MTH443: Quiz 1
Full Marks 20
Time: 17:10-18:00

Consider the state wise crime data of USA for the year 2019 (us_crime_data.csv). The data gives the recorded crimes in various states of USA corresponding to the following different type of crimes: (i) burglary, (ii) larceny, (iii) motor theft, (iv) assault, (v) murder, (vi) rape, (vii) robbery.

(a) Obtain a PCA based 3-dimensional projection of the data. PCA to be done using standardized data.
(b) Detect outliers, if any, from the PCA projection plot obtained in (a).
(c) Obtain scree plot and suggest the number of principal components for efficient data dimensional reduction.
(d) Find the proportion of total sample variation captured by the first 3 principal components.
(e) Find the sample correlation coefficient between the first PC and the variable "assault".
(f) Obtain complete linkage hierarchical cluster dendogram. Partition the states into 5 clusters using the constructed dendogram and list the states in each of the clusters.

Put your code, output/plot, interpretation of the output and answer to the questions in a single file and upload the MSWORD or pdf file through the Assessments-Exam: (Quiz 1) link on mooKIT.

Note: A maximum of 17 mins compensatory time to 231080042 as per GoI rules