

Thrive Africa Data Science Presentation Guidelines

****Presentation Format Guide****

Max total time: 15 minutes for presentation + 2-minute break between sessions

Your team does not need to present for the full 15 minutes; this is simply the absolute maximum time allowed.

****[2 minutes] Project Introduction and Problem Statement****

- Outline the data science problem and its significance.
- Describe the business or social impact of solving this problem.
- Provide relevant background information to help the audience understand the domain.
- Clearly define project goals and expected deliverables.

Outcome: The audience understands the problem being tackled and why it matters.

****[3 minutes] Data Collection and Preprocessing****

- Describe the data sources used (APIs, web scraping, public datasets, internal data, etc.).
- Explain data cleaning, handling missing values, feature engineering, and transformations.
- Discuss challenges faced in data acquisition and preprocessing.

Outcome: The audience understands the quality and characteristics of the dataset.

****[3 minutes] Exploratory Data Analysis (EDA) and Insights****

- Showcase key statistical summaries and visualizations.
- Highlight major patterns, correlations, or anomalies discovered.
- Explain how insights from EDA guided feature selection or model choice.

Outcome: The audience gains insight into the dataset and key analytical findings.

****[3 minutes] Model Development and Performance Evaluation****

- Outline the types of models experimented with (regression, classification, clustering, etc.).
- Describe hyperparameter tuning strategies and performance metrics used.
- Compare different models and justify the final model selection.
- Address overfitting concerns and generalization strategies.

Outcome: The audience understands the modeling approach and performance trade-offs.

****[3 minutes] Deployment and MLOps Considerations****

- Describe the model deployment pipeline (batch, real-time, or edge deployment).
- Discuss infrastructure components (cloud services, containerization, CI/CD pipelines).
- Explain how the model is monitored for drift and performance degradation.
- Address data privacy, security, and ethical concerns if applicable.

Outcome: The audience understands how the model is deployed and maintained in production.

****[2 minutes] Recommendations and Future Work****

- Identify potential improvements (additional data sources, feature engineering, model enhancements).
- Discuss scalability considerations and possible next steps.
- Highlight any limitations or trade-offs encountered in the project.

Outcome: The audience sees the potential for future enhancements and research directions.

****[2 minutes] Questions****

- At 15 minutes, the team will be stopped, and questions will be allowed.
- Each audience member can ask one question.
- Keep responses concise and avoid excessive technical detail.

****Note:**** Thrive Africa data science presentations should be concise and impactful, ensuring clarity and engagement within the shortened timeframe.