1.) **Identify your problem statement** → The user wants Insurance Charge prediction.
   a) Stage 1 – Machine Learning
   b) Stage 2 – Supervised Learning
   c) Stage 3 - Regression

2.) **Tell basic info about the dataset (Total number of rows, columns)**
   a) 1338 rows and 6 columns

3.) **Mention the pre-processing method if you're doing any (like converting string to number – nominal data)**

   a) **Here the dataset we have 2 types of characters value I mean Sex (Male or Female) & Smoker (Yes or No)**
   b) **So, used Get_dummies function to do data preprocessing**

4.) Develop a good model with r2_score. You can use any machine learning algorithm; you can create many models. Finally, you have to come up with final model.

5.) All the research values (r2_score of the models) should be documented. (You can make tabulation or screenshot of the results.)

6.) Mention your final model, justify why u have chosen the same.

→ Please find the below developed models R2 score values and same has been published in Github also. Shared the same information my mentor as well.

Support Vector Machine → Parameters:

Support Vector Machine		
kernel	**gamma**	**R^2 Value**
linear	***scale***	***-0.01***
linear	***auto***	***-0.01***
poly	***scale***	***-0.08***
poly	***auto***	***-0.08***
rbf	***scale***	***-0.07***
rbf	***auto***	***-0.07***
sigmoid	***scale***	***-0.07***
sigmoid	***auto***	***-0.07***

Support Vector Machine		
kernel	**gamma**	**R^2 Value**
linear	***scale***	***-0.01***
poly	***scale***	***-0.08***
rbf	***scale***	***-0.07***
sigmoid	***scale***	***-0.07***
precomputed	***scale***	***Its not working becoz not matching metrix***

Decision Tree Parameters:

- ✓ Highlighted green color for which parameter have good model.

Decision Tree	
criterion	**R^2 Value**
squared_error	***0.69***
friedman_mse	***0.68***
absolute_error	***0.66***
poisson	***0.72***

Decision Tree		
criterion	**splitter**	**R^2 Value**
squared_error	***best***	***0.68***
squared_error	***random***	***0.72***
friedman_mse	***best***	***0.71***
friedman_mse	***random***	***0.71***
absolute_error	***best***	***0.69***
absolute_error	***random***	***0.72***
poisson	***best***	***0.72***
poisson	***random***	***0.69***

Decision Tree			
criterion	**splitter**	**max_features**	**R^2 Value**
squared_error	***best***	***sqrt***	***0.71***
squared_error	***best***	***log2***	***0.70***
squared_error	***random***	***sqrt***	***0.68***
squared_error	***random***	***log2***	***0.58***
friedman_mse	***best***	***sqrt***	***0.75***
friedman_mse	***best***	***log2***	***0.73***
friedman_mse	***random***	***sqrt***	***0.62***

friedman_mse	*random*	*log2*	*0.71*
absolute_error	*best*	*sqrt*	*0.74*
absolute_error	*best*	*log2*	*0.70*
absolute_error	*random*	*sqrt*	*0.71*
absolute_error	*random*	*log2*	*0.73*
poisson	*best*	*sqrt*	*0.64*
poisson	*best*	*log2*	*0.70*
poisson	*random*	*sqrt*	*0.63*
poisson	*random*	*log2*	*0.66*

Random Forest Parameters R2 Value:

- ✓ Highlighted Green color parameter is performed good model.

n_estimators=50	**random_state=0**	**0.84**

criterion	**R^2 Value**
squared_error	**0.85**
absolute_error	**0.85**
friedman_mse	**0.85**
poisson	**0.85**

criterion	**max_features**	**R^2 Value**
squared_error	*sqrt*	**0.86**
squared_error	*log2*	**0.87**
absolute_error	*sqrt*	**0.87**
absolute_error	*log2*	**0.87**
friedman_mse	*sqrt*	**0.86**
friedman_mse	*log2*	**0.87**
poisson	*sqrt*	**0.86**
poisson	*log2*	**0.86**