**DAY -2:**

**Statistics:**

It is the science of collecting and organizing data.

**Types of Statistics:**

1. **Descriptive stats:** it consists of organizing and summarizing the data.
2. **Inferential stats:** used where the data we have measured to form an output.

**Measure of central tendency:**

**Mean:** Avg of the values.

**Median:** Middle value of ordered data( asc || dec)  
**Mode:** frequency(no. of occurences) of the values.

**Measure of dispersion:**

it represents the degree of spread in the data it indicates how much the data point differ from mean.

**Formula – ((x)-mean)^2)/N**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **X** | **mean** | **X-mean** | **(x-mean)^2** | **(x-mean)^2/N** |
| **1** | **2.83** | **-1.83** | **3.3489** | **0.55815** |
| **2** | **2.83** | **-0.83** | **0.6889** | **0.11481** |
| **2** | **2.83** | **-0.83** | **0.6889** | **0.11481** |
| **3** | **2.83** | **+0.17** | **0.0289** | **0.00481** |
| **4** | **2.83** | **+1.17** | **0.0289** | **0.00481** |
| **5** | **2.83** | **+2.17** | **4.7089** | **0.78481** |

**Five Number Summary:**

1. **Minimum**
2. **1st Quartile** = 25/100\*(no of values+1)
3. **Middle** = Median
4. **3rd Quartile** = 75/100\*(no of values+1)
5. **Maximum**

**Example -**

**X** = [1, 2, 2, 3, 3, 5, 5,5, 6, 6 ,6, 6, 7, 8, 8, 9, **27**]

**Min** = 1

**Q1 i**s the 25percentile of the data: 3(index)

**Median** = 5

**Q3** is the 75percentile of the data: 7(index)

**Max** = 27(outlier) || 9 (after removing the outlier)

**Box Plot**

