Statistics worksheet 6

Q1 d

Q2 a

Q3 a

Q4 c

Q5 a

Q6 b

Q7 c

Q8 b

Q9 b

Q10

Histogram is kind of bar graph that shows a bar for a range of data values instead of a single values

Where as

Boxplot provide some information that histogram does not

Min, max, 25th percentile, median(50th percentile),75th percentile and IQR(interquartile range of the data 75th percentile- 25th percentile) and explicitly separates the points that are considered outliers.

Q11

SELECT MERICS using

Classification: this algorithm will predict data type from defined data arrays: for example yes or no

Regression: the algorithm predicts some values

Ranking: based on orders of items.

Q12
Statistical significance is often calculated with statistical hypothesis testing.

Q13
Exponential distributions do not have a log-normal distribution or a Gaussian distribution. In fact, any type of data that is categorical will not have these distributions as well.

Q14
Income is the classic example of when to use the median instead of the mean because its distribution tends to be skewed.

Q15
Likelihood is a strange concept in that it is not a probability but is proportional to a probability.