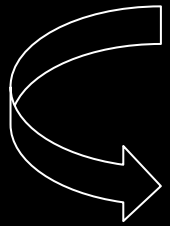
Do Canid Species at Higher Latitudes Have Larger Ranges?

Student: Fangyi Li (1092229)

Course: BINF*6210 Assignment 1

What I Wanted to Find Out

🎯Test whether Canid species that live farther from the equator have wider geographic ranges?



a pattern known as **Rapoport's rule**

- I explored global DNA barcode data (BOLD SYSTEMS) for the dog family (*Canidae*).
- I built this small project to see what real data say.

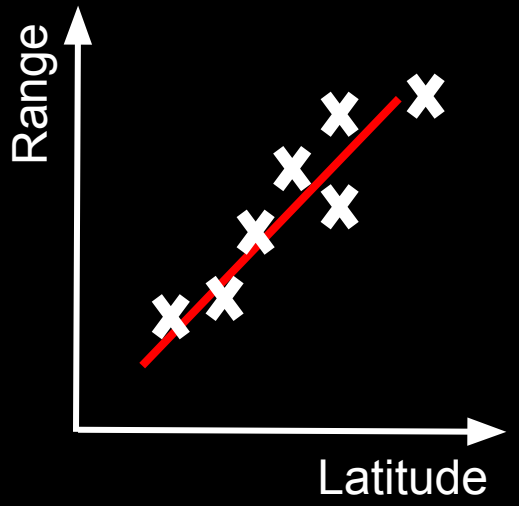
Rapoport's rule describes how species at higher latitudes often have broader geographic ranges (Stevens, 1989).

Does this rule hold true for Canid (dogs, wolves, foxes, jackals, etc.)?

Hypothesis (H_1):
Latitudinal range increases with median latitude (positive relationship).

Hypothesis (H_0):
No relationship exists between latitude and range size.

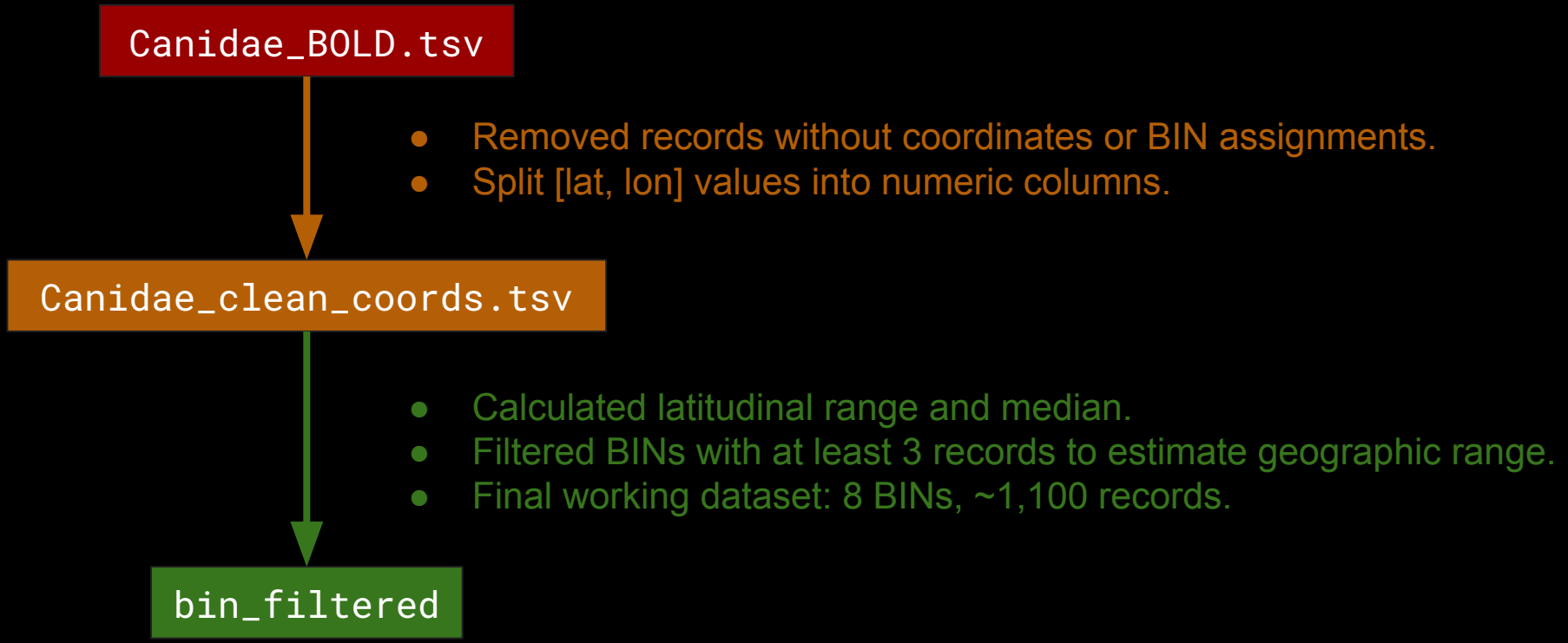
Expectation:
If Rapoport's rule holds, the regression line will slope upward.



Gathered Data from the BOLD SYSTEMS

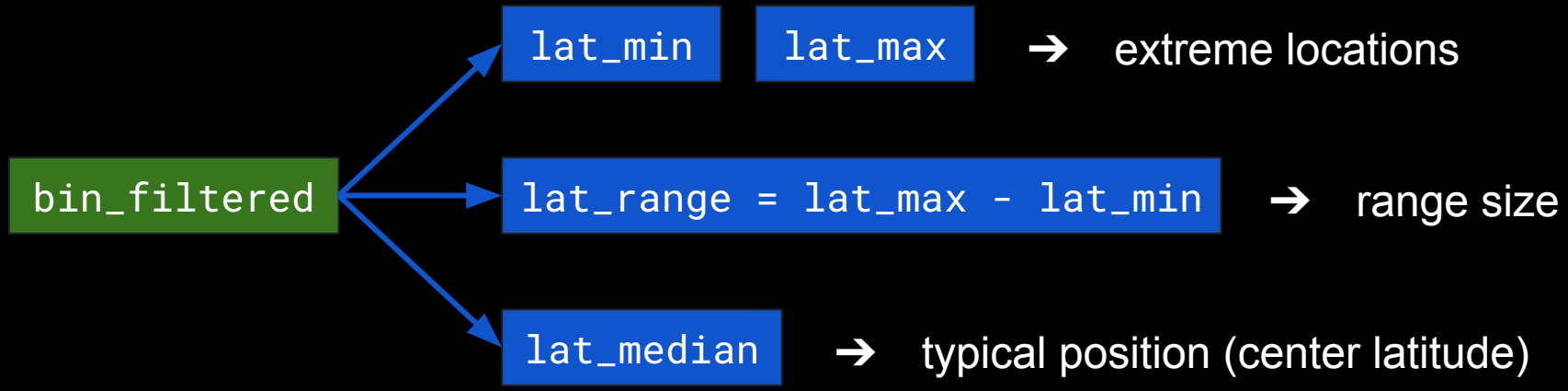
- Downloaded ~3,700 specimen records, 30 BINs (genetic clusters).
- Each record includes latitude, longitude, and species name.
- Dataset looks large, but geographic info was messy; many samples had missing coordinates.

Cleaning the Data to Keep Only Reliable Records



What I Calculated for Each BIN

For each BIN (≈ species), I summarized:



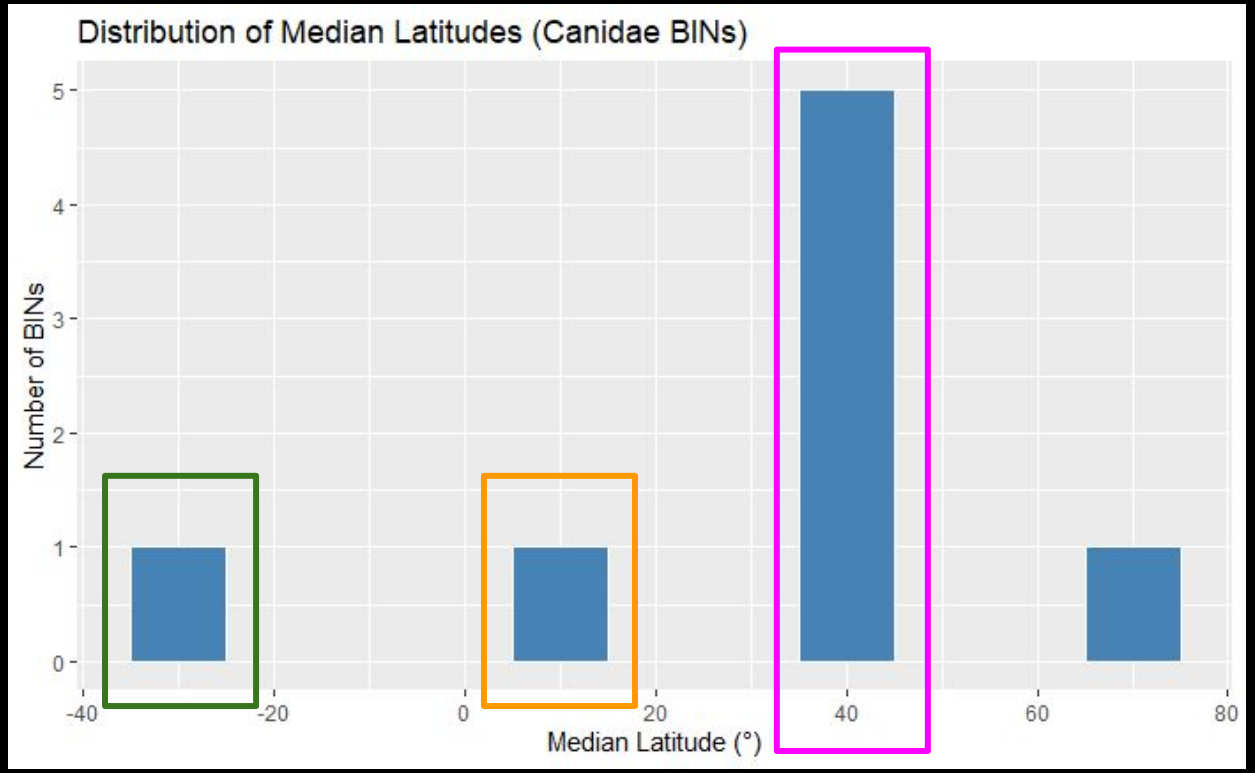
Then tested if BINs farther from the equator (higher latitude) have broader ranges.

8 Canid BINs remained after cleaning mostly from the Northern Hemisphere

	bin_uri	n_records	lat_min	lat_max	lat_range	lat_median	species	hemisphere
1	BOLD:AAA1542	844	-35.73359	65.01579	100.749383	12.89357	Canis lupus	Northern
2	BOLD:AAC5017	20	24.13800	53.20000	29.062000	38.82081	Canis latrans	Northern
3	BOLD:AAC5230	187	16.34300	49.25010	32.907100	38.82081	Urocyon cinereoargenteus	Northern
4	BOLD:AAC5231	58	31.91670	76.53000	44.613300	74.16847	Vulpes velox	Northern
5	BOLD:ACR0824	4	-28.55362	36.40239	64.956006	-28.55362	Lycaon pictus	Southern
6	BOLD:ADC5726	26	32.90602	62.73421	29.828185	38.93208	Vulpes vulpes	Northern
7	BOLD:ADK6164	5	32.90602	38.93210	6.026075	38.93208	Vulpes vulpes	Northern
8	BOLD:ADM0647	4	31.51354	48.57328	17.059738	38.93208	Canis aureus	Northern

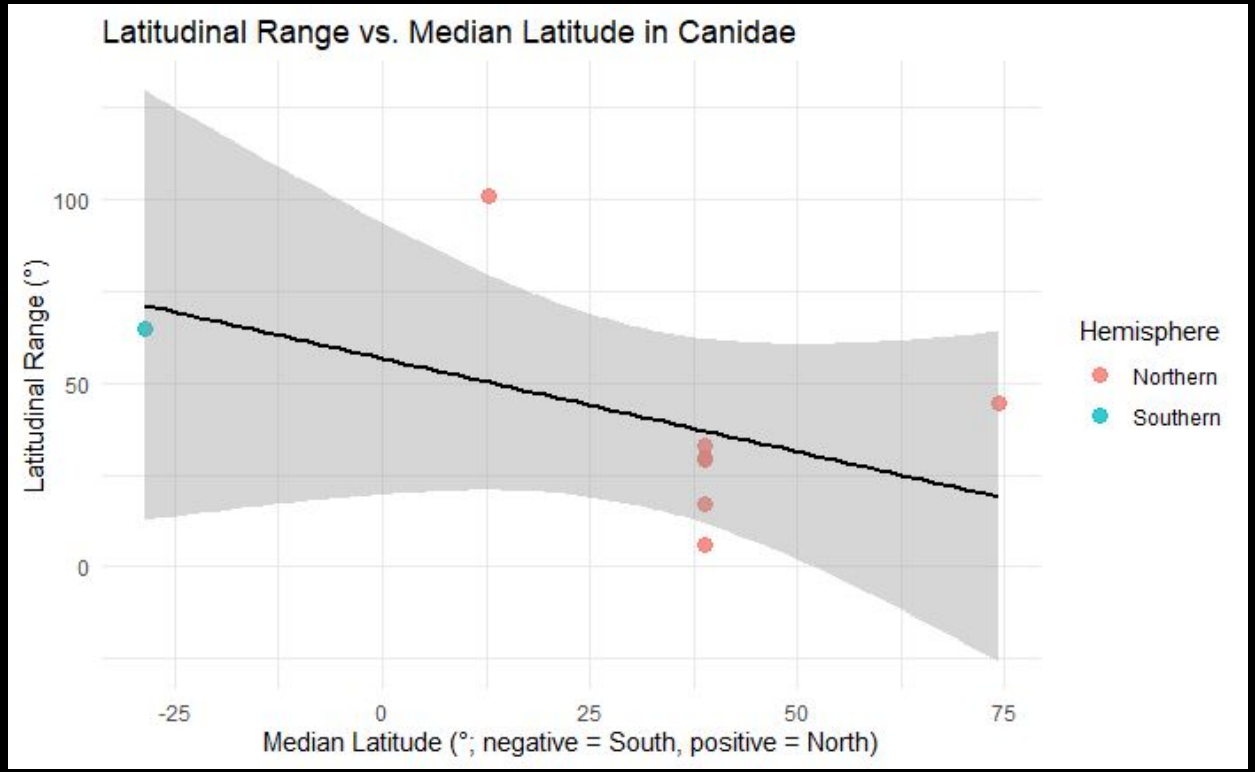
- 8 BINs passed all filters with valid coordinates.
- Each BIN represents 1 Canid species or cluster.
- 7 occur in the north, one (*Lycano pictus*) in the south.
- These data provide the basis for the following analysis

Almost all BINs cluster between 30°–50° North latitude



- Histogram of BIN median latitudes shows concentration in temperate zones.
- Suggests limited coverage in tropical and southern regions.

Do higher-latitude BINs actually have wider ranges?



- This scatterplot shows each BIN's median latitude versus its latitudinal range, with colors indicating hemisphere.
- The regression line slopes slightly downward, showing no visible support for Rapoport's rule.

Statistics show no significant relationship between latitude and range size

```
Call:
lm(formula = lat_range ~ lat_median, data = bin_filtered)

Residuals:
    Min      1Q  Median      3Q     Max
-30.918 -10.925  -6.649   3.310  50.612

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  56.6705    15.1339   3.745  0.00957 **
lat_median   -0.5067     0.3612  -1.403  0.21026
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 28.09 on 6 degrees of freedom
Multiple R-squared:  0.247,    Adjusted R-squared:  0.1215
F-statistic: 1.968 on 1 and 6 DF,  p-value: 0.2103
```

- Linear regression found a weak negative slope ($\beta = -0.51$, $p = 0.21$, $R^2 = 0.25$).
- The p-value is above 0.05, meaning the trend is not statistically significant.
- Therefore, Canid BINs at higher latitude do not have larger geographic ranges in this dataset.

The expected pattern was not observed in Canidae

- The regression showed no significant link between latitude and range size.
- The trend was slightly negative instead of positive, opposite to Rapoport’s rule.
- This likely reflects small sample size and limited southern-hemisphere records.
- One widespread species (Canis lupus) strongly influenced the dataset.
- Latitude alone may not capture the environment factors shaping range size.

Overall, the Canidae data did not support Rapoport’s rule, but the analysis demonstrates how BOLD records can be used to test ecological hypothesis.

How this project could be improved next…

- ➔ Add more BINs by combing Canidae with other carnivore families.
- ➔ Include tropical and southern-hemisphere species to balance the dataset.
- ➔ Incorporate environmental variables such as temperature or habitat type.
- ➔ Apply the same workflow to test other ecological rules.

Reference

[1] Ratnasingham S, Wei C, Chan D, Agda J, Agda J, Ballesteros-Mejia L, Ait Boutou H, El Bastami Z M, Ma E, Manjunath R, Rea D, Ho C, Telfer A, McKeowan J, Rahulan M, Steinke C, Dorsheimer J, Milton M, Hebert PDN . "BOLD v4: A Centralized Bioinformatics Platform for DNA-Based Biodiversity Data." In DNA Barcoding: Methods and Protocols, pp. 403-441. Chapter 26. New York, NY: Springer US, 2024

[2] Ratnasingham, S. & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). Molecular Ecology Notes 7, 355-364. DOI: 10.1111/j.1471-8286.2007.01678.x

[3] Stevens, G. C. (1989). The latitudinal gradient in geographical range: How so many species coexist in the tropics. *The American Naturalist*, *133*(2), 240–256. https://doi.org/10.1086/284913

[4] Wickham H, Hester J, Bryan J (2024). *readr: Read Rectangular Text Data*. R package version 2.1.5, https://github.com/tidyverse/readr, https://readr.tidyverse.org.

[5] Wickham H, François R, Henry L, Müller K, Vaughan D (2025). *dplyr: A Grammar of Data Manipulation*. R package version 1.1.4, https://dplyr.tidyverse.org.

[6] Wickham H, Vaughan D, Girlich M (2025). *tidyr: Tidy Messy Data*. R package version 1.3.1, https://tidyr.tidyverse.org.

[7] Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org