**Problem Statement or Requirement:**

A client's requirement is, he wants to predict the insurance charges based on the several parameters. The Client has provided the dataset of the same.

As a data scientist, you must develop a model which will predict the insurance charges.

1. Identify your problem statement
    - Insurance changes to be predicted and create model based on age, sex, bmi, children and smoking habits.
2. Tell basic info about the dataset (Total number of rows, columns)
    - 1338 rows × 6 columns, in that two nominal data columns
3. Mention the pre-processing method if you're doing any (like converting string to number – nominal data)
    - Used get_dummies function to convert string to number. two columns converted (sex, smoke) to nominal data columns
4. Develop a good model with r2_score. You can use any machine learning algorithm; you can create many models. Have used following machine learning algorithm
    - MultiLinear
    - Support Vector Machine
    - Decision Tree
    - Random Forest
5. All the research values (r2_score of the models) should be documented. (You can make tabulation or screenshot of the results.)

| **Regression Type** | | | **R Score** |
|---|---|---|---|
| MultiLinear | | | 0.789479035 |
| Support Vector Machine | KernelType | | R Score |
| | rbf | | 0.390601625 |
| | linear | | 0.635950358 |
| | poly | | 0.750819225 |
| | sigmoid | | 0.527560378 |
| Decision Tree | criterion | splitter | R Score |
| | squared_error | best | 0.679640893 |
| | squared_error | random | 0.724727807 |
| | friedman_mse | best | 0.691758478 |
| | friedman_mse | random | 0.695108231 |
| | absolute_error | best | 0.659744801 |
| | absolute_error | random | 0.6900408 |
| | poisson | best | 0.735565995 |
| | poisson | random | 0.706167131 |
| Random Forest | criterion | splitter | R Score |
| | squared_error | | 0.855811989 |
| | absolute_error | | 0.854289766 |
| | friedman_mse | | 0.853340077 |
| | poisson | | 0.853100238 |

6. Mention your final model, justify why u have chosen the same
    - Choosing Random Forest method as close fit model.