**Insurance Charges Prediction**

# Problem Statement:

Predict how much could be the insurance charges for a beneficiary based on the data provided using Linear Regression.

# Dataset Information:

|  |  |
| --- | --- |
| **Column** | **Description** |
| Age | Age of Primary Beneficiary |
| Sex | Insurance Contractor gender (female, male) |
| BMI | Body Mass Index, providing an understanding of body, weights that are relatively high or low relative to height, objective index of body weight (kg / m^2) using the ratio of height to weight, ideally 18.5 to 24.9 |
| Children | Number of children covered by health insurance / Number of dependents |
| Smoker | Whether the beneficiary smokes or not |
| Region | The beneficiary’s residential area in the US, northeast, southeast, southwest, northwest. |
| charges | Individual medical costs billed by health insurance (target variable) |

# Scope:

* Exploratory data analysis
* Data Pre-processing
* Training linear regression model with OLS method for prediction
* Tuning the model to improve the performance