Projeto : Análise de Agrupamentos

Prof. Abraham Laredo Sicsu

1. Grupos de 2 a 5 participantes. **Trabalhos individuais não serão aceitos**
2. Data de entrega: **04 /12/2020 até às 23:38 no eclass**
3. ***Dados : MINIMARKET(anexado no eclass)***
4. Forma de entrega: o relatório em WORD deverá ser salvo como .pdf e postado no eclass. Utilize o ícone “envio de atividades” e faça upload do arquivo com o relatório em pdf na pasta correspondente. Cuidado para não trocar de pasta.
5. Não há limite para o número de páginas, mas

* Na capa devem constar os nomes dos participantes que e **contribuíram significativamente** para a elaboração do relatório, em ordem alfabética.
* Formatar com cuidado para que as tabelas não fiquem bagunçadas. Minha sugestão é que formatem o arquivo inteiro em **Lucida Console** que é compatível com a saída do R
* Não devem colocar os scripts
* Não imprimir a tabela de dados no relatório.

segue

1. O roteiro obrigatório para o exercício:

* Análise de cada variável **driver**. Não exagere colocando conclusões ou gráficos não ilustrativos. Descrevam como trataram *outliers e* *missing values* (se houver). Uma variável por página, por favor.
* A Matriz de Correlação das variáveis drivers quantitativas e conclusões;
* Clusters obtidos pelo método hierárquico aglomerativo:
* Caracterização e comparação dos clusters **via drivers** para a caracterização dos clusters
* Tentem diferenciá-los “em palavras” além de fazer a comparação estatística dos mesmos
* Verificar se há relação entre a variável vinícola e os clusters obtidos
* Clusters obtidos pelo método k-medoids (opte por k=2:6)
* Indique qual o número de cluster selecionado pela função pamk; difere do número que Você selecionou acima? Utilize como set.seed o ano de nascimento do mais velho da turma (informe no trabalho, por favor)
* Caracterização e comparação dos clusters **via drivers** para a caracterização dos clusters
* Tentem diferenciá-los “em palavras” além de fazer a comparação estatística dos mesmos
* Verificar se há relação entre a variável vinícola e os clusters obtidos
* Cruzar os resultados obtidos pelo dois tipos de clusters e comentar **. calcular o índice de rand**