Multi_linear regressor:

r_score=-0.07564862243293402

Support Vector Regreesor

S.no	Kernel	Gamma	r_score
1.	rbf	scale	-0.08347748928652265
2.	poly	scale	-0.07564862243293402
3.	linear	scale	-0.00967358055216927
4.	sigmoid	scale	-0.07536003101256816
5.	precomputed	scale	Precomputed matrix must be a square matrix. Input is a 936x5 matrix.
6.	precomputed	auto	Precomputed matrix must be a square matrix. Input is a 936x5 matrix.
7.	sigmoid	auto	-0.07536003101256816
8.	linear	auto	-0.00967358055216927

9.	rbf	auto	`-0.08347748928652265`
10.	poly	auto	`-0.07564862243293402`

Support Vector Regressor,The highest value of r_score is in Column third and eight ie.`-0.0096735805521627.`

Decision tree regressor

S.no	Criterion	splitter	r_score
1.	`squared_error`	best	`0.6921595839380573`
2.	friedman_mse	best	`0.708632443267764`
3.	friedman_mse	random	`0.7182703839242428`
4.	absolute_error	random	`0.750256349727309`
5.	absolute_error	best	`0.6643278835143465`
6.	poisson	best	`0.7314436277252554`

7.	poisson	random	0.7285082092093387
8.	**squared_error**	random	0.7290561071205468

The highest value of r_score is in Column 4 ie.0.750256349727309.

Random Forest Regressor

S.no	Criterion	r_score
1.	squared_error	0.8505585405305971
2.	friedman_mse	0.8531559923531832
3.	absolute_error	0.8495393958770256
4.	poisson	0.8538954994965602

The highest value of r_score is in Column 4 ie.0.8538954994965602.

1.) Identify your problem statement.

The client wants to **predict insurance charges** based on various parameters such as age, bmi, smoking status, and factors. They have provided a dataset containing relevant features.

The client data falls under supervised learning since,input and output are clearly defined and requirement is clear. Under supervised learning, regression is used because it is based on numerical data.Thus, we need to predict insurance charges for values other than this.

2.) Tell basic info about the dataset (Total number of rows, columns)

the dataset contains **1,338 rows** and **6 columns**.

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

Yes,here a pre-processing method is used to convert sex (male, female) into numerical format using **One-Hot Encoding**.
Converted smoker (yes, no) into numerical values (1 for smokers, 0 for non-smokers).

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 the finals model.
The models used are,
1.Multiple linear regressor
2.SVR
3.Decision Tree
4.Random Forest
5.) All the research values (r2_score of the models) should be documented.
(You can make a tabulation or screenshot of the results.)

Various models tabulation are attached in page 1.

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

The final model is _Random Forest Regressor._ Since, it has highest r_score.