1. D
2. A
3. A
4. C
5. A
6. A
7. C
8. B
9. B

10. Boxplot provides us mean, quartiles, median. While histogram only gives us the count.

11. Metrics are selected based on the effect theyre having on the outcome. The higher the apparent effect, higher the need to include that metric in further process. E.g. when predicting salary of an employee. Age and experience are directly proportional to it, so they'll get selected. At the same time, gender may be excluded as it wont directly affect the salary.

12. To assess statistical significance, we would use hypothesis testing. The null hypothesis and alternate hypothesis would be stated first. Next, we have to calculate the p-value, which is the likelihood of getting the tests observed findings if the null hypothesis is true. Then we would select the threshold of significance (alpha) and reject the null hypothesis if the p-value is smaller than the alpha.

13. The datasets that are categorical wont have gaussian distribution. Predicting outcome of coin flip, pulling a card from deck of cards, these outcomes wont have Gaussian or log normal distribution.

14. Datasets where theres lot of outliers, in such cases median would be preferred. Like per capita income.

15. Likelihood and Probability have same meaning but are used in different context. Probability is simply based on the number of possible outcomes, while Likelihood takes past results into consideration where probability doesn't.