### Problem Statement

In order for an insurance company to make money, it needs to collect more in yearly premiums than it spends on medical care to its beneficiaries. As a result, insurers invest a great deal of time and money to develop models that accurately forecast medical expenses.

The goal of this analysis is to use patient data to estimate the average medical care expenses for such population segments. These estimates could be used to create actuarial tables which set the price of yearly premiums higher or lower depending on the expected treatment costs.

### Data

|  |  |
| --- | --- |
| **Column** | **Description** |
| age | Age of the primary beneficiary |
| sex | Gender of the beneficiary |
| bmi | Body mass index |
| children | Number of dependents covered by health insurance |
| smoker | yes/no |
| region | The beneficiary’s residential address in the US: northeast, southeast, southwest, northwest |
| charges | Medical costs billed by health insurance |