**Final Project Proposal: Classifying Retirement Plan Financial Adequacy Using Form 5500 Data**

**Context & Motivation**

Employer-sponsored retirement plans, especially 401(k)s, are the primary savings vehicle for millions of Americans. But not all plans equally support long-term financial security. Some offer strong structural support for participant savings, while others suffer from low engagement, high leakage via early withdrawals, or poor contribution patterns.

The U.S. Department of Labor collects detailed annual filings from these plans via Form 5500, which include financial, demographic, and operational data. These filings offer a unique opportunity to assess whether a retirement plan is structurally enabling participants to save adequately for retirement.

**Dataset Selection**

I propose to use publicly available Form 5500 filings from the Department of Labor’s EFAST2 portal for the year 2023. Each filing contains structured data on plan assets, contributions, distributions, participant counts, administrative expenses, and more. I will filter the dataset to focus on complete filings for single employer defined contribution plans, which are the most common group retirement vehicle in the United States.

**Learning Question**

Can we classify retirement plans as “adequate” or “inadequate” in supporting participant financial readiness, based on structural indicators such as participation rate, average account balance, contribution stability, and leakage volume?

This question is relevant to employers that sponsor these plans, policymakers, and retirement service providers who want to understand whether a plan is structurally supporting participant financial security, especially in the face of economic disruptions like inflation, unemployment, or market volatility.

**Dataset Details**

The dataset includes approximately 12,177 plan filings for the year 2023. Each row represents a plan-year filing and includes 32 variables I will use for analysis.

The dataset, which can be viewed via this link, provides a rich foundation for classifying retirement plans by participant financial readiness. It captures structural, demographic, and financial dimensions for plans, sponsors, and participant activity. Key variables span plan metadata (e.g., effective dates, entity codes), sponsor identifiers (EIN, business codes), and participant counts at both the beginning and end of year—enabling temporal diagnostics and cohort tracking.

Financial flows are well represented: employer and participant contributions, loans, distributions, and asset transfers offer a dynamic view of plan liquidity and engagement. Asset metrics at beginning and end of year support derived measures of growth, volatility, and adequacy.

This is a classification problem, with the target variable defined as a binary label: adequate vs. inadequate, based on engineered thresholds for savings adequacy, engagement, and leakage.

To classify retirement plans as structurally adequate or insufficient, I plan to engineer features that capture how well a plan enables saving, accommodates financial stress, and combats participant inertia. Ratios like participation rate, contribution per participant, and leakage burden offer interpretable signals of plan adequacy, while stratified diagnostics across industries and plan sizes support fairness and stakeholder relevance.

**Model Selection**

To build a principled and interpretable classifier, I plan to explore:

**Logistic Regression:** For baseline interpretability and fairness audits.

**Random Forest:** To capture nonlinear relationships and assess feature importance.

Each model will be evaluated using ROC/AUC, confusion matrices, and fairness overlays across industries and plan sizes. I’ll also use residual diagnostics to identify misclassified plans and explore macroeconomic sensitivity.

**GitHub Repository**

I will share my data, code, documentation, and findings in this Github repository.

**Data Leakage Avoidance**

This project avoids data leakage by using only plan-level features available at the time of filing. No future outcomes, external economic indicators, or derived labels from post-filing behavior are used in training. All engineered features, e.g. participation rate, average account balance, leakage ratio, will be constructed from within-year data.

**Question Significance**

The learning question, “Can we classify retirement plans as structurally adequate or insufficient in supporting participant financial readiness?” is highly impactful. It addresses a real-world challenge faced by sponsors, policymakers, and service providers: identifying which plans are enabling long-term financial security and which may be structurally failing participants. The insights can inform outreach, redesign, and policy interventions.

**Potential for Misuse**

There is a risk that the classifier could be misinterpreted as a judgment of individual participant behavior or investment performance. To mitigate this, the model is explicitly framed as a structural diagnostic tool, not a predictor of individual outcomes. Misuse could also arise if sponsors use the model to justify disengagement from vulnerable plans rather than improving them. Clear documentation and ethical framing are essential.

**Investment Data**

While investment selection and performance are undeniably a component of retirement readiness, this analysis intentionally excludes direct investment-level data, e.g. asset allocation, fund returns, volatility, from the classification model. The rationale is twofold:

**Structural Focus:** This project emphasizes plan-level structural indicators, such as participation, contributions, leakage, and fee burden, that reflect how well a plan enables participants to save. These are actionable, interpretable, and directly tied to sponsor decisions.

**Data Limitations:** Investment details in Form 5500 filings are often inconsistently reported, difficult to normalize across plans, and not reliably attributable to participant outcomes without individual-level data.

By focusing on structural adequacy rather than investment performance, the model aims to provide a principled, diagnostic view of retirement plan readiness that is both reproducible and stakeholder-relevant.

**Lack of Participant-Level Financial Context**

One potential weakness is that this model cannot account for individual participant financial circumstances, such as income, debt, employment status, or household dynamics, that directly influence contribution and withdrawal behavior. While the classifier uses plan-level structural indicators, e.g. participation rate, average balance, leakage volume, it cannot infer whether low savings or high leakage is due to plan design or external financial hardship.

Form 5500 data does not include individual-level financial information, and the model is designed to assess plan adequacy, not participant intent. However, this means the classifier may occasionally flag structurally sound plans as “insufficient” if participant behavior reflects broader economic stress rather than plan design flaws.

**Behavioral and Structural Context**

Another concern is that participation in retirement plans is shaped not only by plan design but by the realities of financial wellness and behavioral inertia.

For many individuals, the ability to access funds through loans or hardship withdrawals is essential. Locking up assets in a retirement plan without this flexibility can deter enrollment, especially in financially vulnerable populations.

Moreover, inertia plays a powerful role: if a participant is not automatically enrolled and contributions are not deducted from payroll, saving requires deliberate action. Most people don’t take that step.

These behavioral dynamics mean that low participation or contribution rates may reflect financial constraints or plan accessibility issues, not simply a lack of retirement readiness. This underscores the importance of interpreting model outputs as structural diagnostics, not judgments of individual behavior.

**Ethical Considerations**

This project will emphasize fairness, transparency, and stakeholder relevance. Classification will be done at the plan level to avoid privacy concerns. The model will avoid attributing readiness to uncontrollable market forces and instead focuses on actionable plan traits. Ethical modeling choices will be documented throughout.