Quiz 1 Statistics for Data Science, 4th Semester Course, RVCE Bangalore Marks 10

1. The average income distribution in an organization of 100 employees is typically
   (a) Left skewed
   (b) Right skewed
   (c) Normal
   (d) Uniform

2. The distribution of scores in a laboratory exam in a class of 100 students in an engineering college is typically
   (a) Left skewed
   (b) Right skewed
   (c) Normal
   (d) Uniform

3. In the first test of Statistics for Data Science course for 4th semester students, Prof A finds that the class average is 12 on 50. On the request of students, Prof A adds 9 marks to each of the students. The new class average is
   (a) The same as the old class average
   (b) Increased by 9 marks
   (c) Increased by 4.5 marks
   (d) Increased by 3 marks

4. In the first test of Statistics for Data Science course for 4th semester students, Prof A finds that the class average is 12 on 50. On the request of students, Prof A adds 9 marks to each of the students. The standard deviation of the new scores is
   (a) The same as the old standard deviation
   (b) Increased by 9 marks
   (c) Increased by 3 marks
   (d) Increased by 4.5 marks

5. A pair of six-faced dice is rolled. We define the following events:
   A: The sum of outcomes of both the dice is 1
   B: The sum of outcomes of both the dice is 13.
Which of the following options is **TRUE** about events A and B?
   (a) They are disjoint
   (b) They are both impossible

(c) They are independent
(d) Only (a) and (b)
(e) All three options (a), (b) and (c) are true.

6. Identify the quantity that reflects skewing the most.
(a) Mean
(b) Median
(c) Mode
(d) None of the above

7. You have rolled a fair, six-faced die seven times and in all the seven times you get the same outcome, say 5. You decide to roll the die for the eighth time. The outcome of the eighth trial
(a) is 5 with probability 1.
(b) Is 5 with probability 1/6.
(c) is 5 with probability 7/8
(d) Is 5 with probability 1/279936 (*Hint:* $6^7 = 279936$)

8. A boxplot is given below.

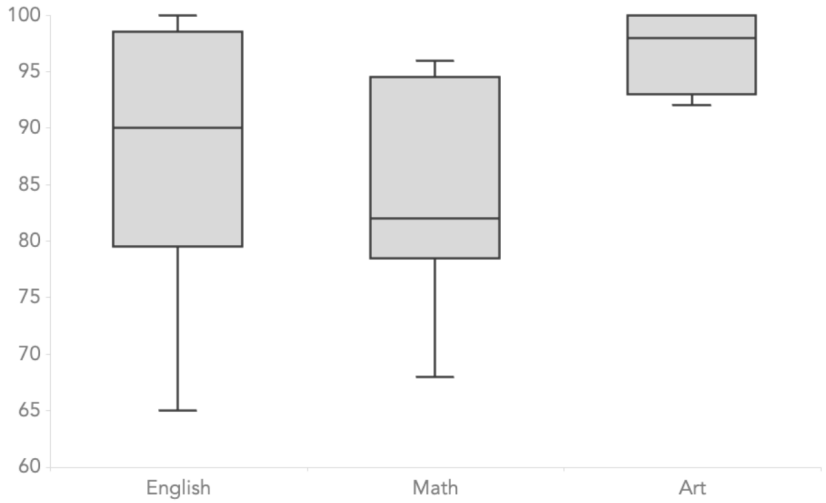

Figure: Question Number 8.

Which of the following cannot be inferred from the boxplot above.
(a) The median scores in each subject
(b) The score range in each subject
(c) The number of students enrolled in each subject
(d) The distribution of marks in each subject

9. Which of the following quantities is the **MOST** affected by outliers?
(a) Mean

(b) Variance
(c) Mean absolute deviation
(d) Median absolute deviation

10. Which of the following is the correct sequence? For this question assume that the data is distributed and is not a constant value.
(a) Median Absolute Deviation > Mean Absolute Deviation > Standard Deviation
(b) Mean Absolute Deviation > Standard deviation > Median absolute deviation
(c) Standard deviation > Median absolute deviation > Mean absolute deviation
(d) Standard deviation > Mean Absolute Deviation > Median Absolute Deviation