Name : Naman Thaker
Roll No : 20BCE529
Subject : Data Mining
Practical 5

Problem Statement : Implement the Natural partitioning (apply 3-4-5 rule) algorithm for generating concept hierarchy (up to two levels) for the any data set values for example, student dataset (attribute 'marks') - 32, 38, 48, 91, 46, 37, 22, 69, 78, 82, 33, 49, 55, 66, 84, 86, 67, 80, 79, 44.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from math import log10, floor, ceil


def Range(sAr):
  return sAr[-1] - sAr[0]


def findNumParts(sAr):
  range_d = Range(sAr)
  no_of_partitions = 0
  msd = floor(range_d / (10**floor(log10(range_d))))
  if msd in [3,6,7,9]:
    no_of_partitions = 3
  elif msd in [2,4,8]:
    no_of_partitions = 4
  elif msd in [1,5,10]:
    no_of_partitions = 5
  print('Range:',range_d)
  print("Most Significant digit:",msd)
  print('No of partitions:',no_of_partitions)
  return [range_d,no_of_partitions]



def makePartitions(sAr):
  parts = []
  data_range,no_of_partitions = findNumParts(sAr)
  print(*np.array_split(sAr,no_of_partitions))
  num_elements = ceil(len(sAr)/no_of_partitions)
  fulls = int(len(sAr)/num_elements)
  for i in range(0,fulls):
    parts.append(sAr[i*num_elements:(i+1)*num_elements])
  if len(sAr[(i+1)*num_elements:]) != 0:
    parts.append(sAr[(i+1)*num_elements:])
  return parts
```

```
def printing(parts):
  print("Output partitions:")
  for li in parts:
    print(li)
  print('\n')

def intial(ar):
  ar.sort()
  print("sorted values:",ar)
  subans = makePartitions(ar)
  print("Answer after level 1 partioning")
  printing(subans)
  print("\n\nAnswer after level 2 partioning")
  for part in subans:
    sub_parts = makePartitions(part)
    printing(sub_parts)



marks = [32, 38, 48, 91, 46, 37, 22, 69, 78, 82, 33, 49, 55, 66, 84, 86, 67, 80, 79, 44]
intial(marks)
```

```
sorted values: [22, 32, 33, 37, 38, 44, 46, 48, 49, 55, 66, 67, 69, 78, 79, 80, 82, 8
Range: 69
Most Significant digit: 6
No of partitions: 3
[22 32 33 37 38 44 46] [48 49 55 66 67 69 78] [79 80 82 84 86 91]
Answer after level 1 partioning
Output partitions:
[22, 32, 33, 37, 38, 44, 46]
[48, 49, 55, 66, 67, 69, 78]
[79, 80, 82, 84, 86, 91]



Answer after level 2 partioning
Range: 24
Most Significant digit: 2
No of partitions: 4
[22 32] [33 37] [38 44] [46]
Output partitions:
[22, 32]
[33, 37]
[38, 44]
[46]


Range: 30
Most Significant digit: 3
No of partitions: 3
[48 49 55] [66 67] [69 78]
Output partitions:
[48, 49, 55]
[66, 67, 69]
[78]
```

```
Range: 12
Most Significant digit: 1
No of partitions: 5
[79 80] [82] [84] [86] [91]
Output partitions:
[79, 80]
[82, 84]
[86, 91]
```