Litle & Co Data Scientist Interview Problem

Background

An *authorization attempt* (or simply *authorization*) is a request by a merchant to hold funds on a customer's payment card for a purchase. If the authorization is declined by the bank that issued the card, the merchant may attempt another authorization (a practice called *recycling*). The merchant is charged a small fee for each authorization attempt.

A *billing series* consists of an *initial authorization* followed by zero or more recycle attempts such that all authorizations in a billing series are declined except possibly the final one. For the purposes of this problem, if a single customer has the approval history over time `001101000`, where `0` is declined and `1` is approved, we may group the customer's authorizations into four billing series as follows: `(001)(1)(01)(000)`.

Recycling can be effective, particularly in recurring and installment billing situations, in which the cardholder may be difficult to contact. As a result, the practice is widespread throughout the industry. However, recycling strategies are far from standard across merchants, and there is large variety in recommendations from payments industry advisors.

This problem asks you to investigate the science behind recycling, and to find *actionable* intelligence for the merchant to inform a recycling strategy.

Data

The file `data.txt` contains a number of authorization attempts submitted to Litle & Co. by an e-commerce merchant. The fields in the file are:

Field name	Definition
`Account_ID`	an identification number representing a single customer
`Transaction_Timestamp`	the time at which the authorization was submitted by the merchant
`Factor_A`	an unknown factor
`Factor_B`	an unknown factor
`Factor_C`	an unknown factor
`Factor_D`	an unknown factor
`Factor_E`	an unknown factor
`Response`	the card issuing bank's response to the authorization
`Transaction_Status`	whether the authorization was approved or declined

Problem statement

Your task is to use this data to recommend ways to improve recycling practices (e.g. to improve the approval rate, decrease the number of authorization attempts, reduce the time to approval). You should assume that the only practice that may be changed is *when* authorizations are retried (including which should be abandoned, if any). The goal is to find actionable recommendations for the merchant with measurable financial impact.

You are encouraged to present other insights you gain from the data over the course of your investigation. For example: Can the success of an authorization or a billing series be predicted based on the factors? Are there any trends over time in the data? What additional information might help you better advise the merchant? Are there any other recommendations you could make to the merchant about authorization practices based on this data?