Capstone Project #1 Proposal By Lin Ma

1. Problem Description

Starting in 2014, the Affordable Care Act(ACA) requires everyone to have health insurance or pay a penalty. The Health Insurance Marketplace is a service that helps people shop for and enroll in affordable health insurance so as to be ACA compliant. The Health Insurance Marketplace puts all the healthcare insurance plans in one place and display the same basic information to consumers, which makes it easier to compare and choose the plans best fit consumers’ need.

While people try to decide which plan to choose, questions may rise on which state(s) has the best rate for plans with the same level of coverage (Metal Level), are there any relationships among the factors to the rate? Can we use age, tobacco, rating area and other factors to predict rate for individuals?

Answers to these questions will help individual consumers or small business owners to make better decision in choosing the healthcare insurance plans.

1. Data

The data I am using for this project is provided by The Centers for Medicare & Medicaid Services (CMS). The Health Insurance Exchange Public Use Files (Exchange PUFs) consist of ten separate files, and I will focus on the following three files:

* Rate PUF (Rate-PUF) – Plan-level data on individual rates based on an eligible subscriber’s age, tobacco use, and geographic location, and family-tier rates.
* Benefits and Cost Sharing PUF (BenCS-PUF) – Plan variant-level data on essential health benefits, coverage limits, and cost sharing.
* Plan Attributes PUF (Plan-PUF) – Plan-level data on maximum out of pocket payments, deductibles, HSA eligibility, formulary ID, and other plan attributes.

For more information about Exchange PUFs, please visit: https://www.cms.gov/cciio/resources/data-resources/marketplace-puf.html

1. Problem Resolution

I plan to marry the PUFs on planID, and then explore how the rates for the same coverage level (Metal level) vary by states. Then I plan to use machine learning to develop a model and predict individual rates from the factors I choose, such as age, tobacco, rating area, etc.

1. Project Deliverables

I will create a notebook, and put my code and reports in it.