Who is buying dental insurance from the health exchanges?

Justin Vena
DS5

Adapting to the changes in the Healthcare Marketplace

ACA (Patient Protection and Affordable Care Act)

- Established in 2010, changed the healthcare industry in many different ways.
- Introduced the Health Insurance Marketplace (or called the health exchanges)
- Health and Dental Products became available for sale in 2013 and came into effect for 2014

Adapting to the changes in the Healthcare marketplace

Insurance Products on the Health Exchange:

- Before the Health Exchange was established, most people received their insurance from (1) their employer and (2) buying directly from the insurer.
- What types of consumers will buy from the exchanges?
- The exchanges make it easy for a consumer to shop around between insurance products.

How will we approach this?

- Let's gather insurance, Medicare, and Census information to solve this.

Project Question:

What characteristics can we successfully use to classify which ZIP codes are more likely to buy an ACA dental insurance product?

Where's the Data?

Kaggle.com Health Insurance Marketplace Dataset:

- Providing us with the Insurance and Medicare Data
- Contains data from 2014 to 2016 from the CMS (Center for Medicare Services)
- Has data on 35 States (No Washington State)
- Several different datasets are included on Kaggle, but will primarily use the Service Area Data (has exchange plan geographic information)

Where's the Data?

Census.gov's public data:

- 2014 ZIP Code/Income data set
- Geographical: State, ZIP
- Income: Counts by Income Bracket, Average Income

Data Exploration and Investigation

Kaggle Data Features:

- State
- Plan Standard: Low or High
- Geographic: State, County, ZIP
- Child/Adult Plan: If the plan is for a child and/or an adult
- Year: 2014-2016

Data Exploration and Investigation

Kaggle Data Clean-Up and Issues:

- Missing ZIP Codes is currently the biggest issue
- How to manipulate data to improve model
  - Exclude the missing ZIP codes from the model
  - Replace the missing ZIP with average income of all the ZIP codes offered per state
- Run using data from all years separately and see how it impacts the results
- Run with and without the Child/Adult Flag
- Run initial program with just Texas since has largest number of plans

Data Exploration and Investigation

Census Data:

- ZIP/Income information still needs to be cleaned.
- Will do a check of average using bracket counts versus the average given

Next Steps: Running the Model

Before running:

- Clean the Census Data
- Look at which way will be the best way to deal with missing data

How will we run the model?

- We want to segment plans into clusters that exhibit similar characteristics.
- First attempt will be to use K-Means Algorithm
- After testing if time, check using DBSCAN clustering as well

Other Considerations:

- Census databases have information about number of small businesses per ZIP Code; a known purchaser of exchange products