Attached is the activity database for old age patients in Chinese hospital.

The variables used are

|  |  |
| --- | --- |
| ACTIVITY | activity classification |
| BP | Blood pressure |
| CIRCLUATION | Blood circulation |
| EEG | EEG monitoring rate |
| HR | Heart beat rate |
| SL | sugar level |
| TIME | monitoring time |

The dependent variable is ACTIVITY and it takes the following values

0- Standing 1- Walking 2- Sitting 3- Falling 4- Cramps 5- Running

Questions:

1) Perform variable exploration esp. with respect to the activity values – are there any outliers? How will you deal with these? – 15 marks

2) Which of the variables explain the activity best? Are there more than a few variables that will correlate to the activity in a combination? How can we best explain this – 15 marks

3) Between Heartrate and BP which gives the best cluster for activities? Explain your answer – 20 marks