Endangered Species and Human Influences

COMP3125 Individual Project

Nico Lattarulo  
Wentworth Institute of Technology

*Abstract*—This project is a look into human environmental interaction and examining the impact we have on various animal species. The data selected for this analysis covers many species and their conservation status, and several quantifiable human impacts such as air quality index and deforestation.

# Introduction

Animal Extinction is a telltale sign of human activity. As we have travelled around the earth, some of the sensitive ecological systems that have existed unrelated to us are not able to handle our influence. Whether it be our direct influence in harming the environment, or the resource demands needed for our society, many animal species become endangered or go extinct due to these factors. For my topic, I would like to investigate how much of an impact we have on the threated species of today. It is important to keep track of things like this because we often forget that we are part of the same environmental system that all of these animals are a part of. If these systems were to weaken, there are many unforeseen circumstances that may arise. There has been a lot of research into the sensitive elements of the ecological systems of many of these places and analyzing data like this help improve our understanding of nature works.

# Datasets and methodology

My Data was mainly acquired from Kaggle compiled from organizations, as well as sets from the World Bank Group. As I had to adjust my project idea, the datasets I submitted previously were difficult to work with. As I was adjusting my datasets for this project, the Wentworth login for Statista was unavailable for some reason.

The dataset containing all of the species from the World Bank Group global species database is the largest. The information I found useful from it was the common/scientific names, the conservation status, and the country of origin. The data was already in a very easy to use format and had many entries, so it was very useful.

The deforestation data was compiled from another World Bank set. It contains the percentage land forest coverage of each country in 2000 and again in 2020, as well as the trend between the two. The deforestation dataset contains the countries 3 letter abbreviations, % land use that is forests in 2000, and the same for 2020. There is also a trend column to represent the difference between the two entries which I used for comparisons.

The air quality dataset is contains various air quality data for many countries and towns in much detail. It contains the Air quality report in AQI, CO AQI, and Ozone AQI for cities and towns within countries. I used only the AQI value to represent air pollution status. I did an average of the data for cities in the same country and created average AQI’s for each country to line up better with the deforestation dataset.

I also downloaded a dataset from the World Bank Groups world development indicators databank. This set contains the percentage of the country’s population that lives in an urbanized area over the span of 20 years, to better align with the deforestation set.

I used matplotlib bar graphs to visualize these datasets. Pandas data frames were also used to process these sets into forms that were much easier to work with and analyze in the way I wanted to for this project. I also used Pandas to represent them numerically and draw the necessary conclusions from.

# Results

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedQuestion #1: Where are the most endangered species?

10 countries with the highest number of endangered species / critically endangered species

These portions the datasets show the regions that host the most species under heavy threat of extinction. Many of these are tropical, Equatorial locations. While these locations are home to a larger density of species, they are also susceptible to many threats similar to places like the Amazon rainforest: logging, fires, unsustainable agriculture, industrial waste, etc. Many of these species are more susceptible to small changes in their environment and these kinds of pressures can easily cause mass population loss.



A small note is that over the past 20 years Madagascar, country with the highest number of endangered species by a large amount, has had a small downward trend in % of land that is forest.

A graph with blue and black text

Description automatically generated

This graph represents all of the countries within this dataset and the number of their endangered species.

Question #2: What human influences affect different animal species?

Many of the countries with the most threatened species also appear higher up on the datasets representing human influence on the environment. Colombia, Mexico, Brazil, Ecuador all appear very high on the endangered species lists as was shown in the previous questions. These four countries appear relatively high on the deforestation table,and considering how high they are on the threatened species list, that implies some meaning.

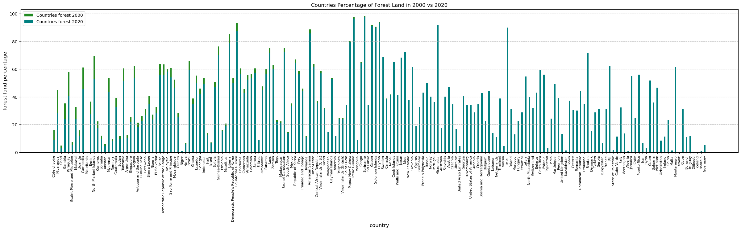


Decrease in % forest land









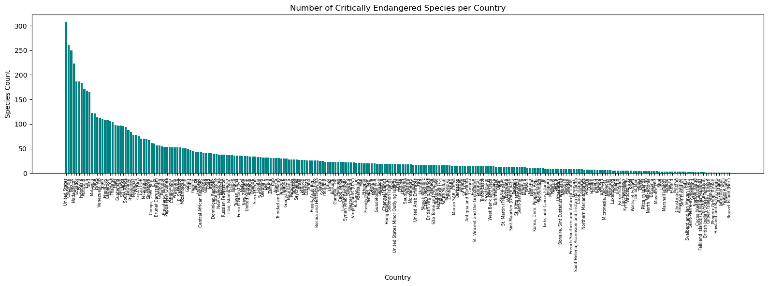
This graph represents all of the countries percent forest land coverage in 2000 and 2020. These South American countries that encompass the Amazon Rainforest represent the damage that humanity has done by way of stripping the land of its resources without considering the ecological systems.

-Mexico 42th on Air quality table and 54th on the urbanization table

Air Quality Index Value % Urban population

The United Staes appears highest in the critically endangered species category. It also appears above average in the percentage of urban population with 80% living in urban areas. This shows that it is possible that the higher rate of urbanization has pushed some sensitive species to a high threat of extinction. At the peak of this graph representing critically endangered species, the U.S. is visible.



Indonesia has 403 endangered species, and 184 critically endangered species, which is very high in the world. It also shows up above average in air quality and deforestation, which it is known for with palm oil production and orangutan populations loss. The Indonesia species of orangutan are critically endangered according to this dataset.

Tanzania hosts 401 species on the endangered list and 88 critically endangered species. Compared to some of these other countries, it also appears fairly high within the deforestation dataset, with 14.9% less forested land in 2020 compared to 2000. The removal of those forests could pose a threat to many of these species.

Question #3: What locations present the most human influence?

Mauritania appears 3rd on the air quality list at 179, and 9th on the deforestation trend list with 25% less forest land over the last 20 years.





Pakistan has the 4th highest air quality index value, at 178.It is also 16th on the deforestation trend list with a downward trend in forest area of -18.6%





These countries show a combination of being high in human ranking, and also appearing across multiple datasets. These show countries with humans influencing the environment in different ways.

Question #4: Do the locations of the human influence usually coincide within this dataset?

The Republic of Korea presents the best example of all three of the human influence datasets converging. It has the highest air quality index, marking the lowest air quality in the set, 66th on the deforestation list with -4% over the last 20 years. As well as 45th on the urbanization list with 81% of the population living in an urban environment.

A screenshot of a computer

Description automatically generated

This graph shows the air quality index of many countries sorted greatest to least. There are a few countries to the left with extreme AQI values including the Republic of Korea as mentioned above.

# Discussion and Conclusion

Given more time with this data and more available time to me at this point in the semester, along with more instruction in class, I would mainly like to model the endangered species dataset with the trends of deforestation and increasing population to estimate future population and conservation status for certain species. This data represents well known, well researched topics like Amazon rainforest loss, and Southeast Asian Orangutan population loss

After spending some time analyzing these datasets and visualizing them in python, I think that there are certain trends within each of them that align in a way that show that impact that humans unsustainable practices have had on our environment.

# References

[1] Nasution, Arfah et al. “Declining Orangutans Population in the Unprotected Forest of Batang Toru.” *Tropical life sciences research* vol. 29,2 (2018): 77-87. doi:10.21315/tlsr2018.29.2.6

[2] Wilson EO, Peter FM, editors. Biodiversity. Washington (DC): National Academies Press (US); 1988. Chapter 3, Tropical Forests and Their Species Going, Going … ? Available from: <https://www.ncbi.nlm.nih.gov/books/NBK219303/>

[3] Zanon, Sibélia. “Deforestation in the Amazon: Past, Present and Future.” *InfoAmazonia*, 22 Mar. 2023, infoamazonia.org/en/2023/03/21/deforestation-in-the-amazon-past-present-and-future/.

[4] https://www.kaggle.com/datasets/konradb/deforestation-dataset

[5] https://databank.worldbank.org/Urbanization/id/f7b5329b#

[6] https://datacatalog.worldbank.org/search/dataset/0063384