

For my computer science final project, I analyzed a dataset of international fishing vessels. My goal was to understand global fishing activity, fleet composition, and efficiency. I used Python as my main programming language. I relied on Pandas for cleaning, filtering, and aggregating data, NumPy for numerical calculations, and Matplotlib for visualizing the data. Pandas helped me prepare the dataset efficiently by removing irrelevant or misleading information. I filtered out vessels that were shorter than 24 meters overall to focus on industrial-scale fishing efforts. Grouping, summing, and sorting operations allowed for meaningful comparisons between countries, vessel types, time spent at sea, and fleet tonnage. Matplotlib enabled me to create clear visualizations from the cleaned datasets, and custom colormaps highlighted the significance across countries and fleets.

Using these tools, I answered four key questions about global fishing. My analysis revealed that fishing effort is highly concentrated, with a small number of countries, led mainly by China, accounting for most vessels and fleet capacity. An analysis of vessel types showed that trawlers dominate global fishing. While they are efficient, they are also among the most environmentally damaging methods. In comparing active hours to fishing hours, I found that vessels spend only about a quarter of their time actually fishing. The majority of their time is spent in transit or engaged in non-fishing activities, indicating a decline in efficiency as fleets move farther from port in search of fish. Lastly, fleet tonnage analysis showed that the number of vessels does not directly correlate with total fishing capacity. Countries with fewer vessels can still have significant fishing power depending on the types of vessels in their fleet and the species that country targets.

If I were to present this analysis to a client, such as an environmental organization, I would recommend focusing oversight and policy efforts on the small group of countries that dominate global fishing activity, rather than trying to apply the same regulations to all nations. Regulations should prioritize vessel type, especially trawlers, because of their significant environmental impact. Fleet tonnage should be considered alongside vessel count as a key measure of fishing capacity. Monitoring should also take into account the increasing non-fishing time, which signals stress on the ecosystem as fishermen are having to move to more foreign grounds that haven't been overfished.

For future work, I would attempt to link my CSV of AIS data with other fishing datasets that display more information on things like targeted species and areas that experience heavy fishing. Using those extra datasets, I can establish a deeper connection not only between the vessels I'm already looking at, but with the real activity that each of these data entries partake in in the real world.