What is the role of statistics in data science?:

Statistics is important in data science for analyzing and interpreting data. Data science involves collecting, processing, and analyzing large datasets to extract insights. Statistics helps to summarize data, make generalizations, test hypotheses, and build predictive models, enabling data scientists to make informed decisions.

What are the mean, median, and mode? - if their order is (Mean > median > mode), so is this data right/ left skewed or normal distribution?:

Mean, median, and mode are measures of central tendency in statistics. The mean is calculated by adding up all values and dividing by the number of observations, while the median is the middle value in a dataset when arranged in order. The mode is the most frequent value. The order of mean, median, and mode indicates skewness in the distribution, with right-skewed distributions having mean > median > mode. The normal distribution has equal mean, median, and mode.

What are outliers? - How do they affect data?:

Outliers are data points that are significantly different from other observations in a dataset, caused by errors or unusual observations. Outliers can affect data by causing skewness, changing measures of central tendency, and increasing variability. Identifying and handling outliers is important in data analysis using visualizations such as box plots or scatter plots. Outliers can be removed or adjusted, but caution must be taken as they can provide valuable information.

What is the probability of tossing a fair coin 10 times and getting heads 9 times? this video might help you:

The probability of getting heads on a single toss of a fair coin is 0.5 or 1/2. The probability of getting tails on a single toss of a fair coin is also 0.5 or 1/2.

The probability of getting exactly 9 heads in 10 tosses can be calculated using the binomial distribution formula:

P(X = k) = (n choose k) \* p^k \* (1-p)^(n-k)

where

n = number of trials (10 tosses)

k = number of successes (9 heads)

p = probability of success on a single trial (0.5)

So, the probability of getting exactly 9 heads in 10 tosses is:

P(X = 9) = (10 choose 9) \* (0.5)^9 \* (0.5)^(10-9)

P(X = 9) = 10 \* (0.5)^10

P(X = 9) = 0.009765625 or approximately 0.01

Therefore, the probability of tossing a fair coin 10 times and getting 9 heads is approximately 0.01 or 1%.

What is the difference between AI and Data Science? (Data Science and AI often come together in many scenarios but are they the same thing?):

Data Science and Artificial Intelligence (AI) are related but distinct fields. Data Science involves extracting insights and knowledge from data, while AI is focused on building intelligent systems that can perform human-like tasks. Data Science provides the data and methods for building AI systems, but AI can also be used outside of Data Science, such as in robotics and decision-making.

Discuss Software 1.0 vs. Software 2.0 and how is machine learning related to software 2.0?:

Software 1.0 is traditional programming where code explicitly tells computers what to do. Software 2.0 is a new approach using machine learning that allows computers to learn from data and adapt to new situations. Machine learning is the key to Software 2.0 as it enables computers to learn and make decisions on their own. This approach can create more powerful and effective software applications that can adapt to changing circumstances in real-time.

How does data science improve business in general? Can you give examples? (why would a business company hire a data science team, what are the business stakeholders expecting?):

Data science can help businesses in various ways. Firstly, it can improve decision-making by providing insights into customer behavior, market trends, and operational inefficiencies. Secondly, it can reduce costs by identifying areas of waste and inefficiency in operations. Thirdly, it can personalize products and services to meet the needs of individual customers. Fourthly, it can detect and prevent fraud by analyzing transaction data. Fifthly, it can help businesses make informed predictions about future events and trends through predictive analytics. To achieve these goals, businesses may hire a data science team to work closely with stakeholders to understand their business needs and develop customized solutions.

Comment on the saying “data is the new oil”.:

"Data is the new oil" metaphor emphasizes the importance of data as a valuable resource. Extracting, processing, and refining data can give companies a competitive edge. However, the metaphor has limitations, as data is not a finite resource and is subject to privacy laws. Additionally, different types of data have varying levels of value. Data is a complex resource that requires careful management and stewardship.