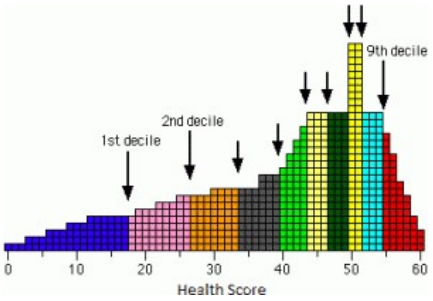
Descriptive Statistics	Descriptive statistics is comprised of those values which explains the spread and central tendency of data. For example, mean is a way to represent central tendency of the data, whereas IQR is a way to represent spread of the data.
Dependent Variable	A dependent variable is what you measure and which is affected by independent / input variable(s). It is called dependent because it "depends" on the independent variable. For example, let's say we want to predict the smoking habits of people. Then the person smokes "yes" or "no" is the dependent variable.
Decile	Decile divides a series into 10 equal parts. For any series, there are 10 decile denoted by D1, D2, D3 ... D10. These are known as First Decile , Second Decile and so on. For example, the diagram below shows the health score of a patient from range 0 to 60. Nine deciles split the patients into 10 groups 
Degree of Freedom	It is the number of variables that have the choice of having more than one arbitrary value. For example, in a sample of size 10 with mean 10, 9 values can be arbitrary but the 10th value is forced by the sample mean. So, we can choose any number for 9 values but the 10th value must be such that the mean is 10. So, the degree of freedom in this case will be 9.
Dimensionality Reduction	Dimensionality Reduction is the process of reducing the number of random variables under consideration by obtaining a set of principal variables. Dimension Reduction refers to the process of converting a set of data having vast dimensions into data with lesser dimensions ensuring that it conveys similar information concisely. Some of the benefits of dimensionality reduction: • It helps in data compressing and reducing the storage space required • It fastens the time required for performing same computations • It takes care of multicollinearity that improves the model performance. It removes redundant features • Reducing the dimensions of data to 2D or 3D may allow us to plot and visualize it precisely • It is helpful in noise removal also and as result of that we can improve the performance of models
Dplyr	Dplyr is a popular data manipulation package in R. It makes data manipulation, cleaning, summarizing very user friendly. Dplyr can work not only with the local datasets, but also with remote database tables, using exactly the same R code. It can be easily installed using the following code from the R console: `install.packages("dplyr")`
Dummy Variable	Dummy Variable is another name for Boolean variable. An example of dummy variable is that it takes value 0 or 1. 0 means value is true (i.e. age < 25) and 1 means value is false (i.e. age >= 25)