1. What is the backbone of data science?

   Data

2. Calculate mean, median, mode, variance and standard deviation for column A

A	B	C	D
7	5	2003	2003
6	8	1976	1976
7	5	2001	2002
7	5	1915	1970
8	5	2000	2000
5	5	1993	1995
8	5	2004	2005
7	6	1973	1973
7	5	1931	1950
5	6	1939	1950
5	5	1965	1965

   Mean: 6.546, Median: 6, Variance: 1.157, S.D: 1.075

3. If the Standard deviation of a particular column is 5, what does it mean?

   It means that there is no much deviation in the data and the deivation from mean is 5.

4. What are the effects of Multi Colinearity?

   Data suffers redundancy, since this redundant data is used variance may differ hence affecting the accuracy of the model.

5. What are the measures of central tendency?

   Measures of central tendency: Mean, Median and Mode.

   Measures of how data is spread: Variance, S.D

6. What are the two different types of populations?

Finite and Infinite Population.

7. Which metric is suitable for analyzing skewed data?

Median

8. What is the difference between correlation and covariance?

Covariance is nothing but measuring the Variance between two variables and Correlation is the normalized form of Covariance.

Covariance:

$$E(XY) - E(X)E(Y)$$

Correlation:

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

9. What is correlation range?

-1 to +1

10. How do you select or reject null hypothesis?

Based on probability value(p-value), if p-value<0.05 then we reject null hypothesis, if p-value>0.05 then we accept it.