

# \* STATISTICS

## INTERVIEW QNAs \*

-BY CLOUDYML

\* What is Mean?

→ Mean is the average of a collection of values. We can calculate the mean by dividing the sum of all observations by the number of observations.

\* What are quantitative and qualitative data?

- 
- Quantitative data is also known as numeric data.
  - Qualitative data is also known as categorical data.



\* What is correlation?

→ Correlation gives the indication of how strong is the relationship between two variables. Ranges from  $-1$  to  $+1$ .

Correlation	Value	Variable 1	Variable 2
Negative	$-1$	↑	↓
Positive	$+1$	↑	↑
Zero	$0$	↑/↓	No effect

\* What is exploratory data analysis?

→ Exploratory data analysis is the process of performing investigations on data to understand the data better using graphs and plots.



Page \_\_\_\_\_  
Date \_\_\_\_/\_\_\_\_/\_\_\_\_

\* How is missing data handled in statistics?

- Prediction of the missing values
- Assignment of unique values
- Deletion of rows having missing data
- Mean or median imputation
- Using random forests, which supports missing values.

\* What is an outlier? How can outliers be determined in a dataset?

- Outliers are data points that vary in a large way when compared to other observations in the dataset. Outliers can be determined using the below methods.
  - Standard deviation/z-score
  - Interquartile range (IQR)





What is the meaning of standard deviation?



Standard deviation represents the magnitude of how far the data points are from the mean. A low value indicates that the data is close to the mean and vice versa.



What is skewness?



Skewness measures the lack of symmetry in a data distribution. It indicates that there are significant differences between the mean, the mode and the median of the data.





What is an alternative hypothesis?



The alternative hypothesis (denoted by  $H_1$ ) is the statement that must be true if the null hypothesis is false. It is the opposite point of view that gets proven right when the null hypothesis is proven wrong.



What are some of the properties of a normal distribution?



Unimodal: has only one mode



Symmetrical: left and right halves of the curve are mirrored.



Central tendency: the mean, median and mode are at the midpoint.



\* What is the law of large numbers in statistics?

→ The law of large numbers is a theory that states that the increase in the number of trials performed will cause a positive proportional increase in the average of results becoming the expected value.

\* What is the meaning of the five number summary?

→ The five number summary is a measure of five entities that cover the entire range of data as shown below.

- LOW EXTREME (MIN)
- FIRST QUARTILE (Q1)
- MEDIAN
- UPPER QUARTILE (Q3)
- HIGH EXTREME (MAX)





# **MAKE YOUR CAREER IN** **DATA SCIENCE AND** **ANALYTICS**

Python, SQL, Statistics, Machine Learning, Deep Learning, AWS Sagemaker, Model Deployment, Tableau, PowerBi, Excel, Case studies, Interview QnAs PDF, Capstone End-to-End Projects of ML and DL

**{Click Here to Visit Our Website}**



**Mr. Akash Raj**

**3+ Yrs Experienced Data Scientist**  
**Founder & CEO: CloudyML**



**35K+**



**31K+**



**44K+**