

# ASSIGNMENT 1 - TOPIC SELECTION

## GROUP- 7

### Tentative Project Title:

Variations in Healthcare Plan Costs and Benefits Across the United States

### Problem Statement:

There are many different plans with varying provider networks, coverage, and costs in the complicated U.S. health insurance market. The *cost-sharing system* is shown by the four metal levels into which plans are divided: Bronze, Silver, Gold, and Platinum. *Selecting the best plan* is made more difficult by the range of plan types, including Preferred Provider Organisations (PPOs) and Health Maintenance Organisations (HMOs).

Many customers find it difficult to choose an appropriate health plan because they lack an understanding of how premiums and benefits are impacted by variables including plan type, cost-sharing, and region. *Access and affordability* differ by state and region, which makes geographic discrepancies even more problematic. By *investigating* how these characteristics *affect rates*, this project aims to assist insurers in *improving their products*, customers in making well-informed decisions, and policymakers in gaining knowledge to *increase healthcare affordability*.

### Brief Description of the Data:

In this project, we use a publicly accessible dataset on health insurance policies in the US. Numerous healthcare plan characteristics, such as premium prices, coverage levels, and geographic availability, are covered in depth by the dataset. We would like to *find patterns and variations* in health insurance affordability by *examining this data* according to plan types, cost-sharing arrangements, and regional characteristics.

### Key Features of the Dataset:

**Data Source:** The dataset is from data.healthcare.gov and contains records of insurance plans offered in different states.

**Number of Records:** The dataset includes approximately 60,345 records, capturing extensive details about healthcare plans.

**Number of Variables:** There are 151 variables, though only a subset of key attributes is used in this study.

#### Key Variables:

##### 1. Geographic Data:

*State & County:* Indicates the area in which the health plan is available, aiding in the analysis of regional variations.

##### 2. Features of the Plans:

*Metal Level:* Sorts plans according to cost-sharing arrangements into Bronze, Silver, Gold, and Platinum.

*Plan Type*: Indicates if the plan falls under the PPO, HMO, or other various classifications.

### 3. Cost Elements:

*Premiums*: The amount that policyholders must pay each month.

*Deductibles*: The sum that customers must pay out of pocket before their insurance starts to pay for medical bills.

*Out-of-Pocket Maximums*: The maximum yearly payment a policyholder must make before their insurance pays for all their medical expenses.

### 4. Additional Cost-Sharing Information:

*Co-payments*: Consumers pay set sums for medical services, such as doctor visits.

*Co-insurance*: The portion of medical expenses that patients are required to pay after fulfilling their deductible.

## Why This Dataset?

- It provides thorough and ***current information about health insurance*** available in the United States.
- ***Geographic comparisons*** are possible using the dataset, which aids in locating differences in healthcare availability and pricing.
- It has sufficient variables to ***enable predictive modeling***, including the use of machine learning techniques to anticipate premium costs.
- This dataset provides the basis for ***examining the regional*** and plan-type variations in healthcare expenditures, which enables us to unearth ***important information for policymakers, insurers, and consumers.***

## Data Source Link:

The dataset is derived from two sources on [data.healthcare.gov](https://data.healthcare.gov):

[Dataset 1 Link](#)

[Dataset 2 Link](#)

The combined dataset, used for this project, is attached to this submission.

## Research Questions:

1. What are the differences in rates for various metal levels and plan types (PPO, HMO) for health insurance?

This question aims to examine ***how plan types (such as PPO and HMO) and cost-sharing arrangements*** (metal levels: Bronze, Silver, Gold, and Platinum) affect premium costs.

2. Do state-by-state variations exist in the availability and cost of healthcare plans?

Finding regional differences in plan affordability and availability helps in ***identifying regions with more expensive*** or more limited healthcare alternatives.

3. Based on the advantages, coverage options, and geographical characteristics of healthcare plans across the US, is it possible to forecast their premiums?

The objective is to ***create a prediction model*** that uses variables like plan benefits, type, and geographic location to forecast healthcare rates.

4. What policy suggestions may be made to enhance consumer decision-making and lessen geographic disparities in the affordability of healthcare plans?

This aims to produce ***useful information*** that government officials, insurers, and regulators may use to guarantee more ***fair healthcare costs***.

5. Is it possible to spot trends in healthcare plan acceptance depending on state laws and income levels?

The purpose of this analysis is to find out how ***state-specific regulations*** and income levels ***impact consumer decisions***.

6. What effects do provider network expansions (the number of hospitals and physicians in the network) have on premium costs for various plan types?

This question examines the ***impact*** of provider network differences on ***healthcare plan affordability***.