**DECLARATION:** I understand that this is an **individual** assessment and that collaboration is not permitted. I have read, understand and agree to abide by the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at http://www.tcd.ie/calendar. I understand that by returning this declaration with my work, I am agreeing with the above statement.

# 1 Visualisation Preview

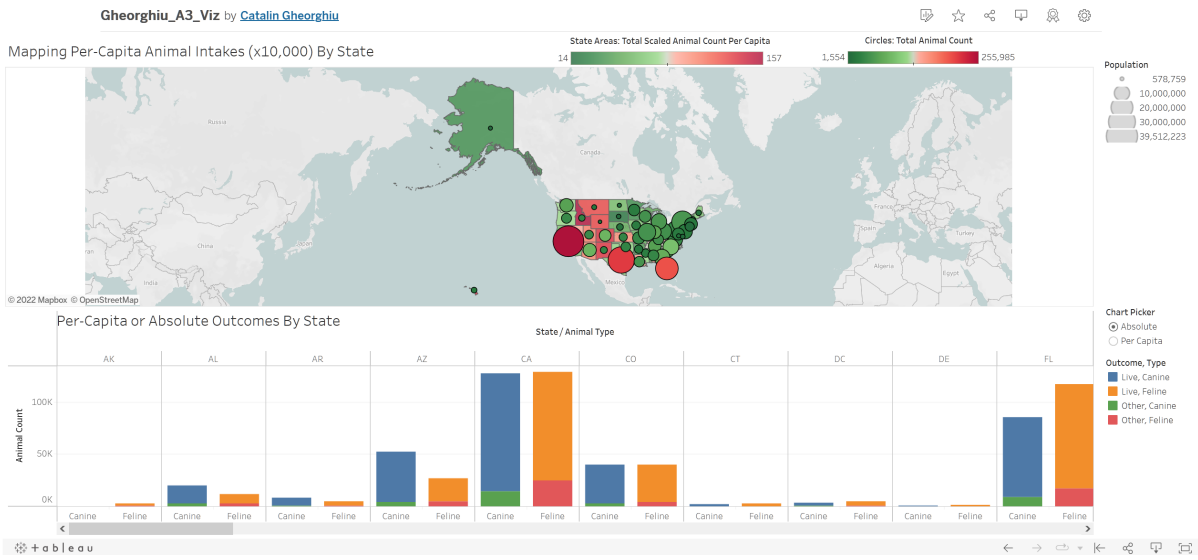



Figure 1: Visualisation of shelter animal counts across the United States, comparing state performances, investigating links with population, and comparing the outcomes for cats and dogs between states.

# 2 Tools, Data, and Pre-Processing

The main dataset is the open data from the Shelter Animal Counts national database (Shelter Animals Count National Database, 2022). The number of feline and canine shelter intakes and a breakdown of outcomes are provided for each US state (including Puerto Rico and the Virgin Islands) on a yearly basis between 2019 and 2021. All attributes are discrete and quantitative, only representing counts for different combinations of animal type and category: 5 intake types, 5 "live" outcome subcategories, and 4 "other" outcome subcategories per animal type yield a total of 28 columns. States and years represent the categorical dimensions that uniquely identify a data entry. All data is freely available for academic purposes on the database's website:

`https://www.shelteranimalscount.org/data-request-free-download/`.

I used a secondary dataset to extract the latest population estimates for each state. This is the US Census Bureau "Population, Population Change, and Estimated Components of Population Change: April 1, 2010 to July 1, 2019" (U.S. Census Bureau, 2022). Only 2019 estimates were extracted for the subset of states present in the shelter animal counts dataset. The data is freely available on the census bureau's website:

`https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html`

The only pre-processing tool I used is Excel. I aggregated the yearly breakdown of the shelter animal counts by averaging all values and rounding the results to keep working with natural numbers only. Furthermore, I reduced the number of dimensions by summing up all the counts in the sub-categories of the "live" and "other" outcomes respectively, resulting in separate total outcome counts for canines and felines. I re-structured this dataset to obtain individual rows for each outcome - animal type combination, in every main US state (thus removing Puerto Rice and the Virgin Islands for a better clustered map view). Finally, the population was appended and I computed per-capita values for all the animal counts.

The next sections only involve the visualisation, made entirely in Tableau Desktop. Further using Tableau's infrastructure, I present the chart by embedding the dashboard in Tableau Public, the result of which is shown in figure 1. At the end of the report, a larger image is shown.

# 3 Idioms, Tasks, and Encoding Channels

There are 3 questions I wanted to explore, which naturally became the tasks of the visualisation:

1. How do states compare in terms of total shelter intakes?
2. Are cats and dogs treated differently in different states?
3. Is population a factor that influences animal shelters?

Note that I deemed intakes unworthy of analysis, as the total intakes often differed from the total number of animals processed (defined as the sum of all "live" and "other" outcomes). The number of animals processed seemed more important, both for the performance of the shelter and the comparison of outcomes between animal types. As a result, the chosen idioms and encoding channels (Munzner, 2014) are tailored to tackle the tasks through the lens of differences in total outcomes.

The top-side map chart addresses the first and third tasks, comparing states by two metrics: the absolute total number of animals processed and the number of animals processed per capita. The former is encoded using the circles in the center of the states, while the latter is encoded in the area of the state itself. Position is, of course, the main channel used to differentiate between states, as the values corresponding to the state are superimposed over its geographical location on the map. I use colour and brightness to encode the differences in absolute and per-capita animal counts, with low numbers tending towards darker shades of green and higher numbers tending to darker shades of red. Furthermore, the size of the circles that reflect absolute animal counts encodes population, leading more populous states being marked by larger circles. I used the same colour palette for both absolute and per-capita animal counts to easily see where they tell the same story - and where they say different things. For instance, while the largest states (CA, TX, FL) also have the most animals taken into shelters, they control the situation quite well given how many people - and therefore pets and animals in total - live there. In contrast, most mid-western states and Hawaii show a large number of sheltered animals per capita, even though they sit comfortably low compared to other states in absolute terms.

The bottom-side grouped stacked bar chart addresses the second and third tasks, comparing the number of "live" and "other" outcomes between cats and dogs in every state. States are ordered alphabetically by their acronyms along the horizontal axis, while the vertical axis shows the number of animals - either absolute or per capita, depending on the version of the chart toggled using the chart picker on the right. Colour encodes the combination of animal type and outcome as per the legend, with complimentary colors contrasting outcomes while colour warmth contrasts animal type. The horizontal scroll bar allows the user to explore the full range of states while keeping the chart readable. By looking at total bar height rather than the individual components, the position channel communicates the difference in total processings of cats and dogs. The possibility to change between the absolute and per-capita version of the chart helps the user understand the importance of considering per-capita values over absolutes, as not only do states that looked problematic in the absolute version turn out to be doing well otherwise, the sheer amount of animals in the largest states creates enough outliers to remove a lot of nuance from the comparisons between smaller states - their values are all dwarfed. Finally, additional information is simply offered through text when hovering an area of the map, a circle, or a bar, respectively.

# 4 Novelty and Critical Analysis

The first point of novelty is the combination of idioms on the dashboard: the map chart and bar charts work together to show how the relative performance of states' animal protection changes when population is taken into account. Secondly, the map chart combines several encoding channels, including position to indicate states, shape to differentiate between per-capita and absolute values, size to express differences in population, colour to differentiate between free and busy states, and brightness to express more nuanced differences in the number of animals processed by the states' shelters. Thirdly, the dashboard is highly interactive: the map chart can be panned and zoomed at will to close in on specific states (Washington D.C. is quite hard to see without it!) and the chart picker allows the user to comfortably look at two different stacked bar charts without squeezing them together.

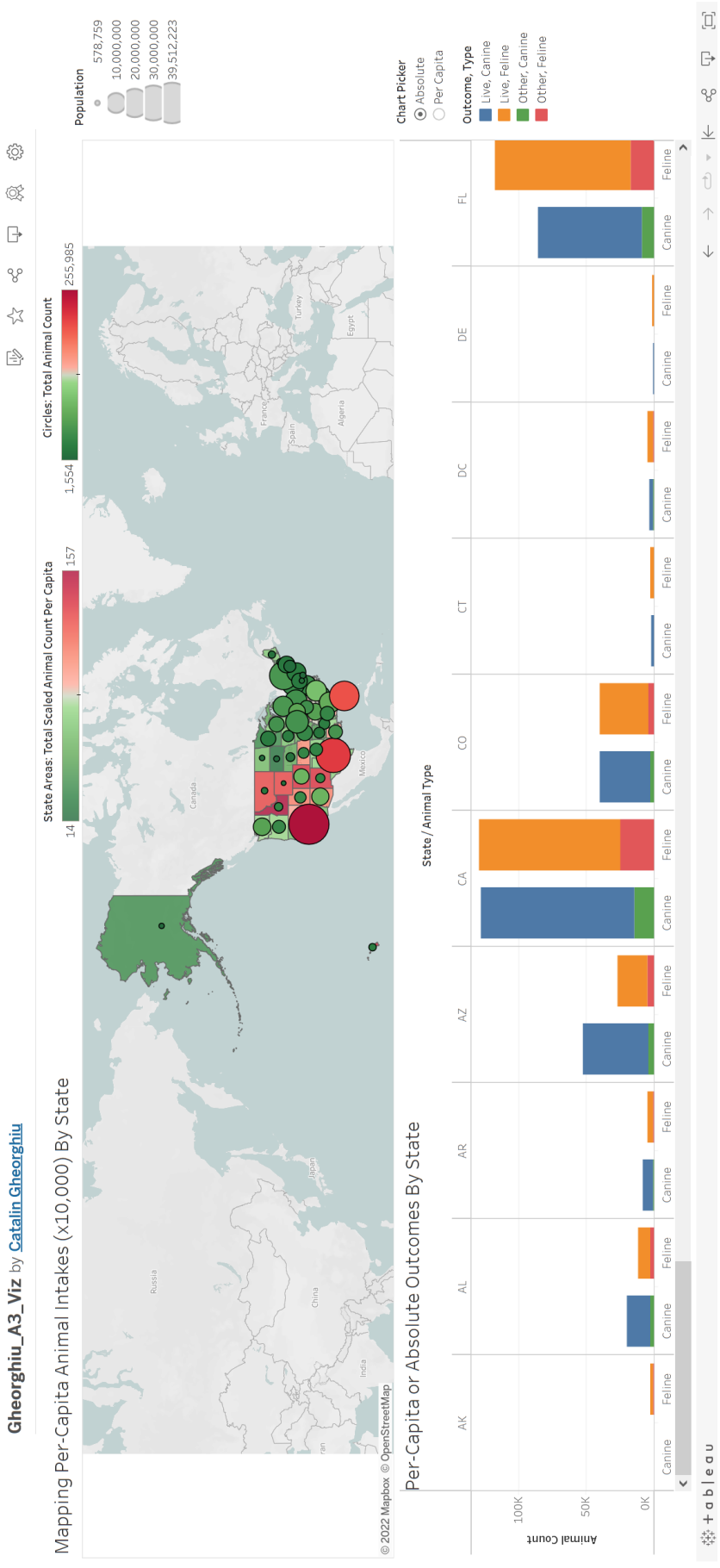
The visualisation's greatest strength is the interplay of the two idioms: the bigger picture of the map chart can be investigated in-depth using the stacked bar chart - for instance, one can see Hawaii's problem with animals processed per-capita on the map and see that the reason is an abnormally high number of cats processed, and unfortunately, the abnormally high of bad outcomes for these cats. The charts' interactivity options play a key part in helping the user look around in their chose level of detail in order to make such observations. Its greatest weakness, on the other hand, is the large amount of aggregation that it is built on. Many possible insights, be they into the differences over years, or more nuanced distinctions of particular outcome subtypes, or even how intakes translate to outcomes are lost as a result. It would be interesting, for instance, to see how often it is that animals brought in for euthanasia end up with a different fate, or if cats are lost by shelters more often than dogs because of their more inquisitive and independent nature. Moreover, the fact that one must scroll the stacked bar chart can be inconvenient; given more space, the possibility to permanently have a big picture that one zooms into rather than scroll through the zoomed-in version may considerably improve readability.

# References

Munzner, T. (2014). *Visualization analysis and design*. CRC press.

Shelter Animals Count National Database (2022). 2019-2021 Intake/Outcome Data. `www.shelteranimalscount.org/data-request-free-download/`.

U.S. Census Bureau (2022). Population, Population Change, and Estimated Components of Population Change: April 1, 2010 to July 1, 2019 (NST-EST2019-alldata). `www2.census.gov/programs-surveys/popest/datasets/2010-2019/national/totals/nst-est2019-alldata.csv`.


Gheorghiu_A3_Viz by Catalin Gheorghiu
Mapping Per-Capita Animal Intakes (x10,000) By State
Circles: Total Animal Count
255,985
1,554
State Areas: Total Scaled Animal Count Per Capita
157
14
Population
578,759
10,000,000
20,000,000
30,000,000
39,512,223
Chart Picker
Absolute
Per Capita
Outcome, Type
Live, Canine
Live, Feline
Other, Canine
Other, Feline
© 2022 Mapbox © OpenStreetMap
Per-Capita or Absolute Outcomes By State
State / Animal Type
Animal Count
100K
50K
0K
AK
AL
AR
AZ
CA
CO
CT
DC
DE
FL
Canine
Feline
+ableau