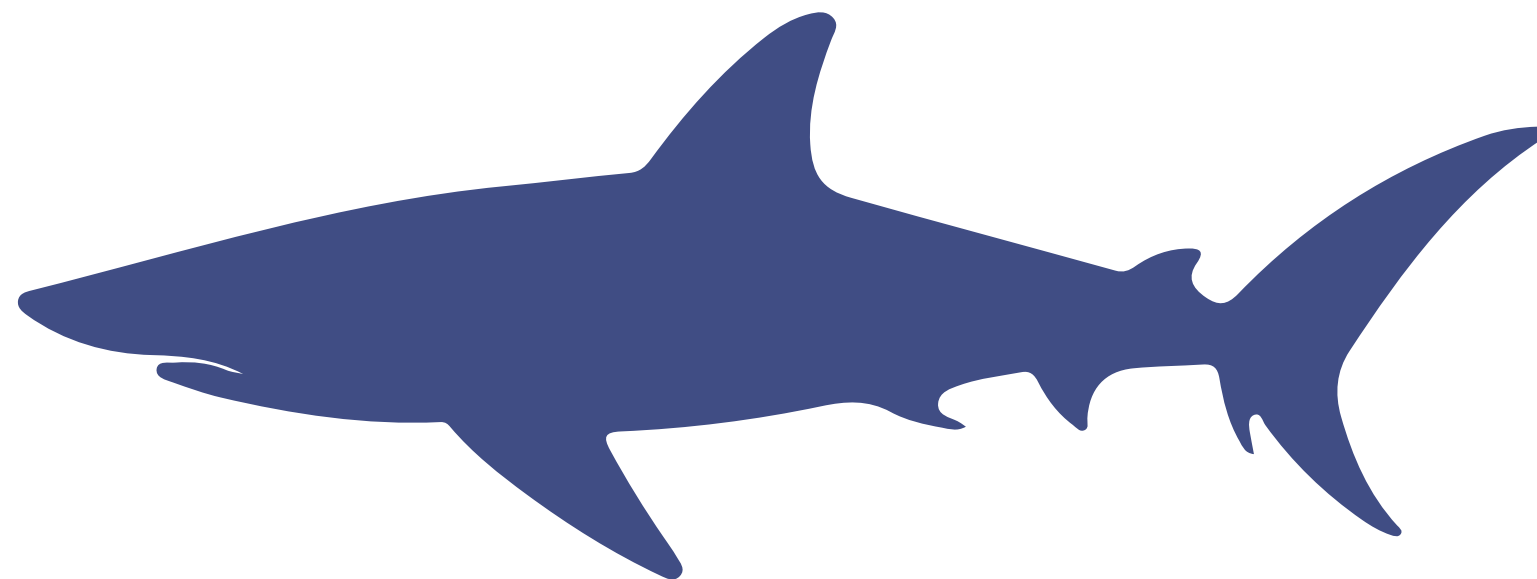


Latitudinal Patterns of BIN Diversity in Requiem Sharks (Family Carcharhinidae)

A global analysis using BOLD DNA barcodes



Requiem sharks are a global warm-sea family; BOLD BINs let us quantify their diversity

- **What are they:** One shark family (Carcharhinidae) of roughly 12 genera and 50 species. Widespread in tropical and warm temperate seas
- **Why they are fit for basin tests:** Many species range along warm currents and across multiple ocean basins
- **Why BOLD/BINs:** DNA barcodes summarized as BINs will provide species like lineage clusters for comparing composition and richness among regions
- **Scope:** Compare ocean basins and latitude bands using BIN counts and overlap

<div><div>Bull shark</div><div>Tiger shark</div><div>Blue shark</div></div> <p>Examples of requiem sharks (Family Carcharhinidae): Tiger (Galeocerdo), Blue (Prionace), Bull (Carcharhinus)</p>	
Specimens:	5924
Sequences:	6206
Records with BINs:	5064
Records with Species:	5876
BINs:	68
Species:	76
Countries/Oceans:	67
Institutions:	57

Tropics tend to be richer and oceans can isolate lineages

- **Tropical Peak:** In aquatic life richness tends to drop with latitude. The Equator is richer than temperate/polar.
- **Basin Effects:** Some sharks follow warm ocean currents and stay in specific basins. Therefore the mixing isn't complete between basins.
- **Sampling Caution:** Global barcode sampling is not evenly distributed. Check for completeness and make sure patterns are not arising due to areas highly sampled.
- **Project Plan:** I will compare how BIN richness varies by region, see how much overlap exists between basins, and will use a global map to display sampling locations by ocean basin

Research question & hypotheses

Research question:

Do tropical oceans have more requiem shark BINs compared to temperate oceans, and do adjacent basins share more BINs than non-adjacent ones?

Hypotheses

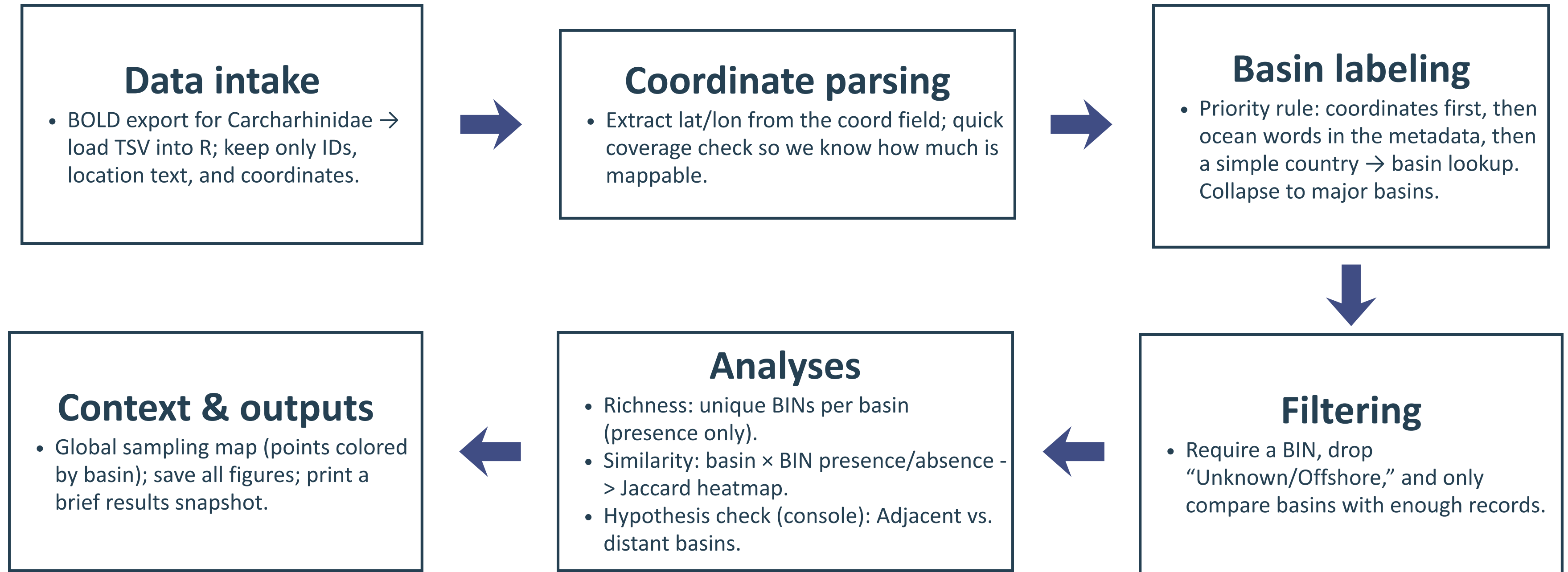
- H1 (latitudinal): Tropical is higher than Temperate in BIN richness.
- H2 (connectivity): Adjacent basins (e.g., Indian-Pacific) share more BINs than distant basins (e.g., Atlantic-Pacific).
- H0 (null): No real differences; patterns could reflect sampling coverage.

Key definitions:

Tropical vs temperate: by latitude (latitude $\leq 23.5^\circ$ = tropical; $> 23.5^\circ$ = temperate).

Adjacent vs distant basins: Adjacent = geographically connected by major warm currents (e.g., Indian-Pacific); Distant = separated by continental or current barriers (e.g., Atlantic–Pacific).

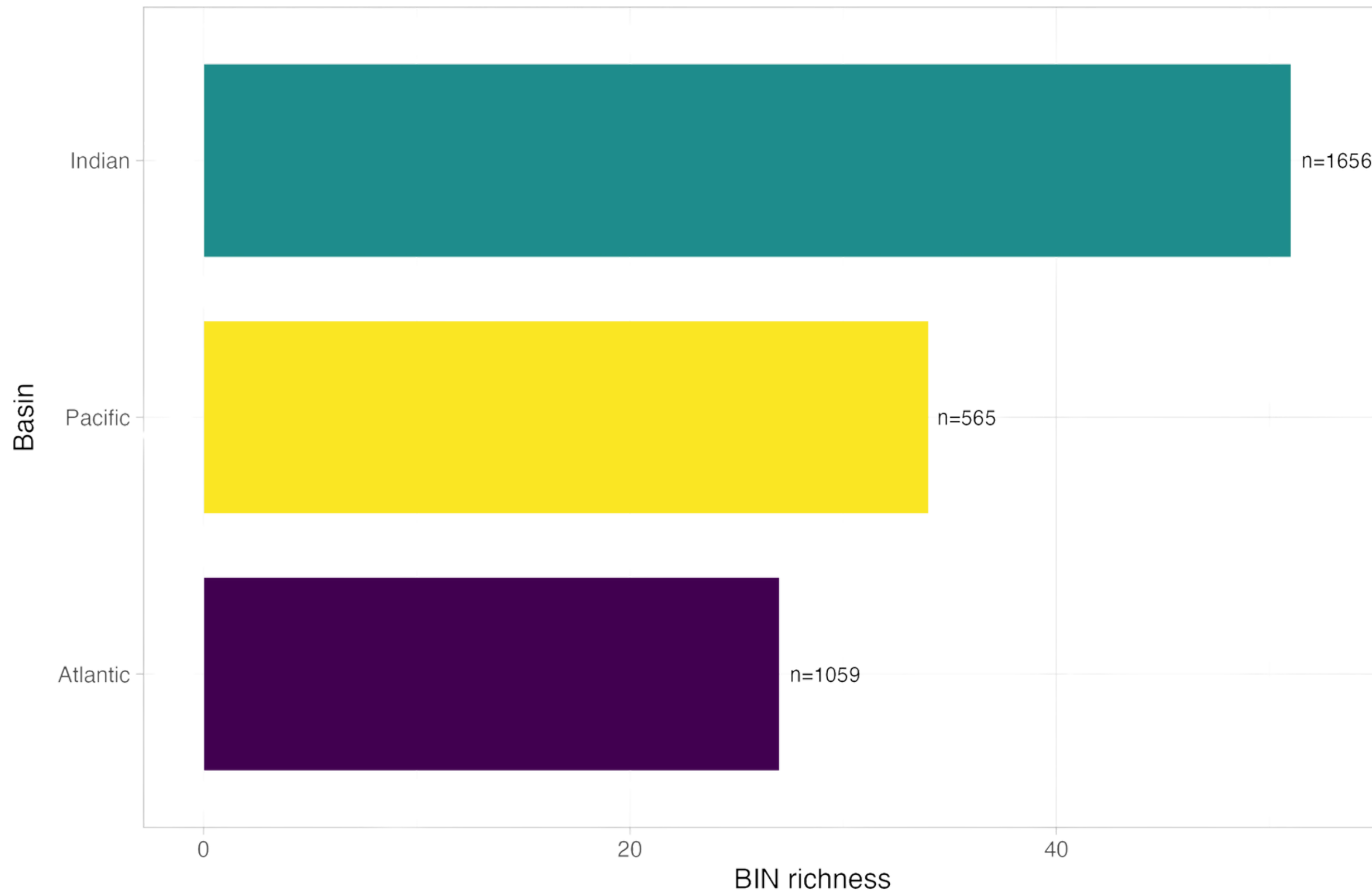
Transparent, reproducible workflow from BOLD to basin results



Notes: R + tidyverse/vegan/viridis, seed set, relative paths, comments inline

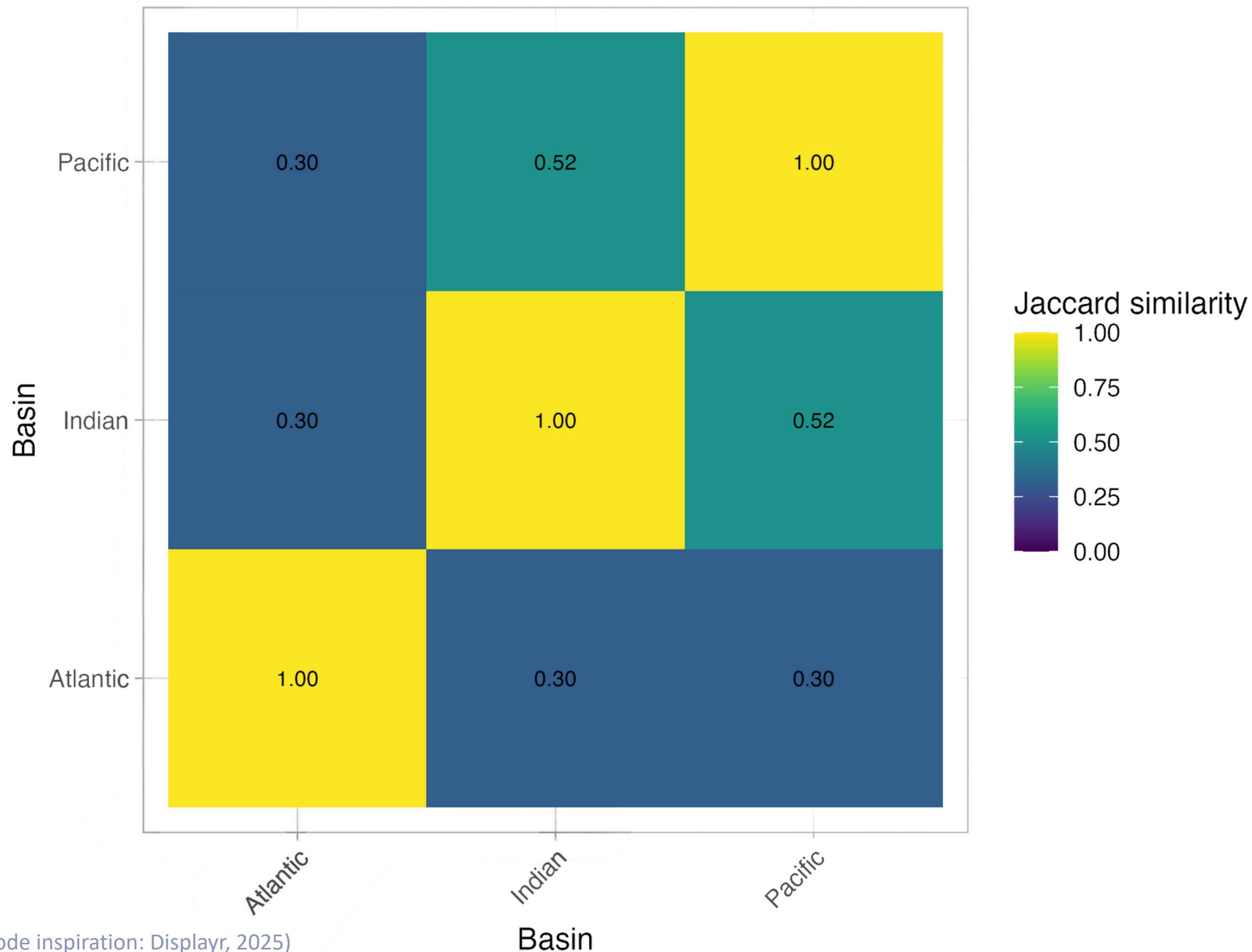
Figure 1: Unique BINs per ocean basin

Presence-only richness (records annotated as n)



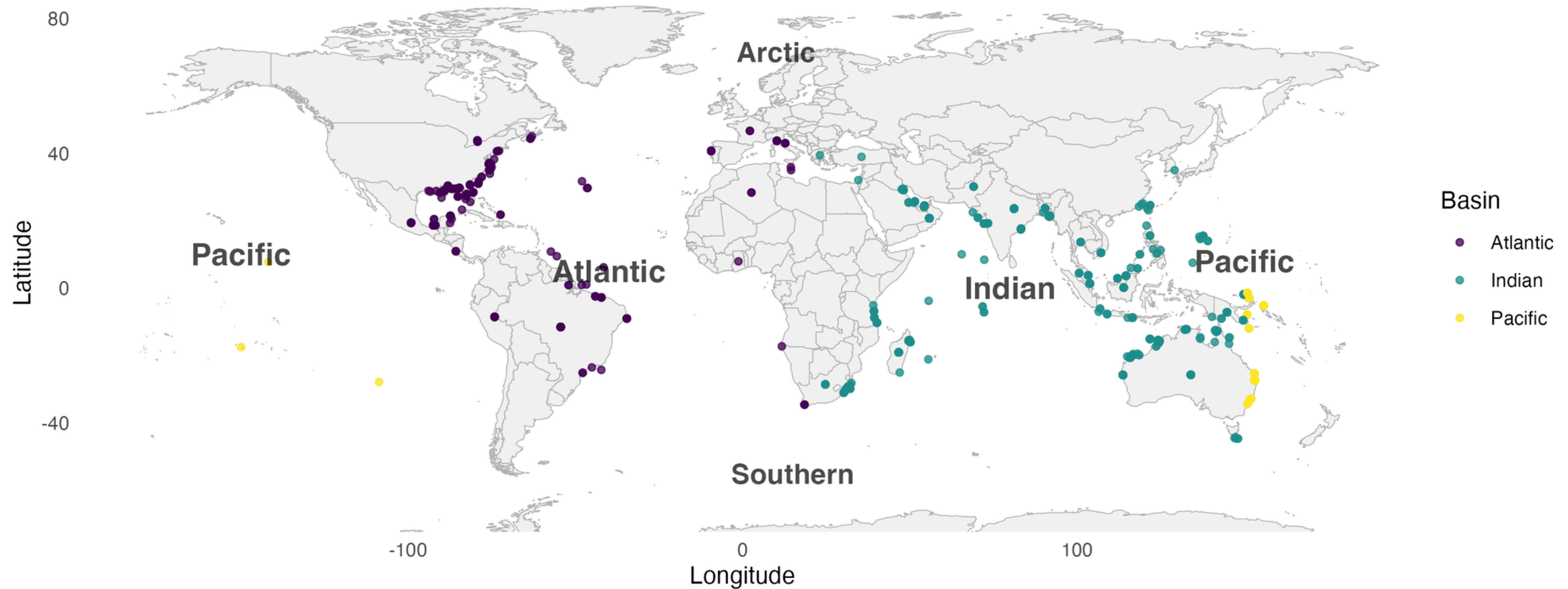
- The bar heights show the number of unique BINs where Indian > Pacific > Atlantic
- Y-axis shows BIN richness (presence only): each bars label (n) shows how many records were available which helps gauge sampling effort
- The Indian basin shows the greatest observed BIN diversity whereas the Atlantic the least
- The sampling isn't even. Some regions such as Indian basin have more records, and many lack exact ocean assignments
- BIN is roughly equal to number of species but not perfect one to one match

Figure 2: BIN Composition similarity among basins



- The jaccard similarities are moderate; no basin pair shows high overlap (close to 1.0)
- Numbers in each box show the Jaccard similarity from 0-1 between basin pairs based on BIN presence and absence
- BIN composition differs across basins and the idea that adjacent basins share more BINS wasn't significantly supported.
- Uneven sampling and missing data likely affect the overlap; presence/absence ignores abundance, and incomplete barcoding or BIN splitting/lumping can change the similarity values

Figure 3: Global sampling locations by ocean basin



Points = records with coordinates; colours = basin assigned from metadata/coords

- Dense point clusters in Indo-West Pacific and Western Atlantic; sparse elsewhere.
- The dataset's spatial footprint is uneven, which likely contributes to Figure 1's richness ordering and Figure 2's only moderate overlap.
- Bias to consider: Records without coords/ocean text are excluded here; country → basin mapping is coarse; coastal bias (near ports/labs) underrepresents offshore regions.

Results

Figure 1: BIN richness by basin

- Indian > Pacific > Atlantic in unique BINs (presence-only).
- Numbers: 65 BINs across 3 basins after filtering (≥ 10 records/basin); records \approx Indian 1,656, Atlantic 1,059, Pacific 565.
- Looks like a warm-sea peak, but richer basins may also be the ones we sampled more.

Figure 2: Jaccard similarity among basins

- Moderate overlap overall; no basin pair near 1.0 similarity.
- Adjacent vs distant: Not significant (perm test $p \approx 0.66$).
- Quick read: At the BIN level, communities differ across basins in this dataset.

Figure 3: Global sampling map

- Dense clusters in the Indo-West Pacific and Western Atlantic; sparse elsewhere.
- The sampling footprint helps explain Figure 1 (richness ordering) and Figure 2 (only moderate overlap).

Limitations & biases

- Sampling is heavily focused in the Indo-West Pacific and Western Atlantic regions and hence under sampled regions show lower richness and overlap
- Many records lack ocean coordinates or labels hence basin assignment based on country or keywords is approximate and can misclassify some samples
- No abundance data or any standardized sampling effort hence the Jaccard similarity can be skewed when sample sizes differ
- BINs are not identical to species and hence splitting or lumping can change both richness and overlap estimates
- After filtering via ≥ 10 record per basin, only 3 basins remained, hence lowering the power for similarity and adjacency tests
- The map excludes non georeferenced data, the actual ranges could be broader than shown
- Records are from different years and with no control for time, newer sampling in one basin could blur temporal patterns

Next steps

Make comparisons fairer

- Subsample to equal record counts per basin (or use coverage-based rarefaction) so richness isn't just tracking effort.
- Report richness with confidence intervals from rarefaction, not just single bars.

Tighten the geography

- Re-run at country or marine ecoregion scale and test distance-decay (does similarity drop with distance?).
- Add ocean currents/fronts to test simple corridor/barrier ideas (e.g., Indian-Pacific gateways).

Improve the data going in

- Push for explicit ocean fields in BOLD; reduce “Unknown/Offshore.”
- Pair BINs with species IDs where reliable; flag obvious BIN splits/lumps.

Stress-test the findings

- Split by ecology (coastal vs oceanic requiem sharks) to see if overlap patterns change.
- Replicate the workflow on another family (e.g., Sphyrnidae) to see if the story holds.

This would give me cleaner, effort-balanced comparisons, clearer geography, and a better sense of whether the weak connectivity result is real or just sampling.

References

BoldSystems. (2016). *BOLD Data Portal - Search Results*. Boldsystems.org.

<https://portal.boldsystems.org/result?query=Carcharhinidae>

Chaudhary, C., Saeedi, H., & Costello, M. J. (2016). Bimodality of Latitudinal Gradients in Marine Species

Richness. *Trends in Ecology & Evolution*, 31(9), 670–676.

<https://doi.org/10.1016/j.tree.2016.06.001>

Diaz-Carballido, P. L., Mendoza-González, G., Yañez-Arenas, C. A., & Chiappa-Carrara, X. (2022). Evaluation

of Shifts in the Potential Future Distributions of Carcharhinid Sharks Under Different Climate

Change Scenarios. *Frontiers in Marine Science*, 8. <https://doi.org/10.3389/fmars.2021.745501>

Displayr. (2025, August 15). *How to calculate jaccard coefficients in displayr using R*. Displayr Help.

[https://help.displayr.com/hc/en-us/articles/4404196636943-How-to-Calculate-Jaccard-Coefficient
s-in-Displayr-Using-R](https://help.displayr.com/hc/en-us/articles/4404196636943-How-to-Calculate-Jaccard-Coefficients-in-Displayr-Using-R)

References

- Gallo, V., Cavalcanti, M., Silva, R., Silva, H., & Pagnoncelli, D. (2010). Panbiogeographical analysis of the shark genus *Rhizoprionodon* (Chondrichthyes, Carcharhiniformes, Carcharhinidae). *Journal of Fish Biology*, 76(7), 1696–1713. <https://doi.org/10.1111/j.1095-8649.2010.02609.x>
- Hillebrand, H. (2004). Strength, slope and variability of marine latitudinal gradients. *Marine Ecology Progress Series*, 273, 251–267. <https://doi.org/10.3354/meps273251>
- Lovelace, R., Nowosad, J., & Muenchow, J. (2019). Chapter 9 making maps with R | geocomputation with R. In *r.geocompx.org*. <https://r.geocompx.org/adv-map>

Acknowledgements

- Saira helped me figure out what topic to do. I originally started off with sea turtles but the data size was too small and she encouraged me to look into oceans and the differences in diversity which inspired the changes in my research question.
- Dr. Cottenie helped me realize what species would work best in my project. I told him about the issue when I was using the data on sea turtles and he suggested I switch to turtles in general. There was much more data but unfortunately, I could not use it since the NA values were not from turtles in sea, there was no data at all for location. This led me to switch to sharks instead which had more data that I could actually use to answer my research question.
- Rameen further discussed how to look into my NA values to see if my data is usable or not as she ran a test to see what values are missing and how much in her data. I ended up using this which led me to change my topic to sharks instead of using turtles.
- I used websites which I cited to try new figures and code for my results section.