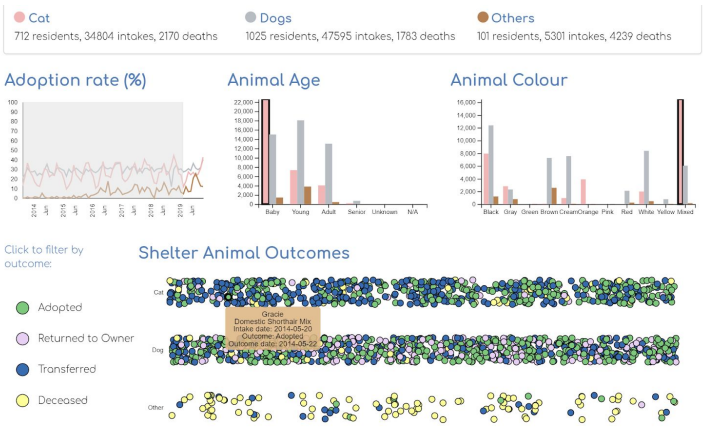
CPSC436V PROJECT MILESTONE 3

Animals of the Austin Animal Centre

Michael Zhang [redacted], Polly Tang [redacted], Eris Lam [redacted]

1. OVERVIEW



This project features data from the Austin Animal Centre, which intakes 20,000 animals annually and whose work is a leading factor behind Austin, Texas' status as the "largest no-kill community"[1] in the US. Since 2011, the Centre has been able to maintain an extremely high survival outcome; its current save rate is 97%[2].

This colour deficiency-safe tool was created with the shelter's managers and communications team in mind. Using three distinct interactive visualizations and a statistics panel, the app presents data on animal outcomes collected from October 2013 to December 2019.

Users can identify trends in historical data, such as fluctuation of adoption rates by time of year or animal type, as well as comparison of animal demographics in terms of age and colour. This gives the users the opportunity to gain a better understanding of adoption preferences and focus on efforts to boost the chance of survival for historically overlooked animals due to characteristics such as age or colour.

2. DATA

2.1 Description of Data

The dataset used in this project can be found on Data World (link: Austin TX Animal Center Stats - dataset by rebeccaclay). The dataset is updated live as the Austin Animal Centre processes new intakes and outbound animals. At the time of writing, it contains 116,079 intakes and 116,264 outbound records. For the purposes of this project, we include data from October 2013 to December 2019. The dataset is in the form of two tables, one for intake records and one for outbound records, which we joined using each tuple's unique animal ID. Attributes relevant to the project are listed below.

Attribute	Type	Scale/Cardinality (Raw)	Scale/Cardinality (Processed)
ID	categorical	116,079	Unchanged
Name	categorical	~ 100,000	Unchanged
Breed	categorical	2565	Unchanged
Colour	categorical	585	Binned to 11 categories: 'Black', 'Gray', 'Green', 'Brown', 'Cream', 'Orange', 'Pink', 'Red','White', 'Yellow', 'Mixed'
Timestamp	quantitative	October 1, 2013 00:00 -	Unchanged

		December 31, 2019 23:59	
Age at intake	quantitative	1 day - 22 years, and unknown	After binning - becomes ordinal: Baby, Young, Adult, Senior
Age at outcome	quantitative	1 day - 22 years, and unknown	After binning - becomes ordinal: Baby, Young, Adult, Senior
Sex	categorical	5 - Intact female, spayed female, intact male, neutered male, unknown	Unchanged
Intake type	categorical	4 - stray, lost, surrender, wildlife	Unchanged
Condition	categorical	10 - aged, behaviour, feral, injured, medical, normal, nursing, other, pregnant, sick	Unchanged
Outcome	categorical	9 - adopted, died, disposal, euthanasia, missing, relocate, return to owner, return to owner - adoption, transfer	Unchanged However, chart 2 contains sampled data and in order to reflect the overall distribution, omits 'missing', of which there were only a few rows 110,000+

2.2 Data Preprocessing

Data preprocessing was done via an R script and some additional Javascript. The preprocessing step included binning the animal types to 'dog', 'cat' and 'other', binning the primary colour type of each animal (see R script for the binning specifics from 585 values to 11), handling missing data and ambiguous categories, and converting age strings (for example -- 4 weeks, or 3 months) into categories (for example -- young or senior).

This preprocessing step meaningfully groups outcomes, ages, and colours into broader categories that can be communicated more expressively in our visualizations. For example, animals recorded as chocolate or liver in colour would be re-recorded as brown to reduce clutter without detracting meaning from our visualization.

3. GOALS & TASKS

Shelter managers are the primary target users of this project. The goal of the project is to enable shelter managers to gain insight into trends and patterns behind shelter animal outcomes such that they may develop strategies to maintain or even increase animal survival rates.

Questions that the tool aims can answer include: what kinds of animals are more likely to be adopted? Are there seasonal patterns that the shelter may better prepare for? Are there specific demographics of animals with exceptionally poor survival rates? Knowing this information allows shelter managers to plan targeted public engagement campaigns, such as "Kitten season coming soon in June - adopt now!" and "Give mixed coloured animals a chance!". In addition, the tool can highlight animal age or colour groups that may need more help than adopters can provide -- in these cases, they can concentrate resources on developing more partnerships with other facilities that can offer specific animal groups sanctuary.

Specific tasks that users may use this app to accomplish includes:

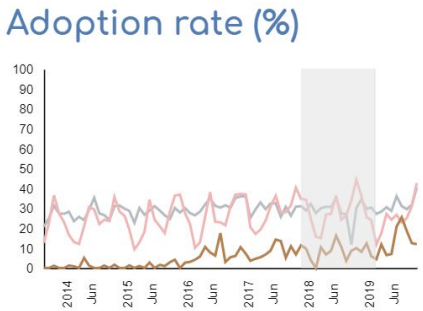
- Discover extremes and trends (are there periods with particularly low adoption rates, or is this seasonal?)
- Query by attributes (timeframe and outcome type)
- Compare attributes (outcome and animal type)
- Derive statistics (adoption rates per animal type)
- Identify dependencies (ie .between adoption rates and animal characteristics)

4. VISUALIZATION

This visualization presents data on the animal outcomes for Austin Animal Centre from October 2013 to December 2019. It contains three views in a dashboard-style format to allow the user to see changes, filtering, and highlighting across the different charts without scrolling.

4.1 Chart 1: Adoption rates line graph

Chart 1 presents an overview of adoption rates over time, with data separated into three series based on animal type: dog, cat and 'other'. The chart also serves as the primary interaction tool for the user to control the time range of interest. Using the brush, the user can choose to select granular monthly data, or brush a large timespan to see more general, overview data in the statistics panel and the other charts.



The brush is locked to one-month increments. This was done to facilitate the filtering of data in other charts; as well, it minimizes the number of x-axis ticks to every six months. This is important due to limited horizontal space.

Data presented in this chart include adoption rate (0-100%) and time (October 2013 - December 2019). Using this chart, users are able to identify how adoption rates have fluctuated over time (task abstraction: present and compare trends). We chose to use a line chart as we are representing temporal and continuous data. We used colour hue to differentiate the categorical attribute of animal type as it is one of the most effective channels for categorical data due to ease of perception and no learning curve.

4.2 Chart 2: unique view of animal outcomes

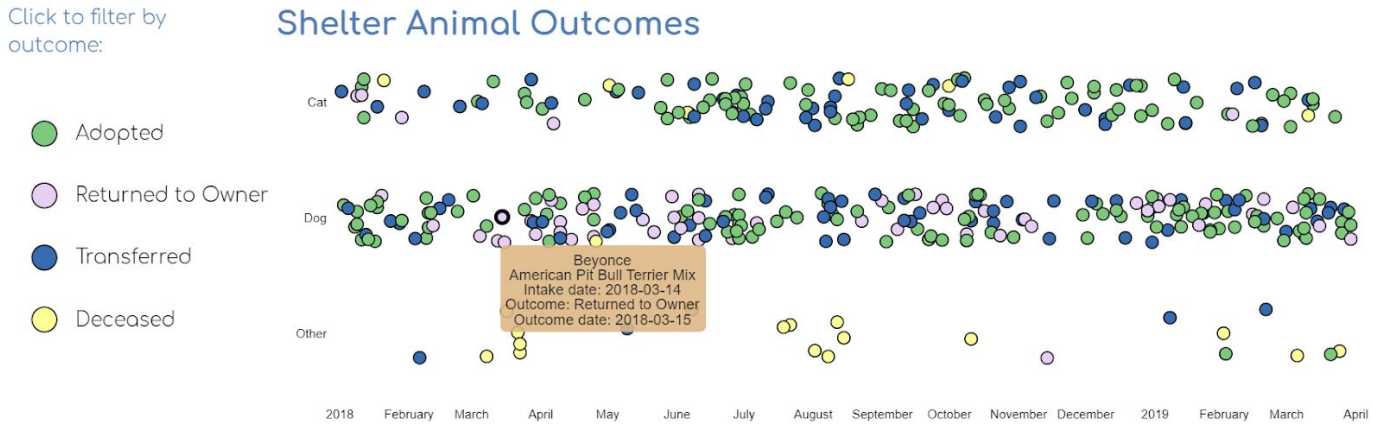


Chart 2 is a faceted small multiple chart. It shows each animal that was in the care of the Animal Centre during the date range in question. This chart shows how the outcome (four categories: adopted, returned to

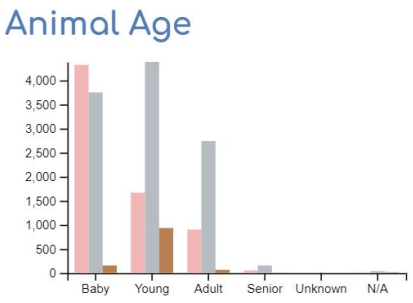
owner, transferred, deceased) for animals varies based on its type (identify trends), and represents each animal as a mark in the chart to visually present the number of lives passing through this centre. By hovering on an individual dot, a tooltip appears to provide details specific to the individual animal.

The legend serves as a filtering widget: by clicking legend label, the corresponding outcome is filtered out. This is extremely useful if the user wishes to make comparisons such as specific outcomes and animal types, or if they wish to analyze trends within one specific outcome type. Having the legend serve double-duty as both a legend and a filtering tool is an effective way to use screen space economically.

Note that this chart contains a subset of the entire dataset, which contained information on over 110,000 animals. This was done to address computational and space limitations. Since the subset reflects the distribution of the overall dataset, the user is still able to achieve the intended tasks of comparing outcomes.

We chose to encode the data with point marks, using hue to differentiate between outcome categories. This is because colour is one of the most effective channels for encoding categorical data. Since hue was used to encode categorical data in charts 1 and 2, replicating this channel provides a sense of continuity throughout the app. Animal type is encoded in the y-axis and outcome date is encoded in the x-axis. Some y-positioning jitter is introduced to minimize overlapping of dots.

4.3 Chart 3: histograms of animal age and colour groups



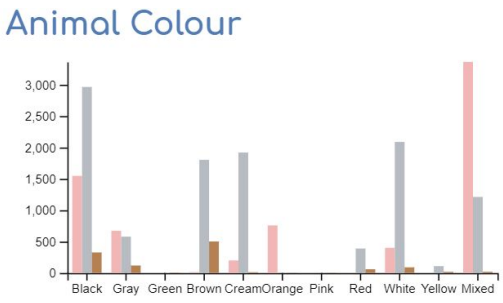
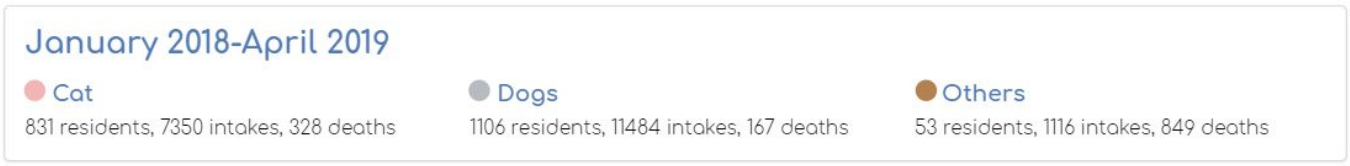


Chart 3 shows the distribution of the animals by age (six categories: baby, young, adult, senior, unknown and not applicable) and colour (11 categories: black, gray, green, brown, cream, orange, pink, red, white, yellow, and mixed). Chart 3 implements linked highlighting with chart 2 so that users can see how each individual animal fits into the distribution of age and colour (present distribution, identify trends) of animals at the shelter. Data timeframe is controlled by the brush in chart 1 and users can see what attributes are more common in general, or during a certain time frame (ie. more baby animals during breeding seasons). The counts for each animal are encoded in a histogram format as it is the simplest to visually present and compare trends across groups, and data from the large number of animals can be aggregated into a single view. This chart utilizes the same colour hue encoding as chart 1 so that there is a unified representation for animal types across the entire visualization.

4.4 Other features



The app also features a statistics panel which supplements the adoption rate presented in Chart 1. It is updated via Chart 1's brushing feature. Positioned directly above Charts 1 and 2, which focus on animal type, the statistics panel also serves as a legend.

This app is colour-deficiency safe. Colours were initially chosen to echo the primary colour of the Austin Animal's Centre's logo and to convey a sense of friendliness and warmth. However, this was changed to prioritize colour-deficiency friendliness. Using the Colorblinding Chrome extension, we rendered the web app in protanopia (red-blind), deuteranopia (green-blind) and tritanopia (blue-blind) modes to ensure that the charts are still differentiable. We reverted the app's background colour to white, rather than cream: this makes the look more compatible with the greater range of distinct colour hues needed to support colour-blind friendliness. As well, we changed the body text from dark brown to black for added contrast, based on web accessibility guideline recommendations[3].

5. REFLECTION

5.1 Overall development

Our goals have remained largely the same since Milestone 1 - to support the shelter managers in identifying trends in the animals' data as they relate to animal attributes such as intake time, age and colour. Arming the shelter managers with this information allows them to distribute resources accordingly to boost chances of survival for as many resident animals as possible.

Although we had initially considered also designing the application with potential adopters in mind by showing animals currently in the shelter, we chose not to pursue this further as our dataset is static and wouldn't be updated frequently with incoming and outgoing animals.

5.2 Changes in visualization goals

Our visualization goals did not undergo much change throughout the project: we had set out to represent shelter resident information by focusing on adoption rates, animal outcomes and relationship between adoptions and animal age/colouring. The main change was using a subset of the data for Chart 2 rather than the full set. From a visual perspective, this is not a significant change since the user is able to achieve the same intended tasks. Furthermore, because we are limiting the data points to render, we were able to justify a more detailed tooltip which includes more details that may be of interest to the user, such as the breed of the animal. Initially, we had limited the tooltip to simply including an animal's name and age, as we were concerned that Chart 2 would be too cluttered.

The other main visual change is incorporating colour-deficiency friendly schemes directly into the design of the visualization, rather than creating it as a togglable feature.

5.3 Changes in technical goals

One of our technical goals was to render the full dataset in Chart 2. This would have involved rendering up to approximately 101,000 data points. The intent was to represent every animal that had come into the care of the shelter by showing specific information such as name, age and breed. However, we were unable to

accomplish this without sacrificing useability, as the resulting visualization was both cluttered and slow to render. Instead of pursuing this, we decided to find a balance between a comfortable user experience that also accomplishes the project's intended goals. See Section 5.4 for details on how this was achieved.

Another technical change was to incorporate UI widgets *within* chart elements. This was done largely to address the limited screen real estate. We replaced a year filter by including a brush in Chart 1. Beyond space saved, this is also an improvement as the user may freely choose different time frames, rather than being limited to year intervals. We replaced an outcome dropdown filter by making Chart 2's legend into a clickable filter.

5.4 Feasibility of original proposal

Overall, the original proposal was achievable. Our main roadblock was the sheer volume of data that needed to be presented in Chart 2. This proved challenging in terms of both screen real estate and performance of calculations, filtering and rendering.

Our team took some time exploring alternatives. One solution involved limiting the brush extent to one fiscal quarter (ie. 3 months). However, we decided against this option as it inhibits the user from analyzing and exploring the data using custom timeframes or brushing over larger time spans for overview data. We decided to simply use a subset of the data for chart 2. This subset reflects the overall distribution of the dataset and therefore allows the user to accomplish the same goal of viewing some of the individual animals that were in the care of the shelter at the timeframe in question.

5.5 Things that could not be implemented

As detailed in Section 5.4, we were not able to include all data points in Chart 2. We are nevertheless satisfied with our solution as it still allows the user to accomplish the project's intended tasks and goals.

None of the optional stretch goals detailed in Milestone 1 were implemented:

- The colour-deficiency safe toggle was no longer necessary as we changed our chart colours to be colour-deficiency friendly
- Filtering/browsing by year UI widget was replaced by the brush functionality in Chart 1
- A glyph-based chart highlighting survival rates of different animal types was replaced by the statistics panel

All stretch goals were replaced by features that reflect our project's evolving design choices. Ultimately, the final set of features produce a cohesive and user-friendly experience.

5.6 Doing things differently

We feel that one design spec we should have clarified at the beginning of the project is the target screen size. This was only considered after Milestone 2. The original design was sketched from the perspective of our laptops, but in reality, many office desktops have larger screens. If we were to confirm this, we would be able to better leverage the full data set and create a more powerful dashboard:

- The current charts would serve as overviews and we would include additional panels for the user to explore the dataset in greater detail
- We would have enough space to render the full dataset in chart 2
- Instead of aggregating birds, rabbits, livestock, wildlife as "other", we would provide separate views for this animals as they likely require very different strategies for improving survival outcomes

- We would include breed information as breed is often another preconceived notion that drives unbalanced adoption rates
- We would not have to aggregate colour combinations as aggressively. Currently, we only look at primary colours, but different breeds have very specific colouring breakdowns that can also affect skewed adoption rates

Since the dataset contains tremendous potential as a public engagement tool to inform the public about overlooked animals, we would also consider a public-facing version. We would achieve this by including data of animals available for adoption and adapt chart 2 so that animals are sorted by their length of residency. In other words, users would be able to browse adoptees by length of stay and, thanks to linked highlighting with Chart 3, gain a larger picture view about the animal's adoptability, thus creating an emotional impact.

This would be an alternative to standard adoption tools which are designed for browsing available animals by age and colour. This model does not challenge preconceived notions of animal age/colouring and can be a factor behind skewed adoption rates in shelters.

6. TEAM ASSESSMENT

6.1 Milestones 1 and 2:

Eris drafted up the majority of Milestone 1 and studied the raw dataset to determine how certain large-cardinality attributes could be binned. She also prepared the layout of the visualization, implemented the statistics panel, implemented the JavaScript data pre-processing and completed the overview line chart.

Michael wrote the R script for data pre-processing, and fleshed out ambiguities in the raw dataset. In addition, he completed the unique animal sample view..

Polly studied the raw dataset and determined the binning for animal type and age. In addition, she implemented the age and colour group histograms and drafted the majority of the writeup for Milestone 2.

6.2 Final Milestone:

The final milestone involved more close collaboration, with most tasks being the responsibility of more than one person. A rough breakdown is as follows:

- Polly and Eris worked on the writeup together.
- Michael and Eris explored different ways to address the unique animal view's performance issues. Part of the work involved additional Javascript preprocessing and testing a cross filter library.
- Michael worked to improve the layout of chart 2, and converted the legend into a UI widget for filtering based on outcome.
- Polly implemented bidirectional linked highlighting between the animal sample view and the age/colour group histograms. She also reorganized the visualization's lead-in text to reflect the latest changes.
- Eris restructured the project so charts and the statistics panel sits within one view.
- The group as a whole debugged the project, performed various QA and testing tasks, and also explored different colour options to make the project colour-blind friendly.

References

1. http://www.austintexas.gov/page/no-kill-plan
2. https://www.austinpetsalive.org/about/our-story
3. https://webaim.org/articles/contrast/