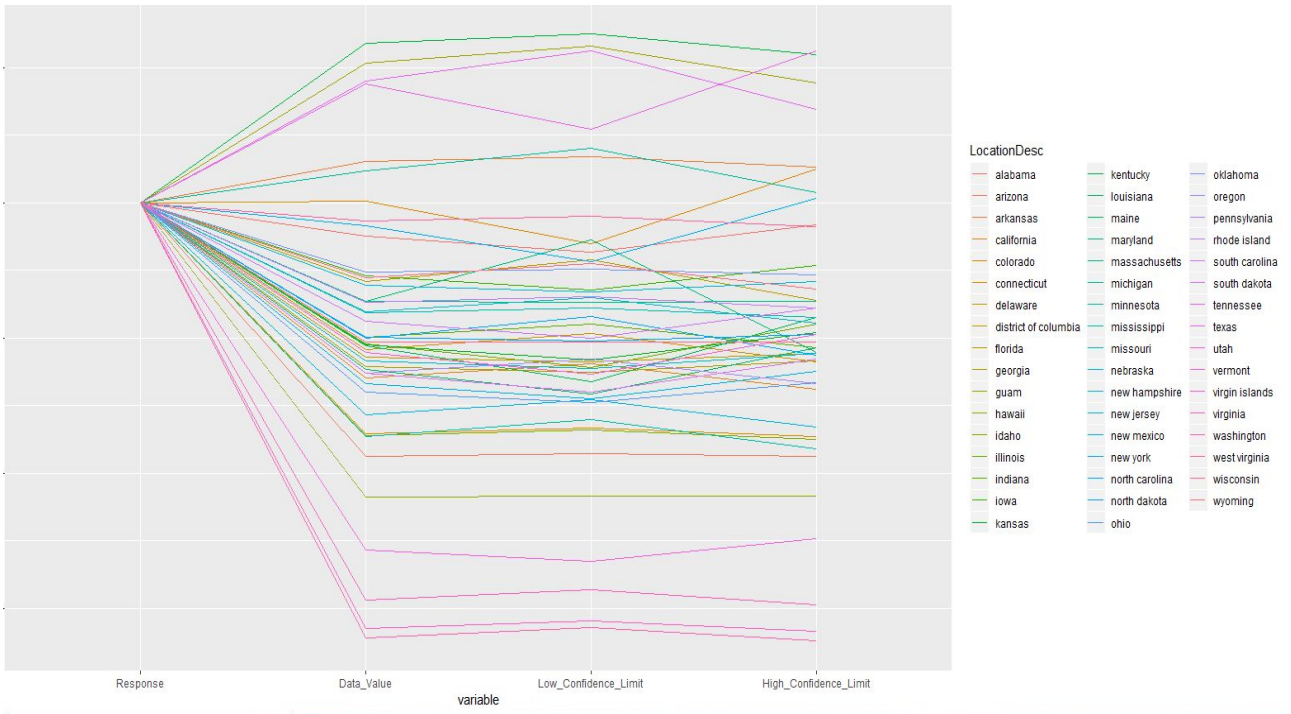
1. **I exported clean data as Tobacco_export.csv**
2. **The clean data bellow:**

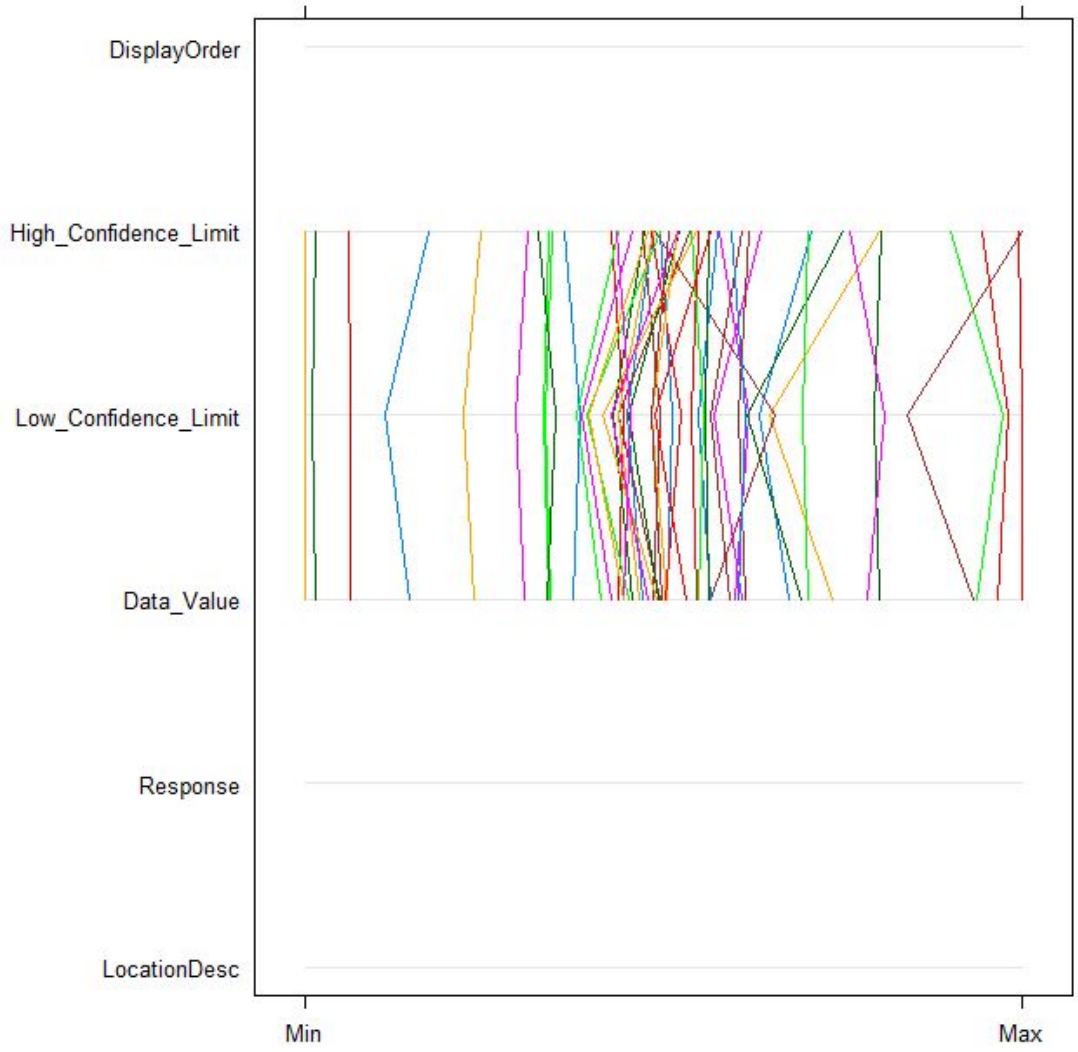
```
> TobaccoModifiedFinalAverage2[1:20,]
          LocationDesc Response Data_Value Low_Confidence_Limit High_Confidence_Limit DisplayOrder
1              alabama        2  15.959259             12.492593              19.425926            7
2              arizona        2   8.800000              6.613889              10.997222            7
3             arkansas        2  18.391667             15.275000              21.512500            7
4           california        2  11.333333              9.233333              13.441667            7
5             colorado        2  17.100000             12.750000              21.450000            7
6          connecticut        2   9.517647              7.341176              11.703922            7
7             delaware        2  12.264583             10.110417              14.414583            7
8  district of columbia        2  12.016667              9.300000              14.733333            7
9              florida        2  14.466667             12.266667              16.666667            7
10             georgia        2  11.717949              8.969231              14.469231            7
11                guam        2  21.566667             18.533333              24.550000            7
12              hawaii        2   7.471429              5.364286               9.566667            7
13               idaho        2   9.466667              7.283333              11.600000            7
14            illinois        2  12.461111              9.113889              15.813889            7
15             indiana        2  12.658974             10.389744              14.930769            7
16                iowa        2  14.666667             11.391667              17.938889            7
17              kansas        2  12.429167              9.350000              15.512500            7
18            kentucky        2  22.236364             18.878788              25.596970            7
19           louisiana        2  12.375000              8.713889              16.033333            7
20               maine        2  11.616667              8.333333              14.933333            7
>
```

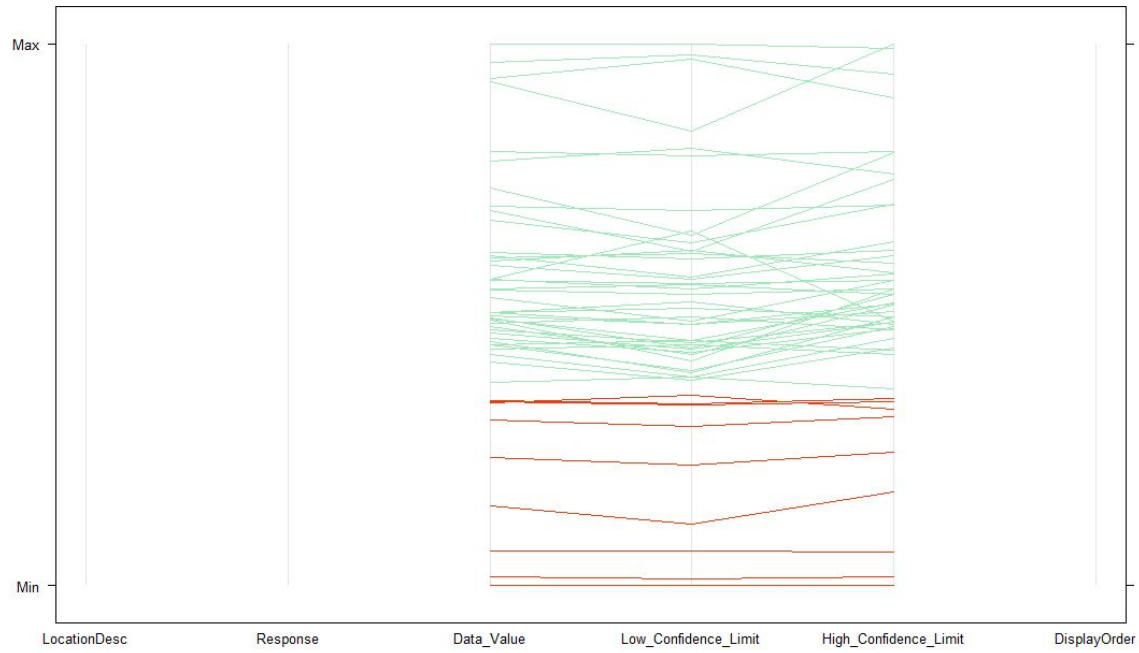
3. I used Parallel Coordinate Plot to plot the data for different states



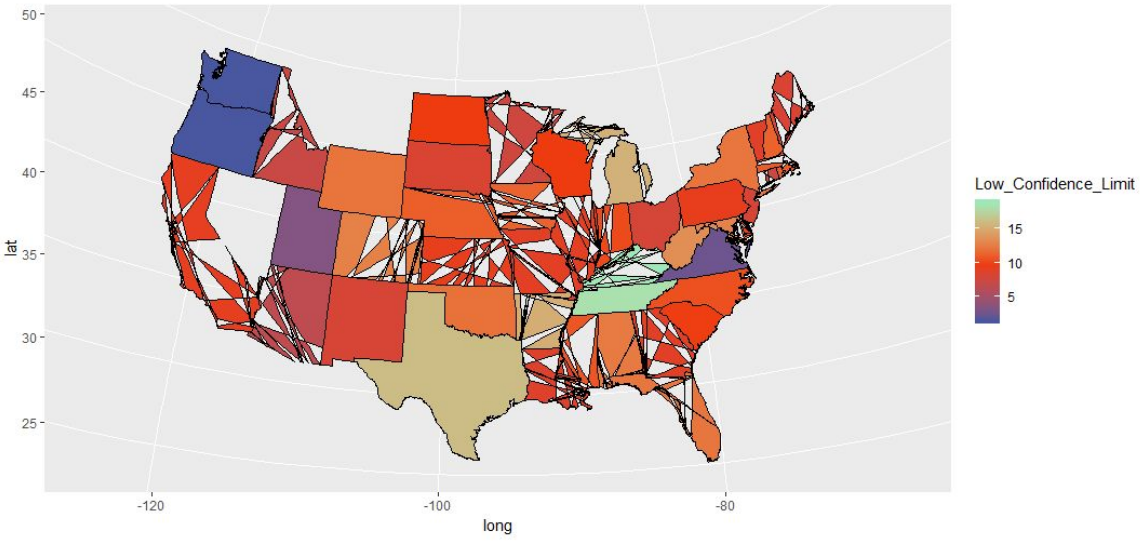
4. I used ‘clustering technique’ to cluster the data

parallelplot(TobaccoModifiedFinalAverage2)





5. I Used map to visualize the variation of a variable across the states



Conclusion:

For this Tobacco Survey, we can make relationship with each variables by parallel plot.

With parallel plot we can easily compare each state. Correlations can be observed as states are plotted on the chart. Each state corresponds to a line drawn through point on each axis corresponding to the value of the variable.

Yes, Clustering help in visualizing information. For example, In this information, if the Data_value is more than 10 then the color is greenish and if less than 10 then the color is reddish.