

Monash University: Assessment Cover Sheet

| | | | |
|-----------------------------|--------------------------------------|--|------------------------|
| Student name | Hee | Zhan Zhynn | |
| School/Campus | | Student's I.D. number | 31989403 |
| Unit name | FIT3179 Data visualisation - S2 2021 | | |
| Lecturer's name | Dr Grace Ting Chai Wen | Tutor's name | Dr Grace Ting Chai Wen |
| Assignment name | Data Visualisation II Report | Group Assignment: No Note, each student must attach a coversheet | |
| Lab/Tute Class: 2 | Lab/Tute Time: 9am - 11am | Word Count: 1003 | |
| Due date: 18-10-2021 | Submit Date: 18/10/2021 | Extension granted <input type="checkbox"/> | |

If an extension of work is granted, specify date and provide the signature of the lecturer/tutor. Alternatively, attach an email printout or handwritten and signed notice from your lecturer/tutor verifying an extension has been granted.

Extension granted until (date):/...../..... Signature of lecturer/tutor:

| Late submissions policy | Days late | Penalty applied |
|--|------------------|------------------------|
| Penalties apply to late submissions and may vary between faculties. Please refer to your faculty's late assessment policy for details. | | |

Patient/client confidentiality: Where a patient/client case study is undertaken a signed Consent Form must be obtained.

Intentional plagiarism or collusion amounts to cheating under Part 7 of the Monash University (Council) Regulations

Plagiarism: Plagiarism means to take and use another person's ideas and or manner of expressing them and to pass these off as one's own by failing to give appropriate acknowledgement. This includes material from any source, staff, students or the Internet - published and unpublished works.

Collusion: Collusion means unauthorised collaboration on assessable written, oral or practical work with another person. Where there are reasonable grounds for believing that intentional plagiarism or collusion has occurred, this will be reported to the Associate Dean (Education) or nominee, who may disallow the work concerned by prohibiting assessment or refer the matter to the Faculty Discipline Panel for a hearing.

Student Statement:

- I have read the university's Student Academic Integrity Policy and Procedures
- I understand the consequences of engaging in plagiarism and collusion as described in Part 7 of the Monash University (Council) Regulations (academic misconduct).
- I have taken proper care to safeguard this work and made all reasonable efforts to ensure it could not be copied.
- No part of this assignment has been previously submitted as part of another unit/course.
- I acknowledge and agree that the assessor of this assignment may, for the purposes of assessment, reproduce the assignment and:
 - i. provide it to another member of faculty and any external marker; and/or
 - ii. submit to a text matching/originality checking software; and/or
 - iii. submit it to a text matching/originality checking software which may then retain a copy of the assignment on its database for the purpose of future plagiarism checking.
- I certify that I have not plagiarised the work of others or participated in unauthorised collaboration or otherwise breached the academic integrity requirements in the Student Academic Integrity Policy.

Date: 17/10/2021 Signature:  *

Privacy Statement:

For information about how the University deals with your personal information go to <http://privacy.monash.edu.au/guidelines/collection-personal-information.html#enrol>

URL: <https://zhanzhynn.github.io/FIT3179-Visualization2/>

Word Count: 1003

Domain, Why and Who

The domain of the visualization is based on the pet dogs and cats' population in the United States.

The visualization is aimed to illustrate and compare the ownership and population of cats and dogs between states in the country and to provide some insights on the popularity of each pet.

The visualization is designed for the average person. Therefore, technical background is not required to understand the visualization.

What

The datasets are Cat vs Dog Popularity in the US (Tableau Public, 2021) and US Census Bureau Regions and Division (Halpert, 2014). The combined datasets are directly related to the domain and have been wrangled by the authors, thus it can be used to construct the visualization with minimal data processing needed. However, extra calculated fields such as normalizing the pet owners to per 100 households for the choropleth map is done directly in VegaLite. The datasets were posted on Tableau Public and GitHub which are a reputable website, deeming them reliable for our visualization.

Why and How

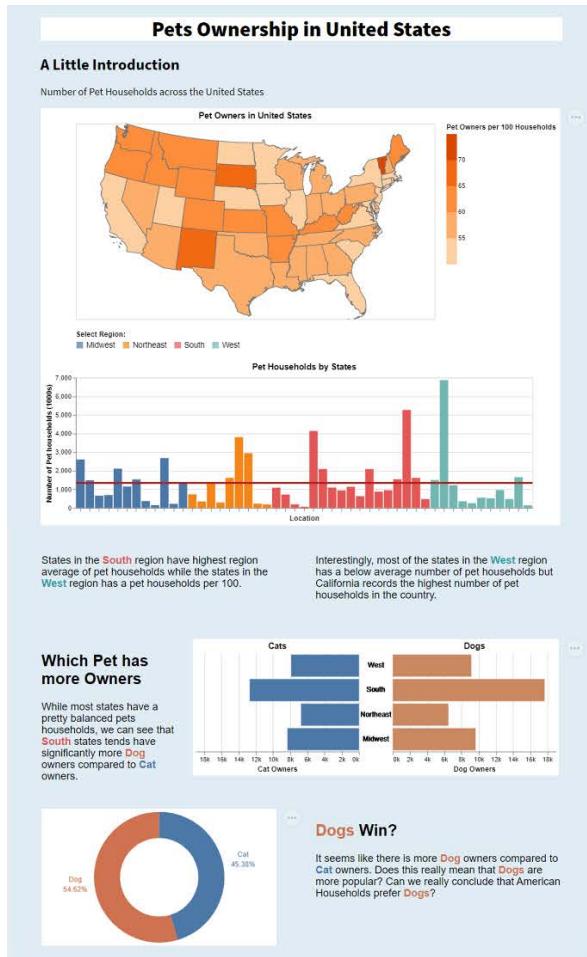


Figure 1.0a: Visualization Screenshot (part 1)

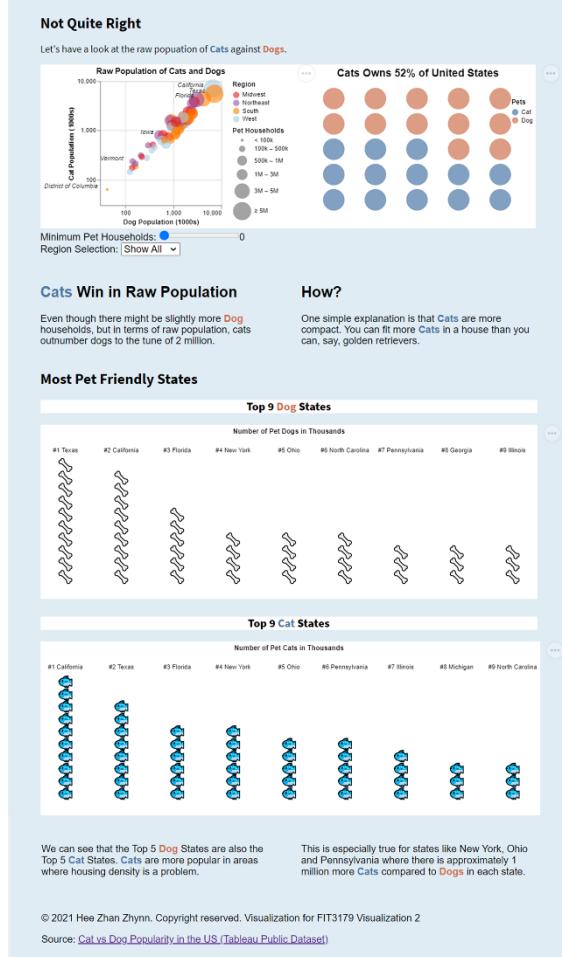


Figure 1.0b: Visualization Screenshot (part 2)

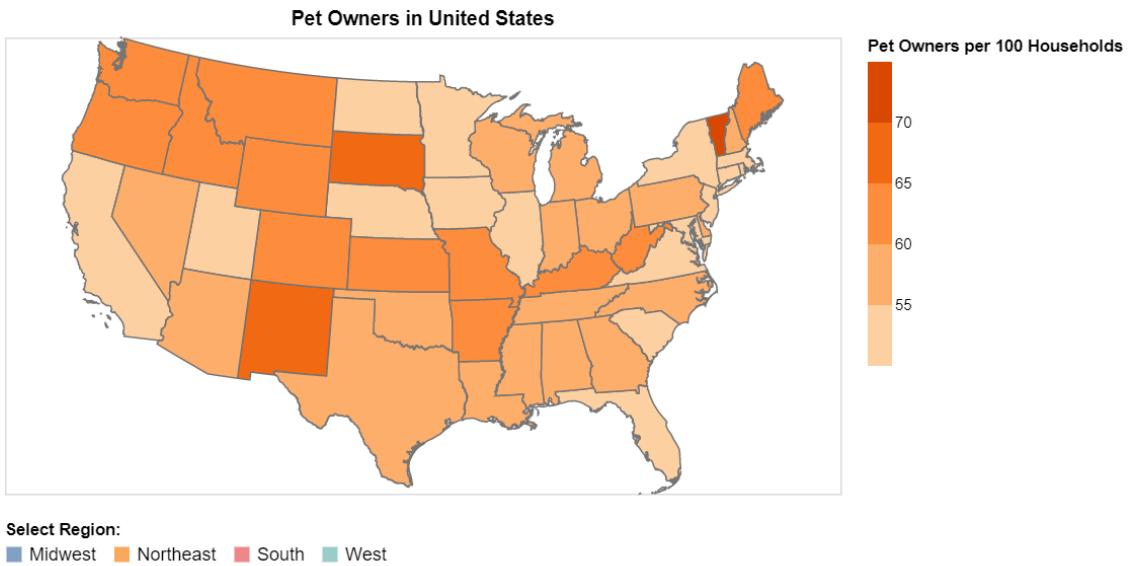


Figure 1.1: Pet Owners in the US

My design of a choropleth map in Figure 1.1 enables me to divide geographical areas and illustrate values over a normalized geographical data. It is ideal as it shows variation across the display location, that is the density of pet owners across the country.

The mark used is area while the channel used is luminance. The luminance represents the density of pet owners where a darker shade indicates a higher density.

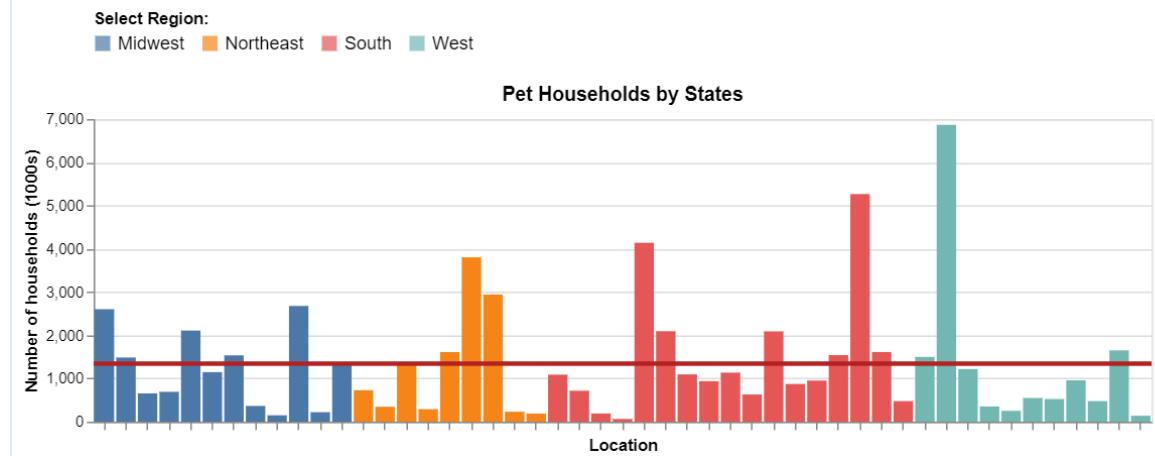


Figure 1.2: Pet Households by States

The bar chart in Figure 1.2 is used to compare the number of pet households in each state. It is suitable as the difference in values is more distinguishable, making the comparison between states easier. A red horizontal line is also used to indicate the average of the selected region.

The mark used is line while the channels are hue and length. The hue represents the region that the state is in while the spatial position indicates the number of pet households.

Figure 1.1 and 1.2 are made interactive by allowing the readers to select and highlight the chosen region, allowing them to filter and compare the numbers for selected region.

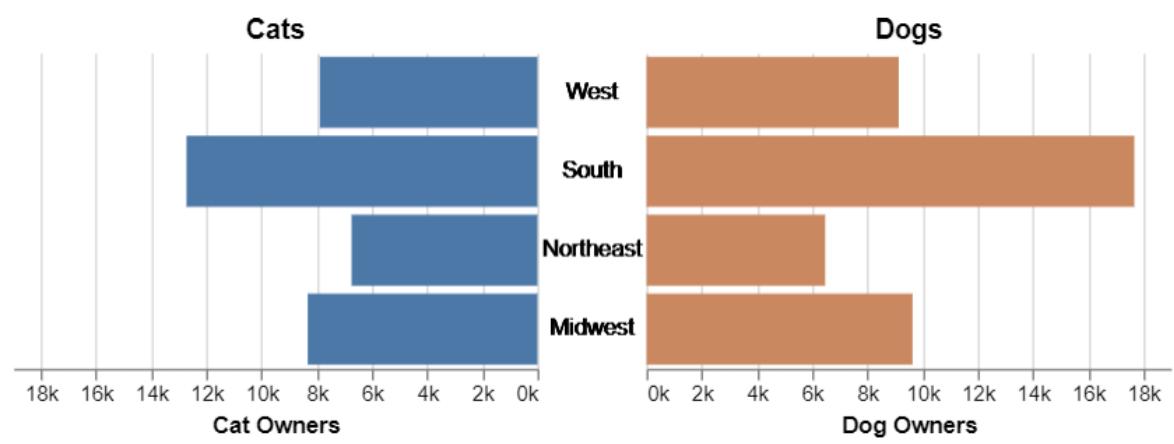


Figure 1.3: Pet Owners by Region

The population pyramid chart in Figure 1.3 compares of cats and dogs' owners based on region. It is appropriate as the focus is to compare the total number of pet owners for each region.

The mark used is line while the channels are hue and length. The hue represents the type of pets while the length represents the number of owners.

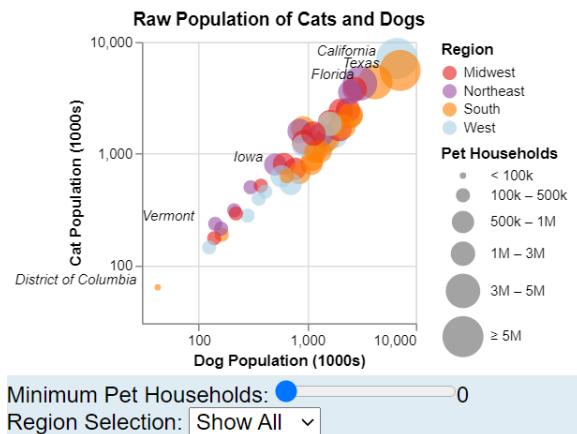


Figure 1.4: Raw Population of Pets

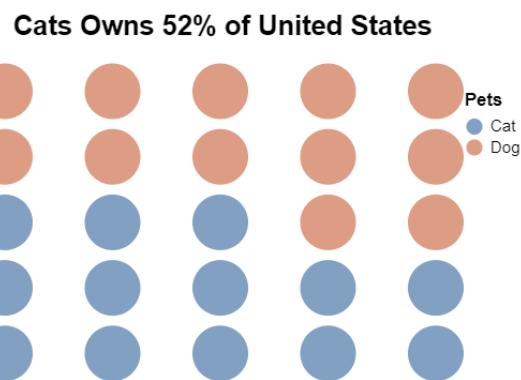


Figure 1.5: Overview Population of Pets

The scatter plot made with reference from Week 9 Studio (Monash University, 2021) in Figure 1.4 enables me to illustrate the raw population of cats and dogs in each state. It can also illustrate the number of pet households, as shown by varying the size of each dot. Interactivity is added to filter the visualization by region and minimum number of pet households.

The mark used is point while the channels are hue, spatial position and size. The hue represents the region that the state is in and the spatial position indicates the population while the size is used to illustrate the number of pet households.

The waffle chart in Figure 1.5 is used to complement Figure 1.4, it gives a representation of the proportion of the raw population of pets briefly.

The mark used is point while the channels are hue and spatial position. The hue represents the pets while spatial position indicates the proportion.

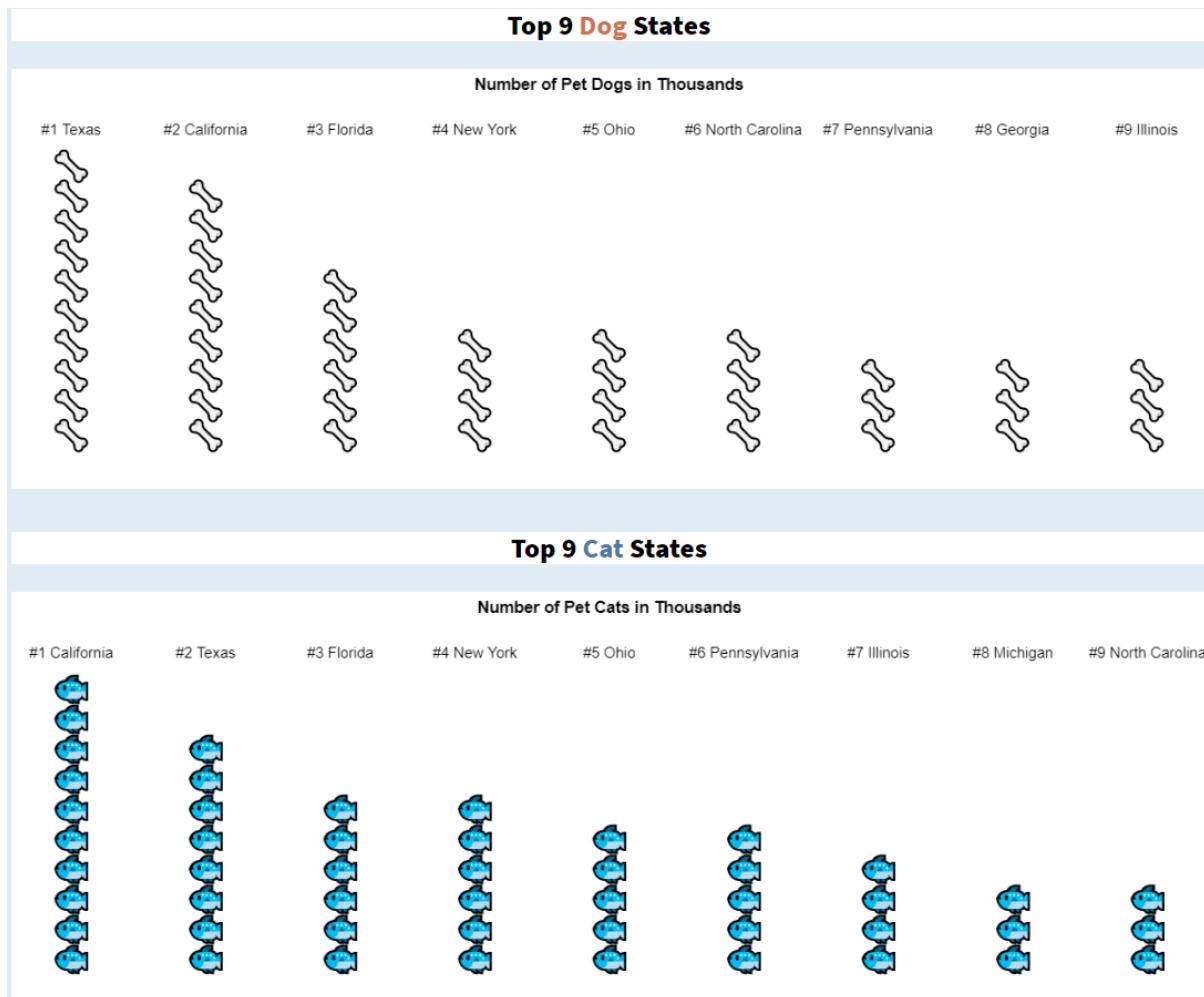


Figure 1.6: Top 9 Dog and Cat States

The chart in Figure 1.6 is known as an isotope graph. It uses symbols to indicate the proportion of the pet population for each state. It is well suited as it allows the readers to compare the population while making the visualization more interesting compared to a regular bar chart.

The mark used is point while the channels are symbols and spatial position. The symbols represent the pet while the spatial position indicates its population.

Design

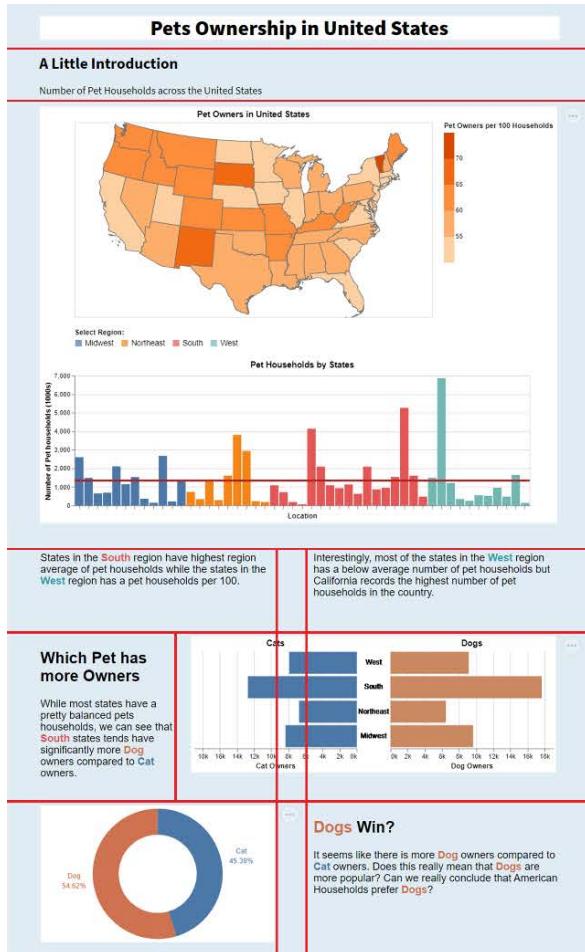


Figure 2.1a: Sightlines (part 1)

The layout is balanced and symmetrical around a central vertical axis where there are mainly two columns for each row except for larger visualizations. The sightlines are minimal and the visual center is the charts on the comparison between cats and dogs as shown by Figure 1.4 and 1.5. They are the important elements of the visualization that illustrate the popularity of the pets to the readers and are placed side-by-side to complement the visual information. Gestalt principle (proximity) is also used to group the elements together using white spaces, eliminating the need for borders.

The colour used is easy on the eyes and consistent to represent the regions and pets across the charts and text. The background is coloured to apply figure-ground technique where the title and charts are emphasized using a white background.

Figures such as charts are given more emphasis by making the size larger with saturated colour, these make them stand out compared to text which is smaller. Background colour is also used to make the charts standout more despite having a lower data-ink ratio. The subheadings are bolded and have a larger font size to indicate a higher hierarchy compared to the content.

The typeface used throughout is Sans-Serif as it is easy to read. The typeface is also suitable for annotations in charts as it is still readable even if the font size is small. The text content is also left aligned to improve sightlines and make the visualization organized.

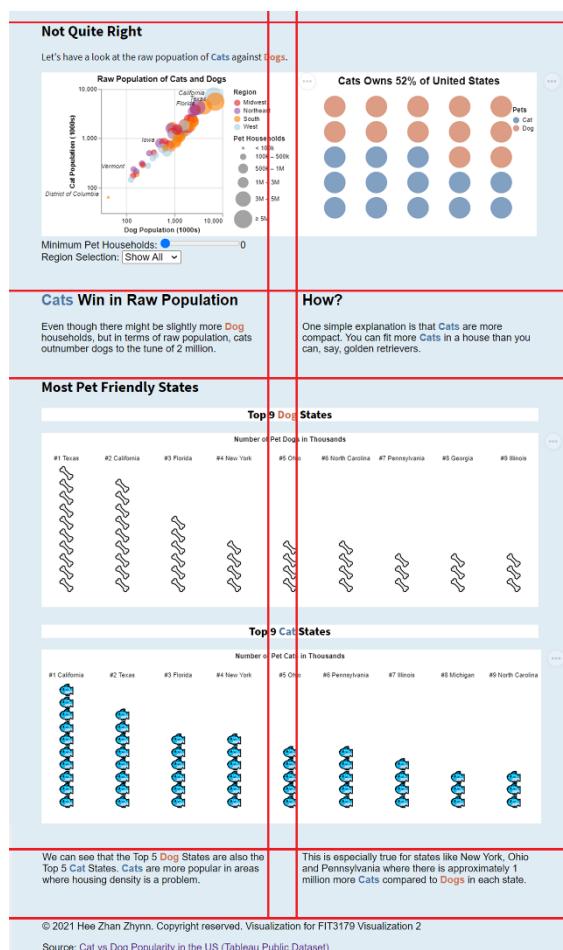


Figure 2.1b: Sightlines (part 2)

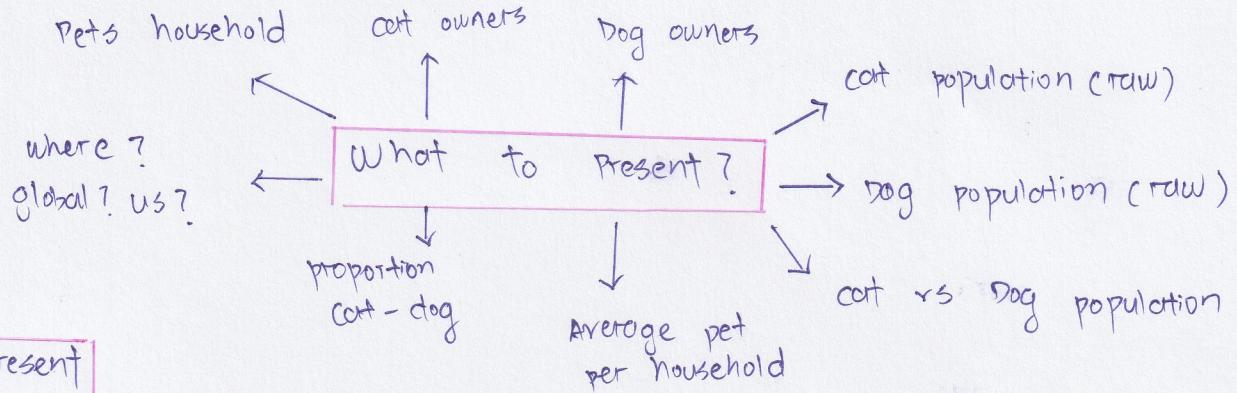
The visualization genre is magazine style and annotated charts where charts with annotations are embedded in text. Readers are guided by text and visualization. Annotations are also used in the chart to highlight important points while text outside the diagrams is used to provide the storytelling elements to the users.

References

- Harper, C. (2014). Census Region. Retrieved from
<https://github.com/cphalpert/census-regions>
- Monash University (2021). FIT3179 Week 9 Studio. Retrieved from
<https://lms.monash.edu>
- Tableau Public. (2021). Cat vs Dog Popularity in the US. Retrieved from
<https://public.tableau.com/en-us/s/resources>

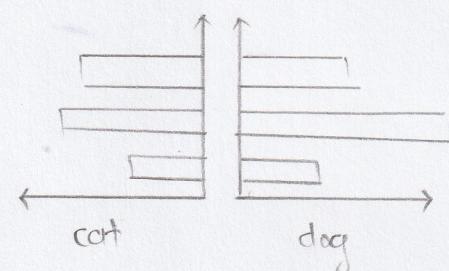
Appendix

1) Ideas



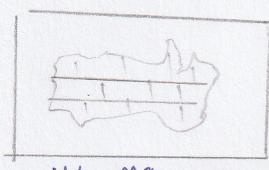
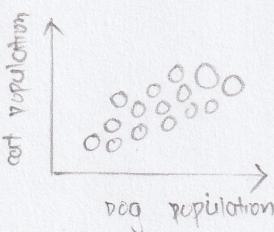
How to present

Pet Population | Owners



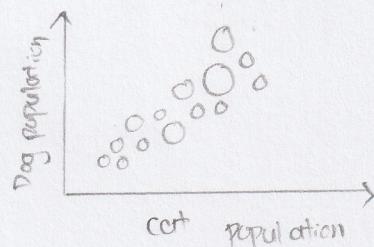
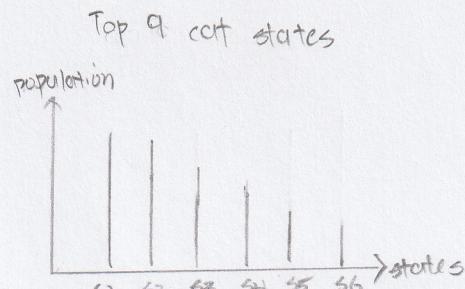
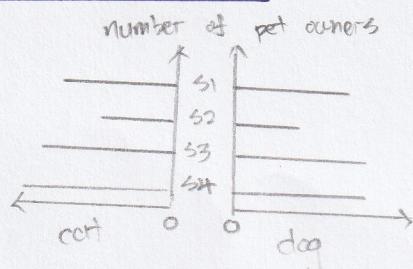
- show each region
- can compare dog & cat

- saves space, can show population & another attribute
- compare numbers of cats & dogs



- area-preserve
- projection = "alberUsa"
- normalised data
- choropleth map

4) combined & refined



2) Filter

- Focusing on a single country (United States)
- map data normalised (pet households per 100 households)
- what to show:
 - normalized pet households per 100 households
 - cat vs dog owners
 - cat vs dog population (raw)
 - proportion cat-dog owners & population
 - Top 9 dog & cat states

3) categorise

Bar chart

- population
- Top 9 states
- pet owners

Pie / waffle

- proportion of total population

Scatter plot

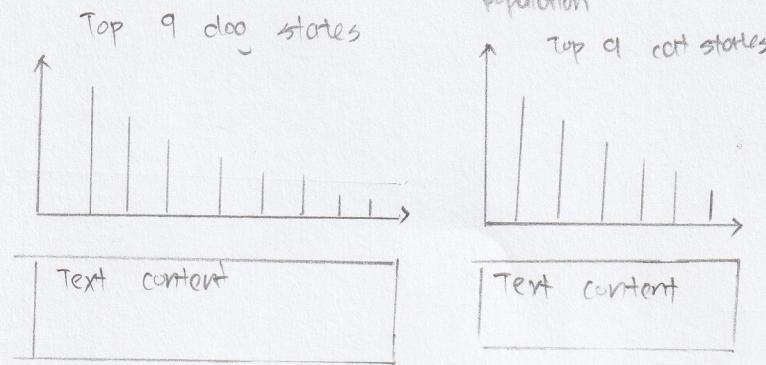
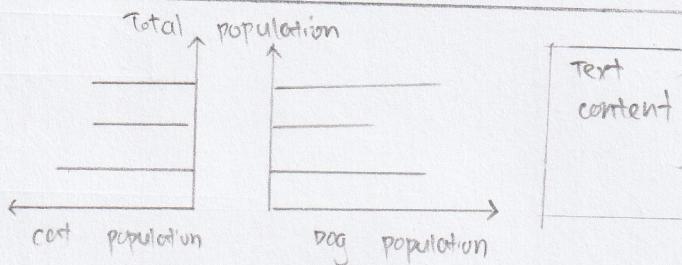
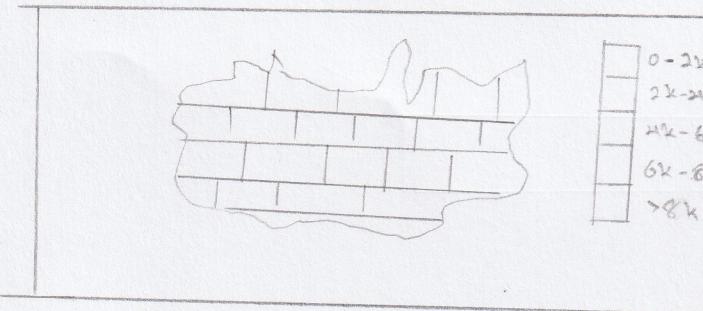
- Dog vs cat population

5) Questions

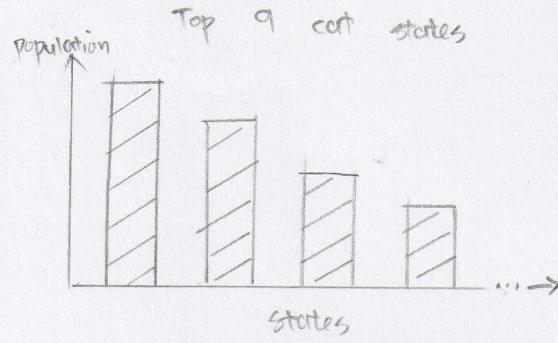
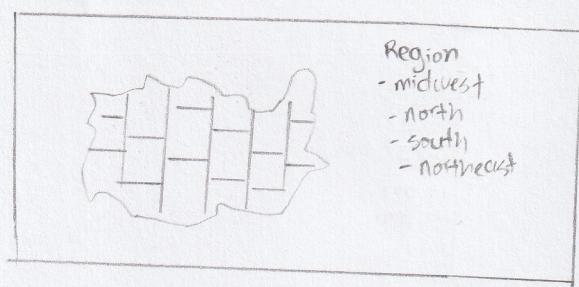
- can these charts inform the readers on the popularity of pets?
- Are they correctly used?
- Are they easy to understand
- Do I need complex technical knowledge to create the charts?

Layout

Pet Household across U.S



FOCUS



- interactivity: filter map based on selected region
- use area-preserving projection
- Data must be normalised

- Arranged in descending order
- colour for cats is consistent

Title: visualization 2

Author: Hee Zhan Zhynn

Date : 09/10/2021

Sheet : 2

Task: cats vs dogs popularity

Operations

- users hover over charts to see annotations
- map can be filtered based on region
- colour will be used to categorise dog and cat
- Read from top-to-bottom

DISCUSSION

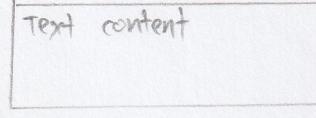
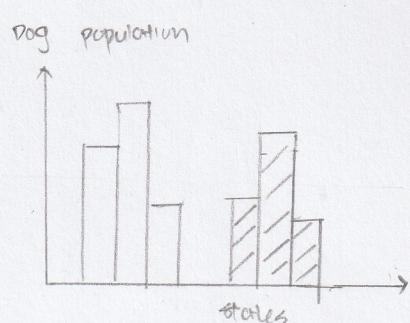
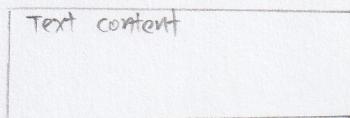
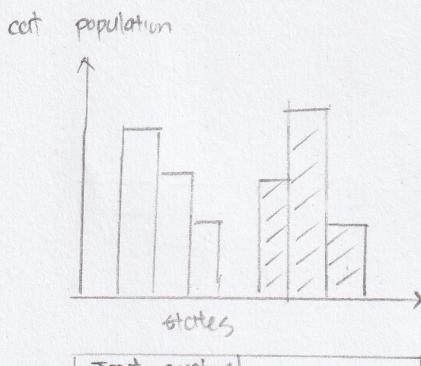
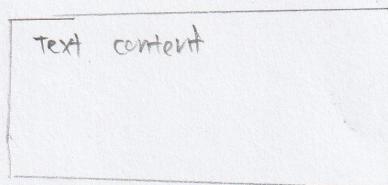
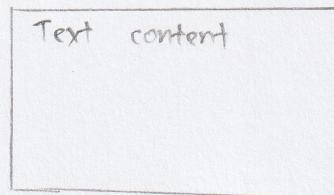
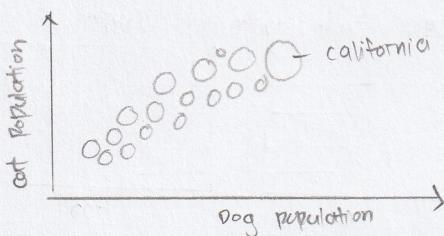
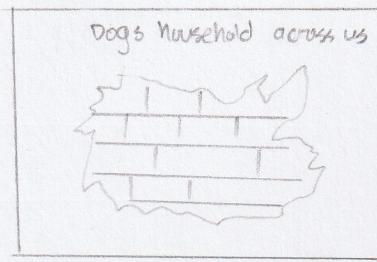
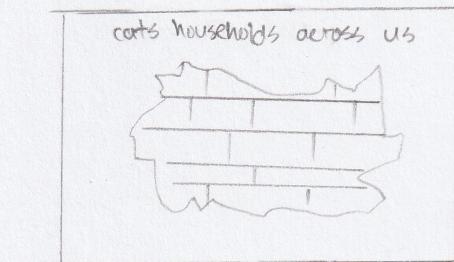
+ve:

- information seeking mantra is applied
- Easy to understand charts with good sightlines

-ve:

- Limited data shown (eg: mean pets per household not shown)

Layout



Title : visualization 2

Author : Hee Zhan Zhynn

Date : 09/10/2021

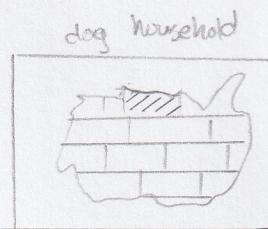
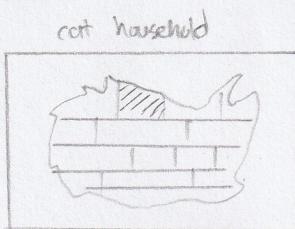
sheet : 3

Task : cats vs dogs popularity

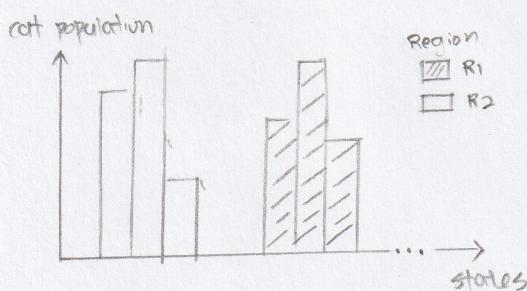
Operations

- Hover charts for annotations
- map & bars can be filtered based on region
- colours for pets & regions will be consistently used
- readers can highlight a single country on first map, then the same country will also be highlighted on the second map.

FOCUS



→ Interactivity :
select 1 country to highlight on 1st map, the same country on 2nd map will also be highlighted



- states in same region will be categorised & grouped together
- Readers can click on region legend to highlight chosen region

Discussion

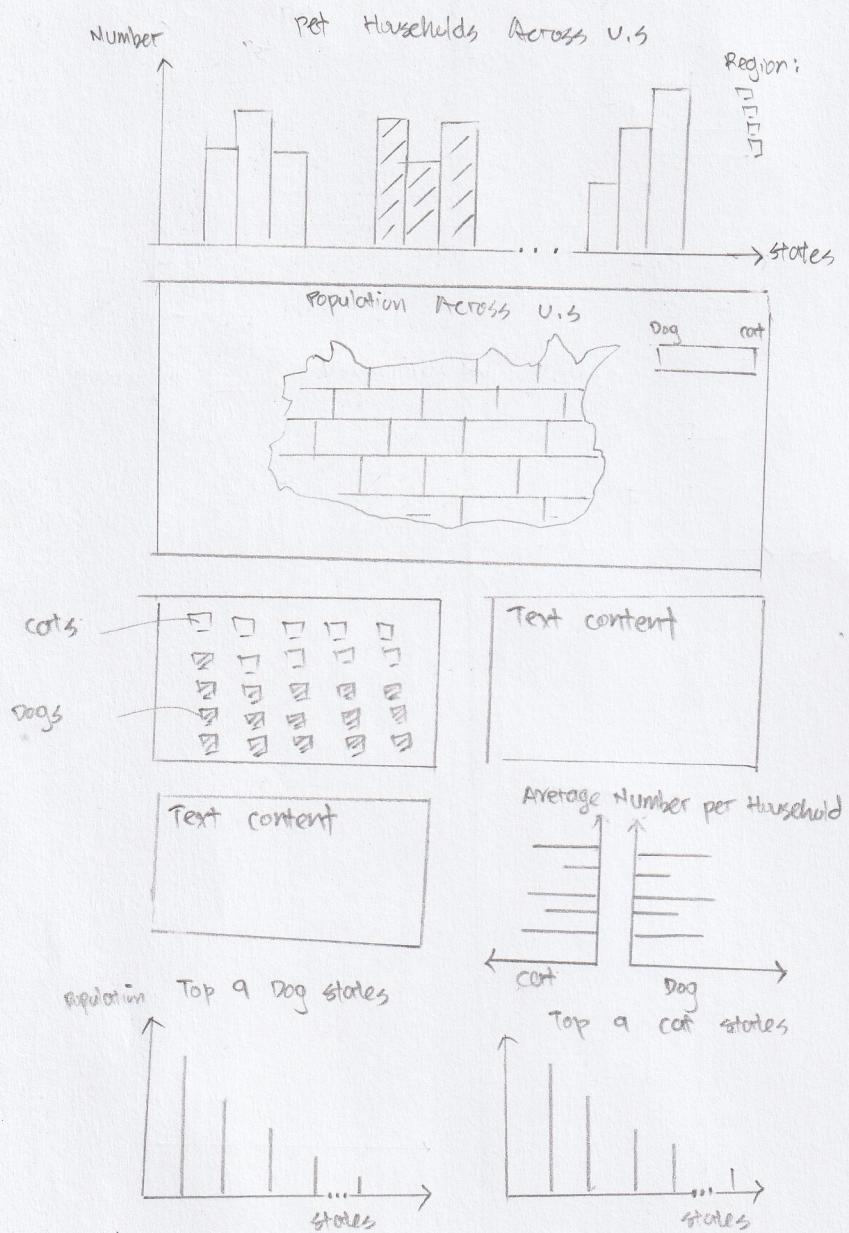
tre :

- Two choropleth maps side-by-side to view distribution of each pet across country
- Donut chart is used instead of regular pie chart
- more interactivity using map & horizontal bar chart

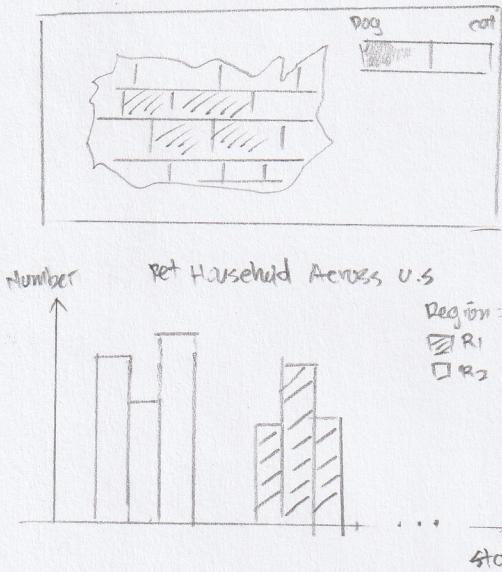
tre :

- pyramid chart showing mean number might take up space. (showing 52 states)

Layout



Focus



Title: visualization 2

Author: Hee Zhan Zhynn

Date: 10/10/2021

Sheet: 4

Task: cats vs Dogs Popularity

Operations

- linked interactivity for first bar chart and map, they can be filtered by region
- colours for pets & regions are consistently used
- Hover over charts for annotations

Discussion

+re:

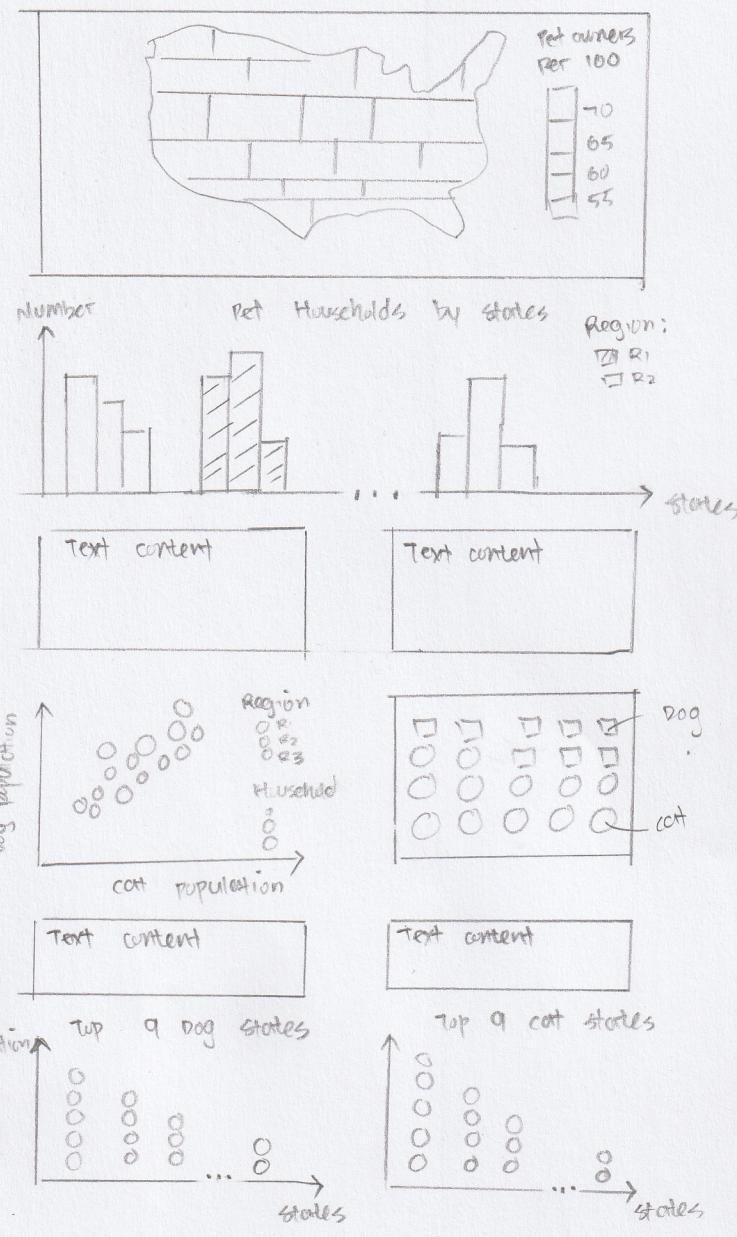
→ Diverging colour scale in choropleth map to show distribution of pets in US

→ waffle charts able to provide more appealing visual
→ good information seeking mantra

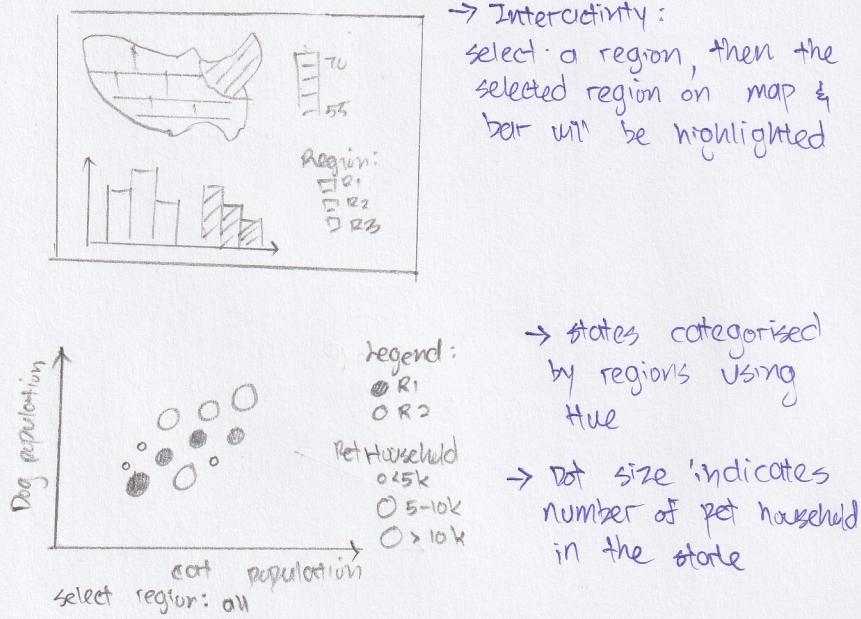
-re:

→ pyramid chart might be too large as it contains too many states

Layout



FOCUS



Title: Visualization

Author: Hee Zhan Zhynn

Date: 10/10/2021

Sheet: 5

Task: cats vs dogs popularity in U.S

operations

- First two visualization idioms are linked. Users can filter them according to regions.
- For scatter plot, readers can filter to show information for selected region
- can hover over charts to view annotations
- consistent colour to be used throughout

Detail

- Data for waffle chart and isopleth graph needs to be reorganized
- first two idioms need to be linked for interactivity - filter option
- Data for choropleth map needs to be normalised
- need to refer to studio worksheet to create interactive scatter plot
- AlbersUSA projection type will be used (area-preserving)