**Prepare support vector machines model for classifying the area under fire for foresfires data.**

**Inferences from the Data Set:**

Data Set talks about the forestfiresof the 517 observations with respect to around 30 Independent variables.

Measures of central tendency explains about mean, median and mode of the data set.

Measures of dispersion explains about the standard deviation variation and range of the data set.

Third moment business decision and Fourth moment business decision explains about the skewness and kurtosis of the data set.

Remove the variables **month & day** which has no effect on the forestfire dataset.

Divide the dataset into train and test dataset with normal splitting of the data.

Training a model on the data begin by training a simple linear SVM before that we have to install the packages called **kernlab**

**Method/ Building Model**

We use the kernlab package and the ksvm() function therein to fit an SVM using a non-linear kernel. We can use the argument kernel = "rbfdot" for a radial basis and " vanilladot" for the complicated sounding hyperbolic tangentsigmoid.

Note the hugh amount of parameter customisation that is possible at this stage. For simplicity we use the default settings which will be far from optimal.

Using the simple defaults, the radial basis non-linear mapping for the SVM appears equivalent to the vanilladot, based on the lower training error; with the model slightly better. We should evaulate the model performance using the **predict()** function. In order to examine how well our classifier performed we need to compare our predicted size of the fire with the actual size in the test dataset.

**Evaluating model performance and predictions on testing dataset:**

We are going to predict the model using the model forest data set and test data.

And by using the confusion matrix table we can predict the **size\_category** of the dataset.

forest\_fires\_predictions large small

large 23 1

small 0 43

We can get the accuracy of the model by using the proportion of the table , predictions of the dataset to the forest dataset of output variable**(size\_catagory)**

We get the 98% of accuracy of model by using **kernel = "vanilladot".**