DATA 23700 Autumn 2023

Assignment 2: Color and cartography

Due Wed, week 5

In this assignment, students will use Altair in a computational notebook to develop a series of maps and provide detailed rationales for their design choices. The purpose of this assignment is to focus on design choices related to color scales and cartography, so special emphasis will be placed on these design choices.

Note: you are required to *work alone*.

Students should submit their assignment as a ipynb file. Be sure to evaluate all code blocks so the desired output you would like to submit is showing. Otherwise, your submission may appear to be incomplete.

Illegal animals kept as pets data

In this assignment students will explore and answer questions about a particular dataset using Altair. Specifically, students will investigate the Illegal Animals represented in the NYC 311-calls dataset. 311 calls are a matter of public record and so are recorded and made available to the public.

This dataset can be found at NYC Open Data and filtering it (using their web interface to include Complaints of type "Illegal Animal Kept as Pet" and "Illegal Animal Sold"). If that doesn't work or for some reason then you can download an older version of the dataset here.

It's a good practice to know how to download things from government websites, so the spirit of this assignment is download the above data yourself. That said, there is no penalty for using the version of the data we prepared.

Note that you will need to supplement this data with geoJSON data for New York City (NYC) and other demographic data for NYC Boroughs, which can be found online. Finding these datasets will be good practice for the project!

Technical specification

Using Altair in an iPython notebook, answer the questions below about the dataset. *Each answer should be expressed as a map or small multiples of maps* rendered using Altair whose meaning is unambiguous:

1. What type of animal is most commonly kept as a pet in each NYC zip code?
2. What part of the NYC Borough of Queens is reported as having the highest number of animal reports (of any complaint type)?
3. What is the rate (ratio) that each species is reported as appearing in each Borough relative to the number of people who live in that Borough?
   (Hint: this will require acquiring an additional dataset).

4. Within each zip code, how does the way that complaints are filed ('Open Data Channel Type') co-locate with the type of animal being reported? For this question, consider only roosters and farm animals. The chart should readily enable to obtain a qualitative answer for each area.
   (Note that there are only three ways of filing a complaint, including 'Unknown').

5. Do illegal animal sales and illegal pet keeping tend to be co-located?

Students will produce one chart (a choropleth map or a small collection of choropleth maps arranged in facets) per question above. One goal is to demonstrate what you have learned about applying color scales to data.

Each map should be preceded by a recap of the question it is answering and should be followed by a markdown text block containing a write-up about the design choices in the map. Write-ups should

1. Provide a rigorous rationale for design decision, especially design choices pertaining to
   - Color scales
   - Label placement
   - Level of geospatial aggregation (e.g., zip code vs county vs state)
   - Map projections
2. Document the visual encodings used and explain why they are appropriate for the data.
3. Briefly explain how these decisions facilitate effective communication.
4. Briefly explain how they answer the question that prompted you to create each map.