Take-Home Assessment: Customer Churn Prediction

Context

You're provided with a dataset containing customer and service-related attributes. Your goal is to build a model that predicts whether a customer is likely to churn.

The dataset reflects real-world data nuances. Your job is to explore, build, and reason through a baseline solution that could be deployed in a production setting.

This exercise is designed to evaluate your ability to:

- Work with data
- Build and evaluate a pragmatic model
- Think through deployment and monitoring
- Communicate your reasoning clearly

How to Approach

Please aim to spend **~90 minutes total** on this exercise.

For each part:

- Show your code or results.
- Include a brief explanation of your reasoning (1–3 sentences per decision).
- Focus on clarity and trade-offs — we care about *why* you chose your approach.

Tasks

Part A — Data Exploration & Feature Design

Explore the dataset and decide how best to prepare it for modeling.

Part B — Modeling

Build a baseline model to predict churn probability.

- Choose an approach that balances simplicity and performance.
- Report a few metrics that you believe are relevant for this task.
- Explain why those metrics are appropriate and how they reflect model quality.
- Describe one modeling decision you made and the trade-offs involved.

You don’t need to optimize for the best possible performance — what matters most is your reasoning and how you justify your choices

Part C — Deployment & Monitoring

Write a short answer covering:

- How would you deploy this model in a real-world system?
- What steps would you take to ensure consistency between training and serving?
- What would you monitor in production (both data and business metrics)?
- How would you decide when to retrain the model?

Part D — Serving

Build a minimal service that exposes your model.

Include a requirements.txt.
Keep it simple and explain any design choices (e.g., schema, threshold).

Deliverables

- notebook.ipynb **or** script.py
    - Contains all code for data exploration, feature engineering, modeling, evaluation, and segment analysis.
    - You can choose between a notebook (for interactivity and narrative) or a script.
- REPORT.md
    - A concise summary of key decisions:
        - Data preparation and feature choices
        - Modeling approach and metrics
        - Segment analysis insights
        - Deployment and monitoring plan
    - Focus is on rationale, not results.