Stat 198: Species Extinction Research Poster Write-Up

Data Collection: Extinction Years

For the data collection for this project, the first thing we knew we needed was a list of all the extinct species as well as each species’ extinction years. The best resource we found that would allow us to get this information was an online database called IUCN Red List (<http://www.iucnredlist.org/>). This database contained a list of all the species they classified as “extinct” and “extinct in wild”. The website allowed us to export these lists as “.csv” files.

However, these files did not contain information about the extinction years for these species. In order to obtain this information, Wideet was able to follow a tutorial created by Professor Carl Boettiger, who taught a Data Science connector course about species extinctions. The method Professor Boettiger described was to look through the “.json” files corresponding to each extinct species. These “.json” files were also stored online by IUCN Red List. In the “.json” files, each extinct species had a section called “rationale”, which described the reasoning for ruling that species extinct. This section usually included a year that the species was officially ruled extinct. By using a regular expression to search for four consecutive digits, Wideet was able to get extinction years for all the species where that information was available.

Unfortunately, only around 50% of the species in the data sets had extinction dates recorded in the “rationale” section of their “.json” files. After talking with Professor Boettiger, he confirmed that this was one of the shortcomings of this method, and that authors of the paper were only able to obtain more extinction dates through various other sources.

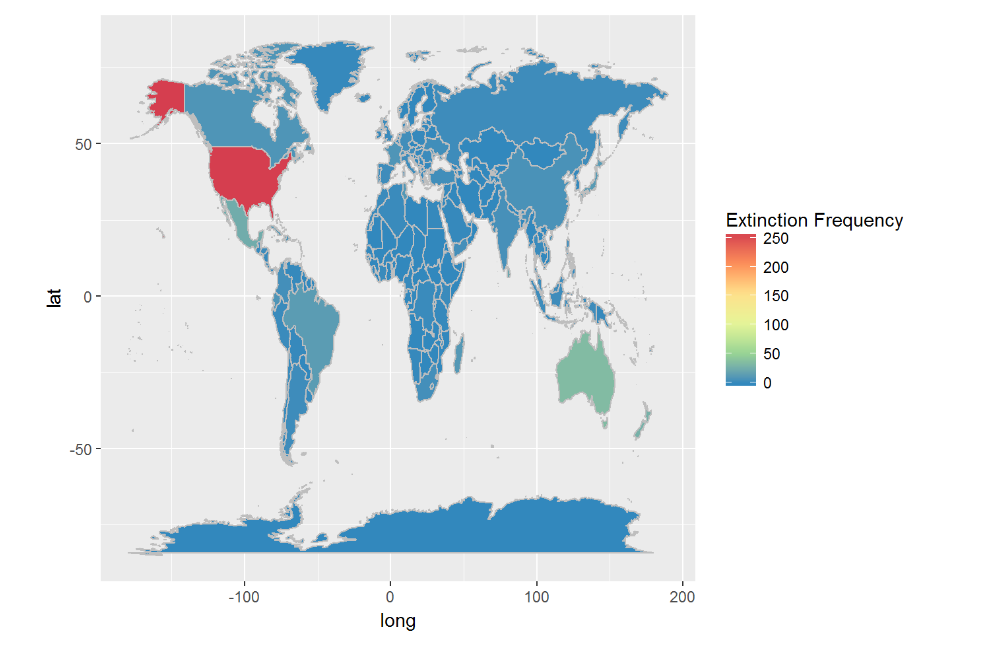
To get more extinction years, we contacted Professor Ceballos, the lead author of the original article, and he was generous enough to provide us with the complete data set that he and his colleagues were able to gather for the original paper.

Data Collection: Locations

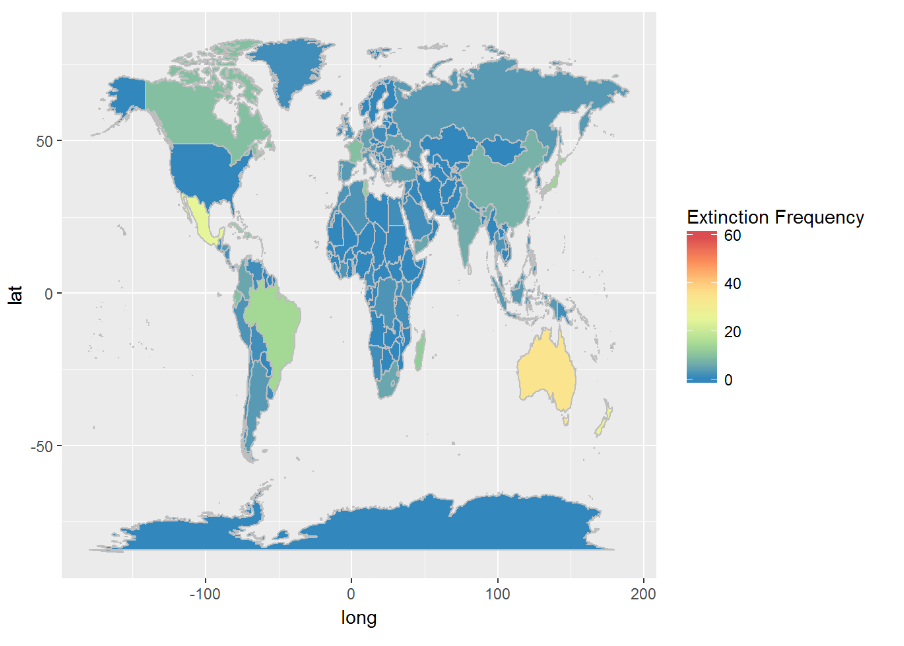
Another set of data that we wanted to explore was the locations where each of the extinct species went extinct. Our idea was to create a map that shows which countries had the most species extinctions. In order to obtain the location data for each extinct species, Wideet created a web scraper that gathered location data from the IUCN web page created for each extinct species. The only complication here was that, since many species were native to more than one location, the location data needed to be cleaned. Wideet was able to manually go through the location data and separate the entries that contained more than one country in them and create duplicate entries in the extinct species data set with separate locations for each of them. From there, Wideet used a package in R-Studio to create an extinction-frequency data frame for each country. This data frame was used to create a “heat map” of extinctions for each country in the world.

Extinction Heat Map

Wideet was responsible for making the extinction heat maps. The initial map he created was to simply represent the frequency of extinctions in each country.



The problem with this map was that the USA had far more extinctions in our data than any other country (257 for USA, next highest is 62). After removing the USA from the data frame, he got a more representative map.



Here we can see that Australia and Mexico had relatively higher extinction frequencies than other countries. The main trend we can see from this map is that developed countries are recording more far more extinctions than under-developed countries. Countries like Canada, China, India, and Brazil are clearly highlighted, while struggling countries such as those in Africa record almost no extinctions. In fact, one of the most developed countries in Africa, South Africa, is the only area of Africa that shows any extinctions. This shows there may be a correlation between the process of becoming a developed nation and the effect on the environment and the native species.