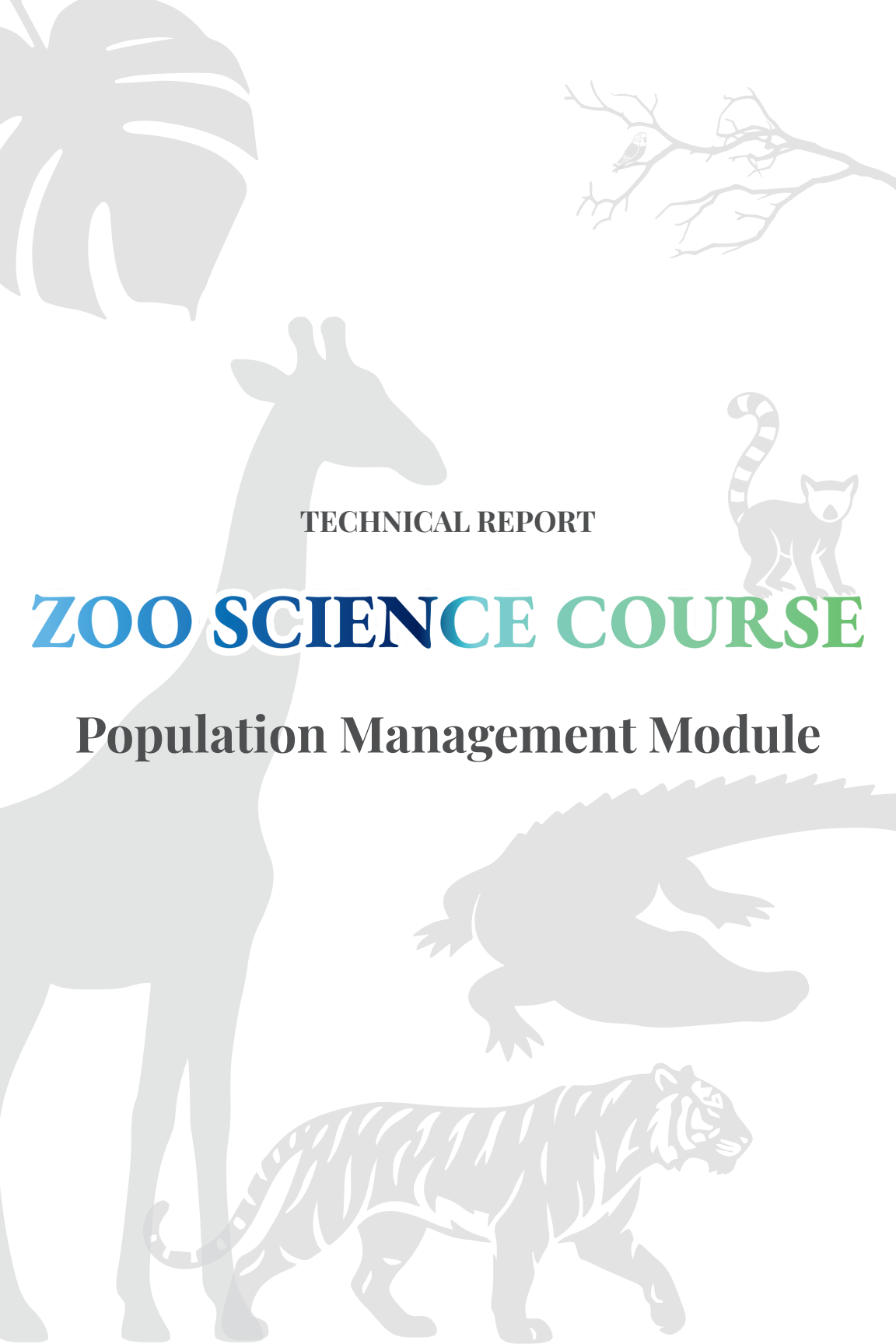




TECHNICAL REPORT

ZOO SCIENCE COURSE

Population Management Module

Zoo Science Course

Population Management Module



ಬನ್ನೇರುಘಟ್ಟ
ಜೈವಿಕ ಉದ್ಯಾನವನ
ಬೆಂಗಳೂರು
**BANNERUGHATTA
BIOLOGICAL PARK**
BENGALURU

Technical report: Zoo Science Course (Population Management Module)



Published by Bannerughatta Biological Park, Bengaluru, Karnataka

Acknowledgements

We express our sincere gratitude to Shri A V Surya Sen, IFS, Executive Director, Bannerughatta Biological Park, whose vision, encouragement, and wholehearted support made this course possible. His active involvement and guidance were instrumental in ensuring the success of this initiative.

We also thank the Zoo Authority of Karnataka and the Karnataka Forest Department for their support, and acknowledge the efforts of Aishwarya K. S. for coordinating logistics and ensuring smooth conduct of the course. Finally, we extend our appreciation to all participants for their enthusiastic engagement and to the speakers for generously sharing their expertise.

Typeset in L^AT_EX

Citation: Lakshminarasimha, R., & Adhikari, P. (2025). *Technical report: Zoo Science Course (Population Management Module)*. Bannerughatta Biological Park, Karnataka.

Foreword

The role of modern zoos is rapidly transforming. No longer seen merely as spaces of recreation, they are now recognized as institutions of conservation, education, and science. At the heart of this transformation lies the ability to manage animal populations with scientific rigour and collaboration across institutions.

It was with this vision that the Population Management Module of the Zoo Science Course was conceived. This initiative represents an important step toward strengthening the technical capacity of Indian zoos, ensuring that conservation breeding and ex situ management are guided by evidence and best practice.

Held from 16th to 20th June 2025 at Bannerghatta Biological Park, Bengaluru, the course brought together zoo professionals from across the country. Through lectures, hands-on exercises, and data-driven case studies, participants engaged deeply with the principles of population management. The energy and enthusiasm displayed throughout the training reaffirmed the urgent need for structured professional development in this field.

This report captures the objectives, structure, and outcomes of the module, while also reflecting the commitment of Indian zoo personnel to move towards evidence-based management. It is my sincere belief that such initiatives will lay the foundation for a stronger, more collaborative zoo community in India — one capable of safeguarding the country's faunal diversity for future generations.

On behalf of Bannerghatta Biological Park, I take this opportunity to thank the Zoo Authority of Karnataka, the Karnataka Forest Department, the course faculty, and all participants for making this initiative possible. I look forward to continued collaboration in building the skills, knowledge, and confidence of zoo professionals across India.

Shri A V Surya Sen, IFS
Executive Director, Bannerghatta Biological Park

Contents

Foreword	4
Executive summary	1
1 Introduction	2
1.1 Zoo Science Course	2
1.2 Background and rationale	2
1.3 Population Management in Zoos	3
1.4 Course overview	3
2 Module framework	5
2.1 Goal	5
2.2 Objectives	5
2.3 Pedagogy and teaching approach	5
2.4 Speakers and participants	6
3 Module schedule	7
4 Session details	9
4.1 Orientation	9
4.1.1 Aim	9
4.1.2 Session details	9
4.2 Fundamentals of population management	11
4.2.1 Aim	11
4.2.2 Session details	11
4.3 Collection planning	17
4.3.1 Aim	17
4.3.2 Session details	17
4.4 Conservation breeding	26
4.4.1 Aim	26
4.4.2 Session details	26
4.5 Data-driven decision making	29
4.5.1 Aim	29
4.5.2 Session details	29
5 Way forward	31
5.1 Pre- and post-course surveys	31
5.2 Future steps	32
5.2.1 Immediate follow-up	32

5.2.2	Short-term (next 6–12 months)	32
5.2.3	Medium-term (1–2 years)	33
5.2.4	Long-term vision	33
5.3	Next steps - Scaling modules nationally	33
Appendices		34
A	People	35
B	Getting Started with R	36
C	Introduction to statistics using R	51
D	Key Variables in a Zoo Collection Plan	64
E	Pedigree construction and analysis	66

Executive summary

The Zoo Science Course (Population Management Module) was developed to strengthen the technical capacity of Indian zoos in managing animal populations through scientific planning, data-driven decision-making, and inter-institutional cooperation. Conducted at Bannerghatta Biological Park, Bengaluru (16–20 June 2025), the module provided zoo biologists with structured training in the principles and practice of ex situ population management.

Key highlights of the module include:

- Captive populations as representative conspecifics: Emphasis was placed on managing zoo populations not simply as animals under care, but as model populations of their wild counterparts. This requires maintaining demographic balance, genetic diversity, and evolutionary potential, while minimising drift, inbreeding, and adaptation to captivity.
- Professional capacity-building: The course introduced participants to demographic and genetic principles, data workflows in R, and cooperative planning methods essential for sustaining zoo populations.
- Applied pedagogy: Through a blend of lectures, hands-on exercises, and case-based simulations, participants engaged directly with real-world datasets and decision-making scenarios.
- Collaborative approach: The module fostered peer learning and cross-institutional dialogue, aligning with Central Zoo Authority (CZA) guidelines and global best practices.
- Long-term vision: The training represents the first step in establishing a recurring framework of professional development under the Zoo Science Course, covering broader themes such as behaviour and welfare, enclosure design, nutrition, and master planning.

By embedding evidence-based practices into daily zoo operations, the module aims to ensure that Indian zoos manage their collections as demographically viable, genetically representative, and evolutionarily resilient populations—contributing meaningfully to both national conservation goals and global ex situ management standards.

1 Introduction

Modern zoos are increasingly recognized not only as spaces for public engagement, but as institutions of conservation, research, and education. Within this evolving role, the scientific management of ex situ populations has become a cornerstone of effective conservation practice.

1.1 Zoo Science Course

The *Zoo Science Course* was conceived as India's first structured professional development programme tailored to the needs of zoo personnel. Developed with the support of the Zoo Authority of Karnataka and hosted by Bannerughatta Biological Park, the course is designed as a modular, recurring framework to build technical and strategic capacity in zoo management. Its long-term vision is to address critical thematic areas—including population management, behaviour and welfare, enclosure design, master planning, and nutrition—through a blend of lectures, applied exercises, and peer learning. The pilot module on *Population Management* represents the first step in this larger initiative, aimed at equipping zoo biologists, curators, and educators with the analytical tools and evidence-based approaches needed to strengthen the scientific foundation of Indian zoos.

This report documents the implementation of the pilot Population Management Module held from 16–20 June 2025 at Bannerughatta Biological Park, Bengaluru, and outlines its rationale, structure, content, and outcomes.

1.2 Background and rationale

India hosts nearly 150 recognized zoos, which increasingly contribute to conservation, education, and species management. With the growing emphasis on conservation breeding, genetic diversity, and the long-term sustainability of ex situ populations, there is a clear need for more structured, data-driven, and cooperative approaches.

Historically, zoo population management in India has been localized and intuitive, often relying on informal exchanges. Challenges such as limited founder numbers, skewed sex ratios, space constraints, and the absence of coordinated planning highlight the need for science-based practices. The Central Zoo Authority (CZA) has advocated the integration of demographic and genetic principles into zoo operations. However, effective implementation depends on building the capacity of zoo personnel to understand and apply these principles in practice.

The *Population Management Module* was therefore developed as a focused training to strengthen technical skills among zoo biologists, curators, veterinarians, and

other key staff. The curriculum introduced concepts of population viability, demographic planning, cooperative breeding, and data analysis using open-source tools. By equipping participants with these skills, the module supports evidence-based decision-making, more effective conservation breeding, and a culture of collaboration across institutions.

1.3 Population Management in Zoos

Population management involves maintaining genetically and demographically viable populations over the long term, minimizing inbreeding, and facilitating cooperative planning across zoos. Importantly, ex situ populations are also recognised as evolutionary units: they carry genetic variation and evolutionary potential, and are subject to forces such as drift, selection, and adaptation to captivity. Without structured management, populations may drift away from their wild counterparts or lose genetic diversity, compromising their future utility for conservation. Sound management seeks to minimize such risks while ensuring adaptability under changing conditions.

To achieve this, modern zoos employ systematic approaches, including:

- i. Studbook-based coordination to guide breeding and transfers
- ii. Demographic planning for age structure, sex ratios, and carrying capacity
- iii. Cooperative management across institutions to treat collections as regional or national meta-populations
- iv. Modelling and decision-support tools to anticipate long-term viability

In India, where many zoos still operate in isolation, capacity-building in population management is critical. It enables zoos to contribute meaningfully to conservation breeding goals, avoid demographic and genetic imbalances, and align with national species management strategies. This training module provided participants with a practical introduction to these concepts, using real datasets, case studies, and open-source tools.

1.4 Course overview

The Population Management Module was conducted from 16–20 June 2025 at Bannerghatta Biological Park, Bengaluru, with sessions hosted at the Jungle Lodges and Resorts campus. The course targeted zoo biologists and technical personnel responsible for studbook keeping, animal exchange coordination, and conservation breeding, with emphasis on participants from southern India alongside select representatives from other regions.

A total of 19 participants from 8 states, representing 17 zoos attended the five-day course (Appendix A). The programme combined lectures with hands-on exercises,

group work, and data-driven case studies using R and PMx. Emphasis was placed on real-world application, contextual relevance, and collaborative problem-solving. The module established a common analytical baseline and provided tools for participants to contribute effectively to cooperative population management and science-based decision-making in their institutions.



Figure 1: Sh. A V Surya Sen, IFS, Executive Director, Bannerughatta Biological Park, addressing the participants during the inaugural session of the Zoo Science Course (Population Management Module), 16 June 2025.

2 Module framework

2.1 Goal

To strengthen the capacity of Indian zoos to manage animal populations effectively through scientific planning, data-driven decision-making, and inter-institutional cooperation, thereby supporting the broader goals of conservation breeding and animal welfare.

2.2 Objectives

The Population Management Module was designed to equip zoo personnel with foundational knowledge and practical skills in the science and application of ex situ population management. Specifically, the module aimed to:

- i. Build conceptual understanding of demographic and genetic principles in population management.
- ii. Introduce tools and techniques for analyzing zoo inventory and studbook data using open-source platforms (e.g., R, PMx).
- iii. Facilitate hands-on learning through case-based exercises, simulations, and decision-making scenarios.
- iv. Encourage inter-zoo collaboration by highlighting the importance of coordinated planning.
- v. Promote long-term, sustainable breeding strategies aligned with national priorities and global best practices.

2.3 Pedagogy and teaching approach

The module adopted a learner-centred pedagogy, delivered through a blend of lectures, guided tutorials, group work, and case discussions. Emphasis was placed on:

- i. Active learning, with participants directly engaged in analysis and decision-making.
- ii. Use of R and Quarto, promoting reproducible and transparent workflows.
- iii. Scenario-based simulations, exploring the outcomes of different breeding or management strategies.
- iv. Collaborative exercises, encouraging cross-institutional thinking and peer learning.
- v. Open discussions, enabling reflection on institutional challenges and opportunities.

This design ensured that participants not only understood theoretical concepts but could also apply them in practical zoo management contexts.

2.4 Speakers and participants

The course brought together a diverse group of participants and resource persons. Detailed profiles are provided in Appendix A.



3 Module schedule

Lectures (L) and hands-on sessions (H) were interspersed throughout the five-day programme to balance conceptual understanding with practical application. The detailed schedule is provided below.

Time	Session
June 16, 2025: Orientation	
09:30–10:30	Inaugural
10:30–11:15	Course introduction (L)
11:15–11:30	Break
11:30–11:45	Getting familiar with the course materials (L)
11:45–12:45	Installation & Set-up (H)
12:45–13:30	Getting Started with R (H)
13:30–14:30	Lunch
14:30–16:00	Getting Started with R (continued) (H)
16:00–16:15	Q&A and Wrap-up
16:15 onwards	Zoo visit
June 17, 2025: Fundamentals of population management	
10:00–11:00	From Intuition to Insight – Role of data in zoo management (L)
11:00–12:15	Introduction to Population Biology (L)
12:15–12:30	Break
12:30–14:30	Zoological Information Management System in zoo management (L & H) [online]
14:30–15:30	Lunch
15:30–16:30	Introduction to statistics using R (H)
16:30–16:45	Break
16:45–17:45	Introduction to statistics using R (H) (continued)
June 18, 2025: Collection planning	
09:00–10:00	Introduction to collection planning (L)
10:00–10:45	Zoo Species Planning (L)
10:45–11:30	Ethics and sustainability in collection planning (L)
11:30–11:45	Break
11:45–12:45	Risk management in collection planning (L)
12:45–13:45	Lunch

Continued on next page

Continued from previous page

Time	Session
13:45–14:30	Integrated Collection Assessment and Planning process (L & H)
14:30–16:30	Developing a collection plan – Group activity (H)
16:30–16:45	Break
16:45–17:45	Developing a collection plan – Group activity (H) (continued)
June 19, 2025: Conservation breeding	
09:00–10:30	Species conservation planning (L)
10:30–11:30	Demographic and Genetic Management of Captive Populations (L)
11:30–11:45	Break
11:45–13:30	Demographic and Genetic Management of Captive Populations (continued) (L)
13:30–14:30	Lunch
14:30–15:30	Drafting a conservation breeding plan (H)
15:30–15:45	Break
15:45–18:00	Demographic and Genetic Analysis of Captive Populations (H)
June 20, 2025: Data-driven decision making	
09:00–10:30	Population modelling (L & H)
10:30–10:45	Break
10:45–11:30	Prioritization frameworks for managing species in captivity (L)
11:30–12:30	Action planning exercise (H)
12:30–13:30	Participant case presentations
13:30–14:30	Lunch
14:30–15:45	Assessment, Learning network, take-home exercises and wrap-up
15:45–16:30	Valedictory

4 Session details



4.1 Orientation

4.1.1 Aim

Day 1 introduced the rationale, structure, and vision of the Zoo Science Course, situating the pilot module on population management within a broader programme of professional development for Indian zoo personnel. Sessions outlined the legal and institutional context of zoos in India, categories under the Recognition of Zoo Rules, 2009, and the need for structured training tailored to national realities. In parallel, participants were introduced to the R environment and basic data management, learning to load datasets, identify data types, perform simple manipulations, and generate initial summaries. This established a common analytical baseline for subsequent modules. The day also set expectations for the course, emphasising evidence-based management, peer learning, and alignment with Central Zoo Authority (CZA) guidelines and global best practices.

4.1.2 Session details

Course introduction

- *Lakshminarasimha R*

The first session set the context for India's first formal professional development initiative focused on strengthening zoo science. Participants were introduced to the overall vision of the Zoo Science Course and the rationale for initiating

a structured, recurring training programme tailored to Indian zoos. The session emphasised the dual goals of building technical capacity and promoting evidence-based management practices.

Legal definitions under the Wild Life (Protection) Act, 1972, and the Recognition of Zoo Rules, 2009, were reviewed to clarify the scope of zoos, rescue centres, and conservation breeding centres. The classification of zoos by size and staffing requirements was outlined, highlighting the diversity of institutions in India and the variation in resources and technical expertise.

Contextual data on the number of zoos, species holdings, and visitor numbers underscored the importance of strengthening technical and strategic capacity. The lack of institutionalised training programmes tailored to national realities was identified as a major gap.

The vision for the Zoo Science Course was presented as a response to this need: a comprehensive and standardised training framework covering modules on population management, behaviour, welfare, exhibit design, nutrition, and master planning. The pilot module on population management was positioned as the starting point for a phased expansion.

Learning outcomes were outlined across knowledge, skills, decision-making, communication, and long-term impact, with emphasis on collaboration, inclusivity, and alignment with Central Zoo Authority (CZA) guidelines and global best practices. Participants were encouraged to engage actively, share perspectives, and contribute to building a community of practice that drives positive change in Indian zoos.

The introductory slide deck for the course can be accessed in the online version of this report.

Getting started with R

- *Lakshminarasimha R; Priyanka Adhikari*

This session introduced participants to the R programming environment with a focus on RStudio as the primary interface for data analysis. The objective was to build familiarity with the workflow of writing, saving, and executing R scripts, and to orient participants to core concepts required for handling zoo data in subsequent sessions.

Participants learned about working directories, creating and saving scripts, and the role of packages in extending R's functionality. A bulk installation script for commonly used packages was provided to ensure readiness for later modules. The session covered key data types in R (numeric, character, and logical) and

introduced the use of variables, assignment operators, and basic arithmetic and relational expressions.

Through guided exercises, participants practised creating vectors, managing variables within the R environment, and applying basic mathematical and statistical functions. The session then progressed to dataframes, including how to construct them, inspect their structure, and select rows, columns, and elements. Practical exercises reinforced good coding practices and encouraged participants to work interactively with their own data.

The latter part of the session introduced the *tidyverse* package, with demonstrations of filtering, mutating, arranging, grouping, and summarising operations. Participants also explored data reshaping using *pivot_wider* and *pivot_longer*, as well as introductory data visualisation with *ggplot2*, including histograms, box-plots, scatterplots, and bar charts. Hands-on exercises enabled participants to practise problem-solving in real time. These activities reinforced the value of reproducible data pipelines and prepared participants to engage with more complex zoo inventory datasets in later sessions.

The R code used during the session is provided in Table A and can also be accessed in the online version of this report.

Further reading

- Wickham, H., & Grolemund, G. (2023). R for Data Science (2nd ed.). O'Reilly Media.
- YCRC. (n.d.). R for Novices. Yale R-novice GAPMINDER. Retrieved from <https://docs.ycrc.yale.edu/r-novice-gapminder/reference.html>

4.2 Fundamentals of population management

4.2.1 Aim

The aim of the sessions conducted on Day 2 was to introduce participants to the fundamental principles of population management in the context of zoological institutions. The sessions aimed to build conceptual understanding of population biology, highlight the importance of high-quality zoo data, and equip participants with introductory-level skills in statistical analysis using R, thereby laying the foundation for evidence-based planning and decision-making.

4.2.2 Session details

From Intuition to Insight – Role of data in zoo management

- *Lakshminarasimha R*

This session introduced participants to the evolving role of data in modern zoo management. It highlighted the shift from experience-based, intuitive decision-making toward structured, evidence-based approaches informed by systematic data collection and analysis.

The presentation began by exploring the limitations of traditional animal care, particularly the challenges of subjectivity and the inability to detect subtle or long-term trends through intuition alone. Participants were then introduced to various categories of data routinely generated in zoos — including life histories, behavioural observations, health records, environmental parameters, and operational data.

Using real-world case studies and comparative visuals, the session demonstrated how data analytics can inform critical decisions such as pairing for breeding, managing surplus populations, and evaluating health and welfare. Tools such as ZIMS, SPARKS, PMx, Vortex, ZooMonitor, and R were introduced as integral components of a zoo’s analytical toolkit. The session also addressed challenges related to data quality, standardization, privacy, and institutional silos.

The session also explained how data flows from individual animals—such as health parameters, behavioural observations, and life history traits—into institutional records, which are then integrated into national and global zoo databases. Through such structured data pipelines, information from multiple sources can be consolidated to support evidence-based decision-making, collaborative breeding programs, and broader conservation planning Figure 2.



Figure 2: Flow of data in zoo population management — Information begins at the level of the individual animal (health, behaviour, physical traits), is recorded within institutional systems, aggregated across multiple zoos at national and global levels, and ultimately contributes to coordinated decision-making in welfare, breeding, and conservation planning.

Emphasis was placed on building a data-centric culture in zoos — one that supports proactive, predictive, and individualized care strategies. The session concluded by advocating for collaborative data networks and emerging technologies

as key enablers for future-ready zoo management.

Persistent challenges such as inconsistent record-keeping, data fragmentation, inadequate staff training, and the lack of standardized protocols were also discussed. Addressing these gaps will require investments in staff capacity building, cross-zoo data integration, and the use of digital tools that support real-time data entry, validation, and sharing. The pathway forward lies in developing institutional commitment to data quality, ethical handling of sensitive information, and the strategic use of analytics for population management, health monitoring, and long-term conservation planning.

The studies referenced during the session — including those related to behavioral monitoring, welfare assessment, and visitor impacts — are listed below for further reading and context.

Further reading

- Galante, J., & Margulis, S. W. (2022). Comparing data collection tools for zoo management decision-making: A case study examining behavioral measures of humboldt penguin bond strength. *Animals*, 12(21), 3031. Marcellini, D. L., & Jenssen, T. A. (1988). Visitor behavior in the national zoo's reptile house. *Zoo Biology*, 7(4), 329–338.
- Montanha, J. C., Silva, S. L., & Boere, V. (2009). Comparison of salivary cortisol concentrations in jaguars kept in captivity with differences in exposure to the public. *Ciência Rural*, 39, 1745–1751.
- Schwitzer, C., & Kaumanns, W. (2001). Body weights of ruffed lemurs (*Varecia variegata*) in european zoos with reference to the problem of obesity. *Zoo Biology: Published in Affiliation with the American Zoo and Aquarium Association*, 20(4), 261–269.
- Wark, J. D., Cronin, K. A., Niemann, T., Shender, M. A., Horrigan, A., Kao, A., & Ross, M. R. (2019). Monitoring the behavior and habitat use of animals to enhance welfare using the ZooMonitor app. *Anim. Behav. Cogn*, 6(3), 158–167.
- Yin, Z., Zhao, Y., Xu, Z., & Yu, Q. (2024). Automatic detection of stereotypical behaviors of captive wild animals based on surveillance videos of zoos and animal reserves. *Ecological Informatics*, 79, 102450.

Introduction to Population Biology

- *Dr. Abhishek Harihar*

This session provided a structured overview of key theoretical concepts in population biology to prepare participants for the applied aspects of zoo population

management. The central emphasis was on understanding how wild populations function—demographically and genetically—and using that knowledge to inform the management of their counterparts in captivity. Participants were encouraged to recognize that effective *ex situ* conservation must be grounded in an appreciation of species-specific biological traits as expressed in natural populations, including life history strategies, social structure, and population regulation mechanisms.

Participants were introduced to basic population models, including exponential and logistic growth, and the role of intrinsic growth rate (r) in determining how populations expand or stabilize over time. The concept of carrying capacity was discussed in the context of both ecological limits and captive space constraints. Life tables were introduced as essential tools for understanding population dynamics, allowing participants to summarize age-specific survival (lx), mortality (qx), and reproductive patterns (mx). Their relevance for identifying demographic bottlenecks and planning long-term breeding strategies was highlighted, particularly in relation to small and structured captive populations.

The session emphasized generation length as a crucial parameter for forecasting breeding cycles and demographic change. The importance of maintaining genetic diversity was addressed through the concept of effective population size (N_e), with discussion on how inbreeding, founder representation, and age distribution can influence genetic health. Participants also gained exposure to population genetics, pedigree development, and basic population projections—all of which are relevant for managing closed or semi-closed zoo populations.

By reinforcing the idea that captive populations should reflect the demographic and genetic realities of their wild counterparts, the session provided a critical scientific context for zoo-based population planning. It laid the theoretical groundwork for subsequent sessions focused on demographic analysis, studbook interpretation, and cooperative breeding program design.

Data management in zoos using ZIMS

- *Sarah Lavin*

This session introduced participants to the Zoological Information Management System (ZIMS), a widely used software platform for animal record-keeping in zoos and aquariums. It highlighted ZIMS as an important tool for evidence-based animal care, population management, and institutional planning.

The session began with an overview of the global Species360 network and how ZIMS enables collaborative data sharing across more than 1,200 institutions. Participants were introduced to the major modules of ZIMS, including those for

animal husbandry, medical records, enclosure tracking, studbook management, and taxonomy. Through live demonstrations and examples, the session illustrated how high-quality, real-time records support informed decisions on welfare, transfers, breeding, nutrition, and veterinary care.

Emphasis was placed on the role of standardized data in ensuring comparability across institutions, supporting regional cooperation, and improving the long-term sustainability of captive populations. Participants learned how ZIMS facilitates tracking of individual animal life histories — including origin, movement, health status, and reproductive events — and how such records are essential for understanding demographic trends and managing genetic diversity.

To strengthen applied learning, each participant was given access to a mock environment *LearnZIMS*, designed to simulate real-world zoo scenarios including a studbook dataset. Under expert guidance, they were taken through the process of creating individual animal records, logging husbandry events, and managing enclosure information. This hands-on engagement helped participants understand the workflow of ZIMS interfaces. To enable continued learning, participants were also granted four weeks of post-course access to the training environment. This encouraged them to independently explore additional modules, reinforce concepts, and practice their data entry and management skills beyond the classroom.

The session also emphasized the importance of consistent data entry, role-based user access, and internal data audits to maintain record integrity. Practical applications discussed included evaluating reproductive success, identifying health issues across taxa, and assessing housing suitability and enrichment outcomes.

By the end of the session, participants had an overview of how applications such as ZIMS serves as the backbone of modern zoo management — transforming animal records into actionable insights, and enabling zoos to contribute meaningfully to global conservation knowledge networks.

The session reinforced that systematic digital record-keeping through platforms like ZIMS is foundational to modern zoo management. By standardizing data across institutions, zoos can make informed decisions, contribute to global knowledge systems, and enhance the quality of care for individual animals and populations.

Introduction to statistics using R

- *Dr. Ashwin Vishwanathan*

Building on the earlier session where participants were introduced to the basics of R programming (see Table A), this session focused on applying statistical tech-

niques to zoo data to derive meaningful biological insights. Using inventory-style datasets—typically the most widely available form of structured data in Indian zoos—participants were guided through a series of hands-on analyses to explore survival trends and identify patterns of population dynamics. This ensured that foundational understanding and analytical skills are developed in tandem.

The session introduced data types (numeric, categorical, text, and factors) and the structure of data frames in R. Participants learned how to calculate key summary metrics such as survival proportions using births, acquisitions, deaths, and collection sizes. Using the *tidyverse* package, they created derived variables, filtered meaningful subsets, and summarised data across species and zoos.

Various types of data distributions—Gaussian (normal), Poisson, and Binomial—were discussed using real and simulated datasets. Participants were introduced to the concept of distribution-specific analysis and the importance of checking assumptions before choosing statistical procedures.

Specific tasks demonstrated included constructing derived variables, creating species- and zoo-level summaries, and applying basic data-quality filters. Participants contrasted Gaussian 95% CIs ($\text{mean} \pm 1.96 \cdot \text{SE}$) with a bootstrap CI (1,000 resamples via a custom *bootst()* function), computed binomial CIs for survival proportions, and visualised results with histograms (Gaussian/Poisson examples), error-bar plots for proportions, and a correlation plot with a *lm()* fit. Factors and their role in modelling were introduced alongside *tidyverse* data pipelines for reproducible analysis.

Key measures of central tendency, including mean, standard deviation, standard error, and confidence intervals, were explained, with both analytical and bootstrapped methods demonstrated. Practical coding examples illustrated how to compute and visualise these metrics using R, especially in the context of zoo datasets. A focus was placed on understanding how these measures can differ depending on the underlying data distribution.

More advanced techniques such as calculating binomial confidence intervals for proportions (e.g., survival rates by species or zoo) and basic linear modelling were also introduced. Using the *lm()* function, participants learned to explore relationships between variables, interpret model summaries, and understand the role of correlation in biological inference.

The session emphasised that even basic zoo records—when systematically analysed—can yield valuable insights about species management, housing adequacy, and mortality factors. Participants were encouraged to think critically about the structure of their data and its implications for statistical analysis. The session was structured as a live demonstration, with participants actively working through

the code examples. The R code used during the session is provided in the appendix for reference and reuse.

The R code used during the session is provided in Appendix B and can also be accessed in the online version of this report.

4.3 Collection planning

4.3.1 Aim

The aim of Day 3 was to strengthen participants' understanding of collection planning as a structured and evidence-based process in zoo management. Sessions introduced the strategic role of collection planning in balancing conservation, welfare, education, and legal considerations, and highlighted international frameworks such as ICAP. The day concluded with a group exercise where participants applied these concepts by developing a mock collection plan, reinforcing the importance of data, risk awareness, and regional cooperation in shaping sustainable zoo collections.

4.3.2 Session details

Introduction to collection planning

- *Lakshminarasimha R*

The opening session provided participants with a comprehensive orientation to collection planning as a core element of modern zoo management. The lecture began by situating collection planning within the broader framework of Indian zoos, where historical practices often centred on displaying charismatic or exotic animals without long-term strategy or consideration of wider conservation impact. Participants were guided to reflect on how this “wishlist” approach, driven by availability and individual preference, has gradually shifted towards structured, evidence-based planning that aligns with institutional missions and global conservation goals Figure 3.

Key concepts were introduced, including the role of institutional mandates, conservation priorities, space and infrastructure limitations, legal requirements under the Wild Life (Protection) Act and Recognition of Zoo Rules, and the ethical responsibilities of zoos as custodians of wild animals. The session highlighted how each of these elements shapes decisions on species acquisition, maintenance, and phase-out. International examples and historical data on trends in species



Figure 3: Collection planning operates at three interconnected levels: institutional, regional, and global. Each level informs the other, creating a feedback loop that links local decisions with wider conservation objectives.

holdings were used to illustrate how planning frameworks—when applied systematically—can reduce redundancy, avoid unsustainable practices, and increase the conservation relevance of living collections Figure 4.

The discussion also underscored the importance of viewing collection planning as a dynamic, living process. Participants were encouraged to consider questions such as: What justifies maintaining a species in captivity? How can institutional space be used most effectively? How should individual institutional priorities connect with regional and global cooperative management frameworks? The session stressed that well-designed collection planning not only reduces dependence on opportunistic wild capture but also maximises conservation, education, and research outcomes.

By the end of the session, participants had a clearer understanding that collection planning is not merely about deciding which species a zoo should house, but about aligning species, space, and institutional vision in a way that ensures long-term sustainability, welfare, and relevance to broader conservation priorities Figure 5.

Zoo Species Planning

- *Amanda Embury*

The second session provided participants with a detailed understanding of how zoos make structured decisions about which species to maintain, prioritise, or

	HIGH	LOW
ROLE VALUE	The species fulfils several clearly defined roles (conservation priority, education, research, historical importance, completes an ecodisplay, etc.).	The species contributes little or nothing to those roles.
HUSBANDRY FEASIBILITY	Existing housing, diet, keeper skill level and veterinary capacity are all adequate and sustainable.	The species is difficult or costly to house, feed or care for at the required standard.

Figure 4: Species evaluation can be guided by two dimensions — their role value (conservation, education, research, display) and husbandry feasibility (capacity to provide sustainable care). This simple matrix helps determine whether a species should be prioritised, maintained, or phased out.

phase out, using Zoos Victoria as a case study. It outlined the institutional processes that connect species planning to wider regional and global objectives, showing how individual zoo decisions are not isolated but part of a coordinated international network.

The lecture emphasised that species planning is not about keeping as many animals as possible, but about ensuring that those chosen are viable, sustainable, and aligned with both conservation and institutional priorities. Participants were introduced to the criteria used by Zoos Victoria, including conservation status, demographic and genetic viability, welfare and husbandry feasibility, visitor engagement value, and institutional capacity. Real-world examples illustrated how these criteria are systematically applied to develop transparent, justifiable species lists.

Attention was drawn to the role of data and evidence-based decision making, including the use of demographic modelling, studbooks, and regional/global population data to forecast sustainability over time. Participants also learned how Zoos Victoria uses structured tools to balance competing objectives — such as conserving space, ensuring welfare standards, maintaining visitor appeal, and supporting broader conservation programs.

The session highlighted that species planning is not a one-off exercise but a dynamic and iterative process. Plans are periodically reviewed, incorporating new data, changing conservation priorities, and updated assessments of institutional resources. This ongoing review process ensures that collection decisions remain relevant and responsible.

**Diverse exhibitry
&
Novelty**



Sustainability

Figure 5: Collection planning requires balancing novelty and diversity of exhibit species with long-term sustainability. Strategic planning helps zoos avoid unsustainable acquisitions while maintaining engaging and relevant collections.

By the end of the session, participants recognised species planning as a disciplined practice that turns complex, and often competing, priorities into clear strategies. They gained insight into how adopting such transparent, criteria-driven approaches can improve institutional credibility, support regional cooperation, and ensure that collections contribute meaningfully to long-term wildlife conservation goals.

Ethics and sustainability in collection planning

- *Lakshminarasimha R*

This session examined the ethical and sustainability dimensions of collection planning, situating them at the heart of how modern zoos justify which species they hold and how they manage them. The lecture began by presenting international data that revealed how many populations in zoos fail to meet minimum demographic or genetic thresholds for long-term viability. Graphs and charts were used to show that, while some species breed to replacement and retain sufficient gene diversity, a large proportion fall short of sustainability targets. The Indian context was then introduced through longitudinal analyses of felid populations, where figures illustrated the persistence of many species as small, scattered groups in a few zoos, contrasted with larger but often skewed populations of big cats and colour morphs. These visuals helped participants recognise how historic decisions have created structural imbalances in Indian collections, with consequences for welfare, management, and conservation impact.

The concept of meta-populations was introduced as an ethical and practical response Figure 6. Participants saw how subdividing populations across zoos, when managed interactively, preserves more genetic variation and reduces the risk of catastrophic loss compared to housing all individuals in a single, isolated group. The principles of effective sub-population management, including determining viable numbers, distribution, and connectivity were demonstrated.

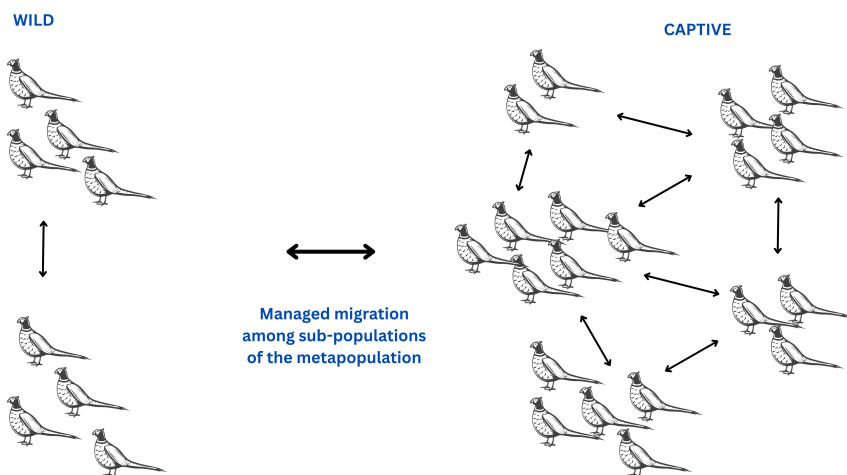


Figure 6: Wild and captive sub-populations are shown as spatially separated groups of individuals. Managed migration (arrows) allows periodic movement of animals between groups, maintaining genetic diversity and reducing extinction risk. This strategy contrasts with maintaining a single isolated population and highlights how interconnected sub-populations can be more resilient over the long term.

Concrete case studies drawn from zoo inventory data reinforced these ideas. Graphs of elephants, leopards, jungle cats, marmosets, rhinos, gaurs, and leopard cats showed how species varied in their distribution across zoos, their breeding performance, and their demographic trajectories. Side-by-side charts comparing births and deaths provided a clear picture of where populations were sustainable and where they were declining, underlining the ethical responsibility of managers to intervene strategically.

The session then shifted to the legal and ethical frameworks governing acquisitions and disposals. Flowcharts on rescue intake pathways and reporting mechanisms helped visualise how animals enter and leave collections, from rescues and confiscations to transfers and reintroductions. This provided a practical grounding in the statutory requirements under the Wild Life (Protection) Act, 1972 and Recognition of Zoo Rules, 2009, while also foregrounding the welfare considerations that must accompany every movement of an animal. Participants were also guided through the ethical paradox of acquiring wild-origin animals, illustrated with decision filters and safeguards. Decision trees were used to demonstrate the narrow circumstances under which such acquisitions can be justified—such as founding a breeding programme, genetic rescue, or compassionate interven-

tions—and the stringent standards of welfare, transparency, and long-term commitment required.

The lecture closed by broadening the discussion to behavioural and ecological considerations. Participants were reminded, with the aid of illustrative diagrams and examples, that ethical planning is not only about numbers or legality but also about aligning husbandry with natural behaviour, ensuring enclosures meet species' ecological needs, and avoiding tokenistic or cosmetic inclusions such as colour morphs. By the end of the session, participants had a holistic understanding that ethical collection planning requires integrating data, law, welfare, ecology, and conservation priorities. The session reinforced that sustainability cannot be assumed—it must be planned and demonstrated through evidence, and zoos carry a clear ethical duty to ensure that every species in their care has a justifiable and sustainable place in the collection.

Risk Management in Collection Planning

- *Dr. Gowri Mallapur*

The session on Risk Management in Collection Planning explored how modern zoos must anticipate and mitigate risks that threaten both animal collections and broader conservation goals. It began by framing risk as the intersection of likelihood and consequence, underscoring that while exact outcomes cannot be predicted, structured analysis allows managers to estimate probabilities and design safeguards. Disease was highlighted as a critical concern, given its multifactorial nature—arising through interactions among agents, hosts, and environmental conditions—and its ability to compromise animal welfare, public health, and ecological integrity.

Participants were introduced to the tools of risk assessment and analysis, including hazard identification, system visualization, and both qualitative and quantitative approaches to evaluating disease threats. The framework of Disease Risk Analysis (DRA) was presented as a structured, evidence-based process that enables decision makers to weigh uncertainties and evaluate the potential impact of infectious and non-infectious hazards on wildlife, domestic animals, ecosystems, and people. Case studies, such as the introduction of bluetongue and African horse sickness in European zoos, illustrated how systematic risk assessments inform quarantine protocols, surveillance, and preventive interventions.

The lecture also emphasized the role of surveillance systems, ranging from wildlife health monitoring and sentinel species to cutting-edge tools like environmental DNA sampling and remote sensing. Examples of global initiatives—including WOA standards, ProMED, and EMPRES-i+ — demonstrated how shared data

strengthens preparedness against emerging zoonotic threats. The session linked these global lessons back to Indian zoos, where effective biosecurity, preshipment health protocols, and quarantine design are critical to reducing cross-infection risks and ensuring compliance with CZA, WLPA, Department of Animal Husbandry & Dairying and CITES regulations.

Beyond disease, other risk domains were explored: reproductive challenges in captivity, cryptic species identification, dietary specialisations, and the coordination required across veterinary, ecological, nutritional, and genetic disciplines. Staff training, robust data management, and adaptive management approaches were identified as essential strategies to address these risks proactively.

The session concluded by stressing that risk management is not a static checklist but a continuous, dynamic process. Periodic audits, incident reporting, and adaptive learning enable institutions to stay responsive to new threats. Ultimately, the key takeaway was that a proactive, evidence-based approach to risk strengthens not only the sustainability of zoo collections but also their credibility as conservation centres capable of balancing animal welfare, biodiversity goals, and institutional resilience.

Collection planning process

- *Lakshminarasimha R*

The session on The Process of Collection Planning built upon the conceptual foundations of the previous lectures by introducing structured frameworks that integrate in situ and ex situ conservation perspectives. It began with the One Plan Approach (OPA), developed by the IUCN-SSC and CPSG, which emphasises that all populations of a species—wild and managed—should be considered together when designing conservation strategies. Participants were shown how this approach overcomes the historic divide between field-based and zoo-based conservation by ensuring that ex situ programmes are evaluated transparently alongside other solutions, avoiding biases either for or against captive management.

The discussion then moved to the Integrated Collection Assessment and Planning (ICAP) process, a structured method widely adopted by regional zoo associations. ICAP was presented not as a replacement for species conservation planning, but as a complementary framework centred on IUCN's five-step planning process Figure 7. Through ICAP, participants saw how pre-work (such as assessing species' wild status, global captive populations, and breeding history) flows into workshops that generate, evaluate, and prioritise potential conserva-

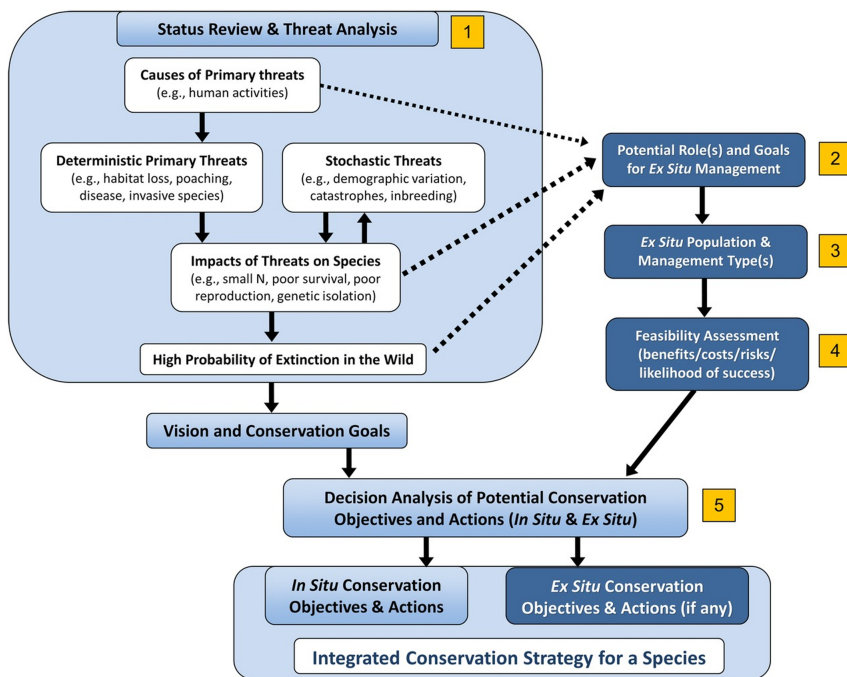


Figure 7: IUCN five-step planning framework – integrates in situ and ex situ roles through threat review, role identification, feasibility assessment, and decision analysis into a unified conservation strategy

tion roles for ex situ management. The emphasis on transparency, repeatability, and scalability was highlighted as key to its growing acceptance.

To contextualise these frameworks, the session outlined how ICAP principles can be adapted to Indian zoos. Here, site-specific considerations such as available space, climatic suitability, and institutional priorities form the foundation of decisions. The workflow demonstrated—beginning with identifying target taxa, conducting cursory analyses of trends and availability, and progressing through decision trees and role evaluations—showed how Indian institutions can build collection plans aligned with both national priorities and global practices.

The lecture also reintroduced reporting mechanisms and acquisition protocols, linking them back to earlier discussions on ethics and sustainability. Participants revisited the statutory requirements for inventories under the Recognition of Zoo Rules, 2009, and examined best practices for acquiring captive-bred specimens, including health checks, studbook verification, and the 75:25 native-to-non-native species guidelines.

Practical considerations in planning collections were further expanded by focusing on behavioural ecology and husbandry needs. Examples highlighted how social structures, diet, breeding biology, and daily rhythms must be incorporated when assigning space and designing enclosures. Special caution was raised against practices like maintaining abundant free-ranging species or including colour morphs without conservation justification.

The session concluded with a look at emerging analytical tools such as the S-Index, which uses demographic metrics to assess the sustainability of institutional collections. Indices like total and intrinsic lambda, producer–consumer balance, and community impact were introduced as ways to quantify the demographic health of captive populations. Together, these tools and frameworks demonstrated that collection planning is not only about selecting species but about embedding those decisions within a transparent, evidence-based system that integrates conservation, welfare, and institutional vision.

Further reading

- Byers, O., Lees, C., Wilcken, J., & Schwitzer, C. (2013). The one plan approach: The philosophy and implementation of CBSG's approach to integrated species conservation planning. *WAZA Magazine*, 14, 2–5.
- Lynch, C. (2019). Sustainability index (s-index): A tool for use in the evaluation and planning of institutional animal collections. *Zoo Biology*, 38(1), 12–23.
- McGowan, P. J., Traylor-Holzer, K., & Leus, K. (2017). IUCN guidelines for determining when and how ex situ management should be used in species conservation. *Conservation Letters*, 10(3), 361–366.
- Traylor-Holzer, K., Leus, K., & Bauman, K. (2019). Integrated collection assessment and planning (ICAP) workshop: Helping zoos move toward the one plan approach. *Zoo Biology*, 38(1), 95–105.

Group exercise: Developing a collection plan

- *Lakshminarasimha R*

The final session of the day shifted to an applied exercise where participants worked in groups to design a hypothetical collection plan for a model zoo. Each group was provided with:

1. Scenarios – outlining zoo characteristics (size, climate, theme, mission focus, and space constraints).

2. Decision Guide – a structured framework with stepwise questions on source, legality, conservation fit, husbandry feasibility, welfare, and long-term responsibility.
 - Each decision point required a clear “Yes,” “No,” or “Caution” response.
 - “No” responses terminated consideration of the species, while “Caution” required documentation and review.
 - This guided participants to think critically about both immediate and long-term implications of acquiring a species.
3. Collection Plan Format – a standardized template to document selected species, including:
 - Founding and target populations (sex-specific numbers)
 - Conservation role/purpose
 - Housing and husbandry requirements
 - Actions and notes (e.g., acquisition source, disposal options, IUCN status, native/non-native status, review timelines).

Groups were tasked with filling out the template based on their assigned scenarios, using the decision guide to justify each inclusion or exclusion. The detailed description of each field in the collection plan template is provided in Appendix C.

4.4 Conservation breeding

4.4.1 Aim

Day 4 aimed to build participants’ understanding of conservation-oriented captive propagation by linking theory with practice. The sessions introduced key demographic and genetic principles for managing viable populations, underscored the role of studbooks and cooperative breeding programmes, and highlighted global case studies of zoo-led conservation successes. The day concluded with hands-on training in SPARKS and PMx, equipping participants to analyse studbook data and apply population management tools to inform evidence-based decisions.

4.4.2 Session details

Conservation breeding in zoos

- *Lakshminarasimha R*

This session introduced participants to the principles and practice of conservation-oriented captive propagation Figure 8, positioning population management as

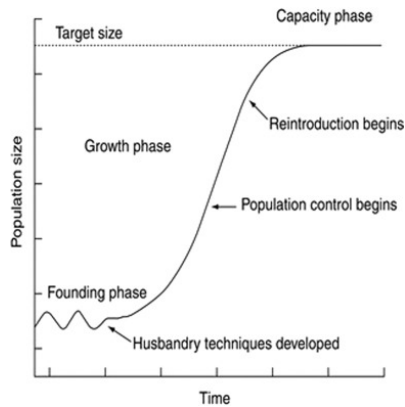


Figure 8: IUCN five-step planning framework – integrates in situ and ex situ roles through threat review, role identification, feasibility assessment, and decision analysis into a unified conservation strategy

the foundation for ensuring the long-term viability and health of species maintained in zoos. The lecture highlighted that while zoos can play a decisive role in ex situ conservation, this contribution depends on the careful design and implementation of scientifically informed population management plans.

Through examples such as the Black-footed ferret (*Mustela nigripes*), California condor (*Gymnogyps californianus*), Golden lion tamarin (*Leontopithecus rosalia*), and Przewalski's horse (*Equus ferus przewalskii*), the discussion illustrated how captive propagation programmes have rescued species from extinction and successfully supported reintroduction to the wild. These examples were contrasted with ongoing challenges, including small founder bases, unstable age structures, and unreliable reproduction, all of which emphasise the importance of demographic stability.

The session distinguished between demographic goals—maintaining viable population sizes and balanced age/sex structures—and genetic goals—retaining founder genetic diversity, avoiding inbreeding, and reducing adaptation to captivity. Concepts such as allelic diversity, heterozygosity, and quantitative variation were explained, underscoring their relevance for long-term population fitness. Statistical models of heterozygosity were used to illustrate how genetic variation can be quantified and monitored over time.

Participants were also introduced to the organisational challenges of regional co-operative programmes such as Species Survival Programs, Australasian Species Management Plan and EAZA Ex situ Programs. These frameworks often re-

quire balancing institutional preferences with science-driven recommendations for transfers, breeding, or even phase-outs. The discussion acknowledged that such processes are resource-intensive—demanding reliable data, studbooks, and coordinated decision-making—but stressed their growing value to conservation. The session further explained the role of studbooks as validated repositories of demographic and genetic data, noting how accurate records of births, deaths, parentage, and transfers are essential for analyses and recommendations. Practical issues of data quality—such as uncertain parentage and missing records—were addressed, with emphasis on documenting uncertainty rather than ignoring it. The lecture concluded with a framework distinguishing between purpose, goals, and targets of population management programmes. An outline of how different programme types—display and education, long-term conservation, release, or husbandry—require tailored demographic and genetic strategies were provided. By the end of the session, participants recognised that conservation-oriented captive propagation is both a scientific and ethical responsibility, requiring evidence-based planning, cooperative implementation, and sustained commitment across institutions.

Further reading

- Ralls Katherine, and Ballou Jonathan D. (2013) Captive Breeding and Reintroduction. In: Levin S.A. (ed.) *Encyclopedia of Biodiversity*, second edition, Volume 1, pp. 662-667. Waltham, MA: Academic Press.

Studbooks and population management

- *Lakshminarasimha R*

The remainder of the day focused on practical training in studbook development and population analysis. Participants were guided through the process of compiling studbooks using SPARKS, with emphasis on systematic data entry, validation, and handling of missing or uncertain information. This exercise highlighted how the integrity of demographic and genetic analyses depends on the reliability of the underlying data, and why documenting uncertainty is preferable to overlooking it.

Building on this foundation, the session introduced the use of SPARKS and PMx as a decision-support tool for demographic and genetic management. Participants explored how demographic indicators (e.g., population size, growth rates, age–sex structure) and genetic measures (e.g., founder representation, gene diversity, inbreeding coefficients) complement each other in evaluating sustainability. Rather than delving into every technical function, the training focused

on demonstrating how PMx outputs can inform real-world recommendations for breeding, transfers, and long-term planning.

The hands-on exercises underscored that studbooks and analytical tools are not abstract exercises but essential instruments for evidence-based management. For Indian zoos, where studbook compilation is still an emerging practice, these skills build vital capacity to support national conservation breeding programmes under CZA. By the end of the session, participants understood how R-based workflows and PMx analyses together form the backbone of population management, ensuring that captive populations are demographically viable, genetically healthy, and aligned with long-term conservation objectives.

4.5 Data-driven decision making

4.5.1 Aim

Day 5 aimed to consolidate participants' learning by applying demographic and genetic principles of population management to practical planning and modelling exercises. The session introduced stochastic and deterministic population models, explored tools such as age pyramids, pedigree charts, and mean kinship analyses, and demonstrated how these outputs can inform breeding recommendations and long-term strategies. Using R workflows alongside PMx, participants gained hands-on experience in integrating demographic and genetic data into coherent population management plans. The day served as a capstone, linking the theoretical foundations of earlier sessions with applied skills for evidence-based conservation breeding.

4.5.2 Session details

Conservation breeding in zoos

- *Lakshminarasimha R*

The final day of the course consolidated the demographic and genetic principles covered in previous sessions by focusing on how they are applied in population planning and modelling. The session began with an introduction to deterministic and stochastic models, explaining how deterministic models provide simplified projections of population growth under fixed assumptions, while stochastic models incorporate the inherent uncertainty of demographic, environmental, and genetic variation. This distinction was used to emphasise why stochastic approaches are essential for realistically assessing extinction risk and population viability.



Figure 9: Changes in age–sex structure of a managed population between 2008 and 2023, visualised as population pyramids. These comparisons highlight demographic shifts over time, including ageing cohorts, recruitment of younger individuals, and sex ratio imbalances.

Participants were then introduced to life tables as a foundation for demographic analysis. These tools summarise age-specific survival and reproduction, allowing managers to calculate mortality rates, survivorship curves, and life expectancy. Worked examples demonstrated how life tables can be constructed from student data, and how they support the identification of vulnerable life stages where interventions may be most effective.

Building on this, the session explored other analytical tools used in population management. Age pyramids were introduced as a way of visualising demographic structure, highlighting issues such as skewed sex ratios, or ageing populations. The importance of monitoring demographic structure was illustrated through a comparison of age–sex pyramids across different years Figure 9. These visualisations demonstrated how population planning must account for ageing cohorts, reproductive bottlenecks, and skewed sex ratios, all of which directly affect long-term viability.

Pedigree charts were used to illustrate how relatedness is tracked across generations, linking them to the concept of mean kinship as the basis for breeding recommendations. The importance of balancing representation among founders and avoiding over-representation of certain lineages was reinforced through practical examples Appendix D.

The session then demonstrated how these demographic and genetic insights come together in population planning exercises. Using R workflows and PMx, participants engaged with worked examples showing how to extract, clean, and analyse

studbook data; generate demographic indicators such as growth rates, survivorship, and reproductive rates; and evaluate genetic parameters including gene diversity, inbreeding coefficients, and kinship values. Rather than focusing on technical detail alone, the training highlighted how outputs such as life tables, age-structure plots, pedigree maps, and mean kinship tables can be interpreted to inform real-world management decisions — whether to recommend new pairings, transfers, or population control.

By the end of the day, participants understood how population models and planning tools translate raw studbook data into actionable strategies for sustainable management. The session reinforced that evidence-based conservation breeding requires not just accurate data, but also the ability to apply modelling approaches to anticipate risks, balance demographic and genetic objectives, and align institutional decisions with long-term conservation goals.

As a capstone to the course, this session brought together the theoretical and practical elements of the previous days, equipping participants with a comprehensive toolkit for managing captive populations within an integrated conservation frame

5 Way forward

This module marked an important step in strengthening the capacity of Indian zoo biologists in collection planning and population management. Over five days, participants engaged with the theoretical foundations of demographic and genetic management, practised the use of studbooks, R workflows, and PMx, and reflected on the ethical and strategic dimensions of zoo collection planning.

The combination of lectures, case studies, and hands-on exercises established a shared baseline of knowledge and skills, while also creating a platform for discussion on the challenges and opportunities in zoo population management in India. The sessions highlighted that while interest and motivation are strong, sustained follow-up and targeted capacity-building will be essential for skills to translate into practice at the institutional level.

5.1 Pre- and post-course surveys

The pre- and post-course surveys provided important insights into participant backgrounds, capacities, and expectations. They revealed that while most participants were engaged in population management or collection planning, relatively few had previously contributed to breeding decisions. Computer literacy varied considerably, with Excel being the most widely used tool, and only a small number familiar with R, SPARKS,

or PMx prior to the training. By the end of the module, however, 85% of participants were able to create basic plots in R, 70% could construct pedigree plots, and much reported confidence in applying PMx for population analysis.

Sessions on collection planning and studbook management were rated most relevant to participants' current roles, with calls for additional depth in demographic and genetic management of captive populations. Importantly, participants also reported a stronger awareness of the ethical dimensions of acquisitions and the long-term viability of species in captivity after attending the course.

The surveys further highlighted systemic challenges. Only half of the zoos represented currently use ZIMS, and two-thirds of participants indicated that data management practices at their institutions need significant improvement. Participants requested more time for group activities, greater emphasis on practical sessions, and opportunities for follow-up to consolidate learning. Encouragingly, all participants expressed willingness to remain connected for post-course support and mentorship.

5.2 Future steps

This module was the first of its kind in India, and as a pioneering effort it must evolve adaptively in response to participant feedback and ongoing assessments by course coordinators. Based on these inputs, a phased roadmap has been outlined to sustain momentum and expand the reach of this initiative.

5.2.1 Immediate follow-up

- Short refresher and catch-up tutorials will be organised over the next months, focusing on practical exercises such as studbook data entry, pedigree plotting, and interpreting PMx outputs.
- Annotated scripts and curated tutorials will be provided to address varying levels of computer literacy and enable participants to revisit concepts at their own pace.
- A peer-support and mentorship group will be established to provide real-time troubleshooting and ongoing guidance.

5.2.2 Short-term (next 6–12 months)

- Participants will be encouraged to prepare pilot collection plans or studbooks for selected species at their institutions, with feedback provided in follow-up modules.
- Structured mentoring will equip biologists to play a more active role in breeding programme design and management, addressing the current gap highlighted in the pre-survey. -

- Support will be extended to zoos exploring ZIMS or similar platforms, with emphasis on best practices in data quality and reporting, tailored to legal and financial feasibility.

5.2.3 Medium-term (1–2 years)

- A second tier of training will cover advanced topics such as viability analysis, stochastic and deterministic modelling, and detailed PMx applications for demographic and genetic management.
- Within 12–18 months, a refresher workshop will revisit core concepts, especially areas identified as challenging during this course (e.g., SPARKS, IUCN Five-Step Process).
- Using the ICAP or similar frameworks, participants will be guided to link institutional collection plans with regional and national priorities, strengthening cooperative species management.

5.2.4 Long-term vision

- With very limited prior training opportunities available, it is essential to build a sustained programme of modules on population management, collection planning, and conservation breeding to develop a skilled pool of Indian zoo biologists.
- Future capacity-building will prioritise embedding data-driven decision-making into routine zoo operations, addressing the gaps in data management identified by participants.
- The collective commitment to stay connected presents an opportunity to establish a long-term community of practice that links Indian zoo biologists with regional and global networks.

5.3 Next steps - Scaling modules nationally

This module was piloted with participants from a limited number of South Indian zoos. The next stage is to expand training to all recognised zoos in India, ensuring that every institution benefits from access to the knowledge, tools, and networks required to contribute to sustainable, science-based population management.

The Zoo Science Course is envisioned as a multi-module programme, with population management successfully piloted in this module. Going forward, additional modules on nutrition, behaviour and welfare, enclosure design and enrichment, and long-term planning should be developed and delivered in a phased manner. Together, these will provide a structured pathway for the professional development of Indian zoo personnel and ensure that training remains comprehensive, relevant, and practice-oriented.

APPENDICIES

A People

Table 2: List of participants

No.	Name	Affiliation
1	Aishwarya K. S.	Biologist, Bannerughatta Biological Park
2	Nisha Majhi	Biologist, Tata Steel Zoological Park, Jamshedpur
3	M. Purushotham	Biologist, Indira Gandhi Zoological Park, Visakhapatnam
4	Dr. Golusu Babu Rao	Biologist, Nehru Zoological Park, Hyderabad
5	T. Lakshmi Harika	Wildlife Biologist, Sri Venkateswara Zoological Park, Tirupati
6	Srinu Srikanta Maharana	Biologist, Wildlife SOS
7	Tejas S	Biologist, Sri Chamarajendra Zoological Garden, Mysuru
8	Krishna Prasad K. S.	Biologist, Pilikula Biological Park, Mangalore
9	Shinoy Rhodes Spadigam	Education Officer & Biologist, MVR Snake Park, Kerala
10	Dr. Chandramouli S. R.	Biologist, Arignar Anna Zoological Park, Chennai
11	Sangeetha Mohan	Curator Grade II, Zoological Garden, Thiruvananthapuram
12	Sinchana J	Education Officer & Biologist, Tiger and Lion Safari, Shivamogga
13	Carol Brunilda Amanna	Biologist & Education Officer, Gadag Zoo, Karnataka
14	Angeline Samuel	Curator, Madras Crocodile Bank Trust, Mamallapuram
15	Apoorva M. P.	Assistant Biologist, Bannerughatta Biological Park, Bengaluru
16	Anuj Kandpal	Biologist & Education Officer, G. B. Pant High Altitude Zoo, Nainital
17	Manjusha S	Curator, State Museum and Zoo, Thrissur

Table 3: List of speakers

No.	Name	Affiliation
1	Lakshminarasimha R	Director (Research & Conservation), GaiaMitra Collective Foundation
2	Abishek Harihar	Director (Tiger Program), Panthera
3	Sarah Lavin	Member Training Specialist, Species360
4	Ashwin Vishwanathan	Scientist, Nature Conservation Foundation
5	Amanda Embury	Senior Manager, Animal Welfare Development, Zoos Victoria
6	Gowri Mallapur	Director, GaiaMitra Collective Foundation
7	Priyanka Adhikari	Teaching Assistant, GaiaMitra Collective Foundation

B Getting Started with R

Install the most recent versions of **R** and **RStudio**. When RStudio is opened, it provides a graphical user interface (GUI) with four main panes:

- Source pane – used to write and save R scripts. A new script can be created from File → New File → R Script or with the shortcut Ctrl + Shift + N.
- Environment/History pane – displays the objects created in the current session (such as vectors and dataframes), the command history, and Git options (if enabled).
- Console/Terminal pane – where commands are run and results appear. The terminal and background jobs are also accessible here.
- Files/Plots/Packages/Help/Viewer pane – provides access to the working directory, generated plots, installed packages, documentation, and outputs.

For reference, see Figure 1.

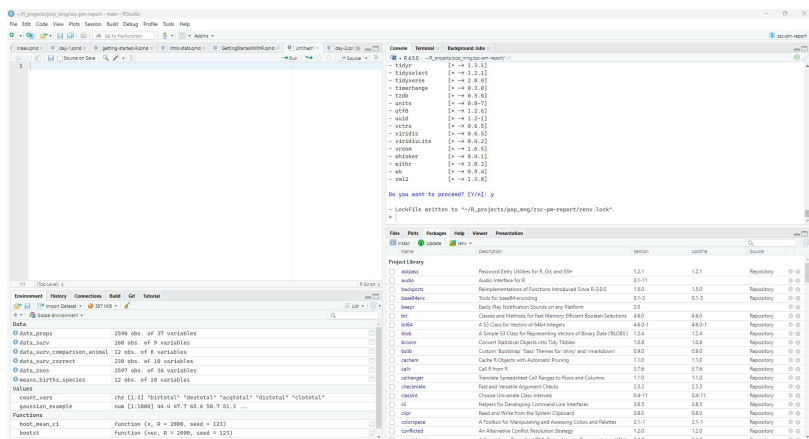


Figure 1: The RStudio interface showing the four main panes: Source (top left), Environment/History (bottom left), Console/Terminal (top right), and Files/Plots/Packages/Help/Viewer (bottom right).

R uses a *working directory* to know where to look for files and where to save outputs. By default, this is set to a system folder, but it can be changed. The function `setwd()` manually sets the working directory to a specific folder. For example, `setwd("D:\R-projects")` sets the working directory to the said path (Use forward slashes / instead of backslashes \ in file paths).

An alternate and more reproducible option is to use the `here` package. It defines file paths relative to the project root, avoiding hard-coded folder locations. This makes scripts easier to share and run on different systems. For example, `here("data",`

"zoo_inventory.csv") looks for the file in the folder data located in the project root. With `here()`, the working directory is automatically detected, so collaborators only need to keep files in the same project structure.

R's core installation comes with a basic set of functions. To extend its functionality, additional packages can be installed and loaded. A package can be installed (done once) using -

```
install.packages("here")
```

An installed package has to be loaded for each session using -

```
library(here)
```

Packages provide new functions, datasets, and tools that are not part of base R. Thousands of packages are available through [CRAN \(the Comprehensive R Archive Network\)](#).

For workshops or projects, it is often convenient to install several packages together using -

```
invisible(lapply(c("tidyverse", "here"), library, character.only = TRUE))
```

R works with different data types. The most common are:

- Numeric – numerical values, e.g., 1, 12.4, -7
- Character – text values, e.g., "thanks", "five", "79"
- Logical – Boolean values, represented as TRUE or FALSE

A variable is a name that stores data. Once assigned, the variable can be used to refer to that data in calculations or other operations.

```
a <- 2  
b <- 2 - 8  
c <- "hello"  
d <- "3"  
f <- TRUE
```

If an undefined object is used, R will return an error. For example -

```
e <- morning # "morning" has not been defined
```

The `str()` function can be used to examine the type and structure of any variable -

```
str(a)
```

```
num 2
```

```
str(c)
```

```
chr "hello"
```

```
str(f)
```

```
logi TRUE
```

In R, it is recommended to use `<-` for assigning values to variables. The `=` operator is generally used for specifying arguments within functions. For example, the function `rnorm()` generates random numbers. Here, the argument `n` specifies the number of values to generate -

```
z <- rnorm(n = 10)
z
```

```
[1] -0.84435835 -0.04844391 -0.17644277 -0.63424656  0.97478865  0.79994039
[7]  0.69038277 -1.36166679  0.26422260  0.54012386
```

In this case, `z` is assigned the generated numbers, and appears in the *Environment* tab of RStudio. The argument `n`, however, is not stored as a separate variable.

Vectors are one of the most common data structures in R. They are sequences of elements, all of the same type (numeric, character, or logical). The function `str()` can be used to check the type and structure of a vector -

```
str(z)
```

```
num [1:10] -0.8444 -0.0484 -0.1764 -0.6342 0.9748 ...
```

Vectors can be also created by combining values with the `c()` function -

```
x <- c(2, 4, 78, 8, 10) # specify elements in a vector
```

Variable names in R are **case-sensitive** (e.g., Name and name are different).

- Names should begin with a letter and may include numbers or underscores.
 - Valid: flower1, flower_petal
 - Invalid: 1flower, _petal

Objects created during a session appear in the *Environment* pane in RStudio. They can be removed when no longer needed -

```
rm(var_name)      # remove a single object
rm(list = ls())    # clear all objects from the environment
```

Alternatively, the *broom* icon in the Environment pane can be used to clear objects through the graphical interface.

Arithmetic expressions can be applied to numeric variables. Either previously declared variables or new ones may be used -

```
a + b # addition
```

```
[1] -4
```

```
a - b # subtraction
```

```
[1] 8
```

```
a * b # multiplication
```

```
[1] -12
```

```
a / b # division
```

```
[1] -0.3333333
```

```
b ^ a # exponentiation
```

```
[1] 36
```

Numeric values can also be compared, with the result returned as a logical value (TRUE or FALSE) -

```
a == b # equal to
```

```
[1] FALSE
```

```
a != b # not equal to
```

```
[1] TRUE
```

```
34 < 6 # less than
```

```
[1] FALSE
```

```
2 <= 8/3 # less than or equal to
```

```
[1] TRUE
```

```
65 + 15 > 89 # greater than
```

```
[1] FALSE
```

```
8 >= 2 # greater than or equal to
```

```
[1] TRUE
```

Membership can be tested with %in% -

```
3 %in% c(2, 5, 3, 6) # returns TRUE
```

```
[1] TRUE
```

A wide range of mathematical operations are available in base R -

```
log(x)          # logarithm of x
```

```
[1] 0.6931472 1.3862944 4.3567088 2.0794415 2.3025851
```

```
round(x, digits = 2) # round to 2 decimal places
```

```
[1] 2 4 78 8 10
```

```
sqrt(x)         # square root
```

```
[1] 1.414214 2.000000 8.831761 2.828427 3.162278
```

Statistical summaries can be applied to vector data -

```
max(z)          # maximum value
```

```
[1] 0.9747886
```

```
min(z)          # minimum value
```

```
[1] -1.361667
```

```
sum(z)          # sum of values
```

```
[1] 0.2042999
```

```
mean(z)         # mean value
```

```
[1] 0.02042999
```

Working with dataframes: Data stored in tabular form is represented in R as a **dataframe**. A dataframe can be created by combining vectors of equal length -

```
col1 <- c("row1", "row2", "row3", "row4", "row5")
col2 <- 1:5
col3 <- 12:16

df <- data.frame(col1, col2, col3)
df
```

```
  col1 col2 col3
1 row1    1   12
2 row2    2   13
3 row3    3   14
4 row4    4   15
5 row5    5   16
```

The `str()` function shows the structure of the dataframe -

```
str(df)
```

```
'data.frame':  5 obs. of  3 variables:
 $ col1: chr  "row1" "row2" "row3" "row4" ...
 $ col2: int   1  2  3  4  5
 $ col3: int  12 13 14 15 16
```

In this example, the dataframe has 5 rows and 3 columns. For selecting elements within a dataframe -

```
df[2, 3]  # element in row 2, column 3
```

```
[1] 13
```

```
df[3, ]   # third row
```

```
  col1 col2 col3
3 row3    3   14
```

```
df[, 1]   # first column
```

```
[1] "row1" "row2" "row3" "row4" "row5"
```

```
df$col2   # column named "col2"
```

```
[1] 1 2 3 4 5
```

The tidyverse is a collection of R packages designed for data manipulation, reshaping, and visualization.

- dplyr provides functions for filtering, mutating, arranging, grouping, and summarising data.
- tidyr provides tools for reshaping data (e.g., pivot_wider() and pivot_longer()).
- ggplot2 enables flexible data visualisation.

If installed, the package can be loaded using library(tidyverse).

The pipe operator %>% is used to link commands into data pipelines, where the output of one function is passed directly to the next:

```
# Creating a sample dataframe
zoo_data <- data.frame(
  species = c("lion", "tiger", "leopard", "hyena", "bear"),
  count   = c(5, 3, 7, 4, 6),
  enclosure = c("A", "B", "B", "C", "C"),
  weight_kg = c(190, 220, 80, 60, 120)
)

zoo_data %>%
  filter(count > 5)
```

	species	count	enclosure	weight_kg
1	leopard	7	B	80
2	bear	6	C	120

Dataset can be explored using -

```
str(zoo_data)
```

```
'data.frame':  5 obs. of  4 variables:
 $ species : chr  "lion" "tiger" "leopard" "hyena" ...
 $ count   : num  5 3 7 4 6
 $ enclosure: chr  "A" "B" "B" "C" ...
 $ weight_kg: num  190 220 80 60 120
```

```
nrow(zoo_data) # number of rows
```

```
[1] 5
```

```
ncol(zoo_data) # number of columns
```

```
[1] 4
```

```
count(zoo_data, enclosure) # how many species per enclosure
```

```
enclosure n
1         A 1
2         B 2
3         C 2
```

```
summary(zoo_data)
```

species	count	enclosure	weight_kg
Length:5	Min. :3	Length:5	Min. : 60
Class :character	1st Qu.:4	Class :character	1st Qu.: 80
Mode :character	Median :5	Mode :character	Median :120
	Mean :5		Mean :134
	3rd Qu.:6		3rd Qu.:190
	Max. :7		Max. :220

Filtering data for specific cases - e.g. Select only species with more than 5 individuals

```
zoo_data %>%
  filter(count > 5)
```

```
species count enclosure weight_kg
1 leopard   7         B      80
2 bear      6         C     120
```

Creating new variables by mutating - e.g. Add a new column showing total biomass in each enclosure (count * weight_kg) -

```
zoo_data %>%
  mutate(total_biomass = count * weight_kg) %>%
  arrange(desc(total_biomass)) # arrange the results in descending order
```

	species	count	enclosure	weight_kg	total_biomass
1	lion	5	A	190	950
2	bear	6	C	120	720
3	tiger	3	B	220	660
4	leopard	7	B	80	560
5	hyena	4	C	60	240

When working with dataframes, it is often useful to perform calculations within categories rather than across the whole dataset. The function `group_by()` allows data to be grouped by the values of one or more columns. Once grouped, any summarising function (such as `sum()`, `mean()`, `min()`, `max()`) is applied within each group separately.

For example, in the `zoo_data` dataset, species are assigned to different enclosures. Grouping by enclosure makes it possible to calculate totals or averages for each enclosure independently -

```
zoo_data %>%
  group_by(enclosure) %>%
  summarise(total_animals = sum(count),
            mean_weight = mean(weight_kg))
```

```
# A tibble: 3 x 3
  enclosure total_animals mean_weight
  <chr>         <dbl>         <dbl>
1 A             5             190
2 B            10             150
3 C            10             90
```

This groups the dataset by enclosure and then produces two new columns:

- `total_animals`: the sum of all individuals in that enclosure
- `mean_weight`: the average weight of species in that enclosure

Grouping is especially powerful when combined with filtering, mutating, and arranging, as it allows detailed summaries to be generated across categories.

The complete pipeline below demonstrates how several steps of data manipulation can be combined into a single, readable workflow. Starting with the `zoo_data` dataframe, the operations proceed in sequence:

1. `filter` – keep only species where the number of individuals is greater than five.
2. `mutate` – create a new column `total_biomass`, calculated as `count * weight_kg`.
3. `arrange` – order the rows in descending order of biomass so the largest contributions appear first.
4. `group_by` – divide the dataset into subsets based on the `enclosure` column.
5. `summarise` – within each enclosure, calculate the total number of animals (`total_animals`) and the mean body weight (`mean_weight`).

This approach avoids creating multiple intermediate objects, makes the sequence of transformations easy to follow, and ensures reproducibility by capturing the entire analysis in a single block of code.

```
zoo_data %>%
  filter(count > 5) %>% # filter for species with
  ↪ >5 individuals
  mutate(total_biomass = count * weight_kg) %>% # calculate biomass
  arrange(desc(total_biomass)) %>% # order by biomass
  group_by(enclosure) %>% # group by enclosure
  summarise(total_animals = sum(count),
            mean_weight = mean(weight_kg)) # summarise totals and
  ↪ mean weight
```

```
# A tibble: 2 x 3
  enclosure total_animals mean_weight
  <chr>      <dbl>      <dbl>
1 B         7         80
2 C         6        120
```

Depending on the type of analysis, data may need to be reorganised so that information is arranged either across more columns or more rows. The following functions can be used for reshaping data -

- `pivot_wider()` – reduces the number of rows and creates more columns, making the table *wider*.
- `pivot_longer()` – reduces the number of columns and creates more rows, making the table *longer*.

Consider a sample dataset -

```
zoo_name <- rep(c("Mysuru", "Delhi", "Nandankanan"), 3)
year <- c(rep("2019", 3), rep("2020", 3), rep("2021", 3))
animal_count <- c(38, 7, 66, 44, 19, 81, 51, 49, 120)

zoo_df <- data.frame(zoo_name, year, animal_count)
zoo_df
```

```
  zoo_name year animal_count
1  Mysuru 2019         38
2   Delhi 2019          7
3 Nandankanan 2019        66
4  Mysuru 2020         44
5   Delhi 2020         19
```

6	Nandankanan	2020	81
7	Mysuru	2021	51
8	Delhi	2021	49
9	Nandankanan	2021	120

This is the dataset in a long format — each row represents a single zoo in a single year. Long tables are preferred in most tidyverse workflows and are essential for many plotting functions in ggplot2. They are flexible because each observation is stored in its own row.

Wide tables are often useful for quick comparisons across years, or when certain statistical functions require each year as a separate column.

```
pivot_wider(zoo_df,
            names_from = year,
            values_from = animal_count)
```

```
# A tibble: 3 x 4
  zoo_name `2019` `2020` `2021`
  <chr>      <dbl> <dbl> <dbl>
1 Mysuru      38     44     51
2 Delhi        7     19     49
3 Nandankanan 66     81    120
```

The ggplot2 package provides functions to create plots in a modular way. Instead of %>% (used for data pipelines), the + operator is used to add layers to a plot.

```
ggplot(data, aes(x = x_col, y = y_col)) +
  geom_point() +
  theme_bw()
```

Consider a sample dataset -

```
zoo_data <- data.frame(
  species = c("lion", "tiger", "leopard", "hyena", "bear"),
  count   = c(5, 3, 7, 4, 6),
  enclosure = c("A", "B", "B", "C", "C"),
  weight_kg = c(190, 220, 80, 60, 120),
  sex      = c("male", "female", "male", "female", "male")
)
```

The ggplot() function sets up the basic plot object, specifying the data frame and mapping variables to axes -

```
ggplot(zoo_data, aes(x = species, y = count))
```

Titles and axis labels are added with `labs()` -

```
ggplot(zoo_data, aes(x = species, y = count)) +  
  labs(title = "Zoo animal counts",  
        x = "Species",  
        y = "Number of individuals")
```

A histogram requires a single numeric variable -

```
ggplot(zoo_data, aes(x = weight_kg)) +  
  geom_histogram() +  
  labs(title = "Distribution of average weights")
```

Box plots are useful for comparing numeric values across categories -

```
ggplot(zoo_data, aes(x = species, y = weight_kg)) +  
  geom_boxplot() +  
  labs(title = "Weight distribution by species")
```

Line graphs are used when both axes are numeric (often for trends across time) -

```
ggplot(zoo_data, aes(x = count, y = weight_kg)) +  
  geom_line() +  
  labs(title = "Counts vs body weight")
```

Scatter plots show the relationship between two numeric variables, with optional colour by category -

```
ggplot(zoo_data, aes(x = weight_kg, y = count, color = species)) +  
  geom_point() +  
  labs(title = "Body weight vs counts by species")
```

Trend lines and jitter can be added -

```
ggplot(zoo_data, aes(x = weight_kg, y = count, color = species)) +  
  geom_point() +  
  geom_smooth(method = "lm") +  
  geom_jitter() +  
  labs(title = "Scatter plot with trend line and jitter")
```

Bar graphs are commonly used for counts across categories -

```
zoo_data %>%  
  count(sex) %>%  
  ggplot(aes(x = sex, y = n)) +  
  geom_col() +  
  labs(title = "Number of individuals by sex")
```

Compare weights of males and females for each species using box plots -

```
ggplot(zoo_data, aes(x = sex, y = weight_kg)) +  
  geom_boxplot() +  
  facet_wrap(~species) +  
  labs(title = "Comparison of weights by sex across species")
```

C Introduction to statistics using R

Importing data and loading libraries

A project-root-relative workflow is used via here package so file paths remain stable across machines and during Quarto renders. The input file `test_data_species.csv` is placed in a `data/` folder at the project root. The tidyverse is loaded for data handling, and the data is imported with `read_csv()` using `here("data", "test_data_species.csv")`.

```
invisible(lapply(c("tidyverse", "here", "extrafont", "Hmisc"), library,
  ↪ character.only = TRUE))

data_zoos <- readr::read_csv(here("data", "test_data_species.csv"))

# To use custom fonts in plots run the following line after loading
↪ `extrafont` package
loadfonts(device = "win")
```

The dataset

The dataset comprises annual inventory records for selected species across multiple zoos. For each zoo-species-year, it includes sex-disaggregated opening and closing balances, births, deaths and acquisitions. These fields are used to derive indicators (e.g., net additions, crude survival ratios) and to summarise trends by zoo and species. Each data-frame column is a vector of a specific type.

Numeric (e.g. counts of individuals) and character (e.g. location, zoo names) columns can be inspected using the following code to confirm how R has parsed the file; this helps prevent downstream errors (e.g., arithmetic on text fields) and informs coercions where needed.

```
# Numeric columns: returns names of columns parsed as numeric
```

```
data_zoos %>%
  dplyr::select(where(is.numeric)) %>%
  names()
```

```
[1] "...1"      "opbalmale"  "opbalfemale" "opbalunsex" "opbaltotal"
[6] "birmale"    "birfemale"  "birunsex"     "birtotal"   "acqmale"
[11] "acqfemale"  "acqunsex"   "acqttotal"    "dismale"    "disfemale"
[16] "disunsex"   "distotal"   "deamale"      "deafemale"  "deahunsex"
[21] "deatotal"   "clomale"    "clofemale"    "clounsex"   "clototal"
```

```
# Character columns: returns names of columns parsed as character
```

```
data_zoos %>%
  dplyr::select(where(is.character)) %>%
```

```
names()
```

```
[1] "zoo"          "FinanceYear"  "Statename"    "Category"
[5] "SpeciesName" "TypeName"     "CityName"     "AnimalNAME"
[9] "ScientificName"
```

```
# Factor columns (if any): returns names of columns parsed as factors. If no
↳ factor variables are present, the last call returns an empty character
↳ vector.
```

```
data_zoos %>%
  dplyr::select(where(is.factor)) %>%
  names()
```

```
character(0)
```

Creating new variables

In the dataset, financial years are recorded as ranges (e.g., *1995–1996*). For longitudinal or time-series analysis, these are normalised by extracting the closing year of the range (e.g., *1996*). The extracted year is then converted to a date format using `as.Date()` within `mutate()`, and the year component is obtained via `lubridate::year()`. This provides a consistent variable (`period`) that can be used in plots or models requiring temporal ordering.

```
data_zoos %>%
  select(FinanceYear) %>%
  head()
```

```
# A tibble: 6 x 1
  FinanceYear
  <chr>
1 1995-1996
2 1995-1996
3 1995-1996
4 1995-1996
5 1995-1996
6 1995-1996
```

```
data_zoos %>%
  select(FinanceYear) %>%
  mutate(FinanceYear = as.Date(FinanceYear, format = "%Y"),
         period = lubridate::year(FinanceYear)) %>%
  head()
```

```
# A tibble: 6 x 2
  FinanceYear period
  <date>         <dbl>
1 1995-09-30    1995
2 1995-09-30    1995
3 1995-09-30    1995
4 1995-09-30    1995
5 1995-09-30    1995
6 1995-09-30    1995
```

The `summarise()` function in the tidyverse provides a convenient way to generate summary statistics. When combined with `group_by()`, it allows data to be aggregated by one or more grouping variables. In this dataset, grouping by `AnimalNAME` (which represents species) enables calculation of totals across all years for each species. For example, the following code calculates the total number of births for each species recorded across all reporting periods:

```
data_zoos %>%
  group_by(AnimalNAME) %>%
  summarise(births_total = sum(birtotal, na.rm = TRUE))
```

```
# A tibble: 12 x 2
  AnimalNAME                births_total
  <chr>                    <dbl>
1 Asiatic Brush-tailed Porcupine      8
2 Binturong                        9
3 Burmese Python                    18
4 Clouded leopard                   36
5 Gee's Golden Langur (Golden Langur)  7
6 Gray Slender Loris                 2
7 Indian Pangolin                   22
8 Leopard (melanistic)              13
9 Marsh Crocodile                   950
10 Red Crowned Roof Turtle           59
11 Rusty-spotted Cat                 9
12 Tawny Eagle                       0
```

The approach can be refined further by combining grouping with filtering. For instance, to examine only one species (e.g., *Marsh Crocodile*) across time, the following code calculates annual totals of births and deaths, derives a new variable (`bd_ratio`), and adds a period variable to support time-series analysis:

```
data_zoos %>%
  filter(AnimalNAME == "Marsh Crocodile") %>%
  group_by(FinanceYear) %>%
  summarise(births_total = sum(birtotal, na.rm = TRUE),
```

```

    deaths_total = sum(deatotal, na.rm = TRUE)) %>%
  mutate(period = lubridate::year(as.Date(FinanceYear, format = "%Y")),
    bd_ratio = births_total/deaths_total) %>%
  head()

```

```

# A tibble: 6 x 5
  FinanceYear births_total deaths_total period bd_ratio
  <chr>          <dbl>         <dbl> <dbl>    <dbl>
1 1995-1996         43           52 1995    0.827
2 1996-1997        110           28 1996    3.93
3 1997-1998         17           48 1997    0.354
4 1998-1999         21          928 1998    0.0226
5 1999-2000          6          114 1999    0.0526
6 2000-2001         32          125 2000    0.256

```

This example illustrates how grouping, summarising, and mutating can be combined within a single workflow to produce derived indicators. Such operations form the foundation of exploratory analyses in zoo data, allowing researchers to assess species-level patterns, calculate key ratios, and prepare the data for more advanced statistical procedures.

Understanding distributions

Zoo inventory variables are typically counts and seldom approximate a Gaussian distribution. Transformations (e.g., log or square-root) and derived rates (e.g., annual growth, survival proportions) can yield approximately Gaussian behaviour, especially after aggregation across years or sites. When Gaussian assumptions are inappropriate, distribution-aware approaches (Poisson for counts, Binomial for proportions) or non-parametric methods are preferred.

To illustrate contrasts: zoo inventory data are naturally modelled by Poisson (counts) and Binomial (proportions), whereas Gaussian (Normal) data are symmetric and unbounded. A simulate Gaussian data for teaching purposes and use the inventory for Poisson/Binomial examples.

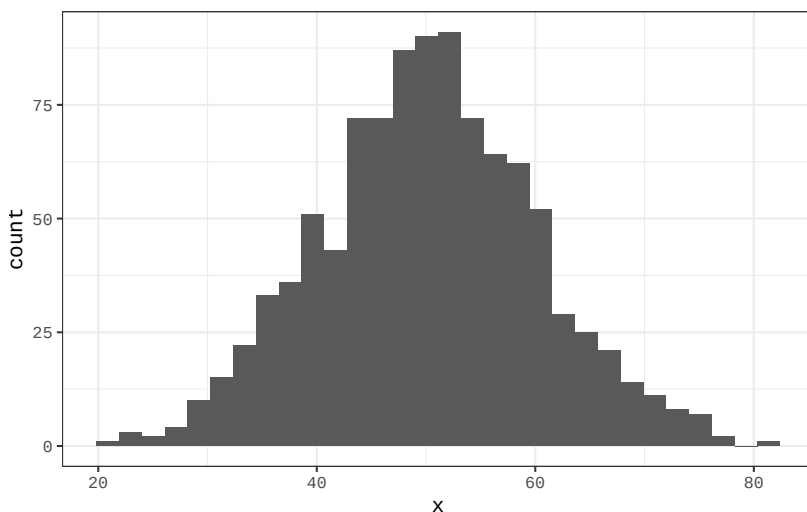
```

set.seed(123)
gaussian_example <- rnorm(1000, mean = 50, sd = 10)

ggplot(data.frame(x = gaussian_example), aes(x = x)) +
  geom_histogram() +
  theme_bw(base_family = "Noto Sans Mono", base_size = 10)

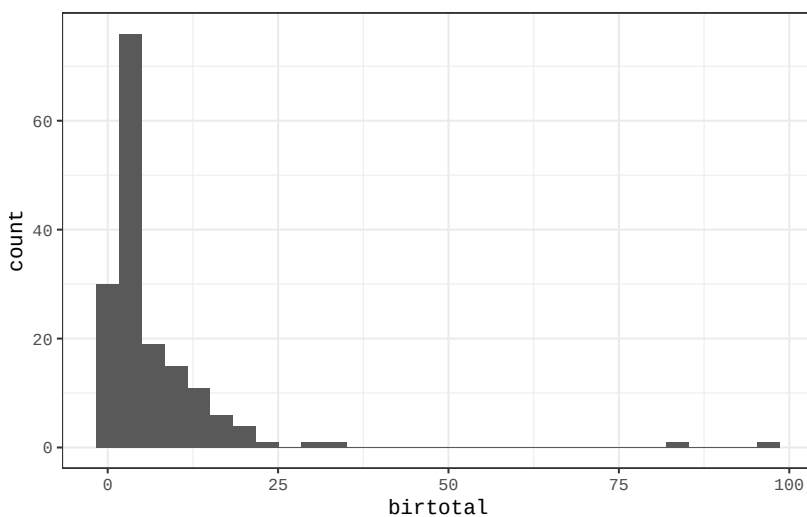
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



Counts such as births or deaths per zoo-year are often right-skewed with variance $\approx$ mean, a characteristic of Poisson-like data.

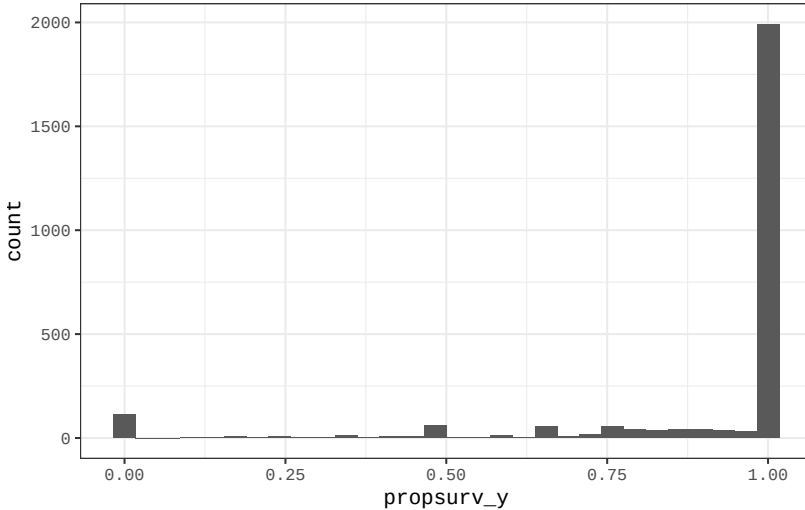
```
data_zoos %>%
  select(birtotal) %>%
  filter(birtotal > 0) %>%
  ggplot(aes(x = birtotal)) +
  geom_histogram() +
  theme_bw(base_family = "Noto Sans Mono", base_size = 10)
```



Binomial outcomes arise when a fixed number of trials each result in either a “success” or a “failure,” and the distribution of successes is bounded between 0 and 1 when expressed as a proportion. In the context of zoo inventory data, a practical proxy is the within-year survival proportion. This is computed as the number of individuals remaining at year-end relative to those present or added during the year (opening stock plus births and acquisitions). The measure is crude but useful as a diagnostic, since it naturally behaves like a binomial proportion (bounded between 0 and 1).

```
data_zoos %>%
  mutate(
    trials_y = opbaltotal + birttotal + acqtotal,
    success_y = pmin(clottotal, trials_y),
    propsurv_y = success_y / trials_y
  ) %>%
  filter(is.finite(propsurv_y), trials_y > 0,
         propsurv_y >= 0, propsurv_y <= 1) %>%
  ggplot(aes(x = propsurv_y)) +
  geom_histogram() +
  theme_bw(base_family = "Noto Sans Mono", base_size = 10)
```

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



Means and Confidence intervals

To illustrate measures of central tendency and uncertainty, species-level birth totals were summarised from the inventory dataset. For each species, mean births per

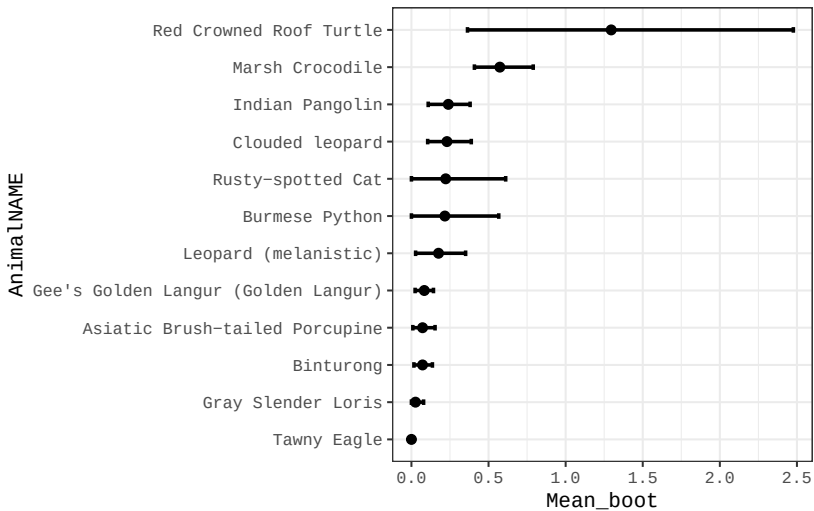
year were calculated along with both *t*-based and bootstrap confidence intervals. The bootstrap method resamples the observed data repeatedly to build an empirical distribution of the mean, which is particularly useful when the underlying data are skewed or contain small sample sizes. The resulting plot shows species ordered by their mean birth rate, with error bars representing the bootstrap 95% confidence interval. This demonstrates how basic inventory records can be summarised and compared across taxa, while also conveying the inherent uncertainty in point estimates.

```
bootst <- function(vec, R = 2000, seed = 123) {
  x <- vec[is.finite(vec)]
  n <- length(x)
  if (n == 0) return(c(NA_real_, NA_real_, NA_real_))
  set.seed(seed)
  mn <- replicate(R, mean(sample(x, size = n, replace = TRUE)))
  c(
    quantile(mn, 0.025, type = 6),
    median(mn),
    quantile(mn, 0.975, type = 6)
  )
}

# Per-species mean births with both t-based and bootstrap CIs
means_births_species <- data_zoos %>%
  group_by(AnimalNAME) %>%
  summarise(
    n      = sum(!is.na(birtotal)),
    Mean   = mean(birtotal, na.rm = TRUE),
    sd     = sd(birtotal, na.rm = TRUE),
    se     = sd / sqrt(n),
    # t-based 95% CI (use t-quantile; falls back to NA if n <= 1)
    lci    = ifelse(n > 1, Mean - qt(0.975, df = n - 1) * se, NA_real_),
    rci    = ifelse(n > 1, Mean + qt(0.975, df = n - 1) * se, NA_real_),
    # bootstrap CI (median of boot means for robustness)
    lci_boot = bootst(birtotal)[1],
    Mean_boot = bootst(birtotal)[2],
    rci_boot = bootst(birtotal)[3],
    .groups = "drop"
  ) %>%
  arrange(Mean_boot) %>%
  mutate(AnimalNAME = factor(AnimalNAME, levels = AnimalNAME))

# Plot: mean births per species with bootstrap 95% CI
ggplot(means_births_species, aes(x = AnimalNAME, y = Mean_boot)) +
  geom_errorbar(aes(ymin = lci_boot, ymax = rci_boot), width = 0.15,
    ↪ linewidth = 0.8) +
  geom_point(size = 1.8) +
  coord_flip() +
```

```
theme_bw(base_family = "Noto Sans Mono", base_size = 10)
```



Survival proportions and binomial confidence intervals

Beyond counts, proportions provide another meaningful perspective. Survival was expressed as the proportion of individuals remaining at the end of a period relative to those present at the beginning (plus acquisitions and births). These proportions are naturally bounded between 0 and 1 and can be analysed using methods suitable for binomial data. Wilson confidence intervals were calculated using the `binconf()` function from the `Hmisc` package, which provides more accurate coverage than simple normal approximations, especially for smaller sample sizes. The plot displays species ordered by their survival proportion, with 95% binomial confidence intervals, enabling comparison of species' outcomes while accounting for sample size effects.

```
data_surv <- data_zoos %>%
  arrange(FinanceYear) %>%
  group_by(AnimalNAME, zoo) %>%
  summarise(
    start_opbal = first(opbaltotal),
    sum_births = sum(birtotal, na.rm = TRUE),
    sum_acq = sum(acqtotal, na.rm = TRUE),
    sum_dists = sum(disttotal, na.rm = TRUE),
    sum_deaths = sum(deatotal, na.rm = TRUE),
    ending_count = last(clototal),
    starting_count = start_opbal + sum_births + sum_acq - sum_dists,
    .groups = "drop"
```



```

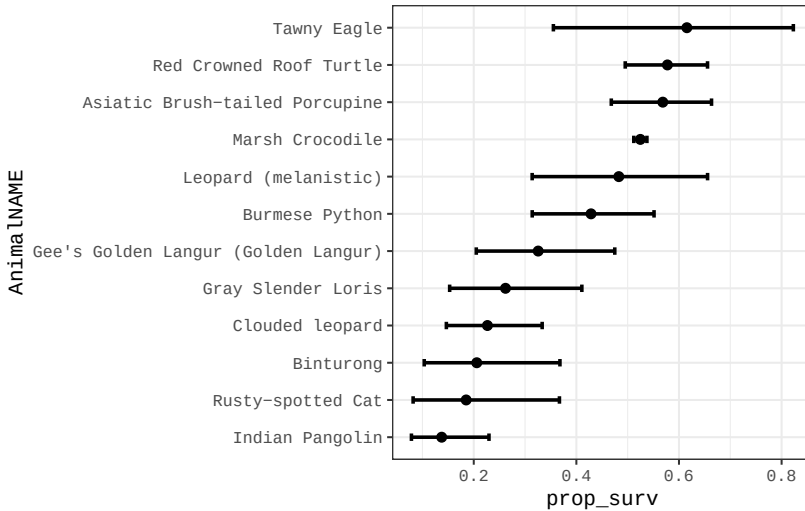
)

data_surv_correct <- data_surv %>%
  filter(
    starting_count > 0,
    sum_deaths <= starting_count,
    ending_count <= starting_count
  ) %>%
  mutate(prop_surv = ending_count / starting_count)

data_surv_comparison_animal <- data_surv_correct %>%
  group_by(AnimalNAME) %>%
  summarise(
    starting_count = sum(starting_count, na.rm = TRUE),
    ending_count   = sum(ending_count,   na.rm = TRUE),
    .groups = "drop"
  ) %>%
  filter(starting_count > 0) %>%
  mutate(
    prop_surv = ending_count / starting_count,
    lci = Hmisc::binconf(x = ending_count, n = starting_count)[, 2],
    rci = Hmisc::binconf(x = ending_count, n = starting_count)[, 3]
  ) %>%
  arrange(prop_surv) %>%
  mutate(AnimalNAME = factor(AnimalNAME, levels = AnimalNAME))

ggplot(data_surv_comparison_animal, aes(x = AnimalNAME, y = prop_surv)) +
  geom_errorbar(aes(ymin = lci, ymax = rci), width = 0.2, linewidth = 0.8) +
  geom_point(size = 1.8) +
  coord_flip() +
  theme_bw(base_family = "Noto Sans Mono", base_size = 10)

```



Institutional comparisons

The same survival analysis could be repeated at the zoo level, aggregating across all species within each institution. This allows exploration of variation between facilities in terms of overall survival outcomes. While such aggregated figures should be interpreted cautiously (given differences in species composition and holding sizes), they provide a starting point for identifying outliers and framing discussions about management practices across institutions.

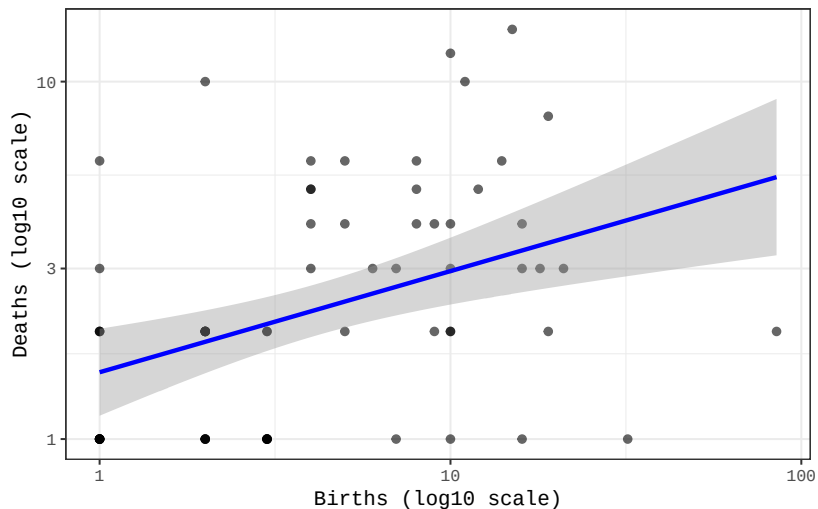
```
data_surv_comparison_zoo <- data_surv_correct %>%
  group_by(zoo) %>%
  reframe(
    starting_count = sum(starting_count, na.rm = TRUE),
    ending_count   = sum(ending_count,   na.rm = TRUE)
  ) %>%
  filter(starting_count > 0) %>%
  mutate(
    prop_surv = ending_count / starting_count,
    lci = Hmisc::binconf(x = ending_count, n = starting_count)[, 2],
    rci = Hmisc::binconf(x = ending_count, n = starting_count)[, 3]
  ) %>%
  arrange(prop_surv) %>%
  mutate(zoo = factor(zoo, levels = zoo))
```

Correlations

To explore relationships between variables, births and deaths were plotted against one another across zoo-year records. Both variables are right-skewed, so a log-log scale was applied to improve visibility of the data distribution. The scatter plot, with a fitted linear trend line, illustrates that higher numbers of births are generally associated with higher numbers of deaths, reflecting the effect of larger underlying populations.

```
data_zoos %>%  
  drop_na(birttotal, deatotal) %>%  
  filter(birttotal > 0, deatotal > 0) %>% # log(0) undefined  
  ggplot(aes(x = birttotal, y = deatotal)) +  
  geom_point(alpha = 0.6) +  
  geom_smooth(method = "lm", se = TRUE, color = "blue") +  
  scale_x_log10() +  
  scale_y_log10() +  
  labs(  
    x = "Births (log10 scale)",  
    y = "Deaths (log10 scale)"  
  ) +  
  theme_bw(base_family = "Noto Sans Mono", base_size = 10)
```

`geom\_smooth()` using formula = 'y ~ x'



In general terms, the positive association between births and deaths indicates that years or institutions with larger populations tend to report both more births and more deaths. This does not necessarily imply a causal link between the two variables; rather, it reflects underlying scale effects, where more animals create more

opportunities for both reproduction and mortality events. Correlation therefore provides a useful first look at how key inventory measures vary together, but it should be interpreted cautiously and supplemented with further analysis to understand drivers of population change.

D Key Variables in a Zoo Collection Plan

Species Common (vernacular) name of the species as it will appear in the collection plan.

Taxa Broad taxonomic group (e.g., Mammal, Bird, Reptile, Amphibian, Fish, Invertebrate). Helps assess balance across major groups.

Scientific name Binomial nomenclature (*italicised* in final reports), ensuring clarity across regions and avoiding misidentification.

Sub-species Subspecific designation, if relevant. Important for avoiding hybridisation and ensuring the plan aligns with recognised taxonomic units.

Founding population (M/F/U) Number of male, female, and unknown/unsexed individuals available at the start of the program. Establishes the genetic and demographic baseline.

Role / Purpose Primary justification for maintaining the species (e.g., conservation breeding, education, research, display). Multiple roles may be indicated but must be documented.

Target population (M/F/U) Desired sex-structured population size required for medium- to long-term viability, reflecting zoo carrying capacity and regional goals.

Housing Information on space allocation, enclosure type, and any special requirements (climatic control, vertical complexity, aquatic features), ensuring compatibility with species biology and welfare standards.

Actions Key management steps (acquisition, breeding recommendations, transfers, phase-out, surplus management).

Source Anticipated origin of animals (rescue intake, captive stock, inter-zoo exchange, imports), compliant with legal frameworks.

Disposal options for surplus animals Mechanisms for responsibly managing individuals beyond carrying capacity (transfers, loans, cooperative programs). Euthanasia only in exceptional, welfare-based cases.

Key husbandry requirements Species-specific needs for diet, group structure, breeding, veterinary care, enrichment, or unique constraints (e.g., nocturnal management).

Timeline / Review year Periodicity for reassessment (commonly every 3–5 years, or sooner if circumstances change).

IUCN Red List status Global conservation status. Helps prioritise species of higher conservation concern.

Native / Non-native Whether the species occurs naturally within India. Aligns with policy of maintaining ~75% native and ~25% non-native taxa.

Notes Free-text field for additional observations, constraints, or justification.

E Pedigree construction and analysis

This appendix illustrates the workflow for constructing pedigrees, calculating kinship coefficients, and identifying potential breeding recommendations. The example below uses a mock dataset and demonstrates the process step by step in R. This workflow complements the main session on population management by demonstrating how pedigree and kinship tools can be applied in practice.

Data

The first step is to load the required packages and import the pedigree dataset:

```
invisible(lapply(c("kinship2", "tidyverse", "here", "gplots",  
  ↪ "wesanderson"), library, character.only = TRUE))  
  
pedigree_file <- read_csv(here("data", "ped_sample.csv"))
```

The dataframe includes five variables:

- id: Unique identifier for each individual
- sire: Unique ID of the father
- dam: Unique ID of the mother
- sex: Sex of the individual
- status: Life status — 0 for deceased, 1 for living

```
head(pedigree_file)
```

```
# A tibble: 6 × 5  
  id sire dam sex status  
  <dbl> <dbl> <dbl> <chr> <dbl>  
1     1    NA  NA  M      1  
2     2    NA  NA  F      1  
3     3     1   2  M      1  
4     4     1   2  M      1  
5     5     1   2  F      1  
6     6     1   2  M      1
```

Pedigree construction

A pedigree object is created using the `pedigree()` function from the `kinship2` package:

```
pedAll_gc <- kinship2::pedigree(id = pedigree_file$id,  
                                dadid = pedigree_file$sire,  
                                momid = pedigree_file$dam,  
                                sex = pedigree_file$sex,  
                                status = pedigree_file$status)
```


The pedigree can be plotted using the `plot.pedigree` function from the `kinship2` package. In the resulting plot (Figure 1), males are represented by squares and females by circles. Where siblings form a mated pair and produce offspring, this is illustrated by two horizontal lines connecting the pair, denoting consanguineous mating. Individuals that are deceased are additionally shown with a diagonal cross line. In some cases, dotted arcs appear in the diagram — these represent duplicate instances of the same individual, introduced automatically to simplify the layout in complex pedigrees where a given individual is paired with multiple individuals.

```
kinship2::plot.pedigree(  
  pedAll_gc,  
  symbolsize = 0.8,  
  cex = 0.7,  
  col = 2)
```

Warning in `autohint(ped)`: Unexpected result in `autohint`, please contact developer

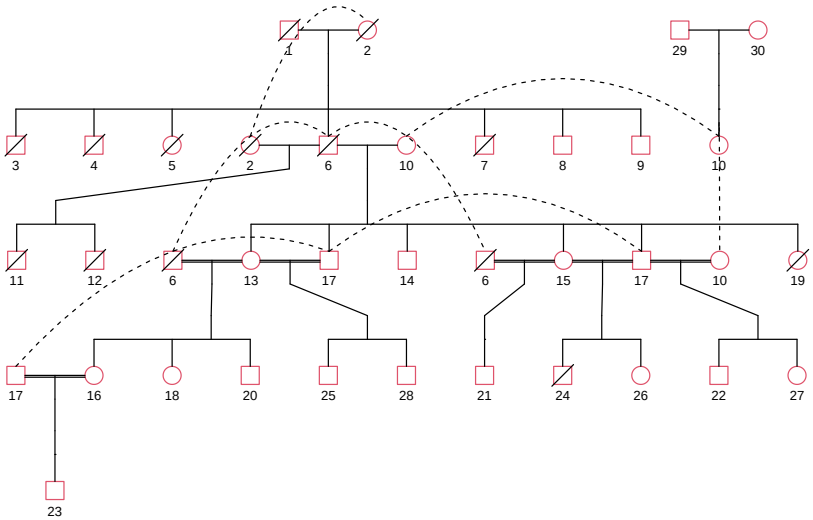


Figure 1: Pedigree plot of the sample population.

Kinship calculations

Pairwise kinship coefficients quantify the probability of two alleles being identical by descent. Values of 0.25 typically indicate full siblings or parent–offspring relationships, while 0.125 corresponds to half-siblings or first cousins.

The full matrix of pairwise kinship coefficients can be visualized as a heatmap (Figure 2), where individuals are ordered by hierarchical clustering to reveal related

clusters. Warmer colors indicate higher relatedness. This graphical representation helps highlight clusters of closely related individuals and identify potential inbreeding risks.

```
kinAll <- kinship(pedAll_gc)
kinAll[1:10,1:10]
```

	1	2	3	4	5	6	7	8	9	10
1	0.50	0.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.0
2	0.00	0.50	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.0
3	0.25	0.25	0.50	0.25	0.25	0.25	0.25	0.25	0.25	0.0
4	0.25	0.25	0.25	0.50	0.25	0.25	0.25	0.25	0.25	0.0
5	0.25	0.25	0.25	0.25	0.50	0.25	0.25	0.25	0.25	0.0
6	0.25	0.25	0.25	0.25	0.25	0.50	0.25	0.25	0.25	0.0
7	0.25	0.25	0.25	0.25	0.25	0.25	0.50	0.25	0.25	0.0
8	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.50	0.25	0.0
9	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.50	0.0
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.5

```
gplots::heatmap.2(kinAll,
  trace = "none",
  col = wesanderson::wes_palette("Zissou1", type = "continuous"),
  dendrogram = "both",
  margins = c(2,2),
  cexRow = 0.5,
  cexCol = 0.5)
```

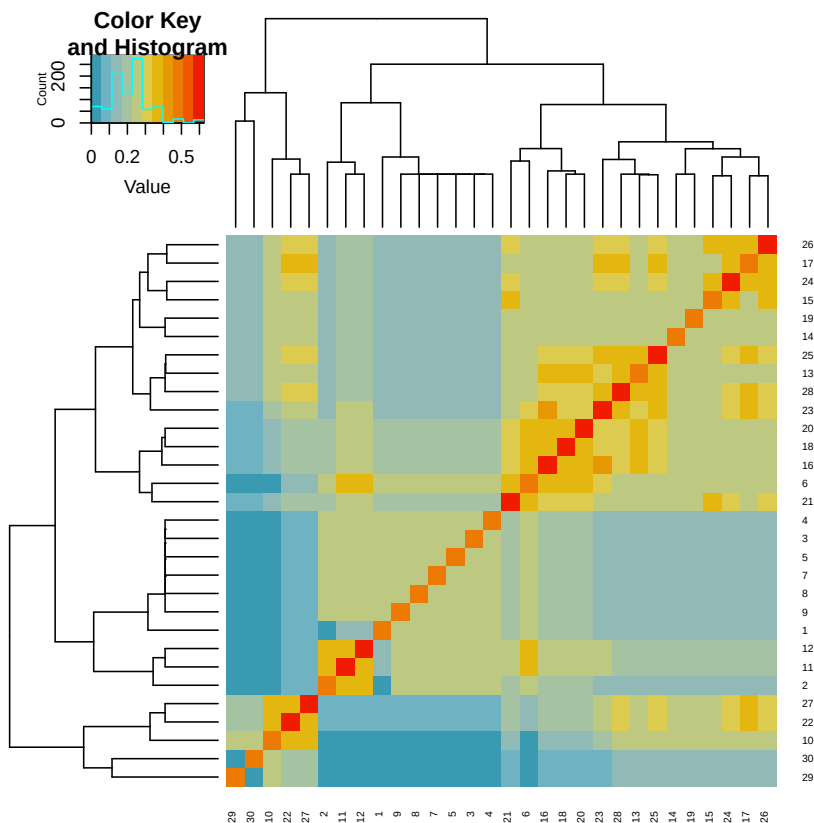


Figure 2: Heatmap of pairwise kinship coefficients in the sample population.

Breeding considerations

For management purposes, potential breeding recommendations often focus on pairs with low relatedness and and to detect and reduce consanguinity risks.. In this example, dyads with kinship coefficients of 0.125 or less (equivalent to first-cousin level or lower) were extracted. Only pairs where both individuals are alive were retained, providing a reference for possible breeding combinations that avoid inbreeding while maintaining genetic connectivity.

```
kin_df <- as.data.frame(as.table(kinAll)) %>%
  rename(id1 = Var1, id2 = Var2, kinship = Freq) %>%
  mutate(id1 = as.numeric(id1),
         id2 = as.numeric(id2)) %>%
  filter(id1 < id2, kinship <= 0.125)
```

```

kin_annotated <- kin_df %>%
  left_join(pedigree_file %>%
    select(id1 = id, sex1 = sex, status1 = status),
    by = "id1") %>%
  left_join(pedigree_file %>%
    select(id2 = id, sex2 = sex, status2 = status),
    by = "id2")

kin_living <- kin_annotated %>%
  filter(status1 == 0, status2 == 0)

kin_living %>%
  select(id1, sex1, id2, sex2, kinship) %>%
  arrange(desc(kinship))

```

	id1	sex1	id2	sex2	kinship
1	8	M	13	F	0.12500
2	9	M	13	F	0.12500
3	8	M	14	M	0.12500
4	9	M	14	M	0.12500
5	8	M	15	F	0.12500
6	9	M	15	F	0.12500
7	10	F	16	F	0.12500
8	8	M	17	M	0.12500
9	9	M	17	M	0.12500
10	10	F	18	F	0.12500
11	10	F	20	M	0.12500
12	10	F	21	M	0.12500
13	8	M	25	M	0.12500
14	9	M	25	M	0.12500
15	8	M	26	F	0.12500
16	9	M	26	F	0.12500
17	8	M	28	M	0.12500
18	9	M	28	M	0.12500
19	13	F	29	M	0.12500
20	14	M	29	M	0.12500
21	15	F	29	M	0.12500
22	17	M	29	M	0.12500
23	25	M	29	M	0.12500
24	26	F	29	M	0.12500
25	28	M	29	M	0.12500
26	13	F	30	F	0.12500
27	14	M	30	F	0.12500
28	15	F	30	F	0.12500
29	17	M	30	F	0.12500
30	25	M	30	F	0.12500
31	26	F	30	F	0.12500

32	28	M	30	F	0.12500
33	23	M	29	M	0.09375
34	23	M	30	F	0.09375
35	8	M	22	M	0.06250
36	9	M	22	M	0.06250
37	8	M	27	F	0.06250
38	9	M	27	F	0.06250
39	16	F	29	M	0.06250
40	18	F	29	M	0.06250
41	20	M	29	M	0.06250
42	21	M	29	M	0.06250
43	16	F	30	F	0.06250
44	18	F	30	F	0.06250
45	20	M	30	F	0.06250
46	21	M	30	F	0.06250
47	8	M	10	F	0.00000
48	9	M	10	F	0.00000
49	8	M	29	M	0.00000
50	9	M	29	M	0.00000
51	8	M	30	F	0.00000
52	9	M	30	F	0.00000
53	29	M	30	F	0.00000



**BANNERUGHATTA
BIOLOGICAL PARK**
BENGALURU

BANNERUGHATTA BIOLOGICAL PARK

Bannerughatta, Bengaluru 560083, Karnataka.
🌐 bannerughattabiopark.org | ☎ +91 80 29776466/67