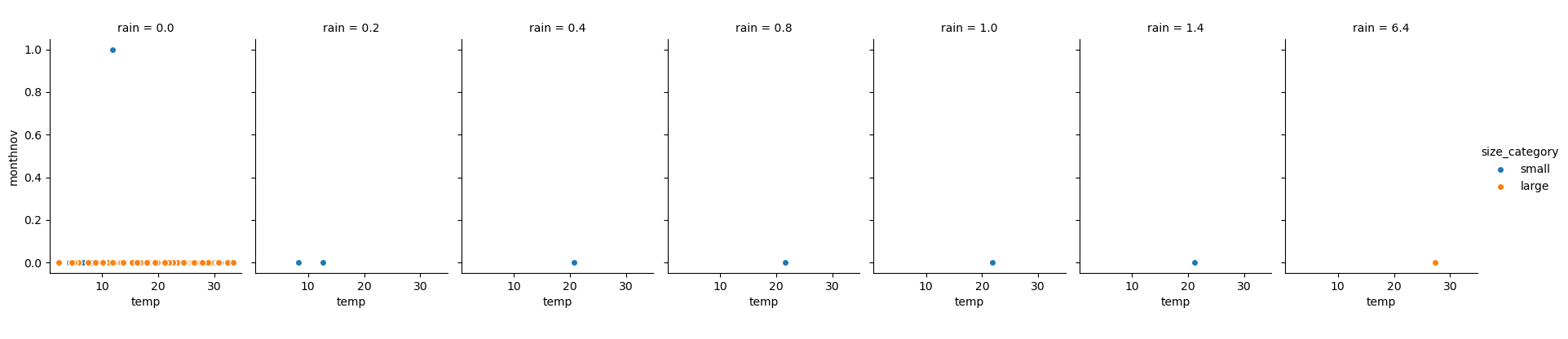
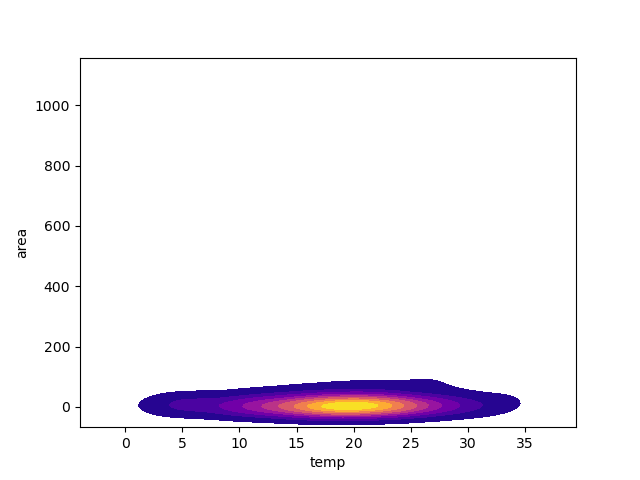
SVM INFERENCES

1. **Forest Fires**

****

The Forest Fire was spreading rapidly in month of November and heat was high when the Rain was 0.0 ( not raining) and temp dropped when the Rain pointer increased.



The Forest fire temp was was hot only upto 25 C. along with area of less then 200km .

Using SVM we deduced whether the area was small or large.

Response variable has 2 category :: small , large

Using Basic SVM with no kernel gave :

[[ 36 11]

[ 6 103]]

Classification report : 71 % overall

precision recall f1-score support

large 1.00 0.02 0.04 47

small 0.70 1.00 0.83 109

accuracy 0.71 156

macro avg 0.85 0.51 0.43 156

weighted avg 0.79 0.71 0.59 156

Using GridSearchCV :

This dict :

param\_grid = {'C': [0.1,1, 10, 100], 'gamma': [1,0.1,0.01,0.001]}

we got the matrix as

[[ 36 11]

[ 6 103]]

Much better result due to GridSearch CV of 89% acc

precision recall f1-score support

large 0.86 0.77 0.81 47

small 0.90 0.94 0.92 109

accuracy 0.89 156

macro avg 0.88 0.86 0.87 156

weighted avg 0.89 0.89 0.89 156

Using Linear Kernel we got the best result so far with result below.

[[ 45 2]

[ 0 109]]

precision recall f1-score support

large 1.00 0.96 0.98 47

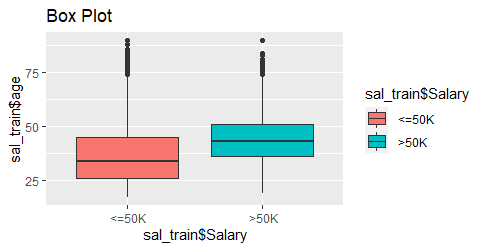
small 0.98 1.00 0.99 109

accuracy 0.99 156

macro avg 0.99 0.98 0.98 156

weighted avg 0.99 0.99 0.99 156

1. **Salary Data**

****

In training data the boxplot of Salary being in 2 category <=50k and >50k

w.r.t to Age .

There seems to many outliers in <=50k since many less Exp might be earning less revenue.



In the Test Data the proportion of <=50 are more and outliers present.

Using Normal SVm with no kernel methods. Model gave results below.

[[10997 363]

[ 2703 997]]

With 80 % accuracy.

With **Radial Bias Kernel** method. Evaluation was slightly worse

As we got 79% od acc