ComfortDelgro Zig Data Assignment

Thanks for expressing your interest in joining the CDG Zig Data team! Please complete the following take home assignment. You will need to prepare a slide deck for a 10-minute presentation as part of the technical interview, covering your approach, insights, and model results. Please submit this deck as well as your code script / notebook the day before your interview.

Problem Statement

You are a data scientist working at a telecommunications service provider and have been provided with the following datasets to 1) develop a model to predict user churn and 2) identify insights and factors influencing churn and suggest possible initiatives or interventions to counter it. You may choose not to use all the datasets for your analysis.

Datasets

1. telco_train: This is the main training dataset. Msisdn is the unique key identifying a customer. Churn is the target variable and is defined by whether the customer has cancelled his service within three months from the reference date.
2. census: Data on household income groups by planning area of the people residing there.
3. cdr: Monthly aggregates of call data records per customer.
4. locations: Top 10 locations that each customer has visited in the month before the reference date.
5. web: Monthly aggregates of website browsing behavior by customer / website combination
6. reviews: Feedback by customers to the support / help portal
7. schools: Locations of schools