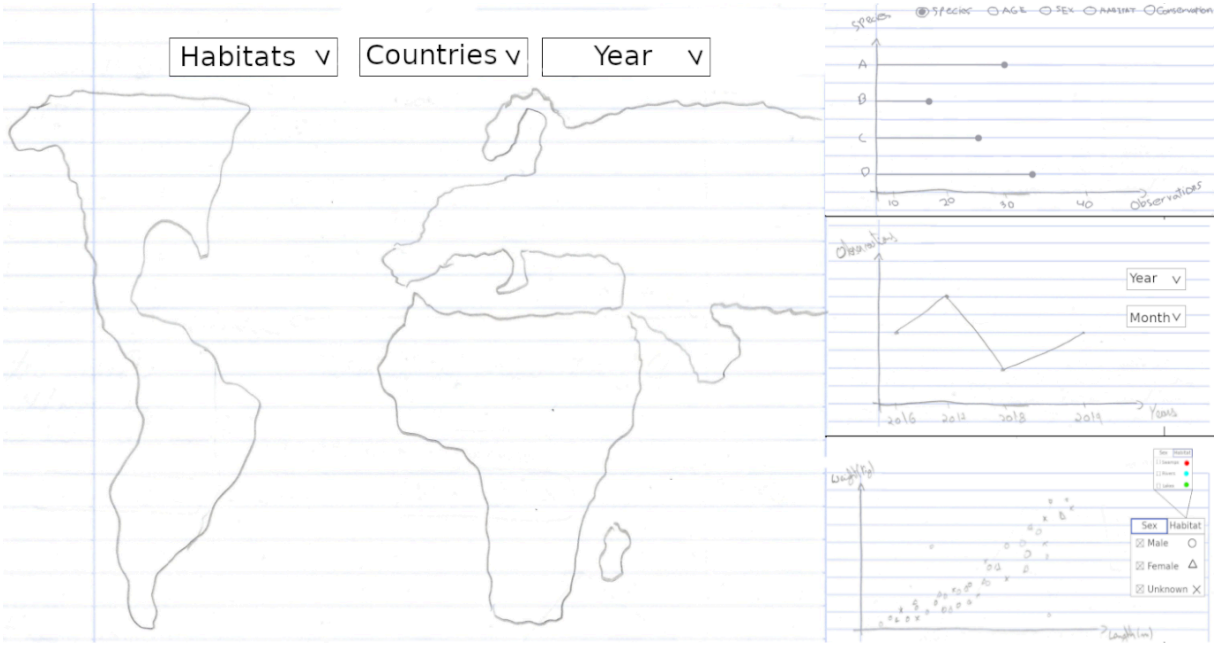
Checkpoint II: Visualization Sketch

Group: 34
Date: 2025/09/22

Dashboard Overview



The dashboard for this project will be composed of a world map with several optional filters. These optional filters will allow viewers to interact with the dashboard and interpret information directly from it. Filtering by average age; by average species vulnerability; by amount of observation dates; or by the amount of different species - these will allow viewers to focus more on any specific questions they may have pertaining to crocodiles. The choropleth map will display the results of the filters using the single sequence single hue colour scale because the data types have, to some degree, a relative 0. The filters, or options, will essentially turn the map into an interactive choropleth map, allowing the user to quickly identify what they want to view. Once the user has identified what country it is they want to view, they will be able to click on it and the three idioms will allow the user to answer any questions they may have about the crocodiles in that situation.

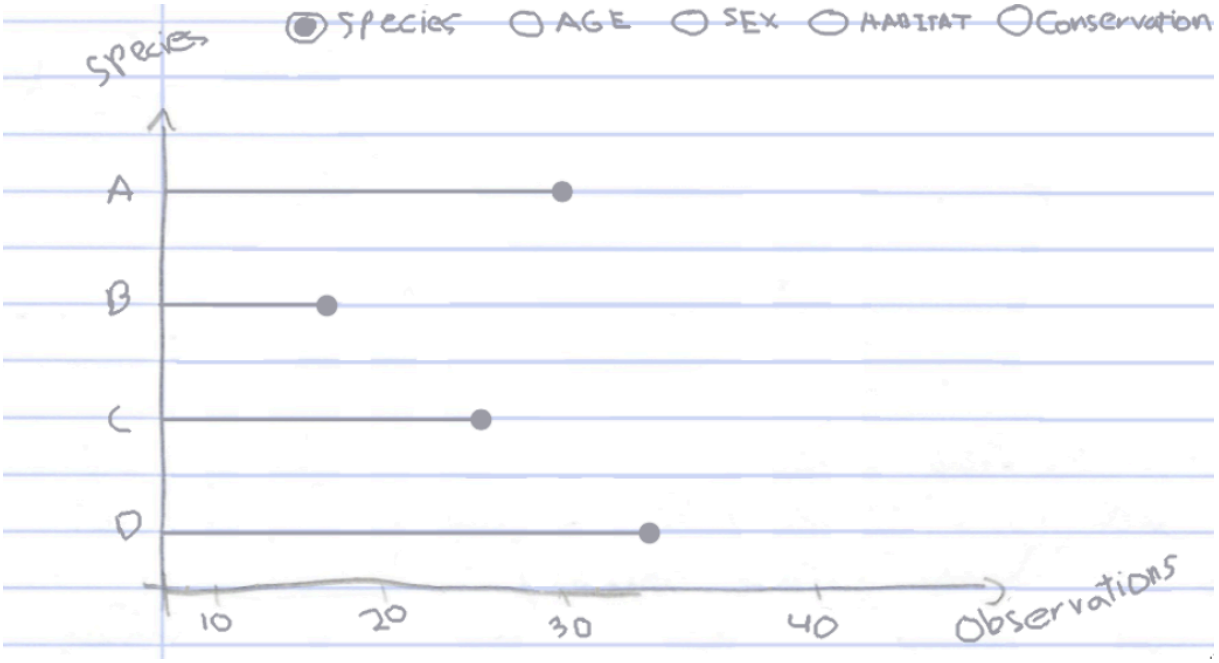
The three charts are displayed to the right of the map, and they will display information relevant to what the user has selected on the map. To the right of each of the charts, the user is shown options pertaining to each of them. There, the user will be able to further customize the results in order to answer any questions with more precision.

The default view of the dashboard will display the main information from the dataset, observed species. The map will display how the amount of observed species varies per country, the Cleveland dot plot will show the distribution of observations per species, the line graph will display the total observations over time and the scatter plot will show the correlation between sex, weight and length.

Charts

Three different idioms were chosen to offer users the answers to their questions. The first idiom is related to distributions; the second to growth or trends; and the third to correlation. So, each question type is answered, respectively, with a Cleveland dot plot, a line chart and a scatter plot.

Chart #1 - Cleveland Dot Plot



The users will want to ask themselves questions about quantities or distributions - questions like "*What is the distribution of species found in each country?*" In these cases, we considered it best to present the information in a Cleveland dot plot.

Marks and Channels

In this idiom, the marks are the stems/lines and the dots/circles , as they are what will represent the values of the items. The channels, which are the visual properties that encode the data, are the positions of the stems and dots on the y-axis, which represent the item of the attribute, as well as their length which represents the number of observations, on the x-axis.

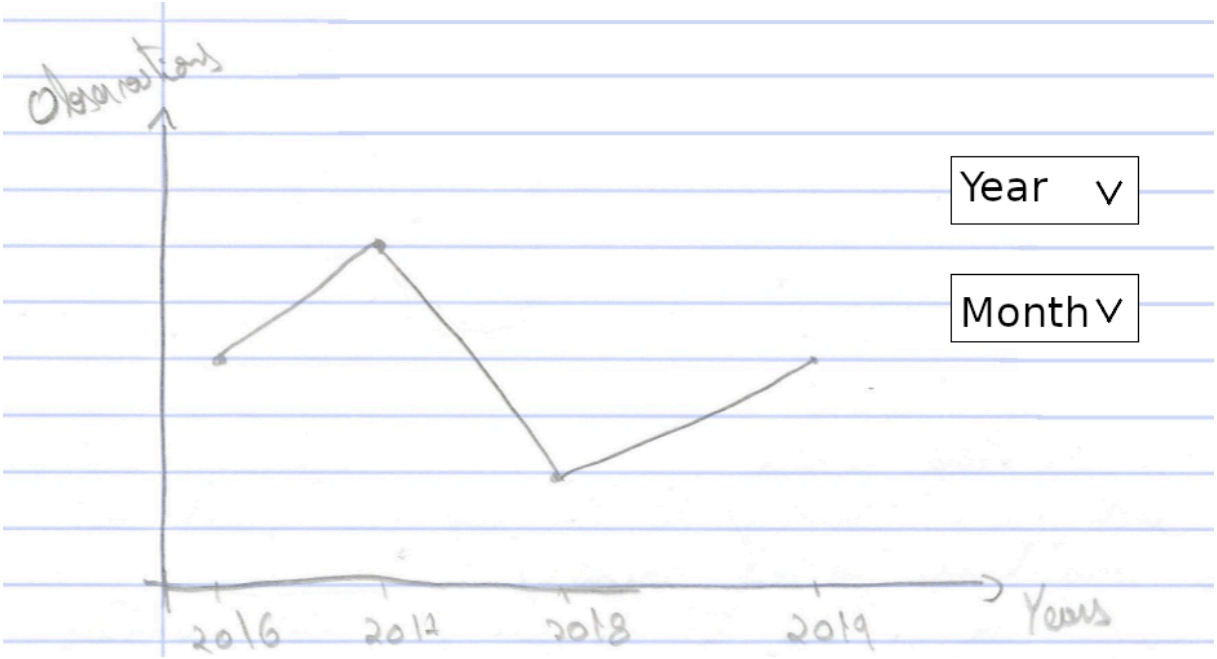
Rationale

We chose to work with a Cleveland dot plot here because we want to represent categorical data in the form of distributions. Each stem and dot clearly represents categorical data (be it nominal or ordinal) and its length represents a quantitative or countable variable. This makes visual comparison between different categories easy, as seeing which categories have the highest or lowest values is quick. This is proven by Steven's Psychophysical Power Law, having that the perception of length is linear to its physical intensity. We considered a bar chart, however, our attributes can be many different values, this would lead to illegibility due to the clutter.

Interaction

This chart has three main points of interaction: changing variables such as "year" on the dashboard Map will filter the results shown in this chart; changing the options above the chart will change the variable that is being processed by the chart; and hovering over each dot will display information about it (the y-axis value and the x-axis value).

Chart #2 - Line Chart



The users will want to ask themselves questions about growth, trends and progress - questions like "Are crocodiles being observed more now than there were a number of years ago?" In these cases, we considered it best to present the information in a line chart.

Marks and Channels

We use points, a 1D mark, to represent the number of observations in a given year or month. This mark's channel is its position, with the x coordinate indicating the year or month and y coordinate the number of observations. These points are connected by a line, a 2D mark. The channel for this mark is the slope of the line itself. The slope is the main point of this chart, and allows users to better understand the evolution of the observations over time. Another channel will also exist if more than one country is selected, that being the hue.

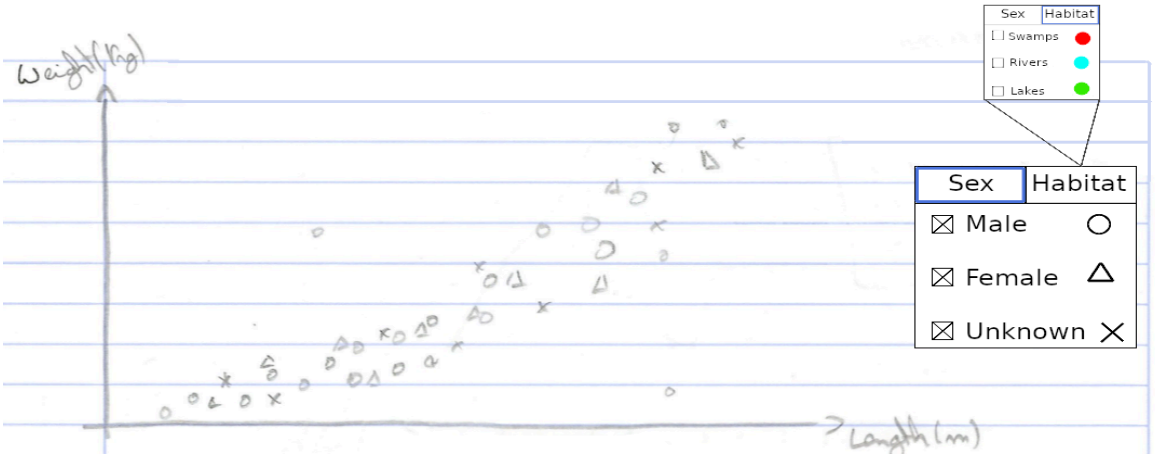
Rationale

One of our main questions was: "*Are crocodiles being observed now more than there were a number of years ago?*". This is a trend question, and line charts are great idioms for answering these types of questions. This chart will allow us not only to answer this question, but to also be able to easily check the number of observations in a determinate year/country or month/country pair. We considered a spiral chart as an alternative to the line chart, but ultimately we chose the line chart because of the granularity of our data, the spiral chart would be too detailed and thus not legible.

Interaction

Changing the year or country on the dashboard will filter this chart. Selecting multiple countries will display multiple lines on the chart. There will be a specific scented widget for this chart that will filter using months, allowing the user to see the evolution of observations for a month throughout the years. While multiple countries are selected, the user can hover a line to highlight the country on the map, and clicking the line will apply the filters just like the scented widget applies them.

Chart #3 - Scatter Plot



Marks and Channels

We are using position as a channel for both length and weight, using the x-axis as the length and the y-axis as the weight. We are using shape as a 0d mark for the gender class. Furthermore, we are also using colour to identify the habitats.

Rationale

We have two main questions: *"Is there a correlation between a crocodile's weight, length and gender?*" and "*Is there a correlation between weight, length and habitat?*". As such, we decided to combine them into the same idiom, using filters to make the distinction between the different questions. There are 29 habitats and only 3 gender classes. Since we intend on showing all habitats at once, we decided to also use colour to group similar habitats together (like the different river types). We considered a bubble chart instead of a scatter plot due to the amount of variables to represent, but ultimately we chose the scatter plot because the human eye is not very good at distinguishing small differences in area, especially if the bubbles are far apart.

Interaction

The chart will have 2 classification buttons (1 for the gender and 1 for the habitat). These classification buttons change which filter buttons are available, changing between the 3 for gender and the necessary ones for the habitats. The filters double as a legend for the marks, and selecting a given filter displays the information pertaining to it on the graph, it is possible to display multiple filters simultaneously. Switching between classification buttons will reset the filters. Hovering the individual observations will show their specific length/weight.

Chart Integration

The dashboard will contain a scented widget of all the countries, with all countries being in alphabetical order and only countries with observations appear, a country which does not have observations for the current time interval will be greyed out and not selectable. Countries can be selected on this list to apply a filter to all idioms. A button will also allow filtering using the year for all idioms.

Answering the Questions

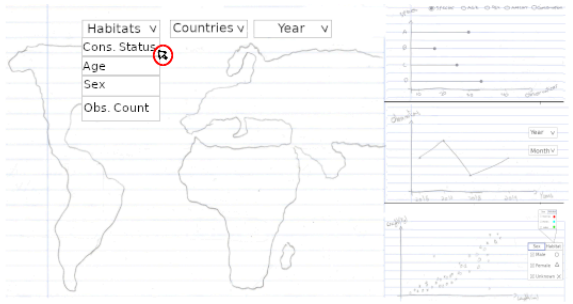
- *"What is the distribution of species found in each country?"*: The user selects the country they

want either on the map, or on the scented widget, then, on the Cleveland dot plot options, select the "species" option. The Cleveland dot plot then shows the result.

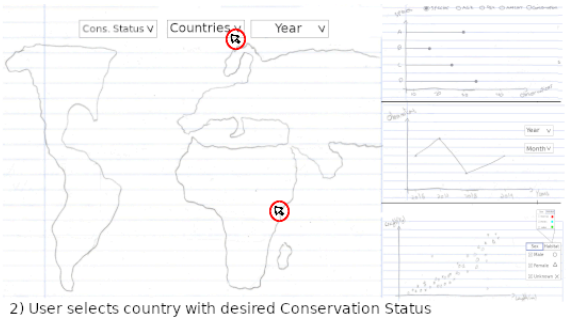
- *"Is there a correlation between a crocodile's weight, length and gender?"*: Here, one can look at the third idiom selecting the sex filter and looking at the chart will answer it. The chart will show dots that symbolize a crocodile's observation and can then plot a regression line to check if the correlation exists, but also if it is positive or negative.
- *"Are crocodiles being observed now more than there were a number of years ago?"*: Here, one can look at the second idiom. Selecting the filters to all countries and years answers the question. The user can then look at the slope of the line chart and determine the answer.
- *"What is the distribution of ages of crocodiles in countries where they are in more vulnerable conservation statuses?"*: The user selects conservation status on the dashboard Map options, and chooses a country with more a vulnerable conservation status after analysing the map. On the Cleveland dot plot options, the user selects "Age". The Cleveland dot plot then shows the result.
- *"What is the distribution of species found in each habitat?"*: The user selects the country they want on either on the map, or on the scented widget, then on the Cleveland dot plot options select the "habitat" option. The Cleveland dot plot will show the distribution of species in habitats in the selected country.
- *"Is there a correlation between weight, length and habitat?"*: Similar to the second question but using the habitat filter instead of the sex filter.

Storyboards

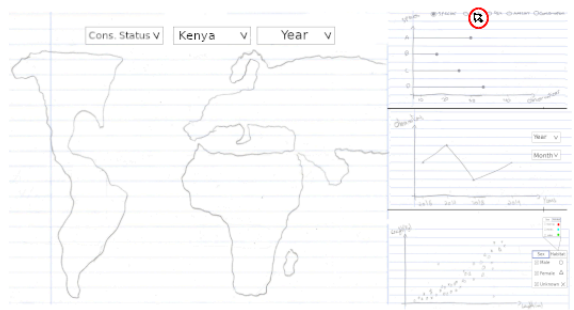
This is the storyboard for question 4: *"What is the distribution of ages of crocodiles in countries where they are in more vulnerable conservation statuses?"*



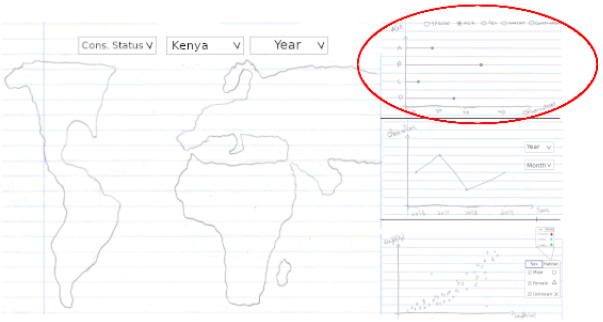
1) User selects Map filter Conservation Status



2) User selects country with desired Conservation Status



3) User selects Age in Bar Chart section



4) User views the new results and takes conclusions from it