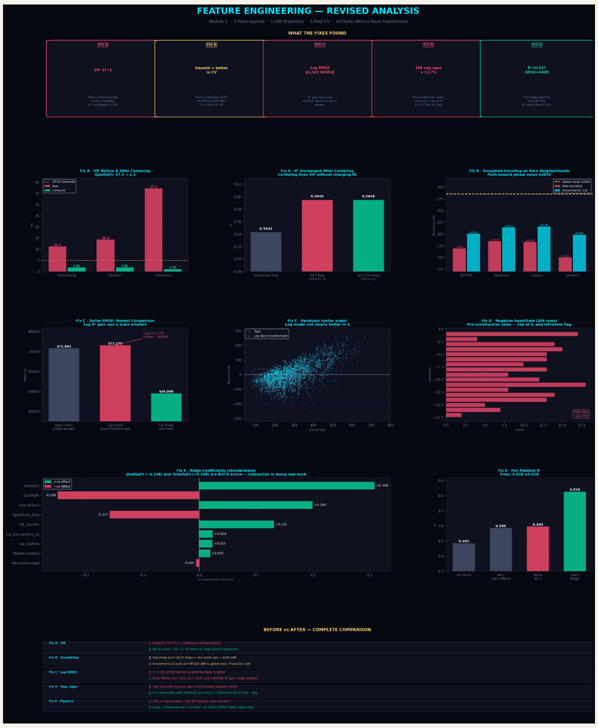
Bolanle Insights Firm

DATA · STRATEGY · PROFIT

Portfolio Catalog

Property Pricing Model · AI Data Analysis

Feature Engineering — Revised Analysis



Overview

Applied five production-grade fixes to a property pricing model across 1,460 properties with 5-fold cross-validation. Final pipeline: R^2=0.827, RMSE=$49K — all dollar metrics back-transformed for honest comparison.

Key Findings

- Fix A — Mean-centred features: VIF 37→1 (multicollinearity resolved).
- Fix B — Smoothed encoding (k=10): rare neighbourhoods pulled toward global mean.
- Fix C — Honest dollar RMSE: exposed log model as $1,501 WORSE, not better.
- Fix D — Capped 186 pre-construction negative ages + added IsPreSale flag.
- Fix E — Ridge + StandardScaler full pipeline: R^2 0.485 → 0.826.

Business Value

Shows how disciplined feature engineering turns a mediocre model into a deployable asset.

WhatsApp: 0902 224 3575 · Email: Omobolanle.analytics@gmail.com

Book a free Data Health Check today.