Name: Abdelrahman Ahmed Hassan Mohamed **Sec : 3**

Exercise: You are given the following dataset representing students' scores in a Statistics exam:

Data = [45, 48, 52, 54, 56, 58, 60, 62, 65, 68, 70, 72, 75, 78, 80]

Discretize the data into 3 bins using the following two methods:

- Equal Width Binning
- Equal Frequency Binning **For each binning method:**

- **Sort the values in each bin in ascending order.**
- **Apply data smoothing using the following techniques:**
  - Smoothing by Mean
  - Smoothing by Median
  - Smoothing by Boundaries

Solution

Discretizing:

1. **Equal Width Binning: Range =** 80 – 45 = 35 ➔ 35/ 3 = 11.67 =12
   Bin 1: {45, 57} ➔ [45, 48, 52, 54, 56 ,58]
   Bin 2: {58, 70} ➔ [60, 62, 65, 68,70,72]
   Bin 3: {71, 83} ➔ [75, 78, 80]
2. **Equal Frequency Binning:** 15/3 = 5 ➔ **so five values per bin**
   Bin 1: [45, 48, 52, 54, 56]
   Bin 2: [58, 60, 62, 65, 68]
   Bin 3: [70, 72, 75, 78, 80]

Smoothing:

1. **By Mean: equal Frequency**
   - **Bin 1: (45 + 48 + 52 + 54 + 56) / 5 = 51 ➔ Bin 1: [51, 51, 51, 51, 51]**
   - **Bin 2: (58 + 60 + 62 + 65 + 68) / 5 = 62.6 ➔ Bin 2: [62.6, 62.6, 62.6, 62.6, 62.6]**
   - **Bin 3: (70 + 72 + 75 + 78 + 80) / 5 = 75 ➔ Bin 3: [75, 75, 75, 75, 75]**

Equal width:

- Bin 1 : (45+48+52+54+56+58)/6 ⇨ bin1 :[52,52,52,52,52,52]
- Bin 2 : (60+62+65+68+70+72)/6 ⇨ bin2 : [66,66,66,66,66,66]
- Bin 3: (75+78+80)/3 ⇨ bin 3 : [78,78,78]

2. **By Median: equal Frequency**
   - **Bin 1 median = 52 ➔ Bin 1: [52, 52, 52, 52, 52]**
   - **Bin 2 median = 62 ➔ Bin 1: [62, 62, 62, 62, 62]**
   - **Bin 3 median = 75 ➔ Bin 1: [75, 75, 75, 75, 75]**

Equal width:

Bin 1 median(52+45)/2 ⇨ bin 1:[53,53,53,53,53,53]

Bin 2 median(65+68)/2 ⇨ bin 2:[66,66,66,66,66,66]

Bin 3 median (78) ⇨ bin3:[78,78,78]

3. **By Boundaries: equal Frequency**
   - **Bin 1: min = 45, max = 56 ➔ Bin 1: [45, 45, 56, 56, 56]**
   - **Bin 2: min = 58, max = 68 ➔ Bin 1: [58, 58, 58, 68, 68]**
   - **Bin 3: min = 70 , max = 80 ➔ Bin 1: [70, 70, 80, 80, 80]**

Equal width:

Bin 1: min = 45, max = 58 ⇨ bin 1:[45,45,58,58,58,58]

Bin 2: min = 60, max = 72 ⇨ bin 2 :[60,60,60,72,72,72]

Bin 3: min = 75, max = 80 ⇨ bin 2 :[75,80,80]