𝗕𝗮𝘀𝗶𝗰𝘀  
- Univariate statistics - mean, median, mode  
- Standard deviation and variance  
- Covariance and correlation  
- Population and sample  
- Nominal, ordinal and continuous, discrete data types  
- Outlines  
- The Simpson’s Paradox  
- Selection Bias  
  
𝗛𝘆𝗽𝗼𝘁𝗵𝗲𝘀𝗶𝘀 𝗧𝗲𝘀𝘁𝗶𝗻𝗴  
- Hypothesis Statements  
- Z-Test  
- T-Test  
- T-Test for sample means  
- T-Test for proportions  
- Paired and unpaired T-Tests  
- Variance test  
- ANOVA  
- Chi-Squared test  
- Goodness of Fit test for categorical data  
- Nominal, ordinal and continuous, discrete data types  
- Pairwise tests  
- T-Test assumptions  
- Non-parametric tests  
- Type 1 & 2 Errors  
  
𝗣𝗿𝗼𝗯𝗮𝗯𝗶𝗹𝗶𝘁𝘆 & 𝗗𝗶𝘀𝘁𝗿𝗶𝗯𝘂𝘁𝗶𝗼𝗻𝘀  
- The Bayes Theorem  
- Conditional probability  
- Normal distribution  
- Uniform distribution  
- Bernoulli distribution  
- Binomial distribution  
- Geometric distribution  
- Poisson distribution  
- Exponential distribution  
- Deriving the mean and variance of distributions  
- Central Limit Theorem  
- The Birthday problem  
- Card probability problems  
- Die roll problems  
  
𝗥𝗲𝗴𝗿𝗲𝘀𝘀𝗶𝗼𝗻 𝗠𝗼𝗱𝗲𝗹𝗶𝗻𝗴  
- OLS regression  
- Confidence vs prediction intervals  
- Logistic regression  
- Regression model assumptions  
- Model diagnostic checks  
- R-Square vs R-Square Adjusted  
- AIC, BIC, CP Statistics  
- Model Interpretation

STATISTICS  
- WHAT IS THE CENTRAL LIMIT THEOREM AND WHY IS IT IMPORTANT?  
- WHAT IS SAMPLING? HOW MANY SAMPLING METHODS DO YOU KNOW?  
- WHAT IS THE DIFFERENCE BETWEEN TYPE I VS TYPE II ERROR?  
- WHAT IS LINEAR REGRESSION? WHAT DO THE TERMS P-VALUE, COEFFICIENT, AND R-SQUARED       VALUE MEAN? WHAT IS THE SIGNIFICANCE OF EACH OF THESE COMPONENTS?  
- WHAT ARE THE ASSUMPTIONS REQUIRED FOR LINEAR REGRESSION?  
- WHAT IS A STATISTICAL INTERACTION?  
- WHAT IS SELECTION BIAS?  
- WHAT IS AN EXAMPLE OF A DATA SET WITH A NON-GAUSSIAN DISTRIBUTION  
  
MACHINE LEARNING  
- WHAT IS BIAS-VARIANCE TRADE-OFF?  
WHAT IS A CONFUSION MATRIX?  
- WHAT IS THE DIFFERENCE BETWEEN “LONG” AND “WIDE” FORMAT DATA?  
- WHAT DO YOU UNDERSTAND BY THE TERM NORMAL DISTRIBUTION?  
-WHAT IS CORRELATION AND COVARIANCE IN STATISTICS?  
- WHAT IS THE DIFFERENCE BETWEEN POINT ESTIMATES AND CONFIDENCE INTERVAL?  
- WHAT IS THE GOAL OF A/B TESTING?  
- WHAT IS P-VALUE?  
- IN ANY 15-MINUTE INTERVAL, THERE IS A 20% PROBABILITY THAT YOU WILL SEE AT LEAST ONE SHOOTING STAR. WHAT IS THE PROBABILITY THAT YOU SEE AT LEAST ONE SHOOTING STAR IN THE PERIOD OF AN HOUR?  
- HOW CAN YOU GENERATE A RANDOM NUMBER BETWEEN 1 – 7 WITH ONLY A DIE?  
- A CERTAIN COUPLE TELLS YOU THAT THEY HAVE TWO CHILDREN, AT LEAST ONE OF WHICH IS A GIRL. WHAT IS THE PROBABILITY THAT THEY HAVE TWO GIRLS?  
- A JAR HAS 1000 COINS, OF WHICH 999 ARE FAIR AND 1 IS DOUBLE HEADED. PICK A COIN AT RANDOM AND TOSS IT 10 TIMES. GIVEN THAT YOU SEE 10 HEADS, WHAT IS THE PROBABILITY THAT THE NEXT TOSS OF THAT COIN IS ALSO A HEAD?  
- WHAT DO YOU UNDERSTAND BY STATISTICAL POWER OF SENSITIVITY AND HOW DO YOU CALCULATE IT?  
-WHY IS RE-SAMPLING DONE?  
-WHAT ARE THE DIFFERENCES BETWEEN OVER-FITTING AND UNDER-FITTING?  
-HOW TO COMBAT OVERFITTING AND UNDERFITTING?  
-WHAT IS REGULARIZATION? WHY IS IT USEFUL?  
-WHAT IS THE LAW OF LARGE NUMBERS?  
-WHAT ARE CONFOUNDING VARIABLES?  
-WHAT ARE THE TYPES OF BIASES THAT CAN OCCUR DURING SAMPLING?  
-WHAT IS SURVIVORSHIP BIAS?  
-WHAT IS SELECTION BIAS? WHAT IS UNDER COVERAGE BIAS?  
-EXPLAIN HOW A ROC CURVE WORKS?  
-WHAT IS TF/IDF VECTORIZATION?  
-WHY WE GENERALLY USE SOFT-MAX (OR SIGMOID) NON-LINEARITY FUNCTION AS LAST OPERATION IN-NETWORK? WHY  
RELU IN AN INNER LAYER?  
- WHAT IS PCA? WHEN DO YOU USE IT?  
- EXPLAIN SVM ALGORITHM IN DETAIL  
- WHAT ARE THE SUPPORT VECTORS IN SVM?  
- WHAT ARE THE DIFFERENT KERNELS IN SVM?  
- WHAT ARE THE MOST KNOWN ENSEMBLE ALGORITHMS?  
- EXPLAIN DECISION TREE ALGORITHM IN DETAIL  
- WHAT ARE ENTROPY AND INFORMATION GAIN IN DECISION TREE ALGORITHM? Gini Impurity and Information Gain – CART, Entropy and Information Gain – ID3