

1. Problem Statement - Predict Insurance Charges
2. Lets identify the 3 stages

Stage 1 - Domain Selection - As we have all the input and output in excel and as the output is numeric, we can go for Machine LEarning.

Stage 2 - As we have clear input and output, this will fall under Supervised Learning.

Stage 3 - Under Supervised LEarning, as the output to be predicted is numeric this will fall under Regression.

3. Preprocessing of data is needed as 2 of the columns are categorical and under categorical they fall under nominal, so we need to use 'One Hot Encoding' method to convert these columns into numeric.
4. As this is having more than one input other than Simple Linear Regression we can try all the other algorithms.

Algorithm	R2_SCore	Hyper Tuning Parameters	
Multiple Linear Regression	.78	Default	
SVM (Before Standardization)	-.11	kernel="linear"	
SVM (After Standardization)	-.01	kernel="linear"	
SVM (After Standardization)	-.08	kernel="rbf"	
SVM (After Standardization)	0.3	kernel=rbf,C=100	
SVM (After Standardization)	0.8	kernel=rbf,C=10000	
RandomForest	0.84	n_estimators = 50, random_state = 0	
RandomForest	0.87	n_estimators = 1000, max_depth = 10, max_features = "log2", bootstrap = "false"	

