Institute of Primate Research

STANDARD OPERATING PROCEDURE (SOP) DOCUMENT

Reproducible coding practices (Git, R Markdown, Jupyter.)

SOP No.	Issue Number	Issue Date	Revision Status	Revision Date
SOP/KIPRE/RPD/DSAS/3.1.76	Version 01	October 2025	-	-

Approvals

	Name	**Signature**	**Date**
Developed by:	_Patrick Waweru Mwaura_	_____	**_6th October; 2025_**
	_____	_____	_____
	_____	_____	_____
Reviewed by:	_____	_____	_____
Approved by:	_____	_____	_____

Table of Contents

1. PURPOSE

To ensure all DS&AS coding practices are transparent, reproducible, version-controlled, and auditable.

2. SCOPE

Applies to all coding activities in DS&AS projects, including statistical analysis, machine learning, bioinformatics pipelines, and reporting.

3. PERSONS RESPONSIBLE:

- **All DS&AS Analysts and Data Scientists:** Follow reproducible coding practices.
- **Data Engineer:** Maintains version-control repositories.
- **Head of DS&AS:** Monitors compliance with reproducibility standards.

4. FREQUENCY

- Applied **throughout project lifecycle**.
- **Quarterly audits** of repositories for compliance.

5. MATERIALS

- Git/GitHub/GitLab for version control.
- R Markdown, Jupyter Notebooks, Quarto for documentation.
- Institutional coding standards and templates.

6. PROCEDURE

1. **Repository Setup:** Create project repository in Git (institutional or GitHub Enterprise).
2. **Version Control:** Commit all scripts with meaningful messages; use branching for feature development.
3. **Reproducible Documentation:** Use R Markdown/Jupyter/Quarto for analyses, embedding code and results together.
4. **Collaboration:** Use pull requests and peer code review before merging.
5. **Archiving:** Tag final project versions; store outputs and code in DS&AS repository.
6. **Audit:** Quarterly review of repositories for compliance with reproducibility standards.

7. REFERENCES