### Specify the business problem

**Project Description:**

health Medical costs are one of the most common recurring expenses in a person's life. Based on different research studies, BMI, ageing, smoking, and other factors are all related to greater personal medical care costs. The estimates of the expenditures of health care related to obesity are needed to help create cost-effective obesity prevention strategies. Obesity prevention at a

Young age is a top concern in global health, clinical practice, and public.

* Medical expenses is one of the major recurring expenses in a human life. Its a common knowledge that one life style and various physical parameters dictates diseases or ailments one can have and these ailments dictates medical expanses. According various studies, major factors that contribute to higher expenses in personal medical care include smoking, aging, BMI. In this study, we aims to find a correlation between personal medical expenses and different factors, and compare them. Then we use the prominent attributes as predictors to predict medical expenses by creating linear regression models and comparing them using ANOVA. In research, we found that smoking, age and higher BMI have a high correlation with higher medical expenses indicating they are major factors in contributing to the charges and the regression can predict with more than 75% accuracy the charges.
* According to WHO, personal expenditure on medical andhealthcare has been increasing faster than the overall economyglobally[1]. This increase in expenditure has been attributedto many causes, major of which include smoking, ageing andincreased BMI. In this study, we aim to ﬁnd a correlation be-tween medical expenses and different factors using insurancedata of different people with attributes such as smoking, age,number of children, region and BMI.
* Here, we ﬁrst ﬁnd the correlation of medical charges witheach of the attributes and use these attributes to predictcharges. The method used is regression analysis, which predictive model to the data and then usethat model to predict an outcome variable from one or moreindependent predictor variables. We used multiple regressionto create different models andthenused ANOVA to comparethe different models and ﬁnd the best-ﬁt model.