**Clustering with R:**

**Accessing Data**

This assignment is related to implementing clustering technique on sample stores data. A Grocery retailer with 515 stores spread across KA and TN wants to create store clusters based on mix of sales by category and average sales per square foot of space. The data needed for this is present on the server in the following location.

Use the store level data provided to you to create clusters using K-Means.

**Z:\Assignments\Class14 - Clustering with SAS**

File name needed for the assignment is **clustering\_data-class14.csv** which has 515 observations and 8 variables. Please consider the following pointers while drafting the required solution for this assignment.

* + - 1. You are free to decide on the number of clusters, but tell us why you decided to go with so many clusters. Remember to always keep the final cluster number to manageable levels so that real time solutions can be implemented based on the insights generated.
      2. Only numerical variables are to be used for building the clusters and be careful not to include any categorical variables in the clustering process. Dummy indicators can be used as substitute for categorical variables, but remember that these variables may not make much sense during the profiling process.
      3. Final deliverables required for this assignment end solution are SAS scripts used to create the clusters, and a PPT that includes details about how you decided upon the final number of clusters, profiling of the end clusters and recommendations observed if any.

***------------------------------End of Assignment Questions-----------------------------***

***Data Dictionary - clustering\_data-class14.csv (column names in sequential order)***

-*Store Num: Unique identifier for each store*

*-Cat 1 -4: Sales by category (in 000 Rs)*

*-Cat 1: Fresh foods category*

*-Cat 2: Frozen Foods*

*-Cat 3: Health and beauty*

*-Cat 4: Tobacco and alcohol*

*-Size: Area of store (in sq ft)*

*-Sale: Total sales of the store*

*-State: Location of store*