Investigating the Correlation between Drosophila Genus Biodiversity and Geographic Regions Using Latitude and BIN Richness

Storyboard by Iroayo Toki

Introduction

- Fruitflies are widely used as research organisms in genetics due to the similarity in their genes with human beings. (Crowley, 2024)

- This makes them great for researching diseases like cancer, diabetes and Alzheimer's Disease. (Crowley, 2024)

- Factors like Temperature, presence of fruit bearing trees and climate affect the survival and reproduction of fruitflies. (Yu et al., 2022 , Roche et al., 2024)

- They have a genus name Drosophila with over 1000 species the most common being *Drosophila melanogaster.* (Tolwinski, 2017)

Hypothesis

- Since environmental factors can affect their survival and reproduction it is plausible to infer that certain ecosystems and by extension certain climatic or geographic regions would improve the chances of survival of fruitflies.

- Tropical regions are described to be warm all year round, with various kind of fruit bearing plants which are just the right conditions for fruitfly survival. (Smith, 2025)

- Since long term survival is important in creating biodiversity as useful mutations only occur once in an extended period of time, we can then generate the following hypothesis. (Kardos et al., 2021)
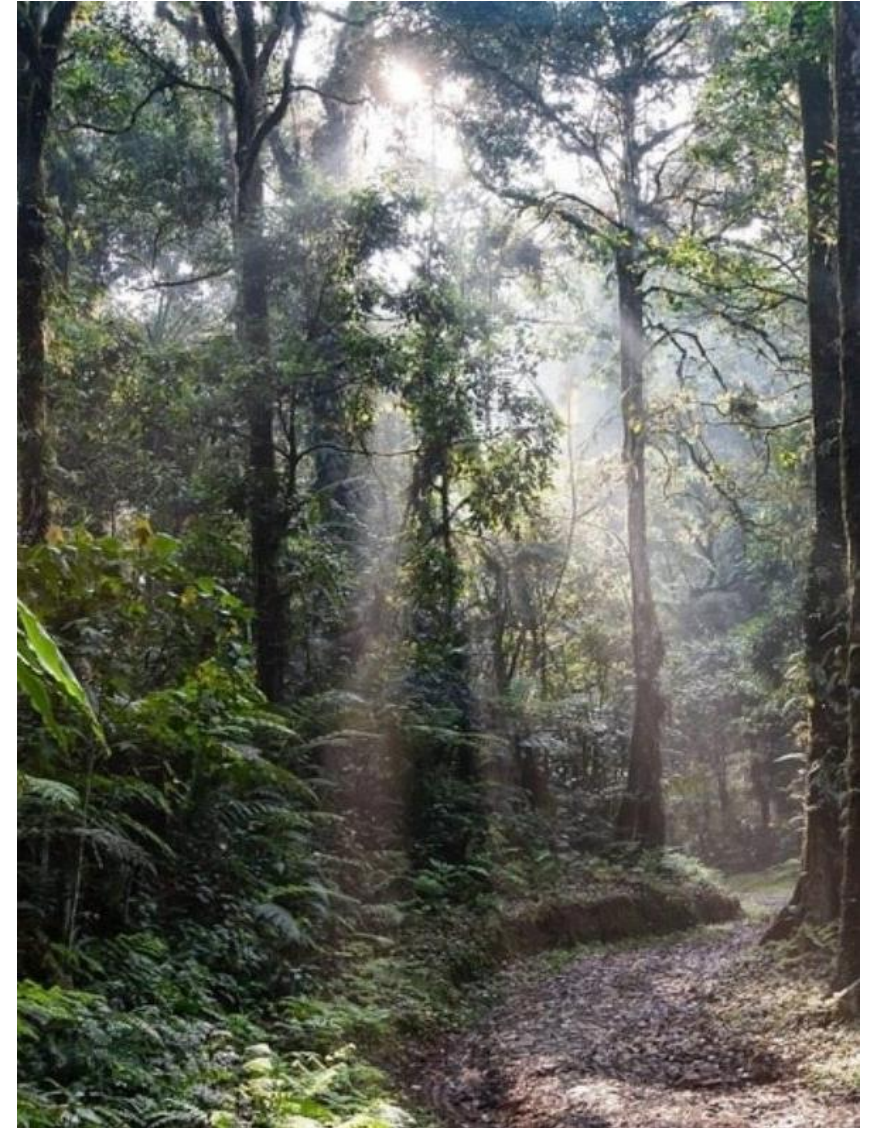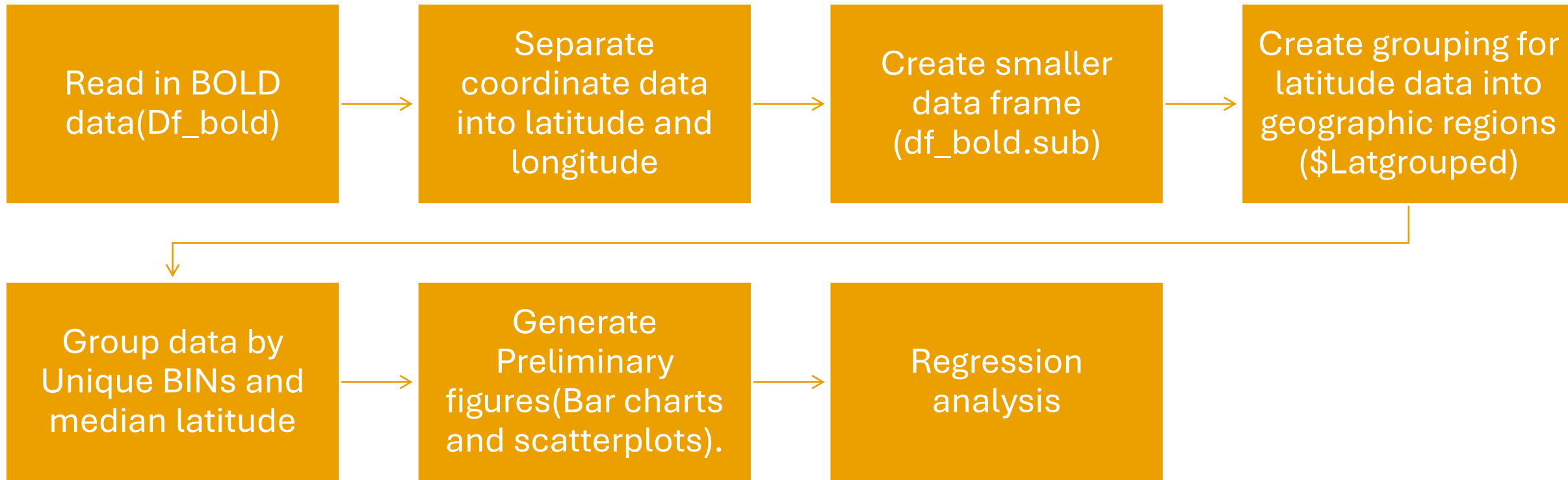
Hypothesis

- Lower environmental variability at lower latitudes as we get closer to the equator leads to increased biodiversity of fruitflies .

- This will be tested by comparing the number of BINs to the average(median) latitude they were found in data from BOLD.

- Studying biodiversity by geographic region can help us understand and improve species conservation by learning the environmental factors best for fruitflies.

Methods

Read in BOLD data(Df_bold) → Separate coordinate data into latitude and longitude → Create smaller data frame (df_bold.sub) → Create grouping for latitude data into geographic regions ($Latgrouped)

Group data by Unique BINs and median latitude → Generate Preliminary figures(Bar charts and scatterplots). → Regression analysis
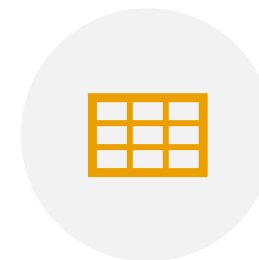
Analytical pipeline for regression analysis

GROUP LATITUDE DATA INTO BANDS OF 10°
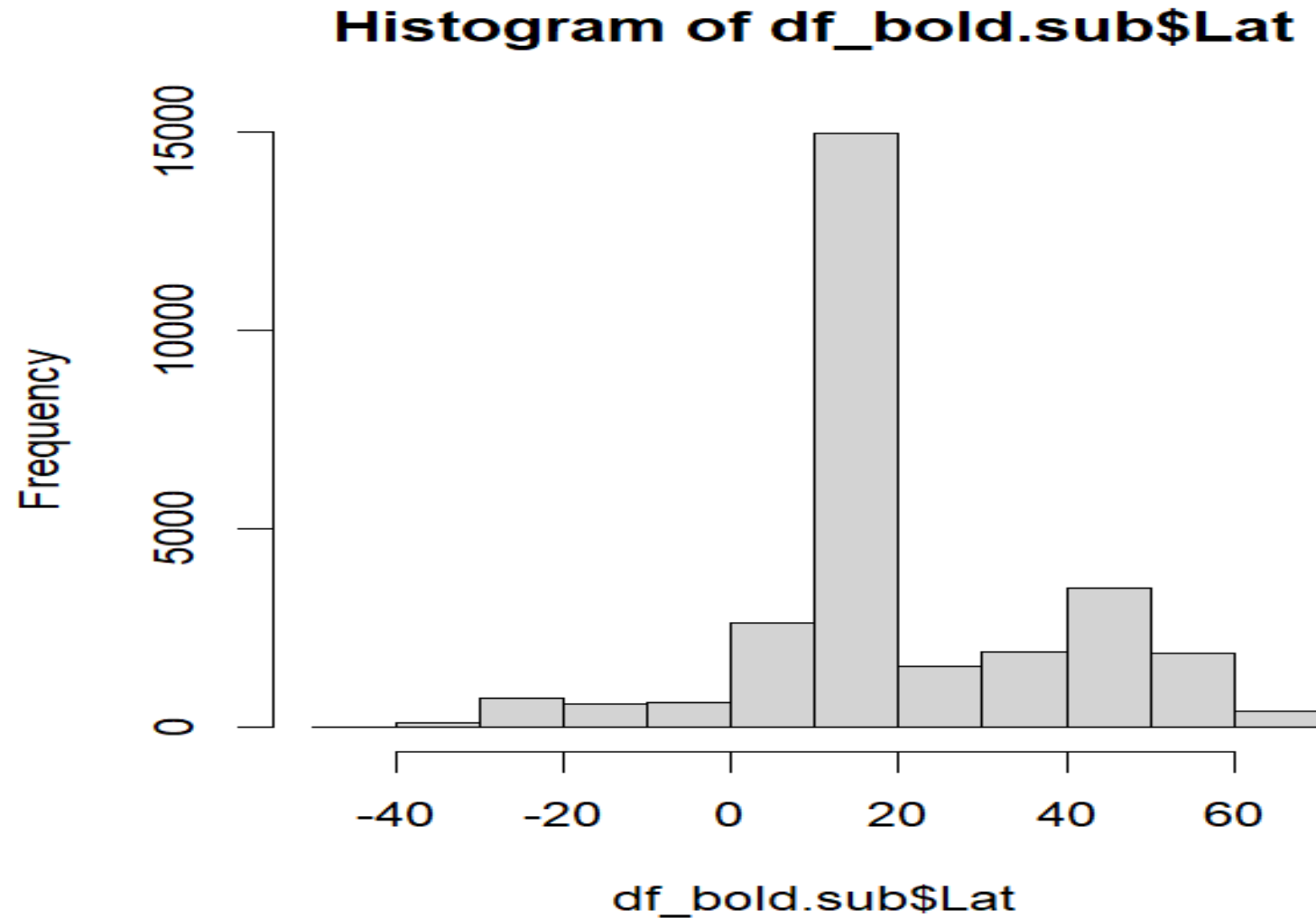
RUN PEARSONS CORRELATION TEST COR()

CREATE SIMPLE REGRESSION MODEL. MODEL <- LM()

GENERATE SCATTERPLOT SHOWING REGRESSION LINE

# Sampling disparity (Over 15000 samples) in the dataset between latitude 10-20°N

Histogram of df_bold.sub$Lat

Unique Drosophila BINs cluster in the tropical regions of the Northern hemisphere

Distribution of Drosophila BINs by latitude

Relationship between Latitude and BIN Richness
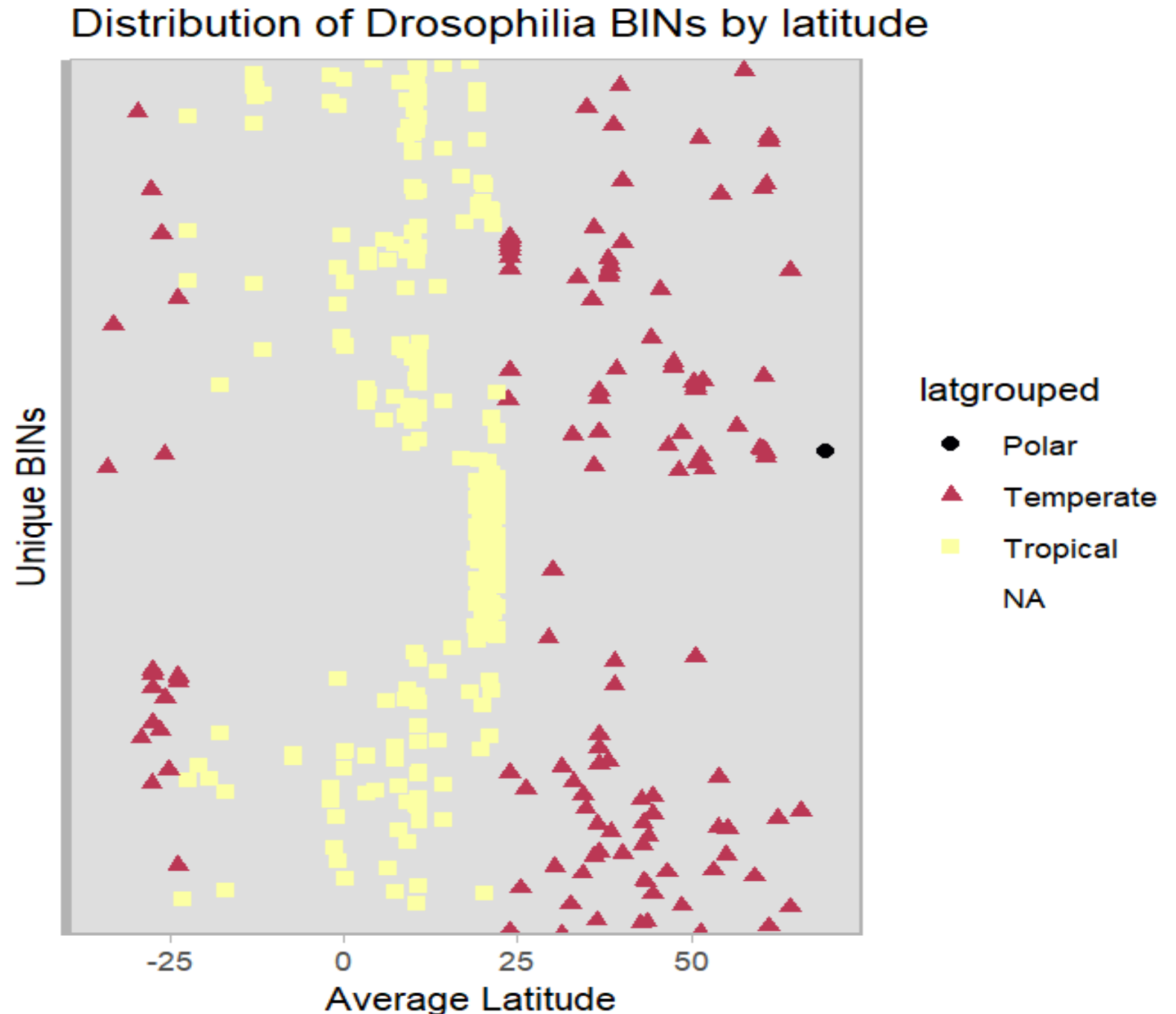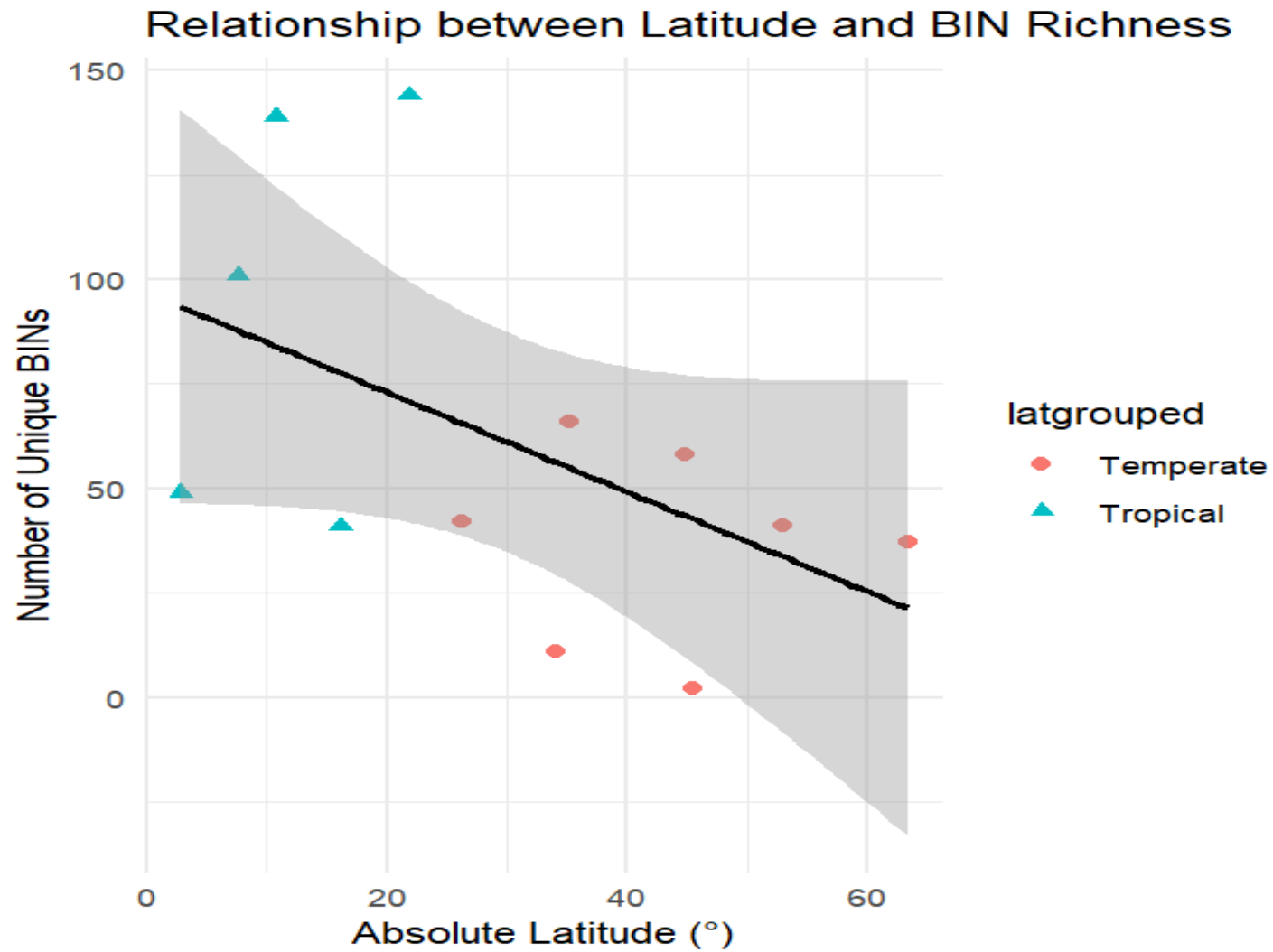
There is increased BIN richness in the Drosophila genus as we get closer to the equator(0°)
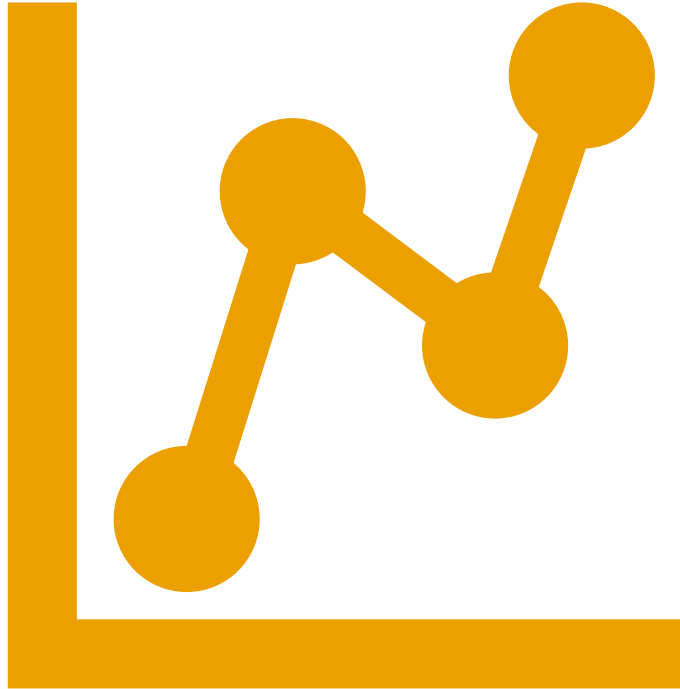
Results

- From the Pearson's correlation test we get a value of -0.5043, showing a moderately negative correlation between our variables i.e as our independent variable which was absolute latitude increased there was a decrease in BIN richness.

- For our Linear regression model we get a p-value of 0.09453 which shows that there is an effect however not statistically significant as it does not cross the $p<0.05$ threshold

Discussion

- Our results show that while there is a moderate negative correlation between latitude and BIN richness but it is not statistically significant. This could be due to a few reasons.

- A. According to a 2023 study there is increased biodiversity for insects in the outer bands of the tropical region and in the subtropical region compared to the inner bands which are close to the equator where they noted precipitation, wettest month and driest month as important factors, it creates this effect where there is first an increase in biodiversity as me move from the equator, then a drop as we move further into the temperate and tropic regions (French et al., 2023).

- B. As seen in the first figure there is a sampling bias in the data with about 15000 samples between 10-20°N while the next highest region has only about 5000 samples which may have affected our analysis. However, we tried to moderate its effect on our analysis by grouping the data into latitude bands before conducting statistical analysis

- The next steps would include

- i. conducting further sampling on other regions to match the samples in the 10-20°N latitude band

- ii. Studying environmental factors within the tropical region to understand how they affect biodiversity

Acknowledgements

References

- Crowley, R. (2024, January 17). *Research organism superheroes: Fruit flies*. BioBeat. National Institute of General Medical Sciences. https://nigms.nih.gov/biobeat/2024/01/research-organism-superheroes-fruit-flies

- French, C.M., Bertola, L.D., Carnaval, A.C. *et al*. Global determinants of insect mitochondrial genetic diversity. *Nat Commun* 14, 5276 (2023). https://doi.org/10.1038/s41467-023-40936-0

- M. Kardos, E.E. Armstrong, S.W. Fitzpatrick, S. Hauser, P.W. Hedrick, J.M. Miller, D.A. Tallmon, & W.C. Funk, The crucial role of genome-wide genetic variation in conservation, Proc. Natl. Acad. Sci. U.S.A. 118 (48) e2104642118, https://doi.org/10.1073/pnas.2104642118 (2021).

- MarinStatsLectures. (2020, October 9). Simple linear regression in R | R Tutorial 5.1 [Video]. YouTube. https://youtu.be/66z_MRwtFJM

- Roche, D. P., Stoleson, S. H., Rota, C. T., Oldland, W. K., & Lituma, C. M. (2024). Host plant species affect the abundance of spotted wing Drosophila and fruit parasitism across the Allegheny National Forest differentially at variable spatial scales. *Forest Ecology and Management, 568*, 122131. https://doi.org/10.1016/j.foreco.2024.122131

- Smith, J.M. (2025, September 16). *tropical rainforest*. *Encyclopedia Britannica*. https://www.britannica.com/science/tropical-rainforest

- Tolwinski NS. Introduction: Drosophila-A Model System for Developmental Biology. J Dev Biol. 2017 Sep 20;5(3):9. doi: 10.3390/jdb5030009. PMID: 29615566; PMCID: PMC5831767.

- Yu C, Zhao R, Zhou W, Pan Y, Tian H, Yin Z, Chen W. Fruit Fly in a Challenging Environment: Impact of Short-Term Temperature Stress on the Survival, Development, Reproduction, and Trehalose Metabolism of *Bactrocera dorsalis* (Diptera: Tephritidae). Insects. 2022 Aug 22;13(8):753. doi: 10.3390/insects13080753. PMID: 36005378; PMCID: PMC9410078.