ANGHELA M. ALIZA
BSIT3B

Frequency Distribution with Datasets

Title of Dataset: Test Scores of 100 Students

Step 1
Source of Data:

78, 85, 92, 88, 74, 69, 95, 81, 76, 84,90, 87, 79, 73, 82, 88, 91, 65, 80, 83,77, 86, 70, 89, 93, 75, 68, 72, 96, 94,85, 88, 90, 97, 71, 83, 79, 66, 92, 98,81, 84, 76, 69, 87, 95, 99, 73, 78, 89,82, 91, 85, 77, 68, 93, 100, 74, 70, 96,88, 79, 67, 80, 82, 90, 84, 97, 75, 64,86, 83, 92, 89, 73, 95, 78, 81, 87, 94,72, 99, 66, 91, 77, 98, 85, 69, 88, 100,93, 70, 96, 83, 75, 89, 84, 90, 76, 92

Step 2 (Determine the number of classes, class intervals, class limits, and class boundaries.)

Lowest Score	64
Highest Score	100
Range	(lowest − highest = range) 36
Number of Classes	K = 1+3.322log10(N) k=1+3.322log10(100) -> 1+3.322(2) -> 7.644 or 8 8 classes
Width	(range/number of classes= width) 4.5 or 5

Class interval:

Class No.	Class Interval (+5 starting to the lowest score) Class Limits (Lower – Upper) (eg. Lower(64) – upper(68)	Class Boundaries 1st (first upper class limit (68) – lower class limit(69) = 1) 2nd (1 / 2) = 0.5 3rd (minus 0.5 to all lower class and add 0.5 to all upper class)
1	**64 – 68**	63.5 – 68.5
2	69 – 73	68.5 – 73.5
3	74 – 78	73.5 – 78.5
4	79 – 83	78.5 – 83.5
5	84 – 88	83.5 – 88.5
6	89 – 93	88.5 – 93.5
7	94 – 98	93.5 – 98.5
8	99 – 103	98.5 – 103.5

Frequency Distribution Table:

Class No.	Class Interval (+5 starting to the lowest score)	Class Boundaries	Class mark	Frequency (f)	Cumulative Frequency (Less Than)	Cumulative Frequency (Greater than)	Relative Frequency (f ÷ total) × 100
1	64 – 68	63.5 – 68.5	66	7	7	100	7%
2	69 – 73	68.5 – 73.5	71	12	19	93	12%
3	74 – 78	73.5 – 78.5	76	14	33	81	14%
4	79 – 83	78.5 – 83.5	81	15	48	67	15%
5	84 – 88	83.5 – 88.5	86	18	66	52	18%
6	89 – 93	88.5 – 93.5	91	18	84	34	18%
7	94 – 98	93.5 – 98.5	96	12	96	16	12%
8	99 – 103	98.5 – 103.5	101	4	100	4	4%

ANGHELA M. ALIZA
BSIT3B

Step 3**:**

Analysis and Interpretation

1. **What is the range of your dataset?**

   The range of the dataset is 36, calculated from the highest score of 100 minus the lowest score of 64. This shows that the scores vary moderately across the dataset.

2. **How many classes did you decide to use, and why?**

   i decided to use 8 classes with a class width of 5 to organize the 100 scores. I use the sturge's formula and This number of classes provides a balance between clarity and detail, allowing patterns to be easily observed.

3. **What patterns or trends can you observe from the distribution?**

   The distribution shows that most scores are clustered between 79 and 93, while fewer scores occur at the lower and higher extremes. This indicates that the dataset is somewhat concentrated in the middle range.

4. **What does the cumulative frequency reveal about your data?**

   The cumulative frequency reveals how many scores fall below or above certain values, helping identify medians, quartiles, and percentiles. For example, 48 scores are below 84, while 34 scores are above 93.

5. **How can this frequency distribution summarize and simplify your dataset?**

   The frequency distribution shows how scores accumulate across the ranges, allowing to see how many students scored below or above certain thresholds. the frequency distribution simplifies the dataset by grouping individual scores into meaningful classes, making trends, patterns, and comparisons easier to understand at a glance.

Most of the scores fall in the middle range (79–93), which suggests that most students performed at a moderate to good level. There are fewer scores at the very low (64–73) and very high (99–100) ends, showing that only a small number of students struggled or excelled. This tells us that overall, the group's performance is fairly consistent.

Looking at the frequency distribution and cumulative frequencies, we can easily see how the scores are spread out. It highlights that most students scored around the middle, with fewer at the extremes. and the distribution gives a clear picture of the data, making it easier to spot trends and understand how the group performed as a whole.