

A Look at Regional Differences in Medicare Coverage Medicare PUF 2017 Dataset

Executive Summary

The purpose of this project is to use clustering to identify whether there are regional differences in Medicare coverage by calculating the average Medicare uncovered cost, as well as seeing how this relates to a medical provider's gender and place of service. Exploratory data analysis revealed that the categorical variables of states and provider types had too many categories to be easily grouped into an "Other" bin, thus states were grouped into census regions to be used for clustering. Distributions for service counts, Medicare costs, and Medicare payments were all severely skewed right, although this was mitigated by removing providers who were Medicare non-participants as well as standardizing and normalizing these numeric features. The metric to determine Medicare coverage was calculated by first finding the difference between the average allowable amount and the standard payment amount, then multiplying by the line service count, and then dividing by the number of unique beneficiaries per provider. Clustering was performed on the average uncovered amount, along with the provider place of service and the provider gender to answer the business question. Results indicate that highly uncovered clusters had a plurality of providers from the southern census region of the US, as well as being located in offices (non-hospital entities) and having male service providers, giving a basis for further investigation to see if and why regional inequalities for Medicare coverage may exist.