1. **Explain the differences between a bar chart, a pie chart, and a histogram. In which scenarios would you use each type of chart?**
2. **Discuss the advantages and disadvantages of using scatter plots to visualize relationships between two variables. Provide an example scenario where a scatter plot would be effective.**
3. **What are the key elements of an effective data visualization? Explain how each element contributes to clear communication of data insights.**
4. **Define and differentiate between mean, median, and mode. How do outliers affect each of these measures of central tendency?**
5. **What is standard deviation, and why is it important in understanding the variability of a dataset? Provide an example to illustrate your explanation.**
6. **What is a box plot, and how does it help in visualizing the distribution of a dataset? Explain the significance of each component of a box plot.**
7. **Define the concept of probability. How is probability applied in data science and analytics? Provide a real-world example.**
8. **Explain the difference between a uniform distribution and a binomial distribution. Provide an example of each.**
9. **What is hypothesis testing, and what are the key steps involved in conducting a hypothesis test? Provide an example of a null hypothesis and an alternative hypothesis.**
10. **Explain the Central Limit Theorem and its importance in inferential statistics. How does this theorem help in making inferences about population parameters from sample data?**
11. **Describe the difference between Type I and Type II errors in hypothesis testing. What are the consequences of each type of error?**
12. **Explain the difference between random sampling and stratified sampling. When would you choose one sampling method over the other in data analysis?**