Tutorial -3

(Data Science-17B11CI611)

-----------------------------------------------------------------------------------------------------------------------------------------

1. Classify the following attributes as binary, discrete, or continuous. Also classify them as qualitative (nominal or ordinal) or quantitative (interval or ratio). Some cases may have more than one interpretation, so briefly indicate your reasoning if you think there may be some ambiguity. Example: Age in years. Answer: Discrete, quantitative, ratio
2. Time in terms of AM or PM. *Binary, qualitative, ordinal*
3. Brightness as measured by a light meter. *Continuous, quantitative, ratio*
4. Brightness as measured by people’s judgments. *Discrete, qualitative, ordinal*
5. Angles as measured in degrees between 0*◦* and 360*◦*. *Continuous, quantitative, ratio*
6. Bronze, Silver, and Gold medals as awarded at the Olympics. *Discrete, qualitative, ordinal*
7. Height above sea level. *Continuous, quantitative, interval/ratio (depends on whether sea level is regarded as an arbitrary origin)*
8. Number of patients in a hospital. *Discrete, quantitative, ratio*
9. ISBN numbers for books. (Look up the format on the Web.) *Discrete, qualitative, nominal (ISBN numbers do have order information, though)*
10. Ability to pass light in terms of the following values: *opaque, translucent, transparent. Discrete, qualitative, ordinal*
11. Military rank. *Discrete, qualitative, ordinal*
12. Distance from the center of campus. *Continuous, quantitative, interval/ratio (depends)*
13. For grouped data, we cannot find the exact Mean, Median and Mode, we can only give estimates.

* ***To estimate the Mean use the midpoints of the class intervals***:

Estimated Mean = Sum of (Midpoint × Frequency)/Sum of Frequency

* ***To estimate the Median use:***

Estimated Median = L + ((n/2) – B)/G × w

Where:

* L is the lower class boundary of the group containing the median
* n is the total number of data
* B is the cumulative frequency of the groups before the median group
* G is the frequency of the median group
* w is the group width
* ***To estimate the Mode use:***

Estimated Mode = L + ((fm − fm-1)/ (fm − fm-1) + (fm − fm+1)) × w

Where:

* L is the lower class boundary of the modal group
* fm-1 is the frequency of the group before the modal group
* fm is the frequency of the modal group
* fm+1 is the frequency of the group after the modal group
* w is the group width

Problem 1: Let’s take example of grouped data given below:

|  |  |
| --- | --- |
| Seconds | Frequency |
| 51 - 55 | 2 |
| 56 - 60 | 7 |
| 61 - 65 | 8 |
| 66 - 70 | 4 |

Compute the estimated mean, median and mode of the group data.

**Solution:**

**For estimated Mean**

|  |  |  |
| --- | --- | --- |
| **Midpoint x** | **Frequency f** | **Midpoint × Frequency f.x** |
| 53 | 2 | 106 |
| 58 | 7 | 406 |
| 63 | 8 | 504 |
| 68 | 4 | 272 |
| Totals: | **21** | **1288** |

And then our **estimate** of the mean time to complete the race is:

**Estimated Mean** = *288/***21** = **61.333...**

**For Median:**

L = 60.5 n = 21 B = 2 + 7 = 9 G = 8 w = 5

Estimated Median = 60.5 + ((21/2) – 9)/8 × 5

= 60.5 + 0.9375

= 61.4375

**For Mode:**

L = 60.5 fm-1 = 7 fm = 8 fm+1 = 4 w = 5

Estimated Mode = 60.5 + ((8 – 7)/(8 − 7) + (8 − 4)) × 5

= 60.5 + (1/5) × 5

= 61.5