of time was invested in data cleaning such as flatten the nested attributes and drop/fill missing values. Throughout this process, we not only reinforced the programming knowledge but understood the significance of data quality.

2) Tableau

We apply gestalt principles by displaying the similar figures together to better explain the results and show differences between topics. In the figures, data-ink is distributed properly by focusing on the data points and trend lines while decreasing the ink used in grid lines. Joining tables in tableau is helpful to connect individual datasets together with the same attributes. It facilitates us to understand the correlation between different factors and discover the data patterns animal intake and adoption.

B. Research and Visualization

The research helps us understand the use cases and scenarios to generate innovative solutions to resolve the real problems. Through visualization, we can help people discover meaningful information effectively and understand the potential root causes. We also witnessed the power of visualization as it could be easily manipulated towards any intended direction, both positive and malicious ones. It is our responsibility, as data scientists, to deliver the correct and unbiased information to ensure data veracity with the literacy of data ethics.

C. Collaboration

We work collaboratively in researching, looking for datasets, cleaning data, conducting Tableau analysis, and drawing conclusions. We meet weekly to report the progress of the project and discuss the problems we have encountered. In the project, each team member utilizes personal background and expertise to facilitate studying.

References

[1] "Characteristics, Challenges of the Shelter Environment," *ASPCApro*, 28-Apr-2020. [Online]. Available: https://www.aspcapro.org/characteristics-challenges-shelter-environment.

[2] "Determining and exceeding capacity," *HumanePro*. [Online]. Available: https://humanepro.org/page/determining-and-exceeding-capacity.

[3] J. Bradley and S. Rajendran, "Increasing adoption rates at animal shelters: a two-phase approach to predict length of stay and optimal shelter allocation," *BMC Veterinary Research*, vol. 17, no. 1, 2021.

[4] M. Hanson, "Animal Euthanasia Statistics [2021]: Shelter Data by Year," *Spots.com*, 03-Jan-2021. [Online]. Available: https://spots.com/animal-euthanasia-statistics/.

[5] M. P. Cab, "Rescue vs Shelter vs Breeder Pet: Which is Right For You?," *Home*. [Online]. Available: https://www.mypetcab.com/blog/rescue-vs-shelter-vs-breeder-pet.

[6] "Pet Statistics," *ASPCA*. [Online]. Available: https://www.aspca.org/animal-homelessness/shelter-intake-and-surrender/pet-statistics.

[7] R. F. and W. and A. Dillingham, "Full Title Name: Table of State Holding Laws," *Animal Law Legal Center*, 01-Jan-1970. [Online].

Available: https://www.animallaw.info/topic/state-holding-period-laws-impounded-animals.

[8] R. Kress, The Doggie in the Window: How One Dog Led Me from the Pet Store to the Factory Farm to Uncover the Truth of Where Puppies Really Come From. Sourcebooks.

[9] R. Woods, “Austin Animal Shelter Adoption Predictions - dataset by rdowns26,” *data.world*, 21-Aug-2017. [Online]. Available: https://data.world/rdowns26/austin-animal-shelter.

[10] S. Normando, C. Stefanini, L. Meers, S. Adamelli, D. Coultis, and G. Bono, “Some factors influencing adoption of sheltered dogs,” *Anthrozoös*, vol. 19, no. 3, pp. 211–224, 2006.

[11] S. Protopopova and C. Wynne, “The effects of social training and other factors on adoption rates of shelter dogs,” *PsycEXTRA Dataset*, 2012.

[12] Soowhan, “Pet (Cat and Dog) Adoption Speed Prediction (Kaggle PetFinderMy) Part 1. EDA,” *Medium*, 28-Mar-2019. [Online]. Available: https://medium.com/@thom42/pet-cat-and-dog-adoption-speed-prediction-kaggle-petfindermy-part-1-eda-631f9538cde6.

[13] “The Humane Society of the United States releases seventh annual Horrible Hundred report,” *The Humane Society of the United States*. [Online]. Available: https://www.humanesociety.org/news/humane-

[14] Nate Nickel, "Animal Care and Control Adopted Animals", bloomington.in.gov, 13-Apr-2018. [online]. Avalible: https://data.bloomington.in.gov/dataset/adopted-animals

[15] Gloria G. Guzman, "Household Income: 2019", census.gov, Sep-2020. [online]. Avalible: https://www.census.gov/content/dam/Census/library/publications/2020/acs/acsbr20-03.pdf

[16] Adam McCann, Financial Writer, "Most & Least Educated States in America", WalletHub, 16-Feb-2021. [online]. Avalible: https://wallethub.com/edu/e/most-educated-states/31075