Lifespan Analysis of Lemurs in Wild and Captive Environments*

Evaluating Effective Strategies and Highlighting Areas for Conservation Focus

Khushaal Nandwani

December 3, 2024

This study investigates the lifespan of lemurs in wild and captive environments using data from the Duke Lemur Center, aiming to identify how factors such as sex, species, genus, and birth month influence longevity. Using Generalized Linear Models, the analysis highlights that lemurs in captivity live significantly longer than their wild counterparts, emphasizing the impact of conservation practices. The findings show species-specific variations in lifespan and the pronounced influence of sex differences in captivity. These insights help identify effective practices in captivity and suggest areas for improvement to optimize conservation strategies, contributing to the survival of this endangered mammalian group.

Table of contents

---

*Code and data are available at: https://github.com/khushaal-nandwani/predicting-lemurs/tree/main.

1 Introduction

Lemurs are the world's most endangered group of mammals (Duke Lemur Center (2024)). Understanding the factors that influence their lifespan is crucial for guiding effective conservation efforts. The Duke Lemur Center (DLC), established in 1966, has played a pivotal role in the study and preservation of lemurs, maintaining the largest population of these species outside their native region Madagascar. Different factors can affect the longevity of lemurs in wild and captive. Longevity is important because the longer lemurs live, the more likely they are to reproduce and hence, contribute to the survival of their species.

My study uses a dataset from the DLC to analyze if and how sex, species, genus and month a lemur is born in can affect its lifespan in the wild and captivity. This study can provide insights into which lemur species we are struggling to conserve in captivity and their performance in the wild. It will highlight areas where our conservation efforts have been effective.

The estimand for our model is the age or the lifespan of the lemur. The predictors are `Sex`, `Species`, `Genus`, and `month_born`.

Using Generalized Linear Models (GLMs) with Gaussian distribution, factors including sex, species, genus, and birth month were evaluated to predict lemur longevity in wild and captive. To do this, the dataset was divided into wild-born and captive-born lemurs to understand the differences in lifespan between the two environments and separate models were built for each group.

The findings reveal that lemurs in captivity exhibit significantly longer lifespans. Males tend to live longer than females in both captive and wild environments, however the difference is more pronounced in captivity. Species-specific variations in lifespan were observed, with some species thriving in captivity while others perform better in the wild. Birth month effects were inconsistent, suggesting complex interactions with environmental factors.

The remainder of this paper is structured as follows: Section 2 details the data collection and processing methods. Section 3 outlines the modeling approach. Section 4 presents the results, highlighting key differences between wild and captive lemurs. The discussion in Section 5 contextualizes these findings, addressing limitations and proposing future directions for research. Finally, Section A is the appendix and includes additional information on data cleaning, the analysis dataset, and the model dataset.

2 Data

The data was taken from the GitHub repository of Cookson (2020), who acquired it from Duke Lemur Center. Details on how data was acquired and processed can be found in Section 2.1. We used R (R Core Team 2023) along with its packages to clean and analyze the data. These packages include `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `patchwork` (Pedersen 2024), `arrow` (Richardson et al. 2024), `knitr` (Xie 2014), `rstanarm` (Goodrich et al. 2024) and `kableExtra` (Zhu 2024). `testthat` (Wickham 2011) was used for testing the simulated and actual datasets.

It is important to know about Lemurs and what affects their lifespan because they are the most endangered mammals on the planet. The Duke Lemur Center (DLC) is a global leader in the research, care, and conservation of lemurs. The DLC hosts the most diverse population of lemurs outside their native habitat in Madagascar.

To do so, I chose the following variables from the dataset, which I believe are the most important factors that can affect the lifespan of a lemur:

- **Sex**: The sex of the lemur on birth. It is a cateogrical variable that can be `M` or `F`. I took sex as a factor because it is common to see females tend to live longer (Lemaître et al. 2020).
- **Species**: The species of the lemur. It is also a categorical variable and can take one of the following values: `GG`, `COL`, `UL`, `RUF`, `MOH`, `MAC`, `CAT`, `FUL`, `ALB`, `AR`, `VV`, `COQ`, `MED`, `MUR`, `COU`, `TAR`, `PYG`, `ZAZ`, `MON`, `RUB`, `COR`, `SAN`, `FLA`, `MAD`, `POT`. The specific or common names of these species can be found in Section G. Species is important as different species have different requirements and threats, which can affect their lifespan.
- **Genus**: The genus of the lemur. It is a categorical variable and can take one of the following values: `O`, `E`, `G`, `H`, `L`, `V`, `P`, `C`, `M`, `N`, `D`. Genus can have the same effect as species.
- **month_born**: The month in which the lemur was born. It is a categorical variable and can take one of the following values: `1`, `2`, `3`, `4`, `5`, `6`, `7`, `8`, `9`, `10`, `11`, `12`, where `1` corresponds to January, `2` to February, and so on. Month born can be important as it can affect the availability of food and other resources during its early stages, which can affect the lifespan of the lemur.

Finally, we have the target variable `Age` which is the lifespan of the lemur in years. It is a continuous variable.

Cookson (2020) also had another dataset which contained the weight of these lemurs, but I did not use it in this analysis. It includes details into pregnancy and weights of lemur during ther lifetime, it can be added to see the how trends in weight, pregnancy and litter size can affect the lifespan of lemurs.

The following figures show the distribution of the lemurs in the dataset based on the variables `Sex` (Figure 1), `Species` (Figure 2 and Figure 3), `Genus` (Figure 4), and `month_born` (Figure 5) for both lemurs in captivity, shown in orange and in the wild, shown in dark green.

For `Sex` (Figure 1), we observe that in the wild, there are more females than males. This might be due to the fact that females tend to live longer than males for most mammals (Lemaître et al. 2020). While in captivity, there are significantly more male lemurs than males. This shows a short coming in the conservation efforts, as we are not able to conserve the females, which are more important for the survival of the species.

Looking at the distribution of `Species` (Figure 11). We observe that in the wild, the distribution is relatively balanced, with a few exceptions of `MAD`, `CAT` and `MED` which appear to be low in number. In captivity, certain species (e.g., MUR, CAT, MED) are more represented, while others are rare or absent. This can be due to captive environments may focus on saving the most endangered species, as we see the ones least in wild like `CAT` and `MED` are more in captivity.

With the exception of genus `E`, genus (Figure 4) representation is balanced in both wild and captive lemurs. `E` might be high in both cases, as it could be a common genus or have a higher conservation priority.

For birth month (Figure 10), births in the wild are concentrated in a few months, suggesting a breeding season. While in captivity, births are more evenly spread across the year, with some peaks in specific months. Seasonal breeding in the wild is influenced by environmental factors like food availability and climate. Captivity can disrupt these natural cycles, as controlled environments and year-round resources allow for more frequent and less seasonal reproduction.

For Age (Figure 6), we observe that in the wild, the age distribution of lemurs shows a higher frequency of younger lemurs, with relatively few individuals reaching older ages. In captivity, the age distribution extends further, with a significant number of lemurs living to older ages. This difference can be attributed to the controlled environments in captivity, which provide consistent food, medical care, and protection from predators, thus increasing longevity. In contrast, the challenges of the wild, such as predation, disease, and fluctuating resources, contribute to shorter lifespans.
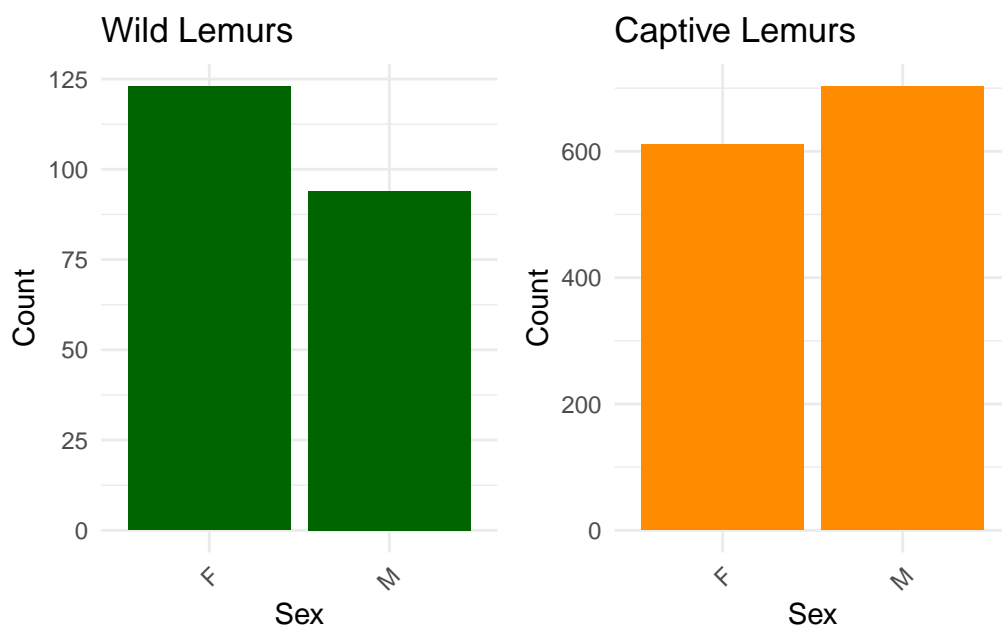
Figure 1: Sex distribution of wild and captive lemurs for female (F) and male (M) individuals.

2.1 Data Collection

The data in these sources was acquired and processed by staff at the Duke Lemur Center (DLC).

Figure 2: Species-wise count of wild lemurs in our dataset.

Figure 3: Species-wise count of captive lemurs in our dataset.

Figure 4: Distribution of lemurs by genus in the wild (on left) and in captivity (on right).

Figure 5: Distribution of lemurs by the month they were born in the wild (on left) and in captivity (on right).

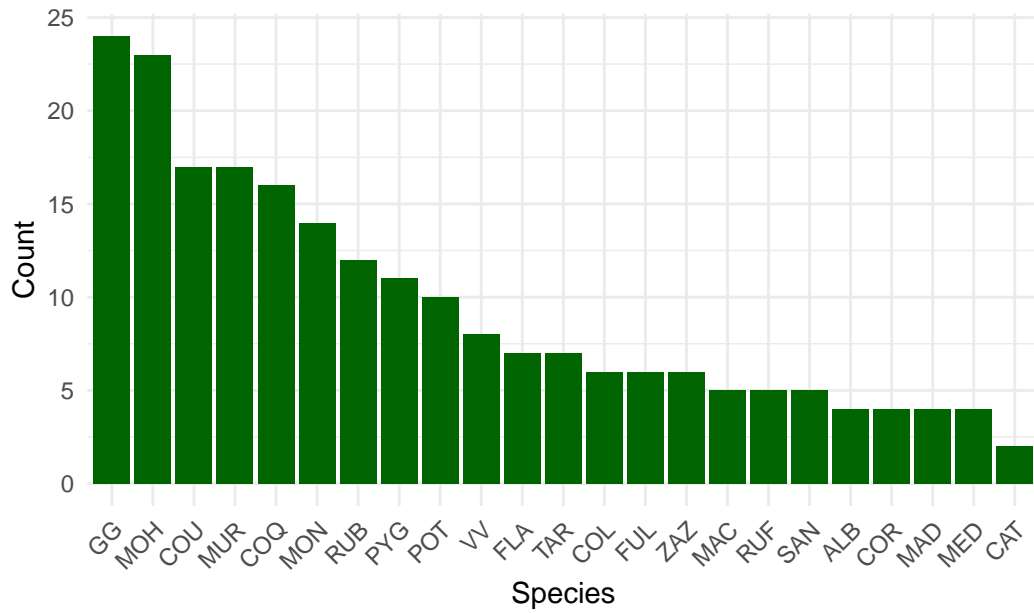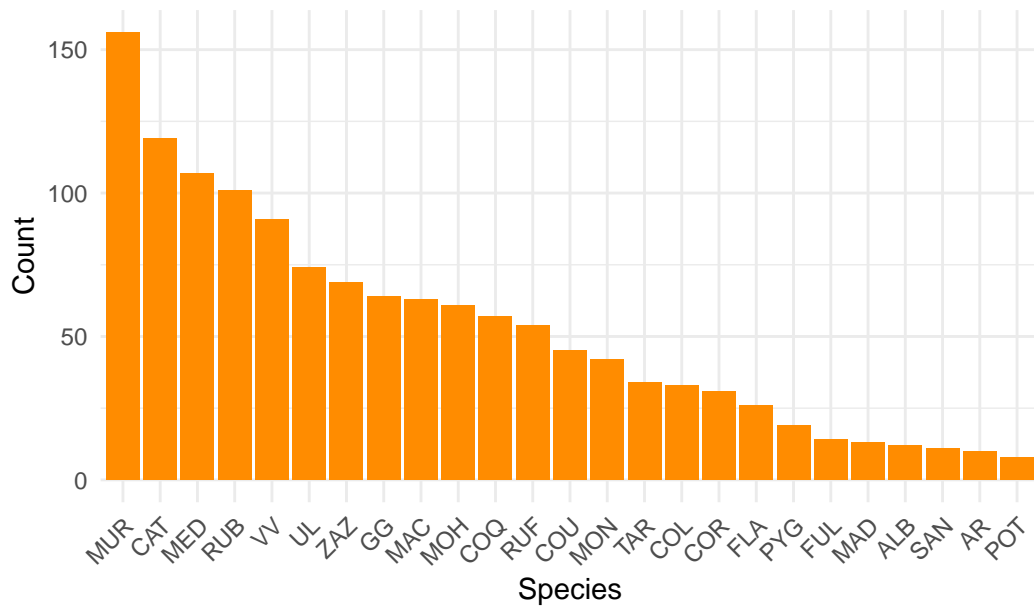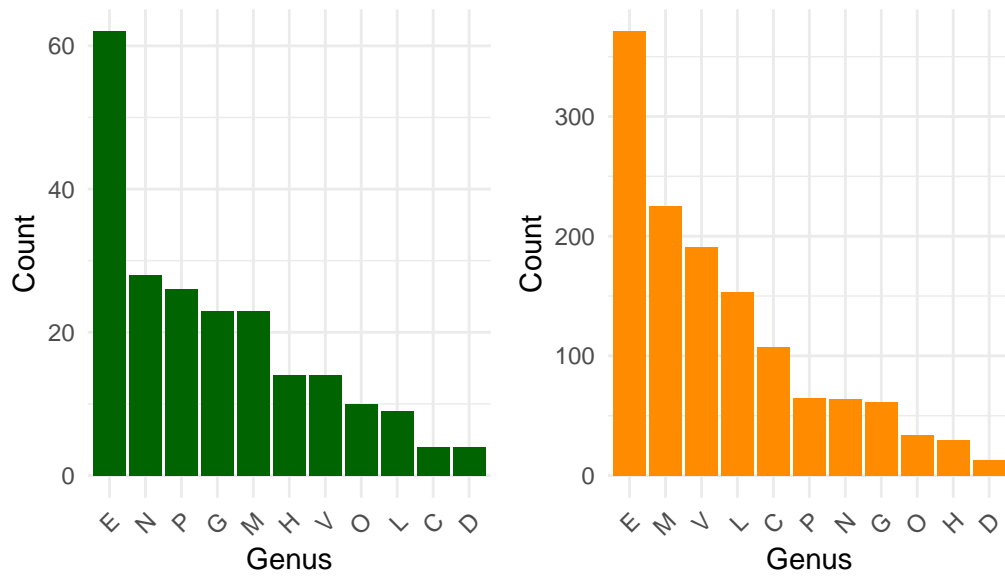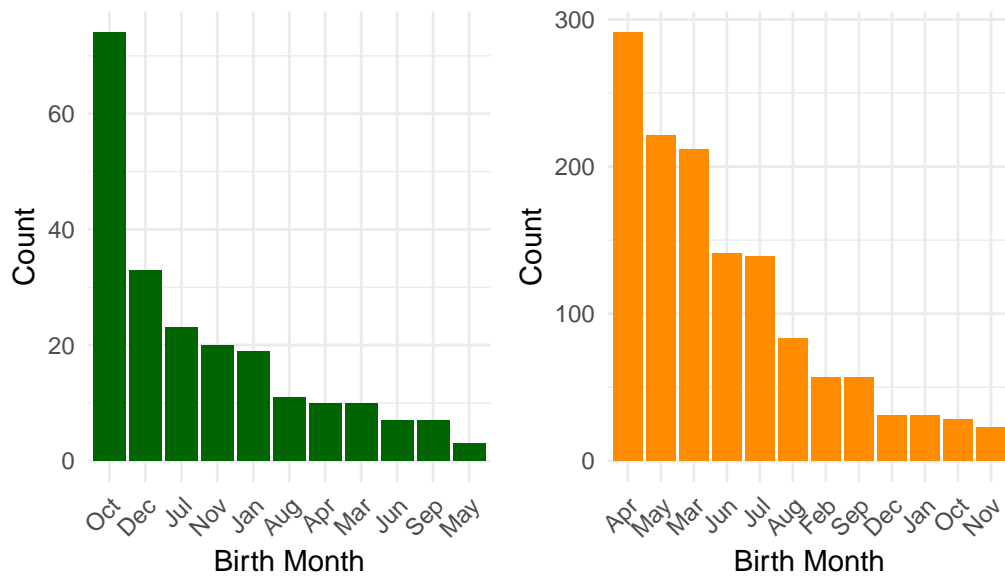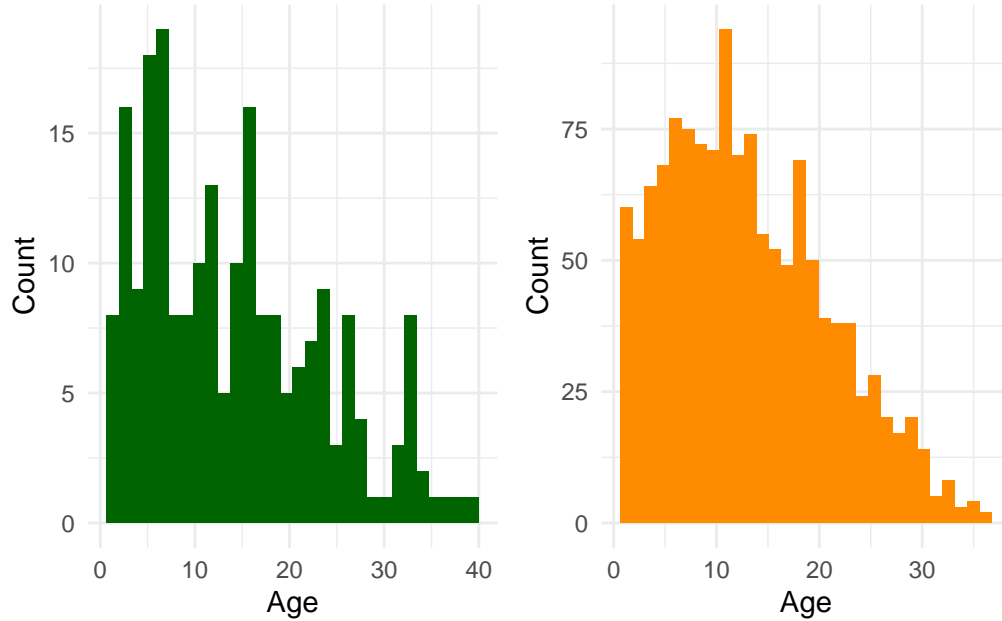Figure 6: Distribution of ages of lemurs in the dataset in wild (on left) and in captivity (on right).

2.1.1 Data Acquisition

As Zehr et al. (2014) points out, DLC staff collected data about the lemurs according to standard operating procedures and USDA, AZA, and IACUC guidelines. They recorded information about births, deaths, weights, enclosure moves, behaviors, and other significant events on a daily basis. Originally, this data was stored in handwritten and typed paper formats. Later, it was computerized.

In the mid-1990s, the DLC started using two databases: the Animal Record Keeping System (ARKS) and MedARKS. These databases allowed the DLC to share information with other organizations through the International Species Information System (ISIS). The DLC is currently transitioning to using the Zoological Information Management System (ZIMS). Data not stored in these databases has been stored in spreadsheets, and the DLC is working on transferring data from older records into these databases.

2.1.2 Data Processing

The DLC used SAS software to build a database for the lemur data. Data from various sources was imported into SAS Enterprise Guide, including ARKS, MedARKS, ZIMS, and spreadsheets. They wrote programs in SAS to extract, match, and join data, calculate new variables, and format the output. They also used tools within SAS Enterprise Guide Projects

for calculations and formatting. The DLC uses a unique ID to match data for individual animals, and the taxonomic name for species-related variables. The data was validated by identifying and locating missing data, standardizing codes and text, investigating outliers, and comparing known values to the database output.Data that could not be verified was excluded from the published dataset.

The DLC created two data files from the database: **the DLC Animal List**, which contains single-copy variables for each animal in the colony's history, and **the DLC Weight File**, which contains all weight measurements for each animal. We used the DLC Animal List for this analysis, which was first cleaned by Cookson (2020) and then I clean it. The steps I took for cleaning the data can be found in Section C.

The data in these sources was updated on February 8, 2019. The DLC plans to update the data on a yearly basis.

3 Model

I used Generalized Linear Model (GLM) framework with a Gaussian family distribution to determine the expected lifespan of a lemur based on the selected variables for both models, wild and captive. `birth type` (captive or wild) was not chosen as a predictor variable because there is significantly less data for wild and I thought it would be easier to visualize the predictions when done seperately. However, in doing so, we lost the oppoturnity of studying interaction effects of `birth_type` with other predictors.

GLM is a statistical model that generalizes linear regression to include non-normal distributions. It is used when the dependent variable is not normally distributed or when the relationship between the dependent and independent variables is not linear. The Gaussian family distribution is used when the dependent variable is continuous and normally distributed, which was in my case.

My model is based on `Sex`, `Species`, `Genus`, and `month_born` as the independent variables and `Age` as the dependent variable.

Mathematically, the model can be represented as:

$$\bar{a} = \beta_0 + \beta_1 \times \texttt{Sex} + \beta_2 \times \texttt{Species} + \beta_3 \times \texttt{Genus} + \beta_4 \times \texttt{month\_born} + \epsilon \qquad (1)$$

where:

- $\bar{a}$ is the expected lifespan of a lemur in the wild or captive.

- β_0 is the intercept term. It represents the expected lifespan of a lemur when all other variables are zero.

- β_1 is the coefficient for `Sex`. It represents the change in the expected lifespan of a lemur for a one-unit change in `Sex`.

- β_2 is the coefficient for `Species`. It represents the change in the expected lifespan of a lemur for a one-unit change in `Species`.

- β_3 is the coefficient for `Genus`. It represents the change in the expected lifespan of a lemur for a one-unit change in `Genus`.

- β_4 is the coefficient for `month_born`. It represents the change in the expected lifespan of a lemur for a one-unit change in `month_born`.

- ϵ is the error term.

I chose a Gaussian family distribution for the model because the dependent variable `Age` is continuous and normally distributed. `glm` was used to create the model which is available in the base `R` package. The model was then saved using `saveRDS` function to be used for predictions and further analysis.

Section E covers coefficients and standard errors and @#sec-appendix-ppc includes posterior predictive checks for the model.

A combined model was tried, where `birth_type` was taken as the predictor variable, but it had worse residual and p values, so I decided to go with two separate models for wild and captive lemurs. The seperate models allow us to better understand the factors affecting the lifespan of lemurs in each environment, making them easy to visualize. At the expense of, not being able to study the interaction effects of `birth_type` with other predictors.

4 Results

Figure 7 and Figure 8 shows the coefficients and their 95% credible intervals for the predictor variables in the respective wild and captive generalized linear regression models. A 95% credible interval indicates a 95% chance that the true parameter value falls within the interval, based on the observed data and model assumptions. Each point represents the estimated coefficient for a predictor, and the horizontal lines show the credible interval around it. Coefficients to the right of zero suggest a positive relationship with the outcome variable, meaning that as the predictor increases, the outcome also increases. Coefficients to the left of zero suggest a negative relationship, meaning that as the predictor increases, the outcome decreases.
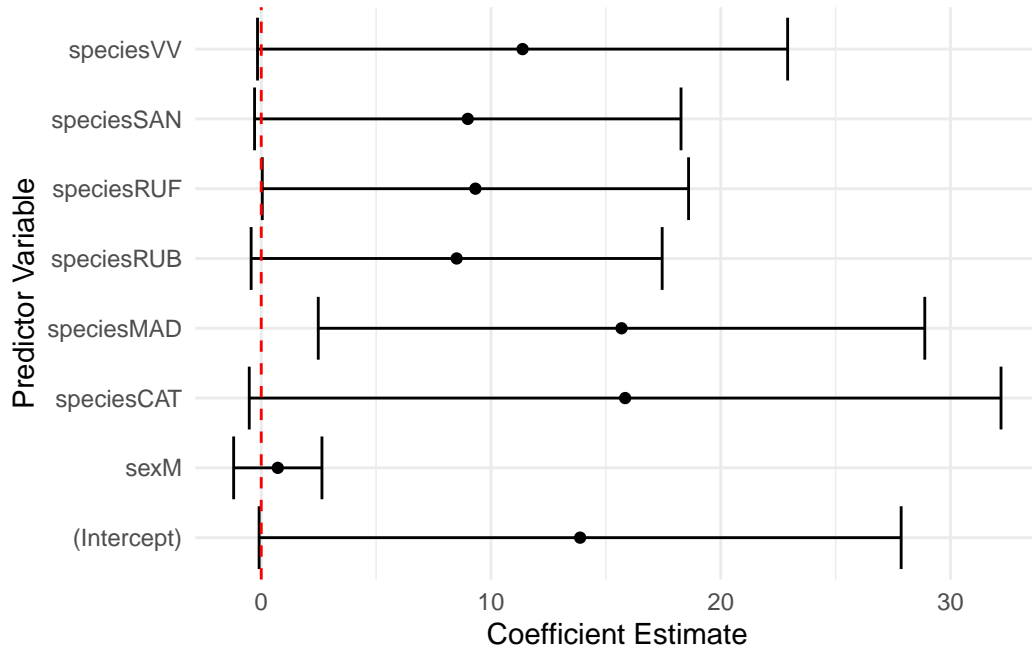
Figure 7: Significant coefficients (p-value < 0.07) with their 95% credible intervals for the wild model.
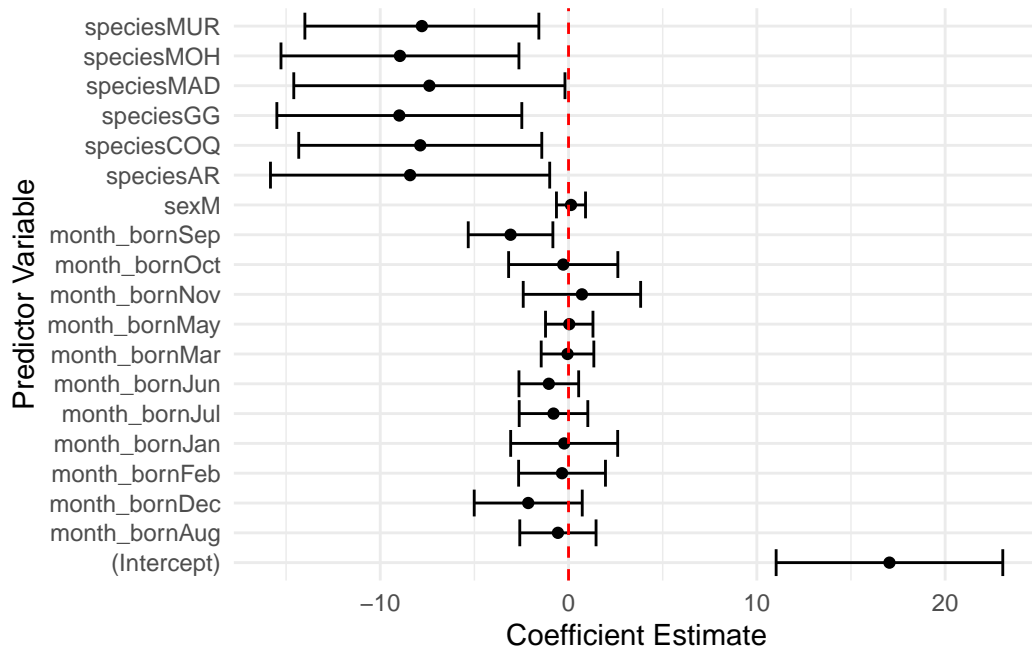

Figure 8: Significant coefficients (p-value < 0.05) with their 95% credible intervals for the captive model.

4.1 Wild Lemurs

Few species in the wild have significant effect on the age, these include VV, SAN, RUF, RUB, MAD and CAT. SAN, RUF and RUB have a negative effect on age, while VV, MAD and CAT have a positive effect.

For sex, in wild, males have a higher lifespan than females. Higher lifespan of males suggest that the more number of females observed in the wild in Figure 1 might have other underlying cause than just females living longer, which can be a good area of research for Duke Lemur Center.

Note that we had to increase our threshold of p-value to 0.07 to get the significant results, as we had limited data in case of wild.

4.2 Captive Lemurs

Few species have a negative effect on age, these include MUR, MOH, MAD, GG, COQ and AR. I did not find any specific species with positive effect on age in captivity. This is a good sign as it shows that the conservation efforts are working for most of the species in captivity.

In case of captivity, sex is not a major factor in age, as its confidence interval is nearly centered at 0. Another good sign towards conservation effects.

Birth month effects appear less consistent, mostly centered around 0. Except December and September that seem to be negatively affecting the age of lemurs in captivity, while November slightly positively.

4.3 Effect of Sex on Age

From the Figure 9, lemurs in captive have a higher average predicted age than those in the wild across sexes. And in the captive, we see males living longer on average than females, while in the wild, the difference is less pronounced. This can be due to the controlled environment in captivity, which provides consistent food, medical care, and protection from predators, thus increasing longevity. In the wild, the challenges of predation, disease, and fluctuating resources contribute to shorter lifespans.

4.4 Effect of Month Born on Age

For months in Figure 10, we see captive lemurs outlive wild lemurs most of the months except August, January, March, Novemeber and September. It is worth exploring the reasons behind these differences by Duke Lemur Center, as they can provide insights into the impact of seasonal factors on lemur lifespan.
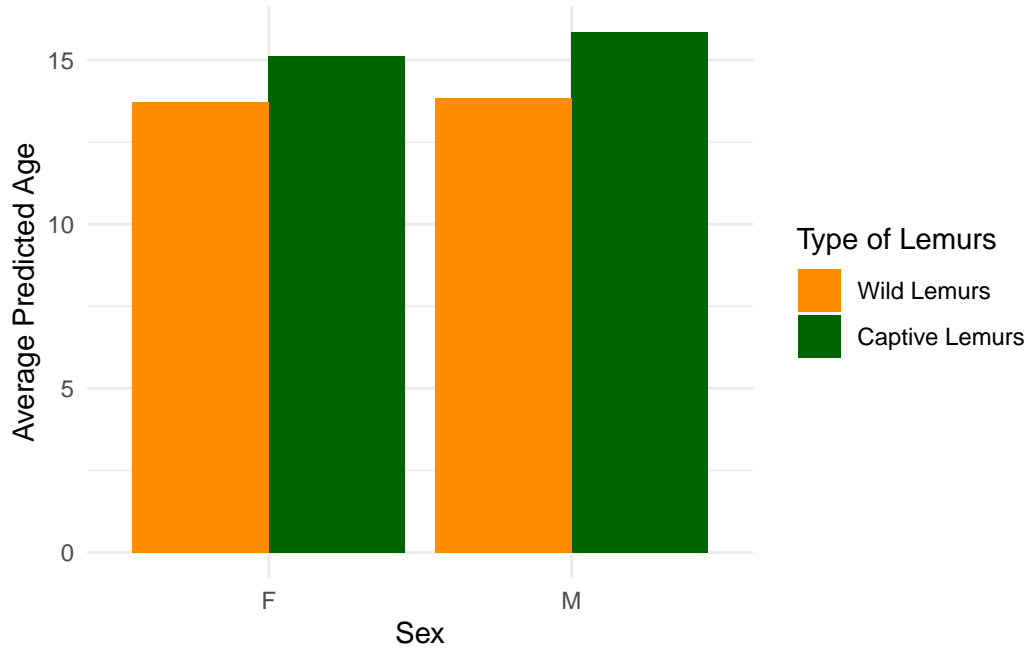
Figure 9: Comparison of average predicted ages of wild and captive lemurs across sexes. Males live longer in both wild and captive environments, but the difference is more pronounced in captivity.
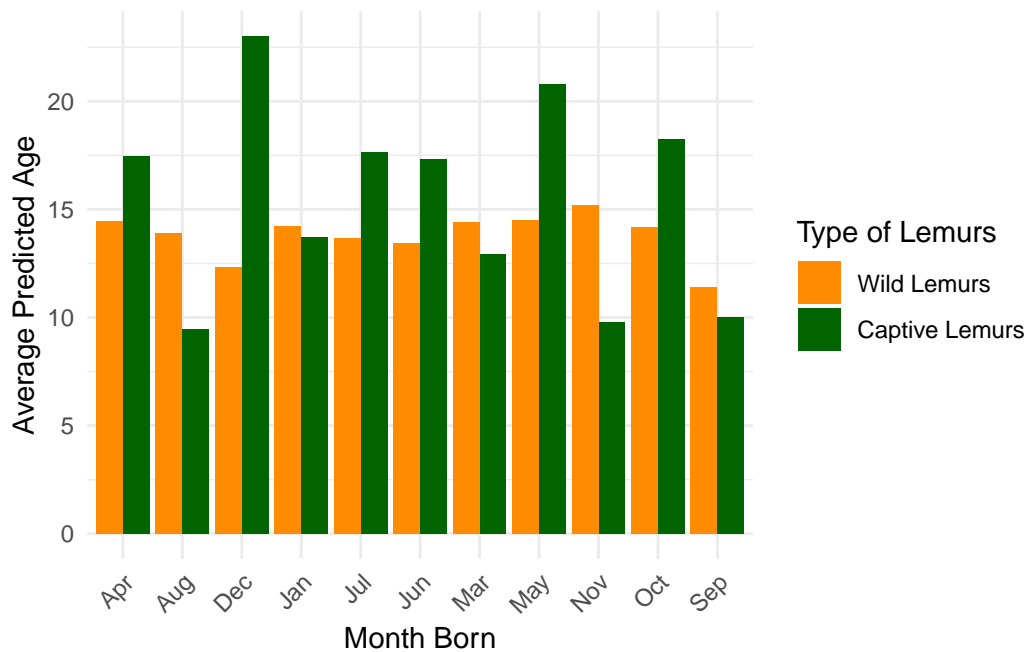
Figure 10: Comparison of average predicted ages of wild and captive lemurs across different months born.

4.5 Effect of Species on Age
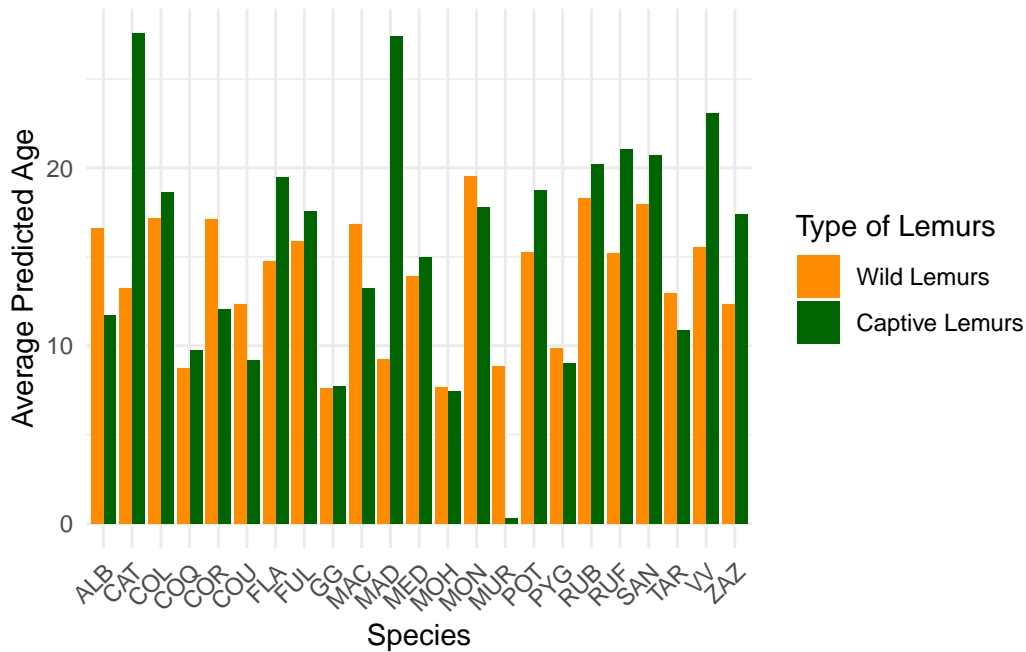
Figure 11: Comparison of average predicted ages of wild and captive lemurs across different species. Captive lemurs generally show a higher predicted age than their wild counterparts for most species.

Once again in species as seen in Figure 11, we see captive lemurs living longer than wild ones in most of them, except `ALB`, `COR`, `MAC` and `MON`. Why these species are living longer in wild, can be a good area of research for Duke Lemur Center.

5 Discussion

This paper investigates the lifespan of lemurs across wild and captive environments, utilizing data from the Duke Lemur Center (DLC). Generalized Linear Models (GLMs) with Gaussian distribution are used to analyze how factors such as sex, species, genus, and birth month impact lemur longevity. Separate models are developed for wild-born and captive-born lemurs to account for environmental differences. The study identifies patterns of increased longevity in captive lemurs and emphasizes species-specific conservation strategies.

5.1 Key Findings

One key insight is that captivity significantly increases lemur lifespan. Controlled environments provide consistent access to food, medical care, and protection from predators, reducing environmental stresses that contribute to shorter lifespans in the wild. This finding underscores the importance of captivity as a conservation tool for endangered species like lemurs.

Species-specific differences in lifespan emerge as a critical factor. Most of the species thrive in captivity due to tailored conservation practices, while some including `ALB`, `COR`, `MAC`, and `MON` perform better in the wild. Understanding the reasons behind these variations can inform targeted conservation efforts to improve outcomes.

Even though no specific trend was observed for birth month, lemurs born in August and November tend to live longer in the wild than in captivity. It is worth exploring the underlying cause for this.

Additionally, the greater influence of sex on lifespan in captivity contrasts with its less pronounced effects in the wild opens up avenues for further research and mitigating any short comings in our conservation efforts.

5.2 Drawbacks of the Study

While the study provides valuable insights, it has its limitations. First, the dataset is skewed towards captive lemurs, limiting the robustness of wild lifespan analysis.

Second, the exclusion of interactions between birth type and other predictors in the combined model reduces the depth of insights.

Third, some predictors, like birth month, have inconsistent or weak effects, indicating potential noise in the data or the need for more refined models.

Finally, the ages of lemurs below one year were excluded, which may have impacted the analysis of early-life mortality and its effects on lifespan.

5.3 Future Directions

Future research should focus on:

1. Expanding datasets to include more wild lemurs for balanced comparisons.

2. Exploring interactions between predictors (e.g., birth type and species) to uncover complex relationships.

3. Investigating why some species perform better in the wild than in captivity, which could inform more effective conservation strategies.

4. Examining seasonal effects and birth month impacts further to identify environmental triggers influencing lifespan.

5. Studying the infant mortality rate and its impact on overall lemur lifespan.

6. Using advanced modeling techniques like Bayesian models or neural networks for more nuanced predictions.

By addressing these areas, conservationists can develop targeted strategies to enhance the survival of lemurs both in the wild and in captivity. Such studies help us understand what are we doing right in conservation efforts, and where should we focus next.

A Appendix

B Survey, Sampling and Observational Data

The study relies on observational data collected by the Duke Lemur Center (DLC), which documents lemur lifespans in wild and captive settings. Unlike experimental data, observational datasets like this are not generated through controlled interventions. This limits the ability to infer causal relationships.

B.1 Data Collection Methodology

The Duke Lemur Center employed standard operating procedures to gather data, adhering to USDA, AZA, and IACUC guidelines. The original data was handwritten and later digitized into two primary databases: the Animal Record Keeping System (ARKS) and MedARKS, and subsequently the Zoological Information Management System (ZIMS). These steps ensured consistency and traceability but may have introduced biases, particularly during digitization. More on this can be found in Section 2.1.

B.2 Sampling Frame

The dataset includes both wild and captive lemurs, but the sampling frame is imbalanced, with far fewer wild lemurs. This is because it is difficult to track wild populations compared to the captive ones. This can be conutered by increasing surveillance and tracking of wild populations by installing cameras and marking them for easy identification. Alternatively, some study can be conducted to distinguish different lemurs through some of their unique biological featrues. For example, tigers are identified by their stripes and pug marks.

To address this, statistical adjustments such as weighting could be applied, though this study opted for separate models to analyze wild and captive populations independently.

B.3 Sources of Bias

1. Survivor Bias: Only lemurs with complete lifecycle information are included, potentially skewing lifespan estimates toward older individuals.

2. Selection Bias: The focus on lemurs cared for by the DLC may not represent broader populations, especially in the wild.

3. Measurement Error: Data inconsistencies could arise from manual transcription and digitization of historical records.

B.4 Literature Context

The methodologies align with best practices in ecological studies, as outlined by Zehr et al. (2014), who emphasize the importance of longitudinal tracking and database integration for species conservation. The use of generalized linear models (GLMs) aligns with approaches recommended in similar studies to account for categorical and continuous predictors.

B.5 Key Takeaways

The study's observational nature limits causal inference but provides valuable insights into factors influencing lemur longevity. Careful data cleaning and validation steps mitigate some biases, but limitations persist, particularly regarding the underrepresentation of wild populations. Future work could employ advanced sampling techniques or collaborative data pooling with other conservation organizations to improve generalizability.

C Data Cleaning

I took the following steps to clean the raw data and prepare it for analysis:

1. Loading the Data: The raw data is loaded from the file data/raw_data/animals.csv.

2. Column Selection: The dataset is reduced to relevant columns: `animal_id`, `taxonomic_code`, `sex`, `birth_date`, `death_date` and `birth_type`.

3. Filtering Rows: Rows where death_date is missing are removed to ensure the data includes only animals with complete lifecycle information.

4. Extracting Genus and Species: The taxonomic_code column is split into two new columns: `genus` (first letter) and `species` (remaining characters). The original taxonomic_code column is removed as it is no longer needed.

5. Age Calculation: A new column, `age`, is calculated based on the difference between `death_date` and `birth_date`, measured in years. Lemurs with an age of less than 1 year are removed to focus on mature animals and ignore infant mortality. This was especially common for captive-born animals because excess breeding can lead to high infant mortality rates.

6. Birth Month Extraction: A new column, `month_born`, is created to indicate the birth month extracted from the `birth_date`. We needed this information to analyze seasonal effects on lifespan.

7. Splitting Data by Birth Type: The data is split into two subsets based on birth_type: one for animals born in captivity (`data_captive`) and another for those born in the wild (`data_wild`).

8. Final Output: NA values are then dropped and the cleaned datasets are saved as CSV files: Wild-born animals: `data/analysis_data/wild.csv` Captive-born animals: `data/analysis_data/captive.csv`

This systematic cleaning ensures the dataset is consistent, reliable, and ready for further analysis.

D Analysis Dataset

Table 1: Analysis dataset for wild lemurs

animal_id	sex	birth_type	genus	species	age	month_born
0002	M	wild	O	GG	3.93	Jan
0007	M	wild	O	GG	7.55	Jan
0030	F	wild	E	COL	5.55	Oct
0046	M	wild	H	GG	1.50	Dec
0202	M	wild	O	GG	3.93	Jan
0251	F	wild	O	GG	2.91	Aug

Table 2: Analysis dataset for captive lemurs

animal_id	sex	birth_type	genus	species	age	month_born
0004	M	captive	O	GG	5.94	Apr
0005	M	captive	O	GG	15.47	Aug
0006	F	captive	O	GG	13.59	Mar
0008	F	captive	O	GG	5.38	Jul
0009	M	captive	O	GG	10.38	Sep

Table 2: Analysis dataset for captive lemurs

animal_id	sex	birth_type	genus	species	age	month_born
0010	M	captive	O	GG	13.46	May

Table 1 provides a glimpse of the dataset used for analysis for wild lemurs, and Table 2 provides a glimpse of the dataset used for analysis for captive lemurs.

E Model Summary

E.1 Wild Lemurs' Model

Table 3: Summary of the model for wild lemurs

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	13.88	7.13	1.95	0.05	-0.09	27.85
speciesCAT	15.84	8.35	1.90	0.06	-0.52	32.20
speciesCOL	6.92	4.56	1.52	0.13	-2.02	15.87
speciesCOQ	-1.95	6.80	-0.29	0.77	-15.29	11.39
speciesCOR	0.31	4.98	0.06	0.95	-9.44	10.07
speciesCOU	-2.54	6.17	-0.41	0.68	-14.64	9.56
speciesFLA	7.76	4.51	1.72	0.09	-1.08	16.60
speciesFUL	5.82	4.56	1.28	0.20	-3.12	14.77
speciesGG	-3.99	6.38	-0.63	0.53	-16.49	8.50
speciesMAC	1.53	4.74	0.32	0.75	-7.76	10.81
speciesMAD	15.68	6.73	2.33	0.02	2.48	28.88
speciesMED	3.24	7.66	0.42	0.67	-11.77	18.25
speciesMOH	-4.27	6.99	-0.61	0.54	-17.98	9.44
speciesMON	6.07	4.04	1.50	0.13	-1.85	13.99
speciesMUR	-11.42	8.16	-1.40	0.16	-27.41	4.57
speciesPOT	7.00	6.57	1.07	0.29	-5.87	19.87
speciesPYG	-2.72	6.96	-0.39	0.70	-16.35	10.91
speciesRUB	8.51	4.56	1.86	0.06	-0.44	17.45
speciesRUF	9.32	4.73	1.97	0.05	0.05	18.60
speciesSAN	8.99	4.74	1.90	0.06	-0.29	18.27
speciesTAR	-0.83	7.16	-0.12	0.91	-14.87	13.21
speciesVV	11.37	5.89	1.93	0.05	-0.16	22.91
speciesZAZ	5.67	7.51	0.76	0.45	-9.05	20.39
sexM	0.72	0.98	0.73	0.46	-1.20	2.64

Table 3: Summary of the model for wild lemurs

term	estimate	std.error	statistic	p.value	conf.low	conf.high
month_bornAug	-8.02	5.75	-1.40	0.16	-19.30	3.25
month_bornDec	5.54	5.86	0.95	0.35	-5.94	17.03
month_bornJan	-3.74	5.59	-0.67	0.50	-14.69	7.21
month_bornJul	0.18	4.70	0.04	0.97	-9.04	9.39
month_bornJun	-0.16	4.67	-0.03	0.97	-9.31	9.00
month_bornMar	-4.53	6.18	-0.73	0.46	-16.64	7.58
month_bornMay	3.30	5.72	0.58	0.56	-7.90	14.51
month_bornNov	-7.70	4.49	-1.71	0.09	-16.50	1.11
month_bornOct	0.78	5.29	0.15	0.88	-9.60	11.16
month_bornSep	-7.48	5.40	-1.38	0.17	-18.06	3.11
genusD	NA	NA	NA	NA	NA	NA
genusE	-0.06	3.97	-0.02	0.99	-7.84	7.72
genusG	NA	NA	NA	NA	NA	NA
genusH	-5.69	6.54	-0.87	0.39	-18.50	7.13
genusL	NA	NA	NA	NA	NA	NA
genusM	NA	NA	NA	NA	NA	NA
genusN	NA	NA	NA	NA	NA	NA
genusO	NA	NA	NA	NA	NA	NA
genusP	NA	NA	NA	NA	NA	NA
genusV	NA	NA	NA	NA	NA	NA

E.2 Captive Lemurs' Model

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	17.04	3.07	5.55	0.00	11.02	23.06
speciesAR	-8.42	3.78	-2.22	0.03	-15.83	-1.00
speciesCAT	-3.36	3.16	-1.06	0.29	-9.55	2.83
speciesCOL	0.56	2.39	0.23	0.82	-4.13	5.24
speciesCOQ	-7.87	3.29	-2.39	0.02	-14.33	-1.42
speciesCOR	0.48	2.40	0.20	0.84	-4.23	5.19
speciesCOU	-4.27	3.25	-1.31	0.19	-10.65	2.11
speciesFLA	-1.84	2.49	-0.74	0.46	-6.73	3.05
speciesFUL	-0.76	2.79	-0.27	0.79	-6.23	4.71
speciesGG	-8.99	3.32	-2.71	0.01	-15.49	-2.48
speciesMAC	0.24	2.25	0.11	0.91	-4.17	4.66
speciesMAD	-7.39	3.67	-2.01	0.04	-14.59	-0.19
speciesMED	-2.69	3.20	-0.84	0.40	-8.96	3.59

term	estimate	std.error	statistic	p.value	conf.low	conf.high
speciesMOH	-8.95	3.22	-2.78	0.01	-15.27	-2.63
speciesMON	2.94	2.31	1.27	0.20	-1.60	7.47
speciesMUR	-7.79	3.17	-2.46	0.01	-14.00	-1.58
speciesPOT	-1.36	4.01	-0.34	0.74	-9.23	6.51
speciesPYG	-6.73	3.50	-1.92	0.06	-13.60	0.14
speciesRUB	1.70	2.96	0.58	0.56	-4.10	7.51
speciesRUF	-1.42	2.28	-0.62	0.54	-5.89	3.06
speciesSAN	1.33	2.95	0.45	0.65	-4.46	7.11
speciesTAR	-3.64	3.32	-1.10	0.27	-10.15	2.87
speciesUL	-4.17	2.20	-1.89	0.06	-8.49	0.15
speciesVV	-1.06	3.14	-0.34	0.74	-7.21	5.09
speciesZAZ	-4.28	3.21	-1.33	0.18	-10.56	2.01
sexM	0.13	0.39	0.33	0.74	-0.64	0.90
month_bornAug	-0.56	1.03	-0.55	0.59	-2.59	1.46
month_bornDec	-2.14	1.46	-1.46	0.14	-5.01	0.73
month_bornFeb	-0.35	1.18	-0.29	0.77	-2.65	1.96
month_bornJan	-0.23	1.45	-0.16	0.87	-3.07	2.61
month_bornJul	-0.80	0.93	-0.86	0.39	-2.62	1.02
month_bornJun	-1.05	0.81	-1.29	0.20	-2.63	0.54
month_bornMar	-0.05	0.71	-0.08	0.94	-1.45	1.34
month_bornMay	0.04	0.64	0.06	0.95	-1.22	1.30
month_bornNov	0.71	1.59	0.45	0.65	-2.41	3.83
month_bornOct	-0.28	1.48	-0.19	0.85	-3.18	2.62
month_bornSep	-3.08	1.15	-2.69	0.01	-5.33	-0.83
genusD	NA	NA	NA	NA	NA	NA
genusE	-0.93	2.27	-0.41	0.68	-5.38	3.53
genusG	NA	NA	NA	NA	NA	NA
genusH	3.04	1.78	1.71	0.09	-0.45	6.53
genusL	NA	NA	NA	NA	NA	NA
genusM	NA	NA	NA	NA	NA	NA
genusN	NA	NA	NA	NA	NA	NA
genusO	NA	NA	NA	NA	NA	NA
genusP	NA	NA	NA	NA	NA	NA
genusV	NA	NA	NA	NA	NA	NA

F Posterior Predictive Checks

F.1 Residual Plots

As show in Figure 12, the residual plots for both captive and wild models show no clear patterns, indicating that the models are capturing the underlying relationships between the predictors and the response variable effectively. The residuals are randomly distributed around zero, with no discernible trends or heteroscedasticity, suggesting that the model assumptions are met.
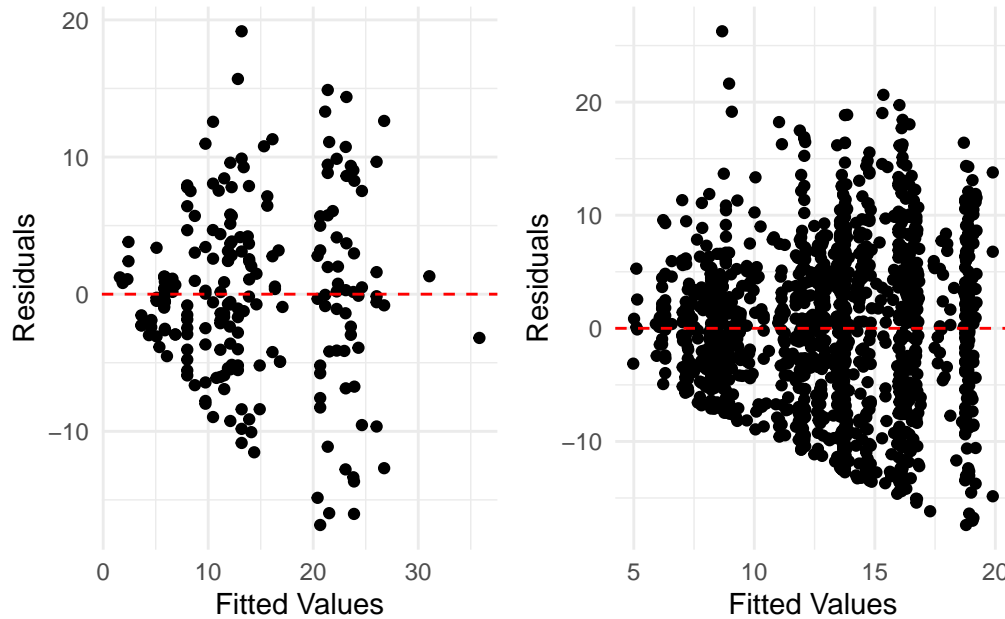

Figure 12: Residual plots for the wild (on left) and captive (on right) models. Both the plots indicate that the residuals are randomly distributed around zero, with no discernible patterns or heteroscedasticity.

F.2 QQ Plots

As show in Figure 13, the QQ plots for both captive and wild models show that the standardized residuals closely follow the expected normal distribution line. The QQ plot for wild lemurs show slight deviations at the tails, but overall, the points are aligned along the line. Overall, the points are mostly aligned along the line, indicating that the residuals are normally distributed. This suggests that the model assumptions of normality are met, and the models are appropriate for the data.
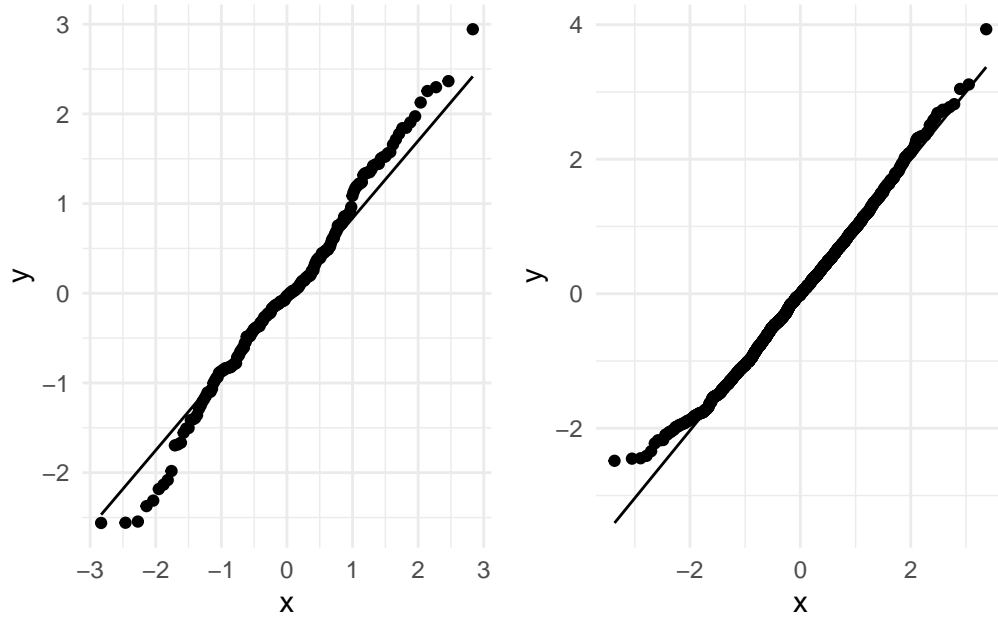
Figure 13: QQ plots of standardized residuals vs the expected normal distribution for the wild (on left) and captive (on right) models. The points closely follow the expected normal distribution line, indicating that the residuals are normally distributed.

G Species of Lemurs

The following are the species of lemurs that are present in the dataset:

Index	Code	Latin Name	Common Name
1	MED	Cheirogaleus medius	Fat-tailed dwarf lemur
2	MAD	Daubentonia madagascariensis	Aye-aye
3	ALB	Eulemur albifrons	White-fronted brown lemur
4	COL	Eulemur collaris	Collared brown lemur
5	COR	Eulemur coronatus	Crowned lemur
6	FLA	Eulemur flavifrons	Blue-eyed black lemur
7	FUL	Eulemur fulvus	Common brown lemur
8	MAC	Eulemur macaco	Black lemur
9	MON	Eulemur mongoz	Mongoose lemur
10	RUB	Eulemur rubriventer	Red-bellied lemur
11	RUF	Eulemur rufus	Red-fronted brown lemur
12	SAN	Eulemur sanfordi	Sanford's brown lemur
13	MOH	Galago moholi	Mohol bushbaby
14	GG	Hapalemur griseus griseus	Eastern lesser bamboo lemur

Index	Code	Latin Name	Common Name
15	CAT	Lemur catta	Ring-tailed lemur
16	TAR	Loris tardigradus	Slender loris
17	MUR	Mircocebus murinus	Gray mouse lemur
18	COQ	Mirza coquereli	Northern giant mouse lemur
19	COU	Nycticebus coucang	Slow loris
20	PYG	Nycticebus pygmaeus	Pygmy slow loris
21	COQ	Propithecus coquereli	Coquerel's sifaka
22	POT	Perodicticus potto	Potto
23	VAR	Varecia	Varecia hybrid
24	RUB	Varecia rubra	Red ruffed lemur
25	VV	Varecia variegata variegata	Black-and-white ruffed lemur

References

Cookson, T. Alexander. 2020. "Duke Lemur Center Dataset." https://github.com/tacookson/data/tree/master/duke-lemur-center.

Duke Lemur Center. 2024. "Duke Lemur Center." https://lemur.duke.edu.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2024. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Lemaître, Jean-François, Victor Ronget, Morgane Tidière, Dominique Allainé, Vérane Berger, Aurélie Cohas, Fernando Colchero, et al. 2020. "Sex Differences in Adult Lifespan and Aging Rates of Mortality Across Wild Mammals." *Proceedings of the Academy of Natural Sciences.* https://doi.org/10.1073/pnas.1911999117.

Pedersen, Thomas Lin. 2024. *Patchwork: The Composer of Plots.* https://CRAN.R-project.org/package=patchwork.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'.* https://CRAN.R-project.org/package=arrow.

Wickham, Hadley. 2011. "Testthat: Get Started with Testing." *The R Journal* 3: 5–10. https://journal.r-project.org/archive/2011-1/RJournal_2011-1_Wickham.pdf.

———. 2016. *Ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC. http://www.crcpress.com/product/isbn/9781466561595.

Zehr, Steven M., Ronald G. Roach, Danielle Haring, Julia Taylor, Fred H. Cameron, and Anne D. Yoder. 2014. "Life History Profiles for 27 Strepsirrhine Primate Taxa Generated Using Captive Data from the Duke Lemur Center." *Scientific Data* 1 (July): 140019. https://doi.org/10.1038/sdata.2014.19.

Zhu, Hao. 2024. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax.* https://CRAN.R-project.org/package=kableExtra.