PROJECT DEFINE

PROBLEM STATEMENT

To identify which specific HR policy areas are most associated with employee turnover and propose targeted, data-driven policy improvements to reduce attrition risk.

DECISION MARKERS

For this portfolio project, I act as a data analyst simulating a recommendation to:

- HR leadership (e.g., CHRO or Head of People Strategy), responsible for setting and refining HR policies
- Business leaders (e.g., unit heads or managers), who implement HR practices at the team level
- Prospective employers reviewing this project to assess my ability to frame business-relevant questions, apply analytics, and derive actionable insights

CRITERIA FOR SUCCESSFUL EFFORT

- Identify which HR policy domains (e.g., compensation, career development, etc.) are statistically associated with higher attrition risk
- Quantify the relative impact of policy-related variables using interpretable models
- Deliver at least 3 practical HR policy recommendations backed by data
- Effectively communicate results via charts, summary briefings, and business language
- Demonstrate business acumen by tying insights to policy levers and organizational goals

KEY FORCES ACTING ON DECISION MARKERS

- High cost of turnover, including recruiting, training, and lost productivity
- Market pressure to retain skilled and high-performing employees
- Limited visibility into which policy levers influence retention
- Push to use data analytics for HR strategy and improve people-related KPIs
- Need for interpretable, evidence-backed recommendations to guide policy decisions

TIME FRAME FOR RESOLUTION

- As a portfolio project (part-time): 4–6 weeks
  - Week 1: Clarify problem, understand HR policy mapping, and clean data
  - Week 2–3: Exploratory data analysis and hypothesis testing by policy areas
  - Week 4: Modeling (logistic regression, decision tree) to identify drivers of attrition
  - Week 5: Translate results into policy-oriented recommendations
  - Week 6: Documentation, visualizations, and final presentation
- In a real-world business setting:
  - 2–3 months for data diagnosis and stakeholder feedback
  - 6–12 months for piloting and evaluating new or revised HR policies

BOUNDARIES

- Analysis is limited to structured, internal historical HR data (no external benchmarks or exit interview data)
- Project is diagnostic and exploratory — no actual HR policy will be implemented
- Assumes generalizability within IBM's historical employee context, not across industries
- Prioritizes clarity and interpretability of insights (e.g., decision trees, logistic regression) over complex models

ACCURACY NECESSARY

- For modeling: Balanced accuracy and interpretability, with ROC-AUC ≥ 0.75 if applicable
- For business actionability: Insights clear and credible enough to guide HR policy review
- Recommendations must be specific to policy decisions and feasible within organizational constraints