Mobile Money Financial Inclusion Analyzer

Prepared by: Eugene Mutembei, Avyn Rabala and Emmanuel Omondi.

Timeline: 3 Days

Start Date: July 28, 2025

1. Purpose

This project addresses financial exclusion in East Africa by identifying key barriers to mobile money adoption and predicting individuals' likelihood of using such services. Despite M-Pesa's popularity, adoption gaps persist. The system will visualize trends, build predictive models, and provide actionable insights.

2. Objectives

- Visualize mobile money adoption across East Africa (2011–2021).
- Predict mobile money usage using demographic features.
- Build an interactive dashboard to explore adoption barriers.
- Deliver a brief technical report with recommendations.

3. Scope

In Scope

- Cleaning datasets and imputing missing values
- Time-series plots for mobile money trends
- Random Forest & XGBoost classifiers
- Feature importance and barrier insights
- Dashboard with filters (age, gender, education, rural/urban)
- Technical summary and recommendation report

4. Data Sources

- World Bank Global Findex (2011–2021)

- GSMA Mobile Money Deployment Tracker
- Kenya FinAccess Survey

5. Functional Requirements

ID	Feature	Description
FR1	Data Preprocessing	Clean and impute missing data from all sources
FR2	Time-Series Graph	Plot adoption rates over time by country
FR3	ML Classification	Predict usage with Random Forest and XGBoost
FR4	Feature Analysis	Analyze which factors (age, gender, etc.) drive adoption
FR5	Interactive Dashboard	Explore barriers by demographic filters
FR6	Final Report	Document methods, results, and recommendations

6. Non-Functional Requirements

Requirement	**Description**
Performance	ML models should train within 5 mins
Usability	Dashboards must be easy to interact with
Accuracy	Minimum 75% accuracy and solid ROC-AUC
Portability	Should run on Jupyter or deployable via Streamlit

7. Timeline (3 Days)

Day	Tasks
Day 1	• Download datasets and clean data • Handle missing values • Create time-series visualizations • Perform exploratory data analysis (EDA)
Day 2	• Train Random Forest & XGBoost classifiers • Evaluate models (accuracy, ROC-AUC) • Extract feature importance • Build interactive dashboard with filters
Day 3	• Finalize dashboard • Write technical report (methods, insights, charts) • Summarize recommendations • Package and present final deliverables

8. Success Criteria

- ✅ Time-series charts showing adoption trends (2011–2021)
- ✅ Classifier with at least 75% accuracy
- ✅ Dashboard with demographic filters
- ✅ 3+ actionable policy or program recommendations in report

9. Risks & Mitigation

Risk	**Mitigation**
Incomplete data	Use imputation and fallbacks across datasets
Model underfitting	Try parameter tuning or alternative classifiers
Limited time	Prioritize core deliverables: visualization, model, dashboard, and report

10. Tools & Tech Stack

- Python (Pandas, scikit-learn, XGBoost)
- Plotly + Streamlit or Dash (for dashboard)
- Jupyter Notebook/Colab
- Google Docs / Word for report