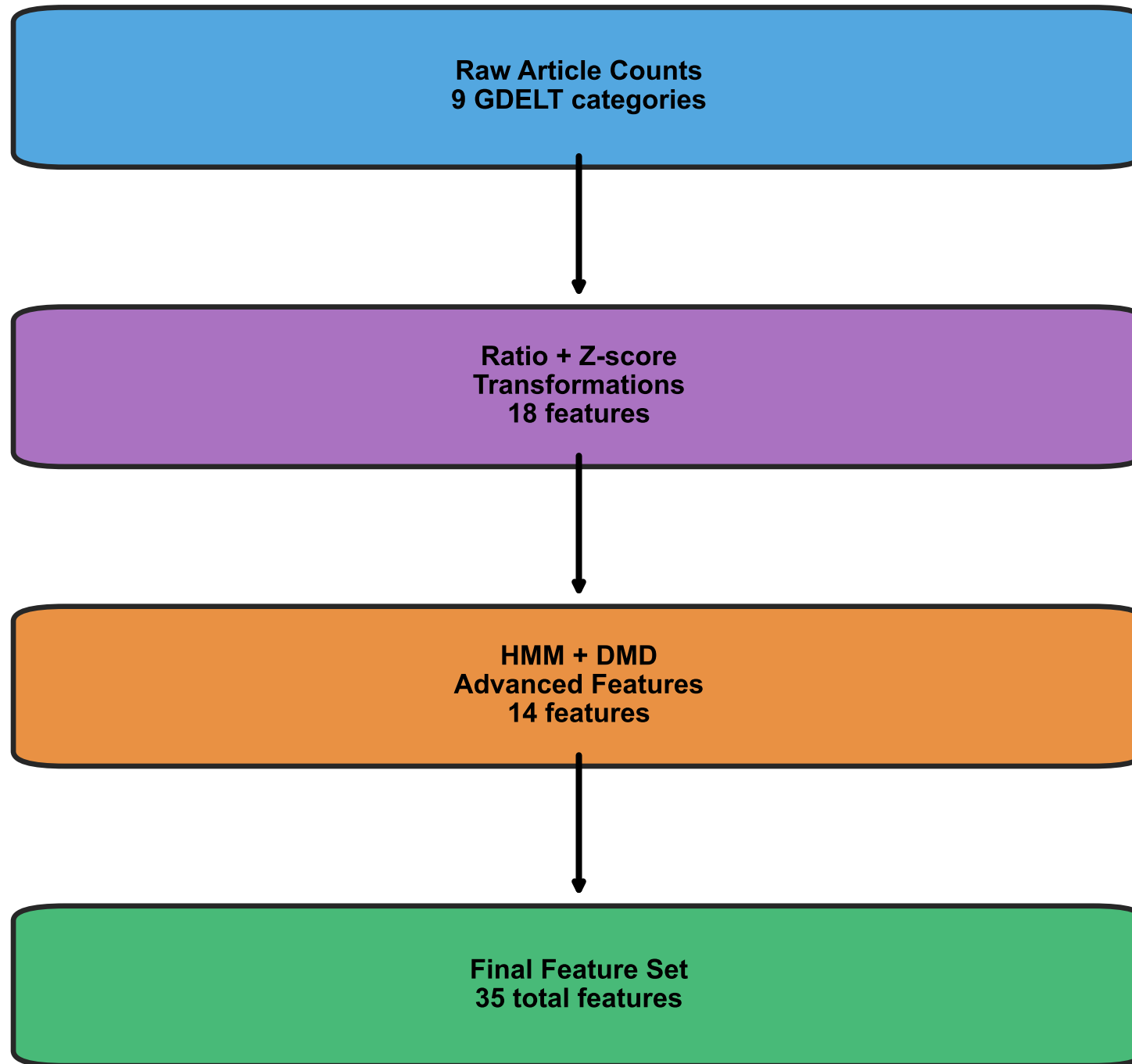


Feature Engineering Pipeline



Feature engineering pipeline transforms 7.6M GDELT articles into 35 engineered features for Stage 2 XGBoost models. Raw article counts (9 GDELT categories) → Ratio features (compositional dynamics) + Z-score features (anomaly detection) → HMM features (regime transitions) + DMD features (temporal dynamics) → Final feature set (3 location + 32 news-derived). Location features dominate importance (country_data_density 13.3%), while HMM transition risk ranks #5 (3.2% importance).