

---

Subject: clustering- yet again

Posted by [nivedita.raghunath](#) on Tue, 09 Oct 2007 21:05:42 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Hi All,

The clustering saga never ended!

I have a large dataset with 5 visibly distinct clusters. Based on advice from y'all, I tried standardize() and it worked to some extent. It finds 4 clusters. I have a good idea about my input variables- included the variable\_wts option and tried, yet with no success.

This is the number of samples it returns in each cluster:

```
IDL> print, clust_num
  6.00000
 28903.0
 25205.0
 37694.0
 42458.0
```

It put only 6 samples into cluster 0, whereas actually I have samples in the range 25000 or so. The remaining samples have been put into cluster 4 (and I can visually see that 0 and 4 are NOT close).

Any ideas on this?

My dataset is very huge to post here, but if anyone is willing to help, I can mail it.

Sincerely,  
Nivedita

---