Visualizing Illegal Cheetah Trade & Seizure

Group Members: Ryan Dillon, Ally Giovenco, Pegah Emdad

Overview & Motivation:

This project aims to visualize data surrounding illegal cheetah ownership, sales, and trade. Poaching and illegal seizure of cheetahs is an issue that is widespread, however, data surrounding this problem has never been collected and concentrated in this manner before. As students interested in and passionate about conservation, we are excited to present the data surrounding this issue in an interactive way.

Questions:

Initial questions that our team asked included:

- What are the differences between illegal seizures and illegal trade?
- Are there concentrations of these different types of activity in specific locations or time frames?
- What correlations exist between the difference in time between incident date to discovery date and the rest of the data?

As the project evolved, questions asked included:

- In what specific scenarios were the greatest number of cheetahs found?

Data:

We utilized a dataset created to document the illegal trade and seizures of cheetahs between the years 2010 and 2019. This data was compiled using a wide range of sources, including field informants, veterinarians, open-source government reports, databases, online articles, and social media platforms. Some of the sources needed to be confidential due to the status of active law enforcement efforts. There are 1887 rows in this dataset. More information, including specific sources and a description of how the data was collected and processed can be found at this link: https://www.sciencedirect.com/science/article/pii/S2352340921001323

As we began to use the dataset, we realized some spelling errors existed in the sheet that may affect using the data in an automated way. It was necessary to clean the data to ensure consistency in spelling (One example was “Possession/trade” vs. “Posession/trade”). Some dates are also considered wrong in the Incident Date/Discovery Date columns where the discovery date is entered as being before the incident date. Based on exploratory visualizations, this error only occurs in about 6 rows in the dataset.

As we used the data in the visualization, we filtered out rows that contained UNK (unknown) numbers of cheetahs. Also, rows that contained Arabic words in the Year column were edited to reflect the numeric year of the incident (about 4 rows in the dataset).

Exploratory Data Analysis:

Location/Incident Type/Country Role

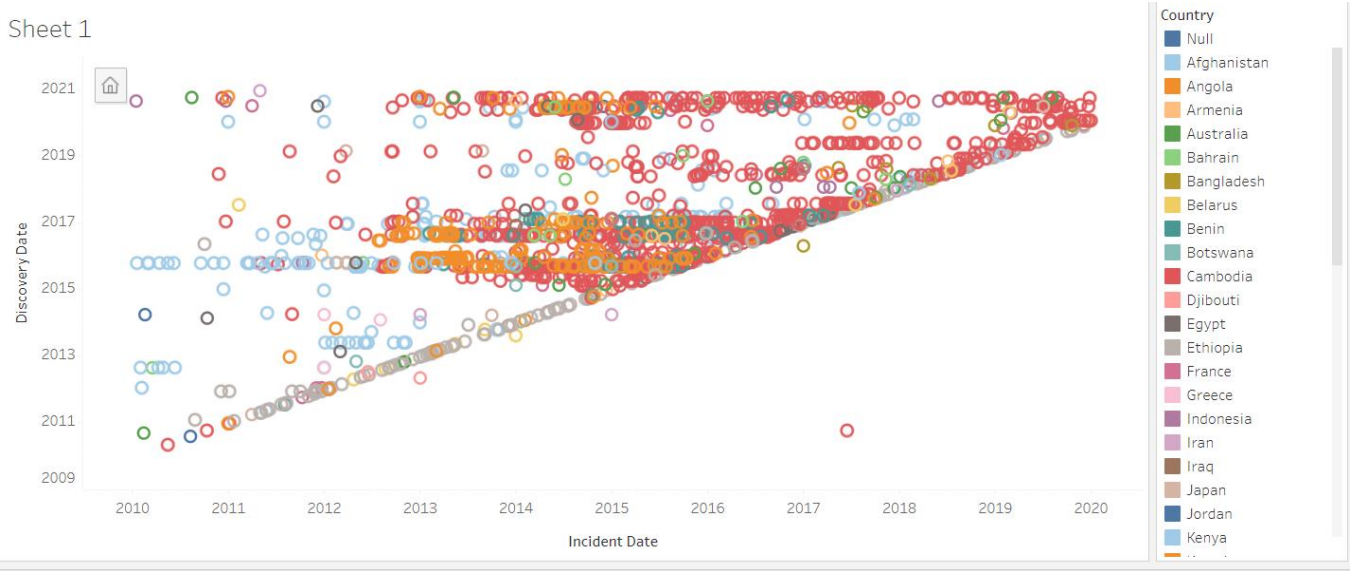
One exploratory visualization we did with the data in Tableau was looking at locations of incidents, in conjunction with the incident type and the country's role in the incident. In this visualization, the marks show the country where the incident occurred. The color shows the incident type (most of them are either seizure or possession/trade). The shape of the mark shows the country's role in the incident, either D (destination), O (origin), T (transit), or a combination of two of them.

This visualization shows that most of the incidents are concentrated in Africa and the Middle East. There are outlying incident locations in countries such as Russia, Australia, USA, Mexico, and others. We can also see that most of the incidents outside this area are usually destinations for the illegal trade, as denoted by the square shape. The visualization utilizes tooltips so the user can mouse over each point to get more information.

Incident Date vs. Discovery Date

This visualization was created to show the difference between incident dates and discovery dates. Each of the dots is colored according to what country the incident was recorded in. This visualization tells us several things about the data: first that there are some outliers in the

data where the incident date was incorrectly recorded as after the discovery date. Additionally, we can see a strong line where the incident and discovery dates are nearly the same, creating a line of slope equal to 1. All the dots that occur above this line represent incidents where time passed in between incident and discovery.

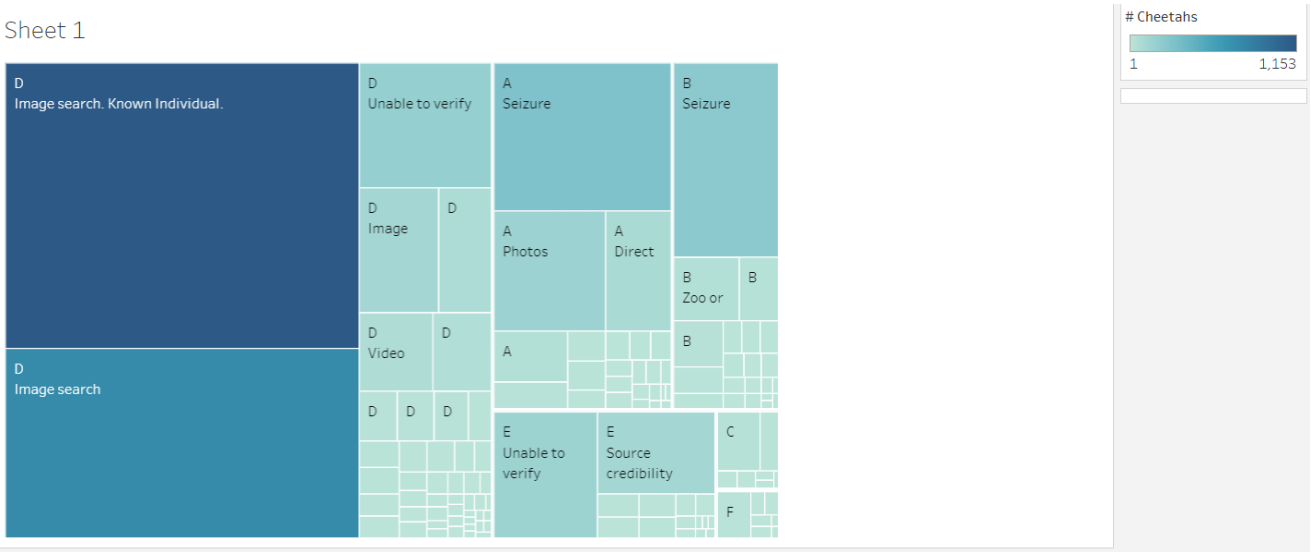


On account of the large clusters of color, we can see in which countries incidents and/or discoveries occurred around the same time. There is a large amount of pink, representing the large number of incidents in Saudi Arabia. Around the 2013-2016 time-period we also see the largest number of orange dots, showing a large uptick in incidents occurring in Angola during that time. However, these incidents seem to have been discovered in varying numbers of years after the incident date. This visualization can be further used to gain similar country specific insights.

Tree Map of Report Grading & Verification Compared to # of Cheetahs

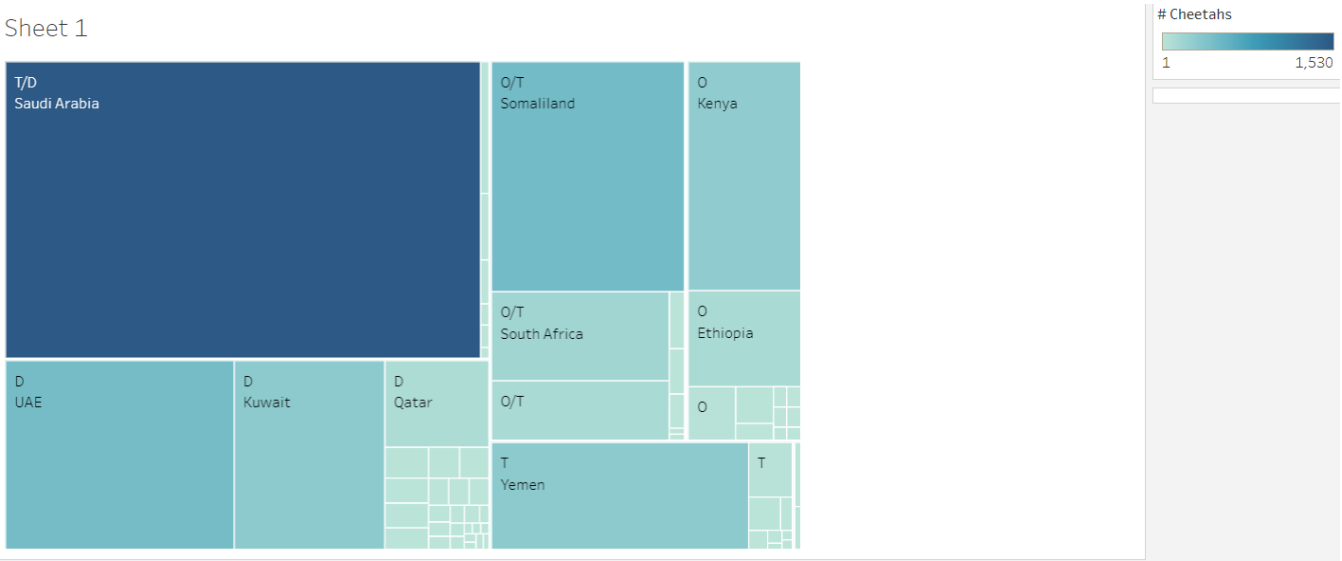
This visualization shows which report grades & verification methods correspond to the greatest number of cheetahs affected. The report grade refers to if the source was official (A), firsthand (B), secondary (C), deemed credible (D), unable to judge (E), or suspected false (F). The verification method refers to how the source was found (ex. By image search, direct, seizure,

etc.) This visualization reveals how the greatest number of cheetahs were found by sources that were deemed credible with a grade of D, but the second largest number were found using official sources that were graded A. I think it's interesting how the data was categorized into different areas based on the type of source, which is what this visualization shows.

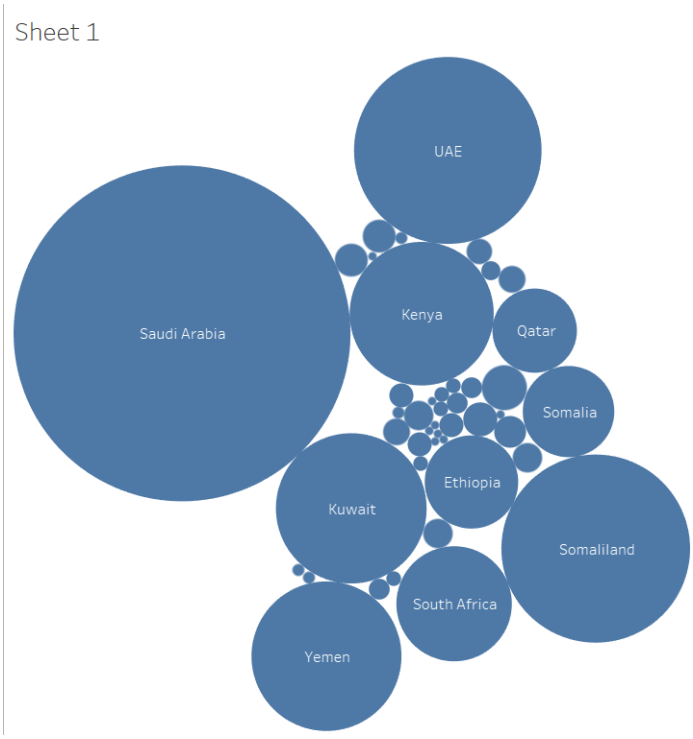


Tree Map of Country/Country Role and Number of Cheetahs Involved

This visualization uses most of the same data from the first exploratory visualization but uses a tree map to show the number of cheetahs involved in each country instead of the type of incident based on location. I think that this visual may come close to what our final design starts out as (in the form of a Voronoi tree map though). It does a good job of illustrating what situations most cheetahs are found in when it comes to location and period of transit. I do think that it may be useful to clarify whether these are seizure or possession/trade incidents, which is something that we may implement in our final designs.

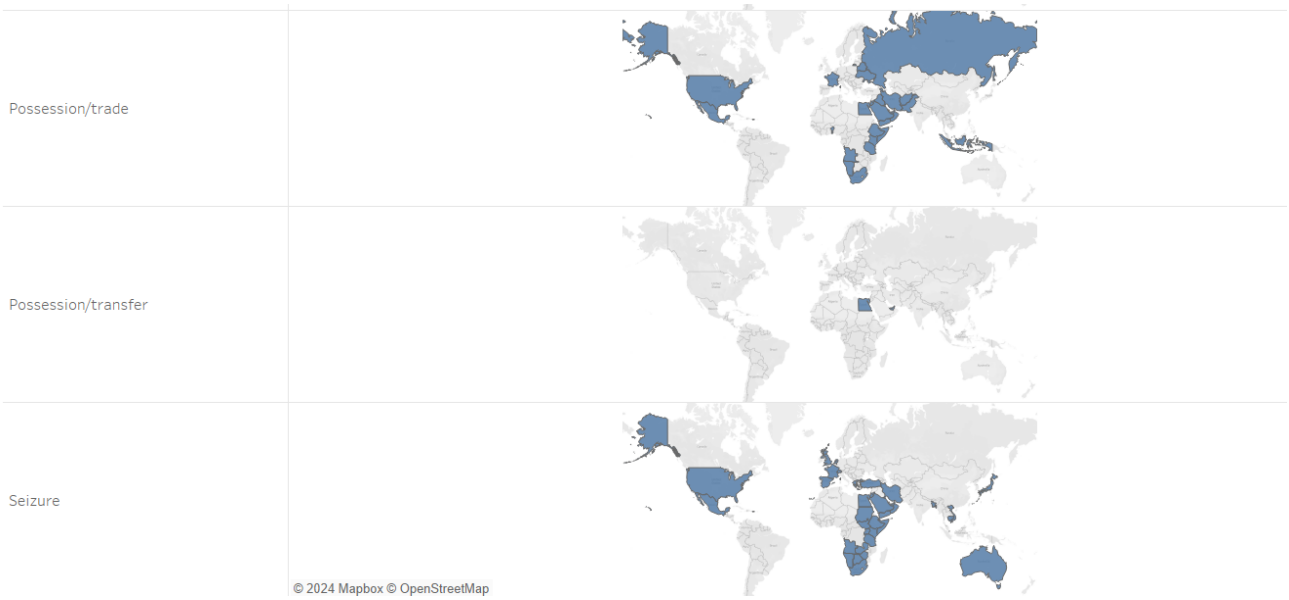


Bubble Chart of Incident Count by Country



In the above graph, we can see the count of incidents based on the number of incidents in each country. It can be concluded that Saudi Arabia has the largest number of cheetahs with the count of 1,559.

Incident Type vs. Country



After recognizing the number of incidents, this graph was visualized to filter the map based on different incident types. So, we can see different incident types based on different countries. The visualization below also shows the incident type based on each country.

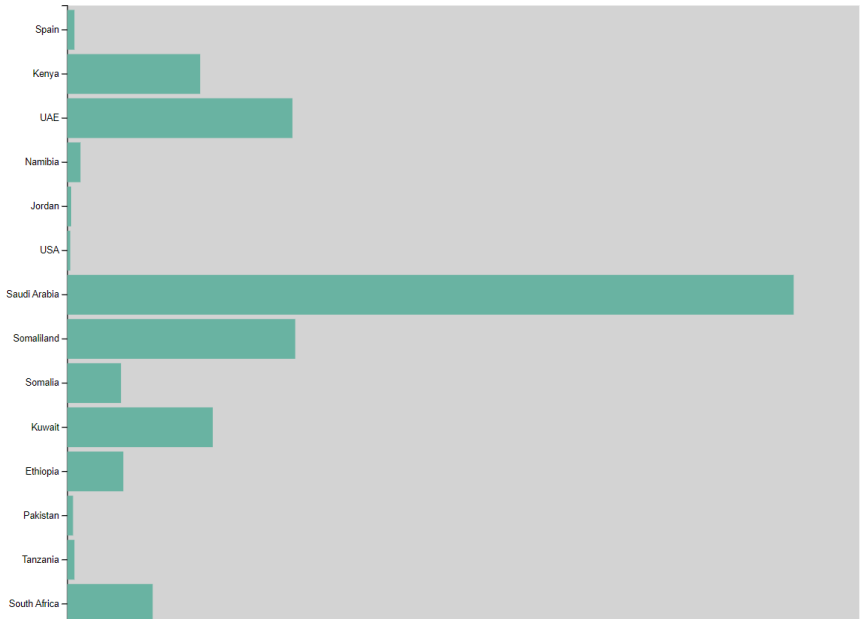
Source: World Bank Official Boundaries

Design Evolution:

Our initial plan was to create a Voronoi Tree Map to showcase parts of the data included in the spreadsheet, and to make it interactive by allowing the user to filter the data either by timeframe, country, incident type, or something along these lines.

Initial Horizontal Bar Chart

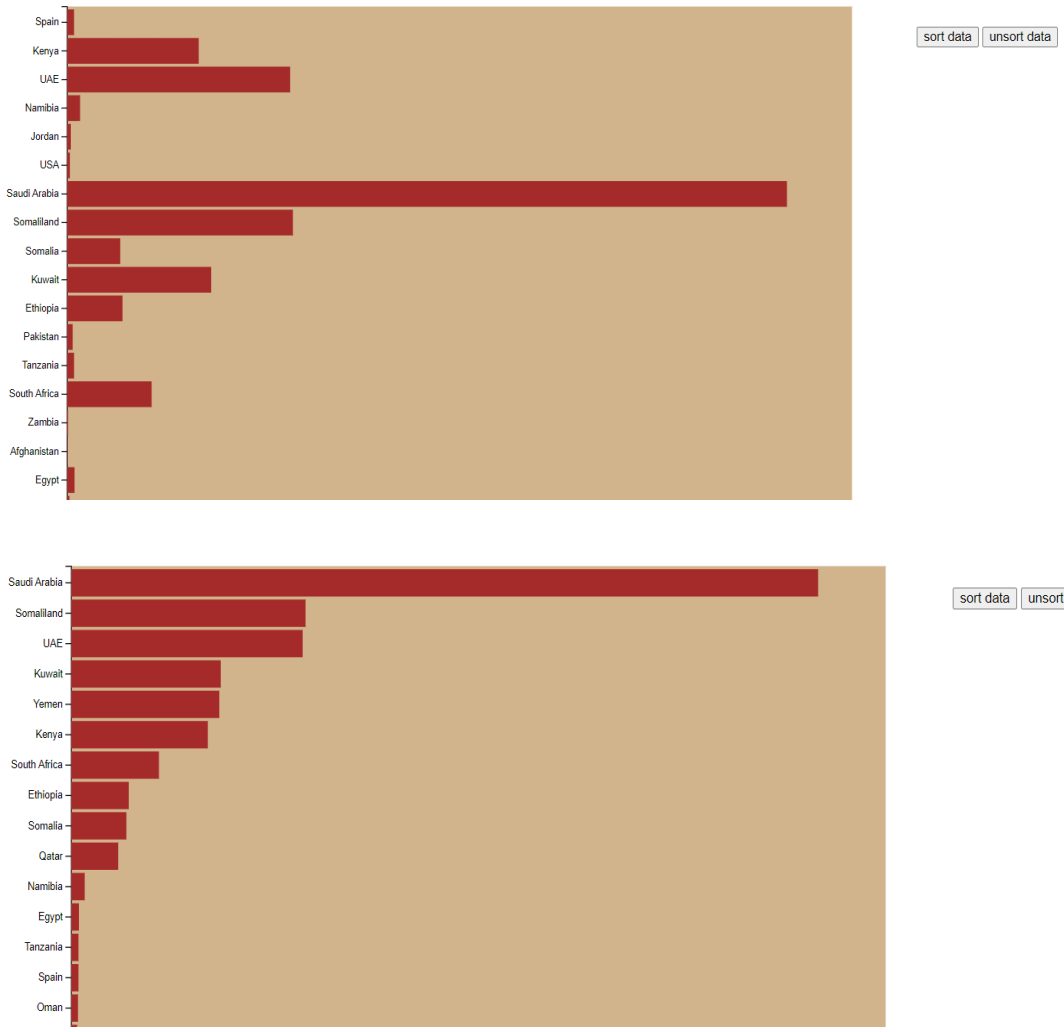
There were issues creating a Voronoi Tree Map. We started experimenting with a horizontal bar chart showing the number of cheetahs involved in incidents in each country contained in the dataset. The thought here is that the user would be able to sort the data by highest to lowest values, as well as filter the data by incident type, source reliability for the incidents, or year. Pictured below is an image of a preliminary visualization example of the horizontal bar chart.



Sort & Tooltips

One of the first things we wanted the user to be able to do was sort the data in descending order, showing the countries with the highest number of cheetahs involved in incidents. As of

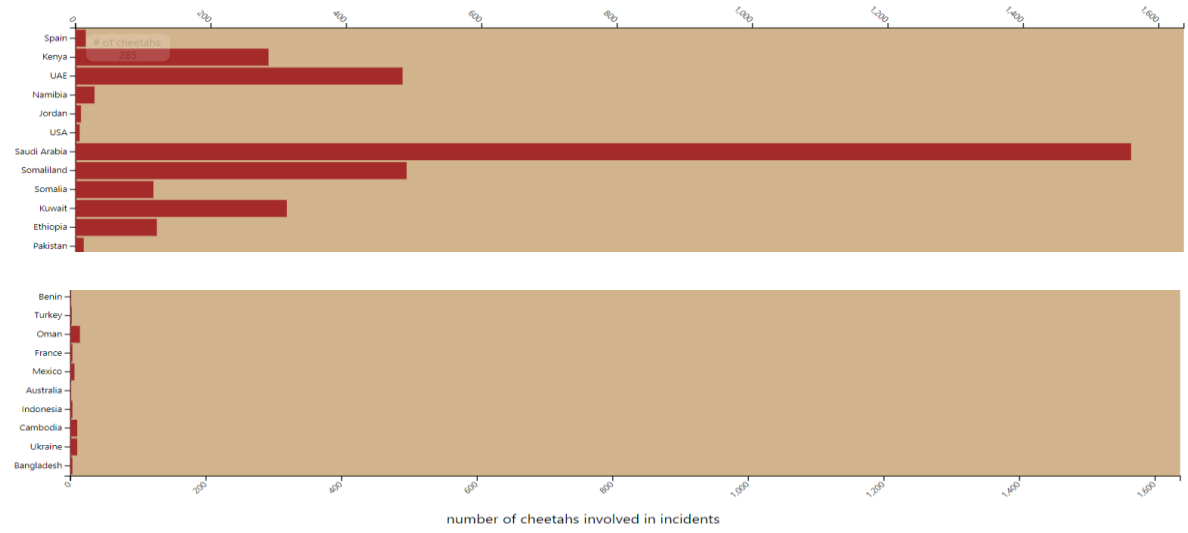
this point in the design evolution, the user had the ability to sort and unsort the data with the push of a button and could mouse over each bar for a tooltip to appear showing the exact number of cheetahs involved in incidents in that country.



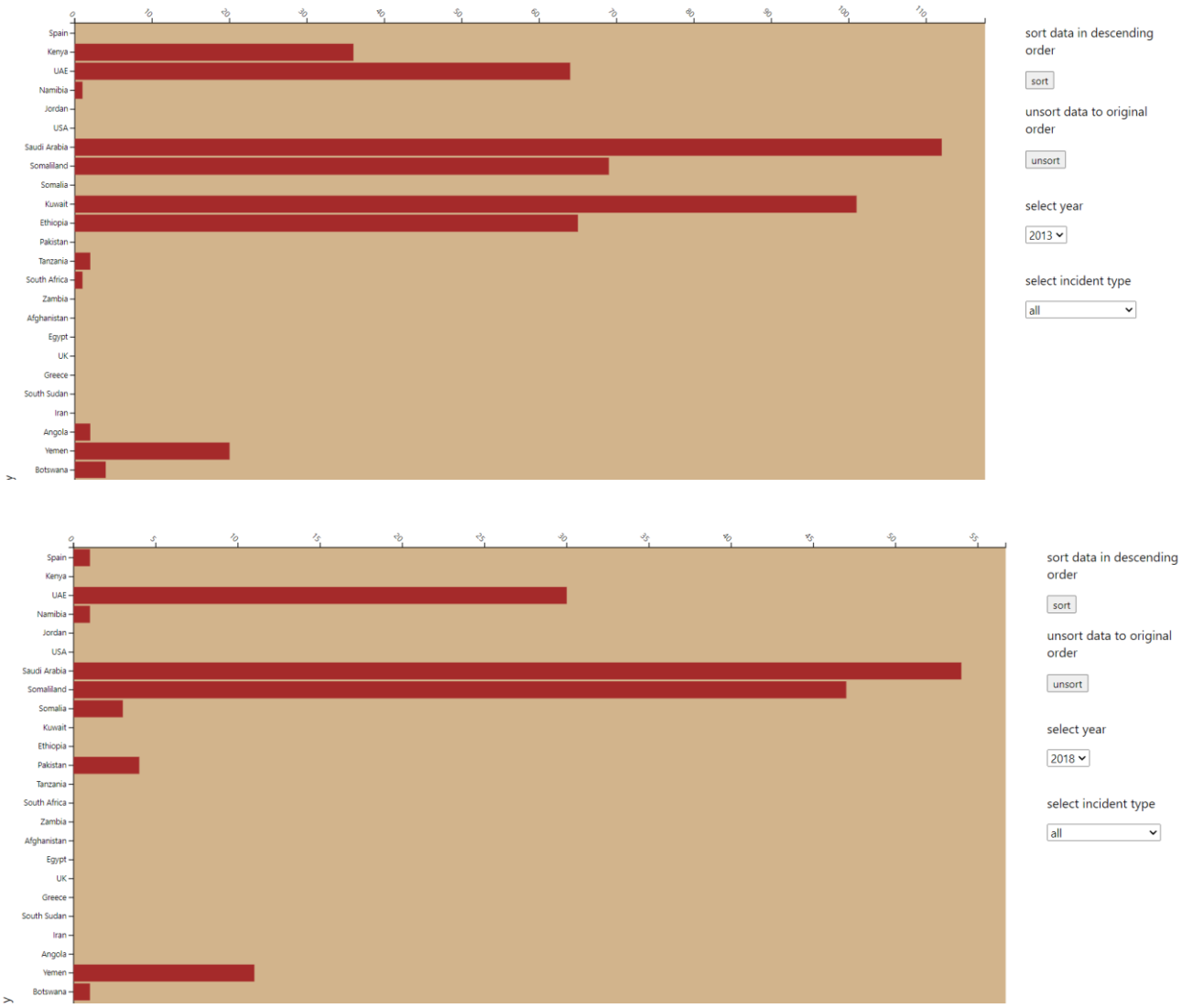
Something that we experimented with was using tooltips vs. having the numbers displayed on load inside or outside the bars. While this was informative, we believe that it made the visualization look too busy and took away from using the size comparison between the bars as the main method of comparing the data to itself.

X-Axis Ticks & Auto Domain Change

The next changes that we developed involved the x-axis. First, since the graph is longer than the initial view, we wanted to include the x-axis ticks on both the top and the bottom of the graph.

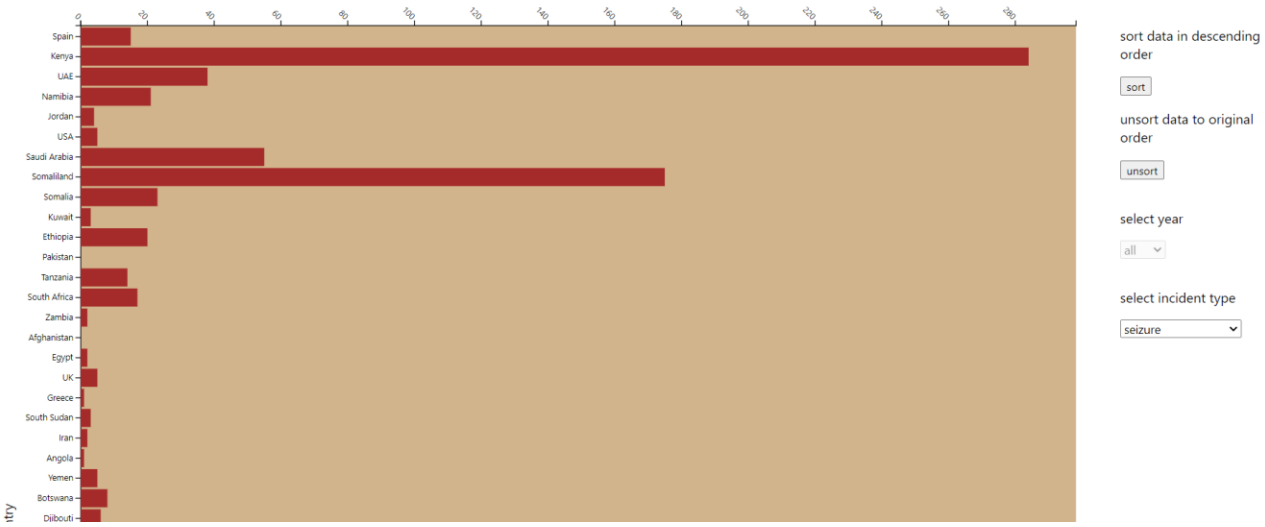


Additionally, when the user filters the data using the select options available, instead of having the x-axis stay the same length when the bars become much shorter, we implemented a change where the x-axis domain is altered when filtering the data to become 105% the length of the longest bar in that dataset.



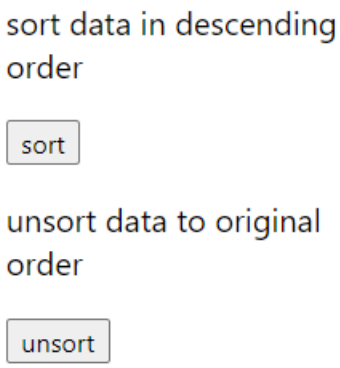
Adding Type Filter/Disabling Selection Options

Next, we wanted the user to be able to filter by another category, this one being incident type. At this point, the data was not in an organized enough place that we could filter by both year and incident type at once, so we added the functionality of disabling either the year or incident type dropdown when the other one was in use.

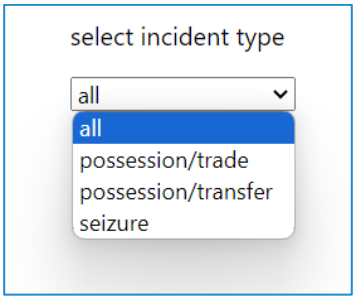

0
20
40
60
80
100
120
140
160
180
200
220
240
260
280
Spain
Kenya
UAE
Namibia
Jordan
USA
Saudi Arabia
Somaliland
Somalia
Kuwait
Ethiopia
Pakistan
Tanzania
South Africa
Zambia
Afghanistan
Egypt
UK
Greece
South Sudan
Iran
Angola
Yemen
Botswana
Djibouti
itry
sort data in descending order
sort
unsort data to original order
unsort
select year
all
select incident type
seizure

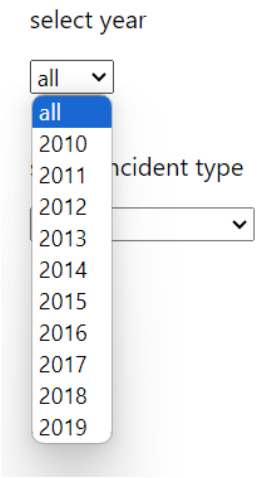
Implementation:

The final result was a filterable and sortable horizontal bar chart. First, we modified the colors from previous designs to utilize colors commonly associated with cheetahs (browns and tans). To make the horizontal chart interactive, we allowed users to sort the countries from highest to lowest number of cheetahs involved in incidents with the press of a button. Additionally, the user can press an "unsort" button to return the data to the original order.

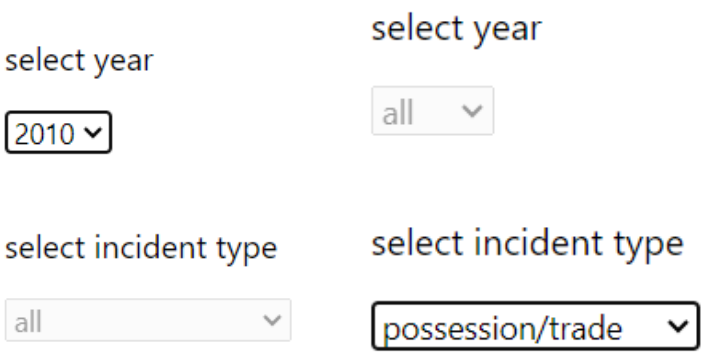


Also, the users could filter the data based on incident type (possession/trade - possession/transfer - seizure) and year (2010-2019). The incident type filter allows users to refine data based on incident categories.





Due to some challenges with data organization, a feature was implemented to disable either the year or incident type dropdown when the other was in use. This decision created a clearer visualization, with an emphasis on size comparison between the bars.



Tooltips were enabled for each bar which provided users with specific information on the number of cheetahs involved in the incidents in a particular country when hovering over the corresponding bar. To keep the visualization clearer, the number of cheetahs were displayed using tooltips instead of displaying numbers inside or outside the bars.

For improving the readability of the graph (due to the length of the y-axis), ticks on the x-axis were included on both the top and bottom of the graph.

The x-axis domain was modified to automatically adjust to 105% of the length of the longest bar in the dataset when filtering the data, ensuring that the bars in the visualization would not become too small to see on the original x-axis scale that ranged from 0-1700.

Evaluation:

There were several insights we gained by the use of this visualization. This horizontal bar chart provides insights into the distribution of illegal chetah trade incidents across different countries. Identifying the countries with high incidents will suggest the regions in which conservation efforts need to be done. These include countries such as Saudi Arabia, Somaliland, and the UAE. In addition, the interactive sorting and filtering of the features, especially regarding years of the incidents, reveals trends of the illegal trade over time. Some examples of these trends include the high numbers for Kenya in 2010 and 2011 compared to other countries, as well as the high numbers for Saudi Arabia that began in 2013 and continued through 2019. Also, being able to filter the graph by the different incident types shows the prevalence of different forms of illegal activities. Most incidents were possession/trade incidents, with much fewer seizure incidents, and virtually no possession/transfer incidents.

There are several strengths and areas for improvement within this visualization. This filterable horizontal bar chart effectively presents the number of cheetahs involved in incidents based on the country and offers users interactive features such as sorting and tooltips for better insights into the data. The inclusion of ticks on both ends of the x-axis enhances readability and the use of adaptive x-axis domain adjustment ensures a consistent display when filtering data.

Despite these strengths which help with the readability and consistency of the data visualization, there are still opportunities for improvement. Tooltips could offer more detailed information about each bar based on the information the dataset provides, including information such as most prevalent source types or the role of the country in the incidents. Additionally, the visualization could utilize color more, either by graying out bars that are not being moused over or using color within the bars to create stacks representing different incident types.

A refresh button could be added for undoing the filtering. This will enhance user experience in exploring the data by providing users with the ability to undo any applied filters and revert the data to its original unfiltered state.

More advanced filtering options could also provide flexibility for the users to explore the dataset even more. For example, trends could be analyzed over time based on months for identifying possible patterns or seasonality related to cheetah incidents. Additionally, being able

to apply more than one filter at a time would be another option that makes the visualization even more user-friendly. Some categorical insights would be exploring incident types and their frequencies to understand the nature of reported incidents.

By implementing these enhancements, the visualization of the cheetah incident dataset could provide a more comprehensive understanding of the patterns associated with the incidents.