

➤ **IMPORTANT TOPICS TO BE LEARNT FOR DATA SCIENCE / TO CRACK INTERVIEWS
(COVERAGE OF ALL INTERVIEW QUESTION ASKED) :**

- **MACHINE LEARNING**
- **SUPERVISED LEARNING**
- **UNSUPERVISED LEARNING**
- **SUPERVISED LEARNING REGRESSION**
- **SUPERVISED LEARNING CLASSIFICATION**
- **GENERALISED FORM OF LINEAR REGRESSION**
- **REGRESSION MODEL EVALUATION**
- **ERROR PARAMETERS**
- **ACCURACY PARAMETERS**
- **R-SQUARE**
- **ORDINARY LEAST SQUARE**
- **ONE-HOT-ENCODING**
- **TRAIN-TEST-SPLIT**
- **VALIDATION METHOD**
- **K-FOLD VALIDATION**
- **VARIANCE INFLATION FACTOR**
- **DIFFERENCE BETWEEN R-SQUARE AND ADJUSTED R-SQUARE**
- **ASSUMPTION OF LINEAR REGRESSION**
- **MULTI-COLLINEARITY**
- **SLC-LOGISTIC REGRESSION**
- **LOGIT FUNCTION**
- **SIGMOID FUNCTION**
- **MAXIMUM LIKELIHOOD ESTIMATE**
- **MODEL EVALUATION FOR LOGISTIC REGRESSION**
- **CONFUSION MATRIX**
- **RECALL/SENSITIVITY/TPR**
- **PRECISION**
- **F-MEASURE**
- **SPECIFICITY/TNR**
- **AUC-ROC CURVE**
- **ASSUMPTION FOR LOGISTIC REGRESSION**
- **SCALING**
- **STANDARDISATION**

- **LABEL ENCODING**
- **K-NEAREST NEIGHBOURS**
- **DECISION TREE**
- **RANDOM FOREST**
- **NAÏVE BAYE'S CLASSIFIER**
- **GRID SEARCH METHOD**
- **CORRELATION COEFFICIENT**
- **SKEWNESS**
- **TRANSFORMATION**
- **GENERALIZATION IN ML**
- **OVERFITTING IN ML**
- **PRUNING**
- **UNDERFITTING IN ML**
- **BIAS ERROR**
- **VARIANCE ERROR**
- **TRADE-OFF**
- **UNSUPERVISED LEARNING**
- **CLUSTERING**
- **K-MEANS CLUSTERING**
- **HIERARCHIAL CLUSTERING**
- **PRINCIPAL COMPONENT ANALYSIS**
- **ENSEMBLE METHODS**
- **BAGGING**
- **BOOSTING**
- **GRADIENT BOOSTING**
- **XGBOOST**
- **REGULARIZATION**
- **RIDGE REGRESSION**
- **LASSO REGRESSION**
- **STATISTICS**
- **TYPES OF STATISTICS**
- **DESCRIPTIVE STATISTICS**
- **INFERENCE STATISTICS**
- **DESCRIPTIVE STATISTICS TYPES**
- **MEASURE OF CENTRAL TENDENCE**
- **MEAN/AVERAGE**
- **MEDIAN**

- **MODE**
- **MEASURE OF SPREAD/DISPERSION**
- **STANDARD DEVIATION**
- **TWO STANDARD DEVIATION**
- **SAMPLE STANDARD DEVIATION**
- **POPULATION STANDARD DEVIATION**
- **VARIANCE**
- **MEAN DEVIATION**
- **RANGE**
- **PERCENTILE**
- **QUARTILES**
- **SKEWNESS**
- **KURTOSIS**
- **CORRELATION**
- **CONFIDENCE INTERVAL**
- **NORMAL DISTRIBUTION**
- **CENTRAL LIMIT THEOREM**
- **HYPOTHESIS TESTING**
- **LEVEL OF SIGNIFICANCE**
- **P-VALUE**
- **ONE TAILED TEST**
- **TWO TAILED TEST**
- **TYPE 1 ERROR**
- **TYPE 2 ERROR**
- **CRITICAL VALUE**
- **Z-TEST**
- **T-TEST**
- **ANOVA**
- **CHI-SQUARE**
- **PROBABILITY**
- **BAYE'S THEOREM**
- **RANDOM VARIABLE**
- **TYPES OF PROBABILITY DISTRIBUTION**
- **BINOMIAL DISTRIBUTION**
- **NORMAL DISTRIBUTION**
- **POISSON'S DISTRIBUTION**
- **BASICS OF PROBABILITY**

- **BASICS OF PYTHON**
- **STRINGS**
- **LISTS**
- **TUPLES**
- **DICTIONARIES**
- **CONDITIONAL STATEMENT**
- **FOR LOOP**
- **WHILE LOOP**
- **FUNCTIONS**
- **LAMBDA FUNCTION**
- **MAP FUNCTION**
- **FILTER FUNCTION**
- **REDUCE FUNCTION**
- **SETS**
- **NUMPY**
- **INTRODUCTION TO SQL**
- **SQL SELECT**
- **SQL SELECT DISTINCT**
- **SQL WHERE**
- **SQL AND,OR,NOT**
- **SQL ORDER BY**
- **SQL INSERT INTO**
- **SQL NULL VALUES**
- **SQL UPDATE**
- **SQL DELETE**
- **SQL SELECT TOP/LIMIT**
- **SQL MIN AND MAX**
- **SQL COUNT,AVG,SUM**
- **SQL LIKE**
- **SQL WILDCARDS**
- **SQL IN**
- **SQL BETWEEN**
- **SQL ALIASES**
- **SQL JOINS**
- **SQL INNER JOINS**
- **SQL LEFT JOIN**
- **SQL RIGHT JOIN**

- **SQL FULL JOIN**
- **SQL SELF JOIN**
- **SQL UNION**
- **SQL GROUP BY**
- **SQL HAVING**
- **SQL NULL FUNCTIONS/IF NULL**
- **SQL CASE**
- **SQL EXISTS**
- **SQL ANY,ALL**
- **SQL SELECT INTO**
- **SQL STORED PROCEDURE**
- **SQL COMMENTS**
- **TEXT MINING**
- **TIME SERIES AND FORECASTING**

DATA ANALYST - Pratik Sonar